Advances in Quantitative Analysis of Finance and Accounting

Volume 5

Editor

Cheng-Few Lee

Advances in Quantitative Analysis of

Finance and Accounting

Volume 5

Advances in Quantitative Analysis of
Finance and Accounting

Volume 5

Editor

Cheng-Few Lee

Rutgers University, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ADVANCES IN QUANTITATIVE ANALYSIS OF FINANCE AND ACCOUNTING
Advances in Quantitative Analysis of Finance and Accounting — Vol. 5

ISBN-13 978-981-270-628-7
ISBN-10 981-270-628-3

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore.

Preface

Advances in Quantitative Analysis of Finance and Accounting is an annual publication designed to disseminate developments in the quantitative analysis of finance and accounting. The publication is a forum for statistical and quantitative analyses of issues in finance and accounting as well as applications of quantitative methods to problems in financial management, financial accounting, and business management. The objective is to promote interaction between academic research in finance and accounting and applied research in the financial community and the accounting profession.

The chapters in this volume cover a wide range of topics including security analysis and mutual fund management, option pricing theory and application, interest rate spread, and electricity pricing.

In this volume there are 15 chapters, 9 of them focus on security analysis and mutual fund management: 1. *Testing of Nonstationarities in the Unit Circle, Long Memory Processes and Day of the Week Effects in Financial Data*; 2. *Equity Restructuring Via Tracking Stocks: Is there any Value Added?* 3. *Do Profit Warnings Convey Information About the Industry?* 4. *Are Whisper Forecasts more Informative than Consensus Analysts' Forecasts?* 5. *Earnings Forecast-Based Returns Predictions: Risk Proxies in Disguise?* 6. *The Long-Run Performance of Firms that Issue Tracking Stocks*; 7. *The September Phenomenon of U.S. Equity Market*; 8. *Identifying Major Shocks in Market Volatility and their Impact on Popular Trading Strategies*; 9. *Performance of Canadian Mutual Funds and Investors*.

Three of other six chapters are related to option pricing theory and application: 1. *The Least Cost Super Replicating Portfolio for Shot Puts and Calls in the Boyle-Vorst Model with Transaction Costs*; 2. *Stock Option Exercises and Discretionary Disclosure*; 3. *On Simple Binomial Approximations for Two Variable Functions in Finance Applications*. Two of other three chapters are related to interest rate spread: 1. *The Prime Rate-Deposit Rate Spread and Macroeconomic Shocks*; 2. *Differences in Underpricing Returns Between Reit Ipos and Industrial Company Ipos*. The remaining one chapter is related to electricity pricing: *Fundamental Drivers of Electricity Prices in the Pacific Northwest*.

This page intentionally left blank

Contents

List of Contributors

Chapter 1

Guan-Yu Chen
Department of Mathematics
Cornell University
Ithaca, NY 14853-4201, USA

Ken Palmer
Department of Mathematics
National Taiwan University
Taipei, Taiwan

Yuan-Chung Sheu
Department of Applied Mathematics
National Chiao Tung University
Hsinchu, Taiwan
Email: sheu@math.nctu.edu.tw

Chapter 2

Guglielmo Maria Caporale
Centre for Empirical Finance
Brunel University
Uxbridge, Middlesex
UB8 3PH, UK
Tel.: +44 (0)1895 266713
Fax: +44 (0)1895269770
Email: guglielmo-maria.caporale@brunel.ac.uk

Luis A. Gil-Alana
Universidad de Navarra
Faculty of Economics
Edificio Biblioteca, Entrada Este
E-31080
Pamplona, Spain
Tel.: +34 948 425 625
Fax: +34 948 425 626
Email: alana@unav.es

Mike Nazarski
Universidad de Navarra
Faculty of Economics
Edificio Biblioteca, Entrada Este
E-31080
Pamplona, Spain
Tel.: +34 948 425 625
Fax: +34 948 425 626
Email: mnazar@sgh.waw.pl

Chapter 3

Beni Lauterbach
School of Business Administration
Bar-Ilan University,
Ramat Gan 52900, Israel
Fax: 972-3-7384040
Email: lauteb@mail.biu.ac.il.

Joseph Vu
Department of Finance
DePaul University
1 E. Jackson Blvd.
Chicago, IL 60604-2287, USA
Fax: 312-362-6566
E-mail: jvu@depaul.edu

Chapter 4

Wei Zhang
School of Business
University at Alabany
1400 Washington Avenue
Alabany, NY 12222
USA

Steven F. Cahan
University of Auckland Business School
Private Bag 92019
Auckland, New Zealand
Tel.: (649) 373-7599 Ext. 87175
Email: s.cahan@auckland.ac.nz

Chapter 5

Dave Jackson
Department of Finance
College of Business
University of Texas-Pan American
1201 West University Drive
Edinburg, Texas 78541-2999, USA
Tel.: (956) 292-7317
Email: dojackson@utpa.edu

Jeff Madura
Department of Finance and Real Estate
Florida Atlantic University
220 SE 2nd ave.,
Fort Lauderdale, FL 33431, USA
Tel.: (561) 297-2607
Email: jeffmadura@bellsouth.net

Judith Swisher
Department of Finance and Commercial Law
Haworth College of Business
Western Michigan University
49008 5420, USA
Tel.: (269) 387-4457
Email: judy.swisher@wmich.edu

Chapter 6

Erik Devos
Department of Finance
College of Business
234 Copeland Hall
Ohio University
Athens, Ohio 45701, USA
Tel.: (740) 593-0819
Email: devos@ohio.edu

Yiuman Tse
Department of Finance
College of Business
University of Texas at San Antonio
501 West Durango Blvd.
San Antonio, Texas 78207, USA
Tel.: 210-458-2503
Fax: 210-458-2515
Email: yiuman.tse@utsa.edu

Chapter 7

Le(Emily) Xu
Whittmore School of Business and Economics
University of New Hampshire
Durham, NH 03824, USA
Tel.: (603) 862-3318
Fax: (603) 862-3383
Email: emily.xu@unh.edu

Chapter 8

Hemantha S. B. Herath
Department of Accounting
Brock University
Faculty of Business, Taro Hall 240
500 Glenridge Ave, St. Catharines,
Ontario, Canada L2S 3A1
Tel.: (905) 688-5550 Ext. 3519
Email: hemantha.herath@brocku.ca

Pranesh Kumar
College of Science and Management
University of Northern British Columbia
3333 University Way, Prince George,
British Columbia, Canada V2N 4Z9
Tel.: (250) 960-6671
E-mail: kumarp@unbc.ca

Chapter 9

Bradley T. Ewing
Rawls College of Business
Texas Tech University
Lubbock, TX 79409-2101, USA
Tel.: (806)742-3939
Email: bewing@ba.ttu.edu

Jamie Brown Kruse
Department of Economics
East Carolina University
Greenville, NC 27858, USA
Tel.: (252) 328-4165

Chapter 10

Charmen Loh
Rider University
2083 Lawrenceville Rd.
Lawrenceville, NJ 08648, USA
Tel.: (609) 895-5508 (office)
Fax: (507) 226-4669
Email: loh@rider.edu

Chapter 11

William Dimovski
School of Accounting, Economics and Finance
Deakin University
Geelong, Victoria 3217, Australia
Tel.: (613) 5227-2560
Fax: (613) 5227-2151
Email: wd@deakin.edu.au

Robert Brooks
Department of Econometrics and Business Statistics
Monash University
PO Box 197
Caulfield East
Victoria 3145, Australia
Tel.: (613) 9903-2178
Fax: (613) 9903-2007
Email: robert.brooks@buseco.monash.edu.au

Chapter 12

Rajeeva Sinha
Odette School of Business
University of Windsor
401 Sunset Ave.
Windsor, Ontario, Canada N9B 3P4
Tel.: 1519-253-3000 Ext. 3124
Fax: 1519-973-7073
Email: rsinha@uwindsor.ca

Vijay Jog
Carleton University
1127 Colonel by Drive
Ottawa, Canada K1S 5B6
Tel.: 1613-520-2600 ext. 2377
Fax: 1613-520-4427
Email: vijay_jog@carleton.ca

Chapter 13

Pauline Shum
Schulich School of Business
York University
4700 Keele St.
Toronto, ON, Canada M3J 1P3
Email: pshum@yorku.ca

Kevin X. Zhu
Ibbotson Associates/Morningstar
225 North Michigan Ave.
Suite 700
Chicago, IL 60601, USA
Email: kzhu@ibbotson.com

Chapter 14

Anthony Yanxiang Gu
Jones School of Business
SUNY College at Geneseo
Geneseo, New York 14454
Email: gu@geneseo.edu

John T. Simon
College of Business and Public Administration
Governors State University
University Park, Illinois 60466
Email: j-simon@govst.edu

Chapter 15

C. K. Woo
Energy and Environmental Economics Inc.
101 Montgomery Street, Suite 1600
San Francisco, CA 94111, USA
and
Hong Kong Energy Studies Centre
Hong Kong Baptist University
Kowloon Tong, Hong Kong

I. Horowitz
Decision and Information Sciences
Warrington College of Business Administration
University of Florida, Gainesville, FL 32611-7169, USA
and
School of Accounting and Finance
Hong Kong Polytechnic University
Hung Hom, Hong Kong

N. Toyama
Sacramento Municipal Utilities District
6201 S Street
Sacramento, CA 95817, USA

A. Olson
Energy and Environmental Economics Inc.
101 Montgomery Street, Suite 1600
San Francisco, CA 94111, USA

A. Lai
ATM/eCommerce Analytics
Bank of America
1755 Grant Street, 4th Floor
Concord, CA 94520, USA
Tel.: +1-925-692-7340
Fax: 925-675-8867
Email: aaron.lai@bankofamerica.com

R. Wan
InfoAtlas Inc.
1441 Franklin Street, Suite 204
Oakland , CA 94612, USA

This page intentionally left blank

Chapter 1

The Least Cost Superreplicating Portfolio for Short Puts and Calls in The Boyle–Vorst Model with Transaction Costs

Guan-Yu Chen
Cornell University, USA

Ken Palmer
National Taiwan University, Taiwan

Yuan-Chung Sheu*
National Chiao Tung University, Taiwan

Since Black and Scholes (1973) introduced their option-pricing model in frictionless markets, many authors have attempted to develop models incorporating transaction costs. The groundwork of modeling the effects of transaction costs was done by Leland (1985). The Leland model was put into a binomial setting by Boyle and Vorst (1992). Even when the market is arbitrage-free and a given contingent claim has a unique replicating portfolio, there may exist superreplicating portfolios of lower cost. However, it is known that there is no superreplicating portfolio for long calls and puts of lower cost than the replicating portfolio. Nevertheless, this is not true for short calls and puts. As the negative of the cost of the least cost superreplicating portfolios for such a position is a lower bound for the call or put price, it is important to determine this least cost. In this paper, we consider two-period binomial models and show that, for a special class of claims including short call and put options, there are just four possibilities so that the least cost superreplicating portfolios can be easily calculated for such positions. Also we show that, in general, the least cost superreplicating portfolio is path-dependent.

Keywords: Option pricing; transaction costs; binomial model; superreplicating.

1. Introduction

Since Black and Scholes (1973) introduced their option-pricing model in frictionless markets, many authors have attempted to develop models incorporating transaction costs. The groundwork of modeling the effects of transaction costs was done by Leland (1985). The Leland model was put into a binomial setting by Boyle and Vorst (1992). They derived self-financing strategies that

*Corresponding author.

perfectly replicate the final payoffs to long and short positions in put and call options, assuming proportional transaction costs on trades in the stocks and no transaction costs on trades in the bonds. Recently, Palmer (2001a) clarified the conditions under which there is a unique replicating strategy in the Boyle–Vorst model for an arbitrary contingent claim. Actually, following Stettner (1997) and Rutkowski (1998), Palmer worked in the framework of asymmetric proportional transaction costs, which includes not only the model of Boyle and Vorst, but also the slightly different model of Bensaid, Lesne, Pages, and Scheinkman (1992). For other recent contributions to this subject, see Perrakis and Lefoll (1997, 2000), Reiss (1999), and Chiang and Sheu (2004). A survey of some related results is given in Whalley and Wilmott (1997).

In arbitrage-free markets in the presence of transaction costs, even when a contingent claim has a unique replicating portfolio, there may exist a lower cost superreplicating portfolio. Nevertheless, Bensaid *et al.* (1992) gave conditions under which the cost of the replicating portfolio does not exceed the cost of any superreplicating portfolio. These results were generalized by Stettner (1997) and Rutkowski (1998) to the case of asymmetric transaction costs. Palmer (2001b) provided a further slight generalization. These results have the consequence that there is no superreplicating portfolio for long calls and puts of lower cost than the replicating portfolio. However, this is not true for short calls and puts. As the negative of the cost of the least cost superreplicating portfolios for such a position is a lower bound for the call or put price, it is important to determine this least cost. Recently, in Chen, Palmer, and Sheu (2004), we determined the least cost superreplicating portfolios for general contingent claims in one-period models and showed that there are only finitely many possibilities for the least cost super replicating portfolios of a general two-period contingent claims. Our result narrows down the search for a least cost superreplicating portfolio to a finite number of possibilities. However, the number of possibilities for the least cost superreplicating portfolios is still large. In this paper, we consider a restricted class of claims for which the number of possibilities can be reduced to a manageable number.

In Section 2, we review some basic results for general n-period models. We also quote two results from Chen *et al*. (2006) about the number of replicating portfolios and the least cost superreplicating portfolios for any contingent claim in a one-period binomial model. In Section 3, we recall the results of Chen *et al.* (2006) for the least cost superreplicating portfolios of a general two-period contingent claim. In Section 4, we show that for a special class of claims including short call and put options there are just four possibilities so that the least cost superreplicating portfolios can be easily calculated for such positions. In Section 5, we show that, in general, the least cost superreplicating portfolio is path-dependent.

2. Preliminaries

We consider an n-period binomial model of a financial market with two securities: a risky asset, referred to as a stock, and a risk-free investment, called a bond. If the stock price now is S, then at the end of the next period it is either Su or Sd, where $0 < d < u$. The bond yields a constant rate of return r over each time period meaning that a dollar now is worth $R = 1 + r$ after one period.

We assume that, on one hand, proportional transaction costs are incurred when shares of the risky asset are traded but, on the other hand, that trading in riskless bonds is cost-free. More precisely, we assume that when the stock price is S, buying one share incurs a transaction cost of λS and that selling one share incurs a transaction cost of μS, where

$$\lambda \geq 0, \qquad 0 \leq \mu < 1.$$

As is usual, we assume throughout this paper that there are no transaction costs when a portfolio is established at time 0. For no arbitrage consideration, we also assume that

$$d < R < u.$$

Let us denote by $\phi = \{(\Delta_i, B_i),\ i = 0, 1, 2, \ldots, n\}$, a (self-financing) portfolio where Δ_i stands for the number of shares and B_i the number of bonds held at time i. Under our assumption, it is natural that the initial value or cost of the portfolio ϕ is $\Delta_0 S_0 + B_0$.

A contingent claim is a two-dimensional random variable $X = (g, h)$ where g represents the number of shares and h the value of bonds held at time n. We say that a portfolio $\phi = \{(\Delta_i, B_i),\ i = 0, 1, 2, \ldots, n\}$ replicates the claim X that is settled by delivery if it is self-financing and $\Delta_n = g$ and $B_n = h$. We say a self-financing portfolio ϕ is a *superreplicating portfolio* for a contingent claim $X = (g, h)$ settled by delivery at time n if at time n we have $\Delta_n \geq g$ and $B_n \geq h$. An upper arbitrage bound for the price at time 0 of a claim

$X = (g, h)$ is given by the cost of a least cost superreplicating portfolio for a long position in the claim X. A lower arbitrage bound for the price of X at time 0 is given by the negative of the cost of a least cost superreplicating portfolio for a short position in the claim X. As pointed out by several authors, in some circumstances, it is possible to find a portfolio which ultimately dominates a given contingent claim and costs less than a portfolio that replicates the claim. Of course, there are circumstances in which no superreplicating portfolio costs less than a replicating portfolio. Theorems 1 and 2 given in Palmer (2001a) generalize results of Bensaid *et al.* (1992), Stettner (1997), and Rutkowski (1998).

Theorem 1. *Suppose that*

$$d(1+\lambda) \leq R(1-\mu) \leq R(1+\lambda) \leq u(1-\mu).$$

Then, for any contingent claim, there is a unique replicating portfolio and no superreplicating portfolio costs less than the replicating portfolio.

Theorem 2. *Consider a contingent claim in an n-period binomial model with holdings (g_j, h_j) when the terminal stock price is $S_0 u^j d^{n-j}$. If these terminal holdings satisfy*

$$g_{j+1} \geq g_j,$$

$$(g_j - g_{j+1}) S u^{j+1} d^{n-j-1}(1+\lambda) + h_j - h_{j+1} \leq 0,$$

and

$$(g_j - g_{j+1}) S u^j d^{n-j}(1+\mu) + h_j - h_{j+1} \geq 0,$$

for $j = 0, 1, \ldots, n-1$, then there is a unique replicating portfolio for such a contingent claim and no superreplicating portfolio costs less than the replicating portfolio.

Clearly long positions in calls and puts satisfy these conditions in Theorem 2. However, short positions in calls and puts do not satisfy these conditions.

Consider a contingent claim in a one-period model with holdings $(\Delta_{\mathrm{u}}, B_{\mathrm{u}})$ in the up state and $(\Delta_{\mathrm{d}}, B_{\mathrm{d}})$ in the down state. Let

$$a_{\mathrm{u}} = \begin{cases} (\Delta_{\mathrm{d}} - \Delta_{\mathrm{u}})Su(1+\lambda) + B_{\mathrm{d}} - B_{\mathrm{u}} & \text{if } \Delta_{\mathrm{u}} \geq \Delta_{\mathrm{d}}, \\ (\Delta_{\mathrm{d}} - \Delta_{\mathrm{u}})Su(1-\mu) + B_{\mathrm{d}} - B_{\mathrm{u}} & \text{if } \Delta_{u} < \Delta_{\mathrm{d}}, \end{cases}$$

and

$$a_{\mathrm{d}} = \begin{cases} (\Delta_{\mathrm{d}} - \Delta_{\mathrm{u}})Sd(1-\mu) + B_{\mathrm{d}} - B_{\mathrm{u}} & \text{if } \Delta_{\mathrm{u}} \geq \Delta_{\mathrm{d}}, \\ (\Delta_{\mathrm{d}} - \Delta_{\mathrm{u}})Sd(1+\lambda) + B_{\mathrm{d}} - B_{\mathrm{u}} & \text{if } \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}. \end{cases}$$

Theorems 3 and 4 are quoted from Chen *et al.* (2004).

Theorem 3. *Consider a contingent claim in a one-period model with holdings* $(\Delta_{\mathrm{u}}, B_{\mathrm{u}})$ *in the up state and* $(\Delta_{\mathrm{d}}, B_{\mathrm{d}})$ *in the down state. Then the contingent claim has a unique replicating portfolio if and only if it satisfies one of the following conditions*:

(a) $\Delta_{\mathrm{u}} \geq \Delta_{\mathrm{d}}$,
(b) $\Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}$, $d(1+\lambda) < u(1-\mu)$,
(c) $\Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}$, $d(1+\lambda) \geq u(1-\mu)$, $a_{\mathrm{u}}a_{\mathrm{d}} > 0$.

The following theorem determines the least cost superreplicating portfolios for any contingent claims in a one-period binomial model.

Theorem 4. *Consider a contingent claim in a one-period model with holdings* $(\Delta_{\mathrm{u}}, B_{\mathrm{u}})$ *in the up state and* $(\Delta_{\mathrm{d}}, B_{\mathrm{d}})$ *in the down state.*

(a) *When the replicating portfolio is unique*, *it is a least cost superreplicating portfolio unless* $R > u(1-\mu)$, $a_{\mathrm{d}} < 0$ *when* $(\Delta_{\mathrm{u}}, B_{\mathrm{u}}/R)$ *are the holdings in a least cost superreplicating portfolio*, *or if* $R < d(1+\lambda)$, $a_{\mathrm{u}} > 0$ *when* $(\Delta_{\mathrm{d}}, B_{\mathrm{d}}/R)$ *are the holdings in a least cost superreplicating portfolio.*

(b) *When the replicating portfolio is not unique, it is necessary that* $\Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}$, $d(1+\lambda) \geq u(1-\mu)$. *Moreover, we have*:

   (i) *If* $R \geq d(1+\lambda)$, *there exists at least one replicating portfolio with share holdings* Δ *satisfying* $\Delta \leq \Delta_{\mathrm{u}}$ *and all such replicating portfolios are least cost superreplicating portfolios.*

   (ii) *If* $d(1+\lambda) \geq R \geq u(1-\mu)$, *there exists at least one replicating portfolio with share holdings* Δ *satisfying* $\Delta_{\mathrm{u}} \leq \Delta \leq \Delta_{\mathrm{d}}$ *and all such replicating portfolios are least cost superreplicating portfolios.*

(iii) *If $R \leq u(1-\mu)$, there exists at least one replicating portfolio with share holdings Δ satisfying $\Delta \geq \Delta_{\mathrm{d}}$ and all such replicating portfolios are least cost superreplicating portfolios.*

Remark 1. As mentioned in the Remarks after Theorem 4.1 in Chen *et al.* (2006), the cost $C(\Delta_{\mathrm{u}}, B_{\mathrm{u}}, \Delta_{\mathrm{d}}, B_{\mathrm{d}})$ of the least cost superreplicating portfolio is a continuous function which is linear in any region in the $(\Delta_{\mathrm{u}}, B_{\mathrm{u}}, \Delta_{\mathrm{d}}, B_{\mathrm{d}})$ space where $\Delta_{\mathrm{u}} - \Delta_{\mathrm{d}}$, a_{u}, and a_{d} are one-signed. In Chen *et al.* (2006), we proved Theorem 4 by considering the contingent claim according to the following cases:

$$
\begin{array}{llll}
\text{Case } 1: & \Delta_{\mathrm{u}} \geq \Delta_{\mathrm{d}}, & a_{\mathrm{d}} \geq a_{\mathrm{u}} > 0 & \\
\text{Case } 2: & \Delta_{\mathrm{u}} \geq \Delta_{\mathrm{d}}, & a_{\mathrm{d}} \geq 0 \geq a_{\mathrm{u}} & \\
\text{Case } 3: & \Delta_{\mathrm{u}} \geq \Delta_{\mathrm{d}}, & a_{\mathrm{u}} \leq a_{\mathrm{d}} < 0 & \\
\text{Case } 4: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) > d(1+\lambda), & a_{\mathrm{u}} > a_{\mathrm{d}} > 0 \\
\text{Case } 5: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) > d(1+\lambda), & a_{\mathrm{u}} \geq 0 \geq a_{\mathrm{d}} \\
\text{Case } 6: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) > d(1+\lambda), & a_{\mathrm{d}} < a_{\mathrm{u}} < 0 \\
\text{Case } 7: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) < d(1+\lambda), & a_{\mathrm{d}} > a_{\mathrm{u}} > 0 \\
\text{Case } 8: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) < d(1+\lambda), & a_{\mathrm{u}} < a_{\mathrm{d}} < 0 \\
\text{Case } 9: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) < d(1+\lambda), & a_{\mathrm{u}} = 0 < a_{\mathrm{d}} \\
\text{Case } 10: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) < d(1+\lambda), & a_{\mathrm{d}} = 0 > a_{\mathrm{u}} \\
\text{Case } 11: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) < d(1+\lambda), & a_{\mathrm{d}} > 0 > a_{\mathrm{u}} \\
\text{Case } 12: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) = d(1+\lambda), & a_{\mathrm{u}} = a_{\mathrm{d}} > 0 \\
\text{Case } 13: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) = d(1+\lambda), & a_{\mathrm{u}} = a_{\mathrm{d}} < 0 \\
\text{Case } 14: & \Delta_{\mathrm{u}} < \Delta_{\mathrm{d}}, & u(1-\mu) = d(1+\lambda), & a_{\mathrm{u}} = a_{\mathrm{d}} = 0.
\end{array}
$$

It follows from Theorem 3 that in Cases 1–8, 12, and 13, there is a unique replicating portfolio (Δ, B) which, as in Chen *et al.* (2004), has cost given by

$$
C = \Delta S + B = \frac{p}{R}[\Delta_{\mathrm{u}} S\overline{u} + B_{\mathrm{u}}] + \frac{1-p}{R}[\Delta_{\mathrm{d}} S\overline{d} + B_{\mathrm{d}}]
$$

where

$$
\overline{u} = \begin{cases} u(1+\lambda) & \text{if } a_{\mathrm{d}} \geq 0, \\ u(1-\mu) & \text{if } a_{\mathrm{d}} < 0, \end{cases} \qquad \overline{d} = \begin{cases} d(1+\lambda) & \text{if } a_{\mathrm{u}} \geq 0, \\ d(1-\mu) & \text{if } a_{\mathrm{u}} < 0, \end{cases} \qquad p = \frac{R-\overline{d}}{\overline{u}-\overline{d}}.
$$

3. General Contingent Claims in the Two-Period Case

In this section, we recall some results of Chen *et al.* (2006) for a general two-period contingent claim with terminal holdings $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{ud}}, B_{\mathrm{ud}}),$

$(\Delta_{\mathrm{dd}},\ B_{\mathrm{dd}})\}$. Write

$$b_{\mathrm{u}}(\Delta_{\mathrm{u}}) = \max\{B_{\mathrm{uu}} + e(\Delta_{\mathrm{u}} - \Delta_{\mathrm{uu}})Su^2, \quad B_{\mathrm{ud}} + e(\Delta_{\mathrm{u}} - \Delta_{\mathrm{ud}})Sud\}$$

and

$$b_{\mathrm{d}}(\Delta_{\mathrm{d}}) = \max\{B_{\mathrm{ud}} + e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{ud}})Sud, \quad B_{\mathrm{dd}} + e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}})Sd^2\},$$

where $e(\Delta) = -\Delta + \mu\Delta^+ + \lambda\Delta^-$. The significance of these two quantities is that $(\Delta_{\mathrm{u}}, B_{\mathrm{u}})$ is a superreplicating portfolio for the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{ud}}, B_{\mathrm{ud}})\}$ with initial stock price Su if and only if $B_{\mathrm{u}} \geq b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R$ and $(\Delta_{\mathrm{d}}, B_{\mathrm{d}})$ is a superreplicating portfolio for the one-period claim $\{(\Delta_{\mathrm{ud}}, B_{\mathrm{ud}}), (\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})\}$ with initial stock price Sd if and only if $B_{\mathrm{d}} \geq b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R$. Denote by $C(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$ the least cost of superreplicating portfolios for the one-period contingent claim $\{(\Delta_{\mathrm{u}}, b_u(\Delta_{\mathrm{u}})/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ with initial stock price S. Then it was proved in Chen *et al.* (2004) that the infimum of the cost of a superreplicating portfolio for the two-period contingent claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{ud}}, B_{\mathrm{ud}}), (\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})\}$ is equal to the infimum over $(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$ of $C(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$. Theorem 5 shows that we need only consider the function $C(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$ in a certain rectangle in the $(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$-plane.

To do this, we consider functions

$$f_{\mathrm{u}}(\Delta_{\mathrm{u}}) = B_{\mathrm{uu}} + e(\Delta_{\mathrm{u}} - \Delta_{\mathrm{uu}})Su^2 - B_{\mathrm{ud}} - e(\Delta_{\mathrm{u}} - \Delta_{\mathrm{ud}})Sud \tag{1}$$

and

$$f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = B_{\mathrm{ud}} + e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{ud}})Sud - B_{\mathrm{dd}} - e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}})Sd^2. \tag{2}$$

Note that the values of Δ_{u} satisfying $f_{\mathrm{u}}(\Delta_{\mathrm{u}}) = 0$ are exactly those for which $(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R)$ is a replicating portfolio for the contingent claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{ud}}, B_{\mathrm{ud}})\}$ with initial stock price Su and the values of Δ_{d} satisfying $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = 0$ are exactly those for which $(\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)$ is a replicating portfolio for the contingent claim $\{(\Delta_{\mathrm{ud}}, B_{\mathrm{ud}}), (\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})\}$ with initial stock price Sd.

Let $[\alpha_{\mathrm{u}}, \beta_{\mathrm{u}}]$ be the smallest closed interval containing all solutions of $f_{\mathrm{u}}(\Delta_{\mathrm{u}}) = 0$ and also Δ_{uu} and Δ_{ud}. Similarly, let $[\alpha_{\mathrm{d}}, \beta_{\mathrm{d}}]$ be the smallest closed interval containing all solutions of $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = 0$ and also Δ_{ud} and Δ_{dd}.

Let Π be the rectangle in the $(\Delta_{\rm u}, \Delta_{\rm d})$-plane given by

$$\Pi = \{(\Delta_{\rm u}, \Delta_{\rm d}) : \alpha_{\rm u} \le \Delta_{\rm u} \le \beta_{\rm u},\ \alpha_{\rm d} \le \Delta_{\rm d} \le \beta_{\rm d}\}.$$

Theorem 5. *For a general two-period contingent claim* $\{(\Delta_{\rm uu}, B_{\rm uu}), (\Delta_{\rm ud}, B_{\rm ud}), (\Delta_{\rm dd}, B_{\rm dd})\}$*, the function* $C(\Delta_{\rm u}, \Delta_{\rm d})$ *takes its minimum in the rectangle* Π *at some point* $(\Delta_{\rm u}, \Delta_{\rm d})$ *and a least cost superreplicating portfolio for the one-period claim* $\{(\Delta_{\rm u}, b_{\rm u}(\Delta_{\rm u})/R), (\Delta_{\rm d}, b_{\rm d}(\Delta_{\rm d})/R)\}$ *with initial stock price S yields a least cost super replicating portfolio for the two-period claim.*

(It is worth noting that in the case $(\Delta_{\rm uu}, B_{\rm uu}) = (\Delta_{\rm ud}, B_{\rm ud})$, there is always a least cost superreplicating portfolio with $\Delta_{\rm u} = \Delta_{\rm uu}$ because in this case $\alpha_{\rm u} = \beta_{\rm u} = \Delta_{\rm uu}$. We consider this special case in more detail in Section 4.)

Consider the two quantities $a_{\rm u}$ and $a_{\rm d}$,

$$a_{\rm u} = a_{\rm u}(\Delta_{\rm u}, \Delta_{\rm d}) = (\Delta_{\rm d} - \Delta_{\rm u})S\overline{u} + \frac{b_{\rm d}(\Delta_{\rm d}) - b_{\rm u}(\Delta_{\rm u})}{R},$$
$$a_{\rm d} = a_{\rm d}(\Delta_{\rm u}, \Delta_{\rm d}) = (\Delta_{\rm d} - \Delta_{\rm u})S\overline{d} + \frac{b_{\rm d}(\Delta_{\rm d}) - b_{\rm u}(\Delta_{\rm u})}{R},$$

where

$$\overline{u} = \begin{cases} u(1+\lambda) & \text{if } \Delta_{\rm u} \ge \Delta_{\rm d}, \\ u(1-\mu) & \text{if } \Delta_{\rm u} < \Delta_{\rm d}, \end{cases} \qquad \overline{d} = \begin{cases} d(1-\mu) & \text{if } \Delta_{\rm u} \ge \Delta_{\rm d}, \\ d(1+\lambda) & \text{if } \Delta_{\rm u} < \Delta_{\rm d}. \end{cases}$$

By using the fundamental theorem of linear programing, the following theorem shows that there are only finitely many possibilities for a least cost superreplicating portfolio.

Theorem 6. *For a general two-period contingent claim with terminal holdings* $\{(\Delta_{\rm uu}, B_{\rm uu}), (\Delta_{\rm ud}, B_{\rm ud}), (\Delta_{\rm dd}, B_{\rm dd})\}$*, there always exists a least cost superreplicating portfolio with initial holdings* (Δ, B) *and holdings* $(\Delta_{\rm u}, B_{\rm u})$*,* $(\Delta_{\rm d}, B_{\rm d})$ *at the end of the first period which represent a least cost superreplicating portfolio for the one-period claim* $\{(\Delta_{\rm u}, b_{\rm u}(\Delta_{\rm u})/R), (\Delta_{\rm d}, b_{\rm d}(\Delta_{\rm u})/R)\}$ *and such that at least two distinct conditions from the following list are satisfied*:

$$\Delta_{\rm u} = \Delta_{\rm uu}, \quad \Delta_{\rm u} = \Delta_{\rm ud}, \quad \Delta_{\rm d} = \Delta_{\rm ud}, \quad \Delta_{\rm d} = \Delta_{\rm dd}, \quad \Delta_{\rm u} = \Delta_{\rm d},$$
$$a_{\rm u}(\Delta_{\rm u}, \Delta_{\rm d}) = 0, \quad a_{\rm d}(\Delta_{\rm u}, \Delta_{\rm d}) = 0,$$
$$f_{\rm u}(\Delta_{\rm u}) = 0, \quad f_{\rm d}(\Delta_{\rm d}) = 0.$$

Note that the condition $a_{\mathrm{u}}(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}}) = 0$ means that $(\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R^2)$ is a replicating portfolio for the contingent claim $\{(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ with initial stock price S. Likewise, the condition $a_{\mathrm{d}}(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}}) = 0$ means that $(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R^2)$ is a replicating portfolio for the contingent claim $\{(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$. We note again that the values of Δ_{u} satisfying $f_{\mathrm{u}}(\Delta_{\mathrm{u}}) = 0$ are exactly those for which $(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R)$ is a replicating portfolio for the contingent claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{ud}}, B_{\mathrm{ud}})\}$ with initial stock price Su and the values of Δ_{d} satisfying $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = 0$ are exactly those for which $(\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)$ is a replicating portfolio for the contingent claim $\{(\Delta_{\mathrm{ud}}, B_{\mathrm{ud}}), (\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})\}$ with initial stock price Sd.

Theorem 6 narrows down the search for a least cost superreplicating portfolio to a finite number of possibilities. However, the number of possibilities is still quite large. In the following section, we consider a restricted class of claims for which the number of possibilities can be reduced to a manageable number.

4. Least Cost Superreplicating Portfolios for Short Puts and Calls in the Two-Period Case

In this section, we determine the initial holdings of the least cost superreplicating portfolios for a claim in the two-period model with

$$\Delta_{\mathrm{uu}} = \Delta_{\mathrm{ud}} < \Delta_{\mathrm{dd}}, \qquad B_{\mathrm{uu}} = B_{\mathrm{ud}}. \tag{3}$$

This includes short calls and puts with the exercise price between Sud and Sd^2. Note that we could treat the case $\Delta_{\mathrm{uu}} < \Delta_{\mathrm{ud}} = \Delta_{\mathrm{dd}}$, $B_{\mathrm{ud}} = B_{\mathrm{dd}}$ similarly. This would include short calls and puts with the exercise price between Su^2 and Sud.

Theorem 7. *Consider a two-period binomial model incorporating transaction costs with parameters S, u, d, R, μ, and λ. For every contingent claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{ud}}, B_{\mathrm{ud}}), (\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})\}$ satisfying Equation* (3), *there always exists a least cost superreplicating portfolio which belongs to one of the following four types (note that in all cases transactions are carried out at the terminal nodes so that the final share holdings are Δ_{uu}, Δ_{ud}, Δ_{du}, Δ_{dd} in* *states uu, ud, du, and dd, respectively):*

(I) *the initial holdings are $(\Delta_{dd}, B_{dd}/R^2)$ and the only additional share transaction is selling $(\Delta_{dd} - \Delta_{uu})$ shares in state u (this type arises only if $R < d(1+\lambda)$ and $B_{uu} - B_{dd} - Sud(1-\mu)(\Delta_{dd} - \Delta_{uu}) < 0$);*

(II) *the initial holdings are (δ, B), where $\delta \leq \Delta_{uu}$ and (δ, B) is such that $BR - B_{uu}/R$ is just enough to carry out the only additional share transaction of buying back $(\Delta_{uu} - \delta)$ shares of stocks in state u; there are two possibilities:*

   (a) *$\delta = \Delta_{uu}$ and the terminal holdings in the du state are (Δ_{uu}, B_{uu}) (this case only arises if $B_{uu} - B_{dd} - Sd^2(1+\lambda)(\Delta_{dd} - \Delta_{uu}) \geq 0$);*

   (b) *$\delta < \Delta_{uu}$ and the terminal holdings in the dd state are (Δ_{dd}, B_{dd}) (this case only arises if $B_{uu} - B_{dd} - Sd^2(1+\lambda)(\Delta_{dd} - \Delta_{uu}) < 0$);*

(III) *the initial holdings are $(\alpha, B/R)$, where $\alpha > \Delta_{uu}$ and (α, B) are the initial holdings in a replicating portfolio for the one-period portion $\{d, du, dd\}$, and the only additional share transaction is selling $(\alpha - \Delta_{uu})$ shares in state u (this case only arises if $R < d(1+\lambda)$);*

(IV) *a replicating portfolio for the whole two-period model.*

Proof. It follows from the remark after Theorem 5 that we need only determine the Δ_d which yields the least cost for the one-period contingent claim $\{(\Delta_{uu}, B_{uu}/R), (\Delta_d, b_d(\Delta_d)/R)\}$ with initial stock price S and then determine a least cost superreplicating portfolio for this one-period claim. For this claim, we have

$$a_u = a_u(\Delta_d) = a_u(\Delta_{uu}, \Delta_d) = (\Delta_d - \Delta_{uu})S\overline{u} + \frac{b_d(\Delta_d) - B_{uu}}{R},$$
$$a_d = a_d(\Delta_d) = a_d(\Delta_{uu}, \Delta_d) = (\Delta_d - \Delta_{uu})S\overline{d} + \frac{b_d(\Delta_d) - B_{uu}}{R},$$

where

$$\overline{u} = \begin{cases} u(1+\lambda) & \text{if } \Delta_{uu} \geq \Delta_d, \\ u(1-\mu) & \text{if } \Delta_{uu} < \Delta_d, \end{cases} \qquad \overline{d} = \begin{cases} d(1-\mu) & \text{if } \Delta_{uu} \geq \Delta_d, \\ d(1+\lambda) & \text{if } \Delta_{uu} < \Delta_d, \end{cases}$$

and

$$b_d(\Delta_d) = \max\{B_{uu} + Sude(\Delta_d - \Delta_{uu}), B_{dd} + Sd^2e(\Delta_d - \Delta_{dd})\}.$$

Also

$$f_d(\Delta_d) = B_{uu} + Sude(\Delta_d - \Delta_{uu}) - B_{dd} - Sd^2e(\Delta_d - \Delta_{dd}).$$

Note that $(\Delta_\text{d}, B_\text{d})$ is a replicating portfolio for the one-period portion $\{d, du, dd\}$ if and only if $f_\text{d}(\Delta_\text{d}) = 0$ and $B_\text{d} = b_\text{d}(\Delta_\text{d})/R$. Further observe that the continuous function $f_\text{d}(\Delta_\text{d})$ is decreasing and linear for $\Delta_\text{d} \leq \Delta_\text{uu}$, $\Delta_\text{d} \geq \Delta_\text{dd}$ and linear and decreasing, constant, or increasing for $\Delta_\text{uu} < \Delta_\text{d} < \Delta_\text{dd}$ depending on the sign of $u(1-\lambda) - d(1+\mu)$. Note also that

$$f_\text{d}(\Delta_\text{uu}) = B_\text{uu} - B_\text{dd} + Sd^2(\Delta_\text{uu} - \Delta_\text{dd})(1+\lambda),$$

and that

$$a_\text{u}(\Delta_\text{uu}) = a_\text{d}(\Delta_\text{uu}) = \frac{[f_\text{d}(\Delta_\text{uu})]^-}{R}.$$

The signs of a_u and a_d: We start by examining the signs of a_u and a_d. First we show that when $\Delta_\text{d} < \Delta_\text{uu}$, then $a_\text{d} > 0$. This follows because

$$b_\text{d}(\Delta_\text{d}) \geq B_\text{uu} - Sud(\Delta_\text{d} - \Delta_\text{uu})(1+\lambda),$$

and so

$$\begin{aligned} a_\text{d} &\geq (\Delta_\text{d} - \Delta_\text{uu})Sd(1-\mu) - \frac{Sud(\Delta_\text{d} - \Delta_\text{uu})(1+\lambda)}{R} \\ &= \frac{Sd(\Delta_\text{d} - \Delta_\text{uu})}{R}[R(1-\mu) - u(1+\lambda)] > 0. \end{aligned}$$

Suppose now that $\Delta_\text{d} > \Delta_\text{uu}$ and $R(1+\lambda) > u(1-\mu)$. Then as

$$b_\text{d}(\Delta_\text{d}) \geq B_\text{uu} - Sud(\Delta_\text{d} - \Delta_\text{uu})(1-\mu),$$

we have

$$\begin{aligned} a_\text{d} &\geq (\Delta_\text{d} - \Delta_\text{uu})Sd(1+\lambda) - \frac{Sud(\Delta_\text{d} - \Delta_\text{uu})(1-\mu)}{R} \\ &= \frac{Sd(\Delta_\text{d} - \Delta_\text{uu})}{R}[R(1+\lambda) - u(1-\mu)] > 0. \end{aligned}$$

Assume next that $\Delta_\text{d} > \Delta_\text{uu}$, $f_\text{d}(\Delta_\text{uu}) \leq 0$, and $R(1+\lambda) \leq u(1-\mu)$. The latter implies that $d(1+\lambda) < u(1-\mu)$ and so $f_\text{d}(\Delta_\text{d})$ is strictly decreasing.

Then as $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) \leq 0$, we have $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) < 0$ for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$ and so

$$b_{\mathrm{d}}(\Delta_{\mathrm{d}}) = B_{\mathrm{dd}} + Sd^2 e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}}),$$

and

$$a_{\mathrm{d}} = (\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})Sd(1+\lambda) + \frac{B_{\mathrm{dd}} - B_{\mathrm{uu}} + Sd^2 e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}})}{R}.$$

It follows that

$$a_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = \frac{B_{\mathrm{dd}} - B_{\mathrm{uu}} - Sd^2(1+\lambda)(\Delta_{\mathrm{uu}} - \Delta_{\mathrm{dd}})}{R} \geq 0$$

and

$$a_{\mathrm{d}}'(\Delta_{\mathrm{d}}) = \frac{Sd}{R} \begin{cases} (R-d)(1+\lambda) & \text{if } \Delta_{\mathrm{d}} < \Delta_{\mathrm{dd}} \\ R(1+\lambda) - d(1-\mu) & \text{if } \Delta_{\mathrm{d}} > \Delta_{\mathrm{dd}} \end{cases} > 0.$$

Hence, in this case, we still have

$$a_{\mathrm{d}} > 0$$

for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$.

Hence we are left with the case $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$, $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) > 0$, and $R(1+\lambda) \leq u(1-\mu)$. In this case, there exists a unique $\gamma > \Delta_{\mathrm{uu}}$ such that

$$f_{\mathrm{d}}(\Delta_{\mathrm{d}}) \begin{cases} > 0 & \text{if } \Delta_{\mathrm{d}} < \gamma, \\ = 0 & \text{if } \Delta_{\mathrm{d}} = \gamma, \\ < 0 & \text{if } \Delta_{\mathrm{d}} > \gamma. \end{cases}$$

Also note that $a_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = 0$. It follows as in the previous case that $a_{\mathrm{d}}'(\Delta_{\mathrm{d}}) > 0$ if $\Delta_{\mathrm{d}} > \gamma$. However, if $\Delta_{\mathrm{uu}} < \Delta_{\mathrm{d}} < \gamma$, then

$$b_{\mathrm{d}}(\Delta_{\mathrm{d}}) = B_{\mathrm{uu}} - Sud(1-\mu)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})$$

and

$$a_{\mathrm{d}} = (\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})Sd(1+\lambda) - \frac{Sud(1-\mu)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})}{R},$$

so that

$$a_{\mathrm{d}}'(\Delta_{\mathrm{d}}) = \frac{Sd}{R}[R(1+\lambda) - u(1-\mu)] \leq 0.$$

Then there exists $\tilde{\delta} \geq \gamma$ such that

$$a_{\mathrm{d}} \begin{cases} \leq 0 & \text{if } \Delta_{\mathrm{uu}} < \Delta_{\mathrm{d}} \leq \tilde{\delta}, \\ = 0 & \text{if } \Delta_{\mathrm{d}} = \tilde{\delta}, \\ > 0 & \text{if } \Delta_{\mathrm{d}} > \tilde{\delta}. \end{cases}$$

Now we examine the sign of a_{u}. First note that if $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$, then

$$\begin{aligned} a_{\mathrm{u}}(\Delta_{\mathrm{d}}) &= (\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})Su(1-\mu) + \frac{b_{\mathrm{d}}(\Delta_{\mathrm{d}}) - B_{\mathrm{uu}}}{R} \\ &\geq (\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})Su(1-\mu) - \frac{Sud(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})(1-\mu)}{R} \\ &= \frac{Su}{R}(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})(R-d)(1-\mu) > 0. \end{aligned}$$

If $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) \geq 0$, then $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) > 0$ for $\Delta_{\mathrm{d}} < \Delta_{\mathrm{uu}}$ and so for $\Delta_{\mathrm{d}} < \Delta_{\mathrm{uu}}$,

$$a_{\mathrm{u}} = \frac{Su}{R}(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}})(R-d)(1+\lambda) < 0.$$

Also as $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) \geq 0$, $a_{\mathrm{u}}(\Delta_{\mathrm{uu}}) = 0$.

If $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) < 0$, there exists a unique $\gamma < \Delta_{\mathrm{uu}}$ such that $f_{\mathrm{d}}(\gamma) = 0$. We show as in the case $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) \geq 0$ that $a_{\mathrm{u}} < 0$ if $\Delta_{\mathrm{d}} \leq \gamma$. Now as $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) < 0$, $a_{\mathrm{u}}(\Delta_{\mathrm{uu}}) > 0$. Then as a_{u} is a linear function of Δ_{d} in the interval $[\gamma, \Delta_{\mathrm{uu}}]$, it follows that there exists a unique δ in $(\gamma, \Delta_{\mathrm{uu}})$ such that $a_{\mathrm{u}}(\delta) = 0$. Note also that $f_{\mathrm{d}}(\delta) < 0$. Thus, if $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) < 0$,

$$a_u \begin{cases} < 0 & \text{if } \Delta_{\mathrm{d}} < \delta, \\ = 0 & \text{if } \Delta_{\mathrm{d}} = \delta, \\ > 0 & \text{if } \delta < \Delta_{\mathrm{d}} \leq \Delta_{\mathrm{uu}}. \end{cases}$$

We now consider four different cases.

1. Suppose first that $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) < 0$. Then $a_{\mathrm{d}} > 0$ for all Δ_{d}, $a_{\mathrm{u}} > 0$ for $\Delta_{\mathrm{d}} > \delta$, and $a_{\mathrm{u}}(\delta) = 0$ and $a_{\mathrm{u}} < 0$ for $\Delta_{\mathrm{d}} < \delta$. Also $f_{\mathrm{d}}(\gamma) = 0$ has at most three solutions. As the function $b_{\mathrm{d}}(\Delta_{\mathrm{d}})$ is linear in any interval not containing Δ_{uu}, Δ_{dd}, or any of the γ's, we see from Remark 1 that the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ is linear in any interval not containing δ, Δ_{uu}, Δ_{dd}, or any of the γ's. So the minimum must be achieved at one of these points.

Suppose the minimum occurs at δ. At δ, $a_{\mathrm{d}} > 0$ and $a_{\mathrm{u}} = 0$ and so the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\delta, b_{\mathrm{d}}(\delta)/R)\}$ is in Case 2 of Remark 1 so that the replicating portfolio is unique and by Theorem 4 is the least cost superreplicating portfolio. However, the condition $a_{\mathrm{u}}(\delta) = 0$ implies that

$(\delta, b_{\mathrm{d}}(\delta)/R^2)$ is a replicating portfolio for this one-period claim. So the initial holdings are $(\delta, B) = (\delta, b_{\mathrm{d}}(\delta)/R^2)$, where $(BR - B_{\mathrm{uu}}/R)$ is just enough to buy back $(\Delta_{\mathrm{uu}} - \delta)$ shares of stocks in state u. Moreover, as $f_{\mathrm{d}}(\delta) < 0$, $b_{\mathrm{d}}(\delta) = B_{\mathrm{dd}} + (\Delta_{\mathrm{dd}} - \delta)Sd^2(1+\lambda)$ and so

$$b_{\mathrm{d}}(\delta) - (\Delta_{\mathrm{dd}} - \delta)Sd^2(1+\lambda) = B_{\mathrm{dd}},$$

that is, the final holdings in the dd state are $(\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})$. This is type (II)(b) of Theorem 7

We now show that in this case the minimum is either not attained at Δ_{uu} or if it is, then it is also attained at Δ_{dd} or at one of the solutions of $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = 0$. Let γ be the least number greater than Δ_{uu} such that $f_{\mathrm{d}}(\gamma) = 0$ (take $\gamma = \infty$ if no such γ exists). Set $\tilde{\gamma} = \min\{\gamma, \Delta_{\mathrm{dd}}\}$. Then in the interval $(\delta, \tilde{\gamma}]$, a_{u} and a_{d} are positive and $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) \leq 0$. So the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ is in one of Cases 4, 7, or 12 of Remark 1 for Δ_{d} in $(\Delta_{\mathrm{uu}}, \tilde{\gamma}]$ and in Case 1 for Δ_{d} in $(\delta, \Delta_{\mathrm{uu}}]$.

If $R \geq d(1+\lambda)$, it follows from Theorem 4 and Remark 1 that the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ in these two intervals is given by

$$C(\Delta_{\mathrm{d}}) = \frac{p}{R}[\Delta_{\mathrm{uu}} Su(1+\lambda) + B_{\mathrm{uu}}] + \frac{1-p}{R}\left[\Delta_{\mathrm{d}} Sd(1+\lambda) + \frac{b_{\mathrm{d}}(\Delta_{\mathrm{d}})}{R}\right],$$

where

$$b_{\mathrm{d}}(\Delta_{\mathrm{d}}) = B_{\mathrm{dd}} - Sd^2(1+\lambda)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}}), \qquad 0 < p = \frac{R - d(1+\lambda)}{(u-d)(1+\lambda)} < 1.$$

We see that for $\delta < \Delta_{\mathrm{d}} < \tilde{\gamma}$,

$$C'(\Delta_{\mathrm{d}}) = \frac{(1-p)Sd(R-d)(1+\lambda)}{R^2} > 0.$$

Hence there is no minimum at Δ_{uu} if $R \geq d(1+\lambda)$.

On the other hand, if $R < d(1+\lambda)$, then it follows from Theorem 4 that

$$C(\Delta_{\mathrm{d}}) = \Delta_{\mathrm{d}} + \frac{b_{\mathrm{d}}(\Delta_{\mathrm{d}})}{R^2},$$

which is linear in $(\delta, \tilde{\gamma}]$. Hence if there is a minimum at Δ_{uu}, there is also one at $\tilde{\gamma}$ and hence at Δ_{dd} or at a solution of $f_{\mathrm{d}}(\gamma) = 0$.

So the conclusion in this case is that the minimum of $C(\Delta_{\mathrm{d}})$ occurs at one of the points δ, giving type (II)(b) of Theorem 7, or at Δ_{dd} or at one of the solutions of $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = 0$.

2. We consider next the case $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = 0$. Then $a_{\mathrm{d}} > 0$ for all $\Delta_{\mathrm{d}} \neq \Delta_{\mathrm{uu}}$, and $a_{\mathrm{u}} > 0$ for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$, and $a_{\mathrm{u}} < 0$ for $\Delta_{\mathrm{d}} < \Delta_{\mathrm{uu}}$ and $a_{\mathrm{u}}(\Delta_{\mathrm{uu}}) = a_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = 0$. If $f_{\mathrm{d}}(\Delta_{\mathrm{dd}}) = 0$, then $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) = 0$ if and only if $\Delta_{\mathrm{uu}} \leq \Delta_{\mathrm{d}} \leq \Delta_{\mathrm{dd}}$. As the function $b_{\mathrm{d}}(\Delta_d)$ is linear in any interval not containing Δ_{uu} or Δ_{dd}, we see from Remark 1 that the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ is linear in any such interval. So the minimum must be achieved at one of these two points. If $f_{\mathrm{d}}(\Delta_{\mathrm{dd}}) \neq 0$, then $f_{\mathrm{d}}(\gamma) = 0$ has at most one more solution in addition to Δ_{uu}. Again the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ is linear in any interval not containing Δ_{uu}, Δ_{dd}, or any of the γ's. So the minimum must be achieved at one of these points.

Suppose it is achieved at $\Delta_{\mathrm{d}} = \Delta_{\mathrm{uu}}$. Then the one-period claim

$$\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\} = \{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R)\}$$

is in Case 2 of Remark 1 so that by Theorem 4 the unique replicating portfolio $(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R^2)$ is the least cost superreplicating portfolio. This is type (II)(a) of Theorem 7.

3. We consider next the case $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) > 0$ and $R(1+\lambda) > u(1-\mu)$ so that f_{d} is strictly decreasing. Then $a_{\mathrm{d}} > 0$ for all $\Delta_{\mathrm{d}} \neq \Delta_{\mathrm{uu}}$, and $a_{\mathrm{u}} > 0$ for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$, and $a_{\mathrm{u}} < 0$ for $\Delta_{\mathrm{d}} < \Delta_{\mathrm{uu}}$ and $a_{\mathrm{u}}(\Delta_{\mathrm{uu}}) = a_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = 0$. Then $f_{\mathrm{d}}(\gamma) = 0$ has exactly one solution γ which is greater than Δ_{uu}. Again we see from Remark 1 that the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ is linear in any interval not containing Δ_{uu}, Δ_{dd}, or γ. So the minimum must be achieved at one of these points.

If the minimum is achieved at Δ_{uu}, we show as in the previous case that it is of type (II)(a) of Theorem 7.

4. We consider next the case $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) > 0$ and $R(1+\lambda) \leq u(1-\mu)$. Then $f_{\mathrm{d}}(\gamma) = 0$ has exactly one solution γ which is greater than Δ_{uu}, and there exists $\tilde{\delta} \geq \gamma$ ($\tilde{\delta} = \gamma$ if and only if $R(1+\lambda) = u(1-\mu)$) such that $a_{\mathrm{d}} > 0$ for $\Delta_{\mathrm{d}} < \Delta_{\mathrm{uu}}$, $a_{\mathrm{d}} \leq 0$ for $\Delta_{\mathrm{uu}} < \Delta_{\mathrm{d}} < \tilde{\delta}$, $a_{\mathrm{d}}(\tilde{\delta}) = 0$, $a_{\mathrm{d}} > 0$ for $\Delta_{\mathrm{d}} > \tilde{\delta}$. Also $a_{\mathrm{u}} > 0$ for $\Delta_{\mathrm{u}} > \Delta_{\mathrm{uu}}$, $a_{\mathrm{u}} < 0$ for $\Delta_{\mathrm{u}} < \Delta_{\mathrm{uu}}$, and $a_{\mathrm{u}}(\Delta_{\mathrm{uu}}) = a_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = 0$. Again we see from Remark 1 that the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ is linear in any interval not containing Δ_{uu}, Δ_{dd}, $\tilde{\delta}$, or γ. So the minimum must be achieved at one of these points.

If the minimum is achieved at Δ_{uu}, we show as in the previous case that it is of type (II)(a).

Suppose a minimum occurs at $\tilde{\delta}$. As $d(1+\lambda) < u(1-\mu)$ and taking into account the signs of a_{u} and a_{d}, the one-period claim

$\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ is in Case 4 for Δ_{d} in $(\tilde{\delta}, \infty)$ and in Case 5 for Δ_{d} in $(c, \tilde{\delta}]$, where we take $c = \gamma$ if $R(1+\lambda) < u(1-\mu)$ and $c = \Delta_{\mathrm{uu}}$ if $R(1+\lambda) = u(1-\mu)$. We also note that $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) < 0$ in (γ, ∞) and $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) > 0$ in $(\Delta_{\mathrm{uu}}, \gamma)$.

If $R \geq d(1+\lambda)$, then the cost function $C(\Delta_{\mathrm{uu}}, \Delta_{\mathrm{d}}) = C(\Delta_{\mathrm{d}})$ in the two intervals $(c, \tilde{\delta}]$ and $(\tilde{\delta}, \infty)$ is given by

$$C(\Delta_{\mathrm{d}}) = \frac{p}{R}[\Delta_{\mathrm{uu}} S\bar{u} + B_{\mathrm{uu}}] + \frac{1-p}{R}\left[\Delta_{\mathrm{d}} Sd(1+\lambda) + \frac{b_{\mathrm{d}}(\Delta_{\mathrm{d}})}{R}\right],$$

where

$$b_{\mathrm{d}}(\Delta_{\mathrm{d}}) = \begin{cases} B_{\mathrm{dd}} + e(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}})Sd^2 & \text{if } f(\Delta_{\mathrm{d}}) < 0, \\ B_{\mathrm{uu}} - Sud(1-\mu)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}}) & \text{if } f(\Delta_{\mathrm{d}}) > 0, \end{cases}$$

$$0 < p = \frac{R - d(1+\lambda)}{\bar{u} - d(1+\lambda)} < 1, \qquad \bar{u} = \begin{cases} u(1-\mu) & \text{if } \Delta_{\mathrm{d}} \in (c, \tilde{\delta}], \\ u(1+\lambda) & \text{if } \Delta_{\mathrm{d}} \in (\tilde{\delta}, \infty). \end{cases}$$

So if $R(1+\lambda) < u(1-\mu)$, there is no minimum at $\tilde{\delta}$, because in (γ, ∞)

$$C'(\Delta_{\mathrm{d}}) = \frac{(1-p)Sd}{R^2} \begin{cases} (R-d)(1+\lambda) & \text{if } \Delta_{\mathrm{d}} < \Delta_{\mathrm{dd}} \\ R(1+\lambda) - d(1-\mu) & \text{if } \Delta_{\mathrm{d}} > \Delta_{\mathrm{dd}} \end{cases} > 0.$$

If $R(1+\lambda) = u(1-\mu)$,

$$C'(\Delta_{\mathrm{d}}) = \frac{(1-p)Sd}{R^2}[R(1+\lambda) - u(1-\mu)] = 0$$

in the interval $(\Delta_{\mathrm{uu}}, \gamma]$ and so if there is a minimum at $\tilde{\delta}$, there is also one at γ or at Δ_{uu} which is type (II)(a) of Theorem 7.

That leaves us with the case $R < d(1+\lambda)$. Then for $\Delta_{\mathrm{d}} > c$ we have the cost function

$$C(\Delta_{\mathrm{d}}) = \Delta_{\mathrm{d}} S + \frac{b_{\mathrm{d}}(\Delta_{\mathrm{d}})}{R^2},$$

which is linear in any interval in (c, ∞) which does not contain Δ_{dd} or γ. Hence the minimum is also attained at γ or Δ_{dd} or Δ_{uu}. Thus, we conclude that if the minimum is attained at $\tilde{\delta}$, then it is also attained at γ or Δ_{dd} or Δ_{uu}, the latter being of type (II)(a) of Theorem 7.

By considering the above four cases, we have shown that there is always a minimum of type (II) or the minimum occurs at Δ_{dd} or at a solution of $f_{\mathrm{d}}(\gamma) = 0$. We now consider the latter two possibilities in detail.

1. Suppose the minimum is assumed at $\Delta_{\mathrm{d}} = \Delta_{\mathrm{dd}}$ but $f_{\mathrm{d}}(\Delta_{\mathrm{dd}}) \neq 0$. If $f_{\mathrm{d}}(\Delta_{\mathrm{dd}}) > 0$, then, as $f_{\mathrm{d}}'(\Delta_{\mathrm{d}}) < 0$ for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{dd}}$, there exists a unique $\gamma > \Delta_{\mathrm{dd}}$ such that $f_{\mathrm{d}}(\gamma) = 0$. This implies that there is a positive number ε such that $\Delta_{\mathrm{uu}} < \Delta_{\mathrm{dd}} - \varepsilon$ and such that $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) \geq 0$ in $(\Delta_{\mathrm{dd}} - \varepsilon, \gamma]$. Also in this interval $a_{\mathrm{u}} > 0$ and throughout the interval either $a_{\mathrm{d}} > 0$ or $a_{\mathrm{d}} \leq 0$. So in the interval we are in one of the Cases 4, 5, 7, or 12 of Remark 1 and as also $b_{\mathrm{d}}(\Delta_{\mathrm{d}})$ is linear in the interval, it follows also that $C(\Delta_{\mathrm{d}})$ must be linear and hence constant if the minimum is at Δ_{dd}. Therefore, if $f_{\mathrm{d}}(\Delta_{\mathrm{dd}}) > 0$ there is also a minimum at γ for which $f_{\mathrm{d}}(\gamma) = 0$, which is the other possibility to be considered presently.

Suppose now that $f_{\mathrm{d}}(\Delta_{\mathrm{dd}}) < 0$. Then there exists $\tilde{\delta} \geq \Delta_{\mathrm{uu}}$ such that $a_{\mathrm{d}} \leq 0$ for $\Delta_{\mathrm{uu}} < \Delta_{\mathrm{d}} \leq \tilde{\delta}$ and $a_{\mathrm{d}} > 0$ for $\Delta_{\mathrm{d}} > \tilde{\delta}$. If $\tilde{\delta} < \Delta_{\mathrm{dd}}$, we choose ε so that $\tilde{\delta} < \Delta_{\mathrm{dd}} - \varepsilon$. Also we choose ε so that $\Delta_{\mathrm{uu}} < \Delta_{\mathrm{dd}} - \varepsilon$ and $f_{\mathrm{d}}(\Delta_{\mathrm{d}}) < 0$ in $(\Delta_{\mathrm{dd}} - \varepsilon, \Delta_{\mathrm{dd}})$. So throughout the latter interval, $a_{\mathrm{d}} > 0$ when $\tilde{\delta} < \Delta_{\mathrm{dd}}$ and $a_{\mathrm{d}} \leq 0$ when $\tilde{\delta} \geq \Delta_{\mathrm{dd}}$. As we also know that $a_{\mathrm{u}} > 0$ for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$, it follows that in the interval $(\Delta_{\mathrm{dd}} - \varepsilon, \Delta_{\mathrm{dd}})$, we are in one of Cases 4, 7, or 12 if $a_{\mathrm{d}} > 0$, and Case 5 if $a_{\mathrm{d}} \leq 0$. Note also that $R < u(1-\mu)$ if $a_{\mathrm{d}} \leq 0$, because we know that $R(1+\lambda) > u(1-\mu)$ implies that $a_{\mathrm{d}} > 0$ for $\Delta_{\mathrm{d}} > \Delta_{\mathrm{uu}}$. Hence, reasoning as we did for the interval $(c, \tilde{\delta}]$ in Case 4 above, we find that $C'(\Delta_{\mathrm{d}}) > 0$ in the interval $(\Delta_{\mathrm{dd}} - \varepsilon, \Delta_{\mathrm{dd}})$ if $R > d(1+\lambda)$. Then we must have $R \leq d(1+\lambda)$, in which case the initial holdings of the least cost superreplicating portfolio are $(\Delta_{\mathrm{dd}}, B_{\mathrm{dd}}/R^2)$. This is of type (I).

2. The final possibility is that the cost function $C(\Delta_{\mathrm{d}})$ has its minimum at $\Delta_{\mathrm{d}} = \gamma$ for which $(\gamma, b_{\mathrm{d}}(\gamma)/R)$ is a replicating portfolio for the one-period contingent claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}), (\Delta_{\mathrm{dd}}, B_{\mathrm{dd}})\}$ with initial stock price Sd. If $\gamma < \Delta_{\mathrm{uu}}$, then $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) < 0$ and so $a_{\mathrm{u}}(\gamma) < 0$. Also we know $a_{\mathrm{d}}(\gamma) > 0$. If $\gamma = \Delta_{\mathrm{uu}}$, then $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) = 0$ and so $a_{\mathrm{u}}(\gamma) = a_{\mathrm{d}}(\gamma) = 0$. Hence, if $\gamma \leq \Delta_{\mathrm{uu}}$, the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\gamma, b_{\mathrm{d}}(\gamma)/R)\}$ is in Case 2 of Remark 1 and so the initial holdings in the least cost superreplicating portfolio are those for the unique replicating portfolio for this one-period claim. This is of type (IV).

In contrast, if $\gamma > \Delta_{\mathrm{uu}}$, then $a_{\mathrm{u}}(\gamma) > 0$ and if $a_{\mathrm{d}}(\gamma) \leq 0$ then $R(1+\lambda) \leq u(1-\mu)$ so that $R \leq u(1-\mu)$. Thus the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\gamma, b_{\mathrm{d}}(\gamma)/R)\}$ is in one of Cases 4, 7, or 12 of Remark 1 if $a_{\mathrm{d}} > 0$ and in Case 5 if $a_{\mathrm{d}} \leq 0$ with $R < u(1-\mu)$. Therefore if $R \geq d(1+\lambda)$, the initial holdings in the least cost superreplicating portfolio are those for the unique replicating portfolio for this one-period claim.

This is again of type (IV). On the other hand, if $R < d(1+\lambda)$, the initial holdings are $(\gamma, b_{\mathrm{d}}(\gamma)/R^2)$. This is of type (III).

So the proof of the theorem is complete.

5. An Example with Path-Dependent Least Cost Superreplicating Portfolios

In a two-period binomial model with parameters S, u, d, R, λ, and μ, we consider a *short position in a put option* with exercise price K satisfying

$$Sd^2 < K < Sud.$$

This is the contingent claim $\{(0,0), (0,0), (1, -K)\}$. It follows from Theorem 7 that there is a least cost superreplicating portfolio and we need only consider the following possibilities for such a portfolio:

(I) the initial holdings are $(1, -K/R^2)$ (only arises if $R < d(1+\lambda)$ and $K < Sud(1-\mu)$);

(II) the initial holdings are (δ, B), where $\delta \leq 0$ and there are two possibilities: $\delta = 0$ which only occurs if $K \geq Sd^2(1+\lambda)$ and then $B = 0$ also; $\delta < 0$ which only occurs if $K < Sd^2(1+\lambda)$ and then δ and B satisfy $BR = -\delta Su(1+\lambda)$ and $BR^2 - (1-\delta)Sd^2(1+\lambda) = -K$;

(III) the initial holdings are $(\alpha, B/R)$, where $\alpha > 0$ and (α, B) are the initial holdings in a replicating portfolio for the one-period portion $\{d, du, dd\}$ (only arises if $R < d(1+\lambda)$);

(IV) a replicating portfolio for the whole two-period model.

Example.
Consider a two-period model with $u = 1.1$, $d = 0.95$, $R = 1.05$, $\lambda = \mu = 0.06$, and $S = 100$. Consider the put with exercise price 93 which is between 90.25 and 104.50. A short position in this put is the claim $\{(0,0), (0,0), (1, -93)\}$.

$$\begin{array}{ccccc}
 & & 121 & & (0,0) \\
 & 110 & & & \\
100 & & 104.50 & & (0,0) \\
 & 95 & & & \\
 & & 90.25 & & (1,-93)
\end{array}$$

Note first that as $R > d(1+\lambda)$, we do not need to consider (I) or (III). We consider (II) first. As $K < Sd^2(1+\lambda)$, the only possibility is that there exist $\delta < 0$ and B such that

$$BR^2 = -\delta SuR(1+\lambda) = (1-\delta)Sd^2(1+\lambda) - K,$$

that is,

$$1.05^2 B = -\delta 110 \times 1.05 \times 1.06 = (1-\delta) \times 90.25 \times 1.06 - 93.$$

We solve the last equation to get $\delta = -0.0996$ and then $B = 12.1940/1.05^2$. Then the initial holdings are $(-0.0996, 12.1940/1.05^2)$, which has cost 1.1003.

The only other possibility is (IV), a replicating portfolio for the whole model. We find that the one-period claim $\{(0,0), (1, -93)\}$ with initial stock price 95 has the unique replicating portfolio $(-0.1764, 18.1627)$. Next we need to determine the replicating portfolio for the one-period claim $\{(0,0), (-0.1764, 18.1627)\}$ with initial stock price 100. It turns out that this has cost 1.16. Hence the least cost is 1.1003. Note that this least cost superreplicating portfolio is path-dependent.

Now we show that the example just given is a special case of a situation in which there is a unique least cost superreplicating portfolio and it is path-dependent.

Theorem 8. *Consider a two-period binomial model with parameters* S, u, d, R, μ, *and* λ *satisfying*

$$d(1+\lambda) < u(1-\mu), \qquad R(1-\mu) < d(1+\lambda) < R.$$

For every contingent claim with terminal holdings $\{(\Delta_{uu}, B_{uu}), (\Delta_{ud}, B_{ud}), (\Delta_{dd}, B_{dd})\}$ *satisfying* $\Delta_{uu} = \Delta_{ud} < \Delta_{dd}$, $B_{uu} = B_{ud}$, *and*

$$B_{dd} - B_{uu} - Sd^2(1+\lambda)(\Delta_{uu} - \Delta_{dd}) > 0, \tag{4}$$

there exists a unique least cost superreplicating portfolio. Moreover, this portfolio is path-dependent.

Proof. Let (Δ_u, B_u) and (Δ_d, B_d) be the holdings in a least cost superreplicating portfolio at the end of the first period. Then we know that the initial holdings (Δ, B) form a least cost superreplicating portfolio for the one-period claim $\{(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ with initial stock price S. We have to show that there is just one possibility for $\{(\Delta, B), (\Delta_{\mathrm{u}}, B_{\mathrm{u}}), (\Delta_{\mathrm{d}}, B_{\mathrm{d}})\}$.

We first show it is necessary that

$$\Delta_{\mathrm{u}} = \Delta_{\mathrm{uu}}.$$

As $d(1+\lambda) < u(1-\mu)$, the one-period claim $\{(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ is in one of Cases 1–6 of Remark 1. Denote by $C(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$ the cost of its least cost superreplicating portfolio. As $R > d(1+\lambda)$, it follows from Theorem 4(a) that the least cost superreplicating portfolio for the one-period claim is either the unique replicating portfolio with corresponding p satisfying $0 < p < 1$ or $(\Delta_{\mathrm{u}}, b_{\mathrm{u}}(\Delta_{\mathrm{u}})/R^2)$. Hence, referring to the proof of Theorem 5.1 of Chen *et al.* (2004), where here we observe that $\alpha_{\mathrm{u}} = \beta_{\mathrm{u}} = \Delta_{\mathrm{uu}}$, we find that for fixed Δ_{d},

$$\frac{\partial C}{\partial \Delta_{\mathrm{u}}} \begin{cases} < 0 & \text{if } \Delta_{\mathrm{u}} < \Delta_{\mathrm{uu}}, \\ > 0 & \text{if } \Delta_{\mathrm{u}} > \Delta_{\mathrm{uu}}, \end{cases}$$

and so we must have $\Delta_{\mathrm{u}} = \Delta_{\mathrm{uu}}$ at a minimum. This also means that $b_{\mathrm{u}}(\Delta_{\mathrm{u}}) = b_{\mathrm{u}}(\Delta_{\mathrm{uu}}) = B_{\mathrm{uu}}$.

Next we determine the zeros of the function f_{d}. As $d(1+\lambda) < u(1-\mu)$, $f_{\mathrm{d}}(\Delta_{\mathrm{d}})$ is strictly decreasing and Equation (4) says that $f_{\mathrm{d}}(\Delta_{\mathrm{uu}}) < 0$. Hence there is a unique γ such that $f_{\mathrm{d}}(\gamma) = 0$ and $\gamma < \Delta_{\mathrm{uu}}$. So

$$b_{\mathrm{d}}(\Delta_{\mathrm{d}}) = \begin{cases} B_{\mathrm{uu}} - Sud(1+\lambda)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{uu}}) & \text{if } \quad \Delta_{\mathrm{d}} \le \gamma, \\ B_{\mathrm{dd}} - Sd^2(1+\lambda)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}}) & \text{if } \gamma \le \Delta_{\mathrm{d}} \le \Delta_{\mathrm{dd}}, \\ B_{\mathrm{dd}} - Sd^2(1-\mu)(\Delta_{\mathrm{d}} - \Delta_{\mathrm{dd}}) & \text{if } \quad \Delta_{\mathrm{dd}} \le \Delta_{\mathrm{d}}. \end{cases}$$

Next it follows from the proof of Theorem 7 that there is a δ with $\gamma < \delta < \Delta_{\mathrm{dd}}$ such that

$$a_{\mathrm{u}}(\Delta_{\mathrm{d}}) = \begin{cases} < 0 & \text{if } \Delta_{\mathrm{d}} < \delta, \\ = 0 & \text{if } \Delta_{\mathrm{d}} = \delta, \\ > 0 & \text{if } \delta < \Delta_{\mathrm{d}}, \end{cases}$$

and $a_{\mathrm{d}}(\Delta_{\mathrm{d}}) > 0$ for all Δ_{d}.

Hence if $\Delta_{\mathrm{d}} \le \delta$, we have $a_{\mathrm{d}} > 0$ and $a_{\mathrm{u}} < 0$ and the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\Delta_{\mathrm{d}}, b_{\mathrm{d}}(\Delta_{\mathrm{d}})/R)\}$ is in Case 2 of Remark 1. If $\delta < \Delta_{\mathrm{d}}$, we have $a_{\mathrm{d}} > 0$ and $a_{\mathrm{u}} > 0$ and the claim is in Case 1 or 4 of Remark 1. In all cases, it follows from Theorem 4 and Remark 1 that the least cost superreplicating portfolio for the claim is the unique replicating portfolio so that the least cost C is given by

$$C = \frac{p}{R}\left[\Delta_{\mathrm{uu}} Su(1+\lambda) + \frac{B_{\mathrm{uu}}}{R}\right] + \frac{1-p}{R}\left[\Delta_{\mathrm{d}} S\bar{d} + \frac{b_{\mathrm{d}}(\Delta_{\mathrm{d}})}{R}\right],$$

where

$$p = \frac{R - \bar{d}}{u(1+\lambda) - \bar{d}}, \quad \bar{d} = \begin{cases} d(1-\mu) & \text{if } \Delta_{\mathrm{d}} \le \delta, \\ d(1+\lambda) & \text{if } \Delta_{\mathrm{d}} \ge \delta. \end{cases}$$

Hence

$$\frac{\partial C}{\partial \Delta_{\mathrm{d}}} = \frac{1-p}{R^2} \begin{cases} SRd(1-\mu) - Sud(1+\lambda) < 0 & \text{if } \Delta_{\mathrm{d}} < \gamma, \\ SRd(1-\mu) - Sd^2(1+\lambda) < 0 & \text{if } \gamma < \Delta_{\mathrm{d}} < \delta, \\ SRd(1+\lambda) - Sd^2(1+\lambda) > 0 & \text{if } \delta < \Delta_{\mathrm{d}} < \Delta_{\mathrm{dd}}, \\ SRd(1+\lambda) - Sd^2(1-\mu) > 0 & \text{if } \Delta_{\mathrm{d}} > \Delta_{\mathrm{dd}}. \end{cases}$$

It follows that the unique minimum is achieved at δ.

Thus, we have shown that $C(\Delta_{\mathrm{u}}, \Delta_{\mathrm{d}})$ achieves its unique minimum at $(\Delta_{\mathrm{uu}}, \delta)$. This means that a least cost superreplicating portfolio for our two-period claim has share holdings Δ_{uu} and δ at the end of the first period and initial holdings which constitute a least cost superreplicating portfolio for the one-period claim $\{(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R), (\delta, b_{\mathrm{d}}(\delta)/R)\}$. Morever, as seen above, the least cost superreplicating portfolio for this one-period claim is the unique replicating portfolio $(\delta, b_{\mathrm{d}}(\delta)/R^2)$.

Now we show that this portfolio, consisting of initial holdings $(\delta, b_{\mathrm{d}}(\delta)/R^2)$ and end of first-period holdings $(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R)$ and $(\delta, b_{\mathrm{d}}(\delta)/R)$, is path-dependent. If it were path-independent, there would be terminal holdings (Δ, B) in the ud state with $\Delta \ge \Delta_{\mathrm{uu}}$, $B \ge B_{\mathrm{uu}}$ such that when the stock price moves from Su to Sud we could rebalance the holdings $(\Delta_{\mathrm{uu}}, B_{\mathrm{uu}}/R)$ in state u to get (Δ, B), and when the stock price moves from Sd to Sud we could rebalance the holdings $(\delta, b_{\mathrm{d}}(\delta)/R)$ in state d to get (Δ, B) also. Thus, we would need

$$B = B_{\mathrm{uu}} - (\Delta - \Delta_{\mathrm{uu}})Sud(1+\lambda) = b_{\mathrm{d}}(\delta) - (\Delta - \delta)Sud(1+\lambda).$$

As $\Delta \ge \Delta_{\mathrm{uu}}$ and $B \ge B_{\mathrm{uu}}$, the first equation implies that $\Delta = \Delta_{\mathrm{uu}}$, $B = B_{\mathrm{uu}}$ and so the second equation can be written as

$$B_{\mathrm{uu}} = b_{\mathrm{d}}(\delta) - (\Delta_{\mathrm{uu}} - \delta)Sud(1+\lambda).$$

However, $a_{\mathrm{u}}(\delta) = 0$. Thus,

$$(\delta - \Delta_{\mathrm{uu}}) Su(1+\lambda) + \frac{b_{\mathrm{d}}(\delta) - B_{\mathrm{uu}}}{R} = 0.$$

From the last two equations we find that $R = d$. This contradiction shows that the least cost superreplicating portfolio is path-dependent.

References

Bensaid, B., J. P. Lesne, H. Pages and J. Scheinkman, "Derivative Asset Pricing with Transaction Costs." *Mathematical Finance* 2, 63–86 (1992).

Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81, 637–659 (1973).

Boyle, P. and T. Vorst, "Option Replication in Discrete Time with Transaction Costs."*Journal of Finance* 47, 271–293 (1992).

Chen, G.-Y., K. Palmer and Y.-C. Sheu, "The Least Cost Super Replicating Portfolio in the Boyle–Vorst Model with Transaction Costs." Submitted to *International Journal of Theoretical and Applied Finance* (revised) (2006).

Chiang, T. S. and S. J. Sheu, "*A Geometric Approach to Option Pricing with Transaction Costs in Discrete Models*." Preprint (2004).

Leland, H. E., "Option Pricing and Replication with Transaction Costs." *Journal of Finance* 40, 1283–1301 (1985).

Palmer, K. J., "A Note on the Boyle–Vorst Discrete-Time Option Pricing Model with Transactions Costs." *Mathematical Finance* 11, 357–363 (2001a).

Palmer, K. J., "*Replicating and Super Replicating Portfolios in the Boyle–Vorst Discrete-time Option Pricing Model with Transactions Costs*." Research Report, Department of Finance, University of Melbourne (2001b).

Perrakis, S. and J. Lefoll, "Derivative Asset Pricing with Transactions Costs: An Extension." *Computational Economics* 10, 359–376 (1997).

Perrakis, S. and J. Lefoll, "Option Pricing and Replication with Transactions Costs and Dividends." *Journal of Economic Dynamics & Control* 24, 1527–1561 (2000).

Reiss, A., "Option Replication with Large Transactions Costs." *OR Spektrum* 21, 49–70 (1999).

Rutkowski, M., "Optimality of Replication in the CRR Model with Transaction Costs." *Applicationes Mathematicae* 25, 29–53 (1998).

Stettner, L., "Option Pricing in the CRR Model with Proportional Transaction Costs: A Cone Transformation Approach." *Applicationes Mathematicae* 24, 475–514 (1997).

Whalley, A. E. and Wilmott "Key Results in Discrete Hedging and Transaction Costs." In *Frontiers in Derivatives*, Konishi, A. and R. E. Dattatreya, eds., Irwin, Chicago, 183–195 (1997).

Chapter 2

Testing of Nonstationarities in the Unit Circle, Long Memory Processes, and Day of the Week Effects in Financial Data

Guglielmo Maria Caporale*
Brunel University, London

Luis A. Gil-Alana and Mike Nazarski
Universidad de Navarra, Spain

This paper examines a version of the tests of Robinson (1994) that enables one to test models of the form $(1 - L^k)^d x_t = u_t$, where k is an integer value, d may be any real number, and u_t is I(0). The most common cases are those with $k = 1$ (unit or fractional roots) and $k = 4$ and 12 (seasonal unit or fractional models). However, we extend the analysis to cover situations such as $(1–L^5)^d x_t = u_t$, which might be relevant, for example, in the context of financial time series data. We apply these techniques to the daily Eurodollar rate and Dow Jones index, and find that for the former series the most adequate specifications are either a pure random walk or a model of the form $x_t = x_{t-5} + \varepsilon_t$, implying in both cases that the returns are completely unpredictable. In the case of Dow Jones index, a model of the form $(1 - L^5)^d x_t = u_t$ is selected, with d constrained between 0.50 and 1, implying nonstationarity and mean-reverting behavior.

Keywords: Fractional integration; seasonality; long memory; day of the week effects; Monte Carlo simulations.

JEL Classifications: C22, C15.

1. Introduction

Nonstationarity is a characteristic of many economic and financial time series. The unit root polynomial is the most natural way of modeling this behavior: once a time series is first-differenced, it is assumed to be stationary, or, more precisely, integrated of order 0 (denoted by $I(0)$). For the purpose of this paper, we define an $I(0)$ process in its most general form, i.e., as a covariance stationary process with a spectral density function that is positive and finite not only at zero, but also at any frequency on the spectrum.

*Corresponding author.

Following the work of Box and Jenkins (1970), the unit root model became very popular, especially after the seminal paper of Nelson and Plosser (1982). Using the tests of Fuller (1976) and Dickey and Fuller (1979), these authors were unable to reject the presence of a unit root in 14 US macroeconomic series. Such tests and others that have been proposed later (Phillips, 1987; Phillips and Perron, 1988; Kwiatkowski *et al.*, 1992)[1] are based on autoregressive (AR) alternatives, in the sense that they are nested in a model of the form

$$(1 - \alpha L)x_t = u_t, \quad t = 1, 2, \ldots, \tag{1}$$

where L is the lag operator (i.e., $Lx_t = x_{t-1}$) and the unit root corresponds to the null $\alpha = 1$. However, a problem with these procedures is that their limit distributions are nonstandard and, therefore, the critical values have to be obtained numerically by performing simulations. This is a consequence of the abrupt change in the asymptotic behavior of the distributions around the unit root. Note that x_t in Equation (1) is stationary for $|\alpha| < 1$, nonstationary but nonexplosive for $|\alpha| = 1$, and becomes explosive for $|\alpha| > 1$. This has motivated the use of other models for testing unit roots. In particular, fractional processes have been considered as an alternative to AR models. In this context, the unit root hypothesis is tested within the model:

$$(1 - L)^d x_t = u_t, \quad t = 1, 2, \ldots, \tag{2}$$

where d may be any real number and the unit root case corresponds to $d = 1$. Here, the limit behavior is smooth around the unit root null, and $d = 0.5$ becomes the crucial parameter to distinguish between stationarity and nonstationarity. Examples of testing procedures using the fractional model (2) are, among others, Sowell (1992), Robinson (1994, 1995a,b), Tanaka (1999), and Dolado *et al*. (2001).

Many series also contain seasonal fluctuations, and stochastic models based on seasonal unit roots have been proposed in recent years. As in the previous case, most of the available procedures (Dickey *et al.*, 1984; Hylleberg *et al.*, 1990; etc.) are based on AR alternatives and, therefore, face the same problem concerning the limit behavior of the distribution. New approaches

[1] Note that, strictly speaking, KPSS is not a unit root test, because the null hypothesis is stationarity (around either a level or a linear trend), whereas the alternative is a unit root.

based on (seasonal) fractional integration have been proposed by Porter-Hudak (1990) and Robinson (1994). They consider processes of the form

$$(1 - L^k)^d x_t = u_t, \quad t = 1, 2, \ldots, \tag{3}$$

where k is the number of time periods within the year and L^k the seasonal lag operator ($L^k x_t = x_{t-k}$). Note that the polynomial on the left-hand side of Equation (3) can be expressed in terms of its binomial expansion, such that, for all real d,

$$(1 - L^k)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^{kj} = 1 - dL^k + \frac{d(d-1)}{2} L^{2k} - \cdots .$$

Clearly, if $d = 0$ in Equation (3), $x_t = u_t$, and a weakly autocorrelated x_t is allowed for. If $d > 0$, the process is said to be a long memory one, so named because of the strong association between observations far apart in time, and the higher the value of d, the stronger will be the association.

The literature on seasonal processes has usually concentrated on the cases of $k = 4$ (quarterly) or $k = 12$ (monthly observations). In this paper, we examine different versions of the tests of Robinson (1994) that enable us to test models like Equation (3) for any integer k and any real value d. Thus, we consider unit (and fractional) orders of integration at the zero and the seasonal frequencies, but also at other frequencies in the interval $(0, \pi]$. Specifically, we analyze the following processes:

(i) $I(1)$ processes: when $k = d = 1$ (see Dickey and Fuller, 1979; Phillips and Perron, 1988; Kwiatkowski *et al.* 1992).
(ii) $I(d)$ processes: when $k = 1$ and d is a real value (see Diebold and Rudebusch, 1989; Baillie, 1996; Gil-Alana and Robinson, 1997).
(iii) Seasonal unit roots: when $k = 4$ and $d = 1$ (see Dickey *et al.*, 1984; Hylleberg *et al.*, 1992; Beaulieu and Miron, 1992, for $k = 12$).
(iv) Seasonal fractional models: when $k = 4$, 12 and d is a real value (see Porter-Hudak, 1990; Ray, 1993; Sutcliffe, 1994),

but also processes of the form $(1 - L^3)^d x_t = u_t$, or $(1 - L^5)^d x_t = u_t$.

Let us focus on the last type of model. If $d = 1$, this implies that the present value of the series (x_t) depends exclusively on its value five periods before (x_{t-5}), and if d is real, it will depend not only on x_{t-5} but on all past observations which are backwards multiples of 5, i.e., $x_{t-10}, x_{t-15}, \ldots$.

This type of model is relevant, for example, in the context of daily financial data, where the value of an asset on a given day of the week may be strongly influenced by its value on the same day of the previous week. There is in fact an extensive literature documenting the presence of calendar anomalies (such as the weekend effect, the day of the week effect, and the January effect) in financial series, both in the US and in other developed markets, dating back to Osborne (1962). Negative Monday returns were found, *inter alia*, by Cross (1973), French (1980), and Gibbons and Hess (1981), the former two analyzing the S&P 500 index, the latter the Dow Jones Industrial Index. Similar findings have been reported for other US financial markets, such as the futures, bond and Treasury bill markets (Cornell, 1985; Dyl and Maberly, 1986), foreign exchange markets (Hsieh, 1988), and for Australian, Canadian, Japanese, and UK financial markets (e.g., Jaffe and Westerfield, 1985; Jaffe, Westerfield, and Ma, 1989; Agrawal and Tandon, 1994). Effects on stock market volatility have also been documented (Kiymaz and Berument, 2003).

Various explanations have been offered for the observed patterns. Some focus on delays between trading and settlement in stocks (Gibbons and Hess, 1981): buying on Fridays creates a 2-day interest-free loan until settlement; hence, there are higher transaction volumes on Fridays, resulting in higher prices, which decline over the weekend as this incentive disappears. Others emphasize a shift in the broker–investor balance in buying–selling decisions which occur on weekends, when investors have more time to study the market themselves (rather than rely on brokers); this typically results in net sales on Mondays, when liquidity is low in the absence of institutional trading (Miller, 1988). It has also been suggested that the Monday effect largely reflects the fact that, when daily returns are calculated, the clustering of dividend payments around Mondays is normally ignored; alternatively, it could be a consequence of positive news typically being released during the week, and negative ones over the weekend (Fortune, 1998). Additional factors which could be relevant are serial correlation, with Monday prices being affected by Friday ones, and a negative stock performance on Fridays being given more weight (Abraham and Ikenberry, 1994), measurement errors (Keim and Stambaugh, 1984), size (Fama and French, 1992), and volume (Lakonishok and Maberly, 1990).

Further empirical evidence on weekday effects is provided in this study using fractional integration techniques. In particular, we use a methodology that allows us to consider fractional processes where the dependence between the observations is a function of a specific day of the week, not only Monday as in earlier studies, but any day of the week. Of particular interest is the order of (fractional) integration of the series. If it is smaller than 1, shocks affecting the weekly structure will be mean reverting, and their effects will disappear in the long run. On the other hand, if it is 1 or higher, shocks will persist forever, and strong measures will be required to bring the variable back to its original level.

The outline of the paper is as follows: Section 2 describes the different versions of the tests of Robinson (1994) we employ. Section 3 presents a Monte Carlo simulation study, examining the size and power properties of these tests in finite samples. Section 4 discusses two empirical applications based on financial data and Section 5 offers some concluding remarks.

2. Testing of Nonstationarities in the Unit Circle

Following Bhargava (1986), Schmidt and Phillips (1992), and other studies in the parameterization of unit roots, Robinson (1994) considers the regression model

$$y_t = \beta' z_t + x_t, \quad t = 1, 2, \ldots, \tag{4}$$

where y_t is the time series we observe, β a $(kx1)$ vector of unknown parameters, z_t a $(kx1)$ vector of regressors, and x_t are the regression errors, taking the form

$$\rho(L; \theta)x_t = u_t, \quad t = 1, 2, \ldots, \tag{5}$$

where ρ is a scalar function that depends on L and the unknown parameter θ that will adopt different forms as below, and u_t is $I(0)$. The function ρ is specified in such a way that all its roots should be on the unit circle in the complex plane, and, therefore, it includes polynomials of the form $(1-L^k)^{d+\theta}$, where k is an integer and d may be a real value. Thus, in what follows, we assume that

$$\rho(L; \theta) = (1 - L^k)^{d+\theta}. \tag{6}$$

Robinson (1994) proposed a Lagrange multiplier (LM) test of the null hypothesis

$$\mathrm{H_o} : \theta = 0, \tag{7}$$

in a model given by Equations (4)–(6). Based on H_o (Equation (7)), the estimated β and residuals are

$$\hat{u}_t = (1 - L^k)^{d_o} y_t - \hat{\beta}' w_t, \quad w_t = (1 - L^k)^{d_o} z_t,$$
$$\hat{\beta} = \left(\sum_{t=1}^{T} w_t w_t' \right)^{-1} \sum_{t=1}^{T} w_t (1 - L^k)^{d_o} y_t.$$

The functional form of the test statistics is given by

$$\hat{r} = \frac{T^{1/2}}{\hat{\sigma}^2} \hat{A}^{-1/2} \hat{a}, \tag{8}$$

where T is the sample size and

$$\hat{A} = \frac{2}{T} \left(\sum_{j=1}^{*} \psi(\lambda_j)^2 - \sum_{j=1}^{*} \psi(\lambda_j) \hat{\varepsilon}(\lambda_j)' \times \left(\sum_{j=1}^{*} \hat{\varepsilon}(\lambda_j) \hat{\varepsilon}(\lambda_j)' \right)^{-1} \right.$$
$$\left. \times \sum_{j=1}^{*1} \hat{\varepsilon}(\lambda_j) \psi(\lambda_j) \right),$$
$$\hat{a} = \frac{-2\pi}{T} \sum_{j=1}^{*} \psi(\lambda_j) g(\lambda_j; \hat{\tau})^{-1} I(\lambda_j), \quad \hat{\sigma}^2 = \sigma^2(\hat{\tau}) \tag{9}$$
$$= \frac{2\pi}{T} \sum_{j=1}^{T-1} g(\lambda_j; \hat{\tau})^{-1} I(\lambda_j),$$
$$\hat{\varepsilon}(\lambda_j) = \frac{\partial}{\partial \tau} \log g(\lambda_j; \hat{\tau}), \quad \lambda_j = \frac{2\pi j}{T}, \quad \hat{\tau} = \arg\min_{\tau \in T^*} \sigma^2(\tau),$$

and the sums over $*$ in the above expressions are over $\lambda \in M$ where $M = \{\lambda : -\pi < \lambda < \pi, \lambda \notin (\rho_l - \lambda_1, \rho_l + \lambda_1), l = 1, 2, \ldots, s\}$ such that $\rho_l, l = 1, 2, \ldots, s < \infty$ are the distinct poles of $\psi(\lambda)$ on $(-\pi, \pi]$. Also,

$$\psi(\lambda_j) = \mathrm{Re} \left[\log \left(\frac{\partial}{\partial \theta} \log \rho(e^{i\lambda_j}; \theta) \right) \right] \theta = 0, \tag{10}$$

and $I(\lambda_j)$ is the periodogram of u_t evaluated under the null. The function g above is a known function coming from the spectral density of u_t,

$$f(\lambda; \sigma^2; \tau) = \frac{\sigma^2}{2\pi} g(\lambda; \tau), \quad -\pi < \lambda \leq \pi.$$

Note that these tests are purely parametric and, therefore, they require specific modeling assumptions about the short memory specification of u_t. Thus, if u_t is a white noise, then $g \equiv 1$ (and thus, $\hat{\varepsilon}(\lambda_j) = 0$), and if it is an AR process of the form $\phi(L)u_t = \varepsilon_t$, $g = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, so that the AR coefficients are a function of τ.

Based on H$_o$ Equation (7), Robinson (1994) showed that under certain very mild regularity conditions:[2]

$$\hat{r} \to_d N(0, 1) \text{ as } T \to \infty. \tag{11}$$

Hence, we are in a classical large sample-testing situation: an approximate one-sided $100\alpha\%$ level test of H$_o$ Equation (7) against the alternative: H$_a$: $d > d_o (d < d_o)$ will be given by the rule: "Reject H$_o$ if $\hat{r} > z_\alpha (\hat{r} < -z_\alpha)$," where the probability that a standard normal variate exceeds z_α is α.

Note that given the functional form of ρ in Equation (6),

$$\begin{aligned}\psi(\lambda_j) &= \mathrm{Re}\left[\left(\frac{\partial}{\partial\theta}\log\rho(\mathrm{e}^{\mathrm{i}\lambda_j};\theta)\right)\right]\theta = 0 = \mathrm{Re}\left[\frac{\partial}{\partial\theta}(d+\theta)\log(1-\mathrm{e}^{\mathrm{i}\lambda_j k})\right] \\ &= \mathrm{Re}\left[\log(1-\mathrm{e}^{\mathrm{i}\lambda_j k})\right] = \mathrm{Re}\left[\log(1-\cos\lambda_j k - \mathrm{i}\sin\lambda_j k\right] \\ &= \log\left|(1-\cos\lambda_j k - \mathrm{i}\sin\lambda_j\right| = \log\left(2-2\cos\lambda_j k\right)^{0.5}.\end{aligned}$$

In some simple cases, the above formula simplifies. Thus, for example, if $k = 1$,

$$\psi(\lambda_j) = \log\left|2\sin\frac{\lambda_j}{2}\right|.$$

If $k = 2$, and noting that $(1-\mathrm{e}^{\mathrm{i}2\lambda}) = (1-\mathrm{e}^{\mathrm{i}\lambda})(1-\mathrm{e}^{-\mathrm{i}\lambda})$,

$$\psi(\lambda_j) = \log\left|2\sin\frac{\lambda_j}{2}\right| + \log\left(2\cos\frac{\lambda_j}{2}\right).$$

and similarly, if $k = 4$,

$$\psi(\lambda_j) = \log\left|2\sin\frac{\lambda_j}{2}\right| + \log\left(2\cos\frac{\lambda_j}{2}\right) + \log\left|2\cos\lambda_j\right|.$$

[2] These conditions are technical and refer to the functional form of $\psi(\lambda_j)$ in the specification of the test statistics. They are satisfied by the model given by Equations (4)–(6).

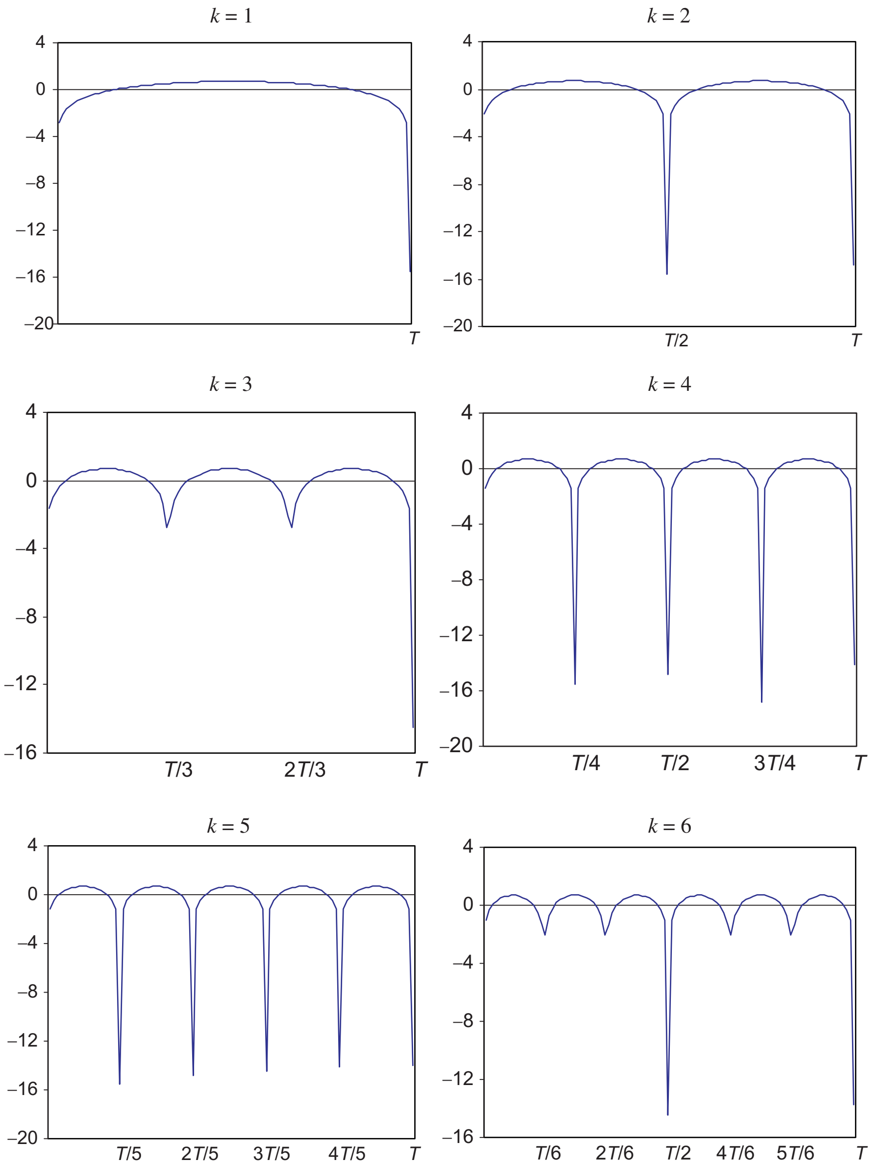


Figure 1. $\psi(\lambda)$ functions for different values of k and $T = 100$.

A common feature of all these expressions is that they have a finite number (k) of poles across the spectrum, but they are all squared integrable. Figure 1 displays plots of the $\psi(\lambda)$ functions for $k = 1, 2, 3, 4, 5$, and 6. The poles are clearly identified in the plots.

Various forms of the test statistics described above have been applied in empirical studies. For example, Gil-Alana and Robinson (1997) used a version of the above tests with $k = 1$ and d being a real number, and Gil-Alana (1999) and Gil-Alana and Robinson (2001) extended the analysis to the case of $k = 12$ and 4, respectively. However, there are no previous empirical studies considering, as in the present paper, the case of $k = 5$, which is particularly relevant in the context of financial daily data for the reasons outlined in Section 1.

3. A Monte Carlo Simulation Study

We start by examining the size of the above versions of the tests of Robinson (1994). We consider six null models of the form:

$$(1 - L^k)x_t = u_t, \quad t = 1, 2, \ldots,$$

with $k = 1, 2, 3, 4, 5$, and 6, with white noise u_t. We generate Gaussian series using the routines GASDEV and RAN3 of Press, Flannery, Teukolsky, and Wetterling (1986), with 10,000 replications in each case. The sample size is $T = 50$, 100, 300, 500, and 1000 observations, and the nominal size is 10% in Table 1 and 5% in Table 2.

We can see from these two tables that if the sample size is small (e.g., $T = 50$), there is a significant positive bias, especially for the cases of $k = 3, 4$, and 6. However, as the sample size increases, the empirical sizes approximate the nominal ones. Thus, for example, if $T = 1000$, the sizes range between 10.6% and 11.8% in Table 1, and between 5.3% and 6.8% in Table 2.

Table 1. Sizes of the different versions of the tests with a nominal size of 10%.

k/T	50	100	300	500	1000
1	0.153	0.132	0.110	0.104	0.106
2	0.299	0.214	0.143	0.128	0.113
3	0.377	0.233	0.191	0.168	0.116
4	0.367	0.261	0.152	0.156	0.114
5	0.251	0.204	0.139	0.132	0.110
6	0.328	0.230	0.143	0.124	0.118

Note: The nominal size is 10%, and 10,000 replications were used in each case.

Table 2. Sizes of the different versions of the tests with a nominal size of 5%.

k/T	50	100	300	500	1000
1	0.073	0.063	0.056	0.053	0.053
2	0.102	0.097	0.074	0.065	0.057
3	0.145	0.111	0.069	0.094	0.066
4	0.135	0.105	0.098	0.094	0.068
5	0.131	0.111	0.091	0.077	0.056
6	0.161	0.112	0.089	0.078	0.055

Note: The nominal size is 5%, and 10,000 replications were used in each case.

Table 3. Rejection frequencies of the different versions of the tests of Robinson (1994), with $T = 100$.

k/θ	−1.00	−0.75	−0.50	−0.25	0.25	0.50	0.75	1.00
1	1.000	1.000	0.997	0.795	0.898	0.999	1.000	1.000
2	0.995	0.994	0.979	0.619	0.943	0.999	1.000	1.000
3	1.000	1.000	0.991	0.726	0.515	0.816	0.932	0.998
4	0.917	0.893	0.741	0.649	0.950	0.999	1.000	1.000
5	0.982	0.973	0.903	0.526	0.541	0.878	0.959	0.998
6	0.828	0.731	0.511	0.443	0.992	1.000	1.000	1.000

Note: 10,000 replications were used in each case.

Tables 3–5 display the rejection probabilities of the above versions of the tests when looking at alternatives of the form given in Equation (6) with $d = 1$ and $\theta = -1, -0.75, -0.50, -025, 0.25, 0.50, 0.75,$ and 1, assuming that k is correctly specified, that is, it is the same under both the null and the alternative hypotheses.

Table 3 reports the results for $T = 100$, whereas Tables 4 and 5 refer to $T = 300$ and 500, respectively. We observe that even for the smallest sample size ($T = 100$, in Table 3), the rejection probabilities are high even for small departures from the null. The lowest values are obtained for the cases of $k = 3$ and 5 (with $\theta > 0$) and $k = 6$ (with $\theta < 0$). However, if $\theta \leq -0.50$ or $\theta \geq 0.50$, the values are higher than 0.90 in all cases. If $T = 300$ (Table 4), the rejection probabilities exceed 0.90 even for $\theta = \pm 0.25$, and they are all 1 with $|\theta| > 0.50$. Finally, if $T = 500$ (Table 5), the values are equal to 1 in all cases.

Table 4. Rejection frequencies of the different versions of the tests of Robinson (1994), with $T = 300$.

k/θ	−1.00	−0.75	−0.50	−0.25	0.25	0.50	0.75	1.00
1	1.000	1.000	1.000	0.998	0.999	1.000	1.000	1.000
2	1.000	1.000	1.000	0.997	0.999	1.000	1.000	1.000
3	1.000	1.000	1.000	0.996	0.994	1.000	1.000	1.000
4	1.000	1.000	1.000	0.962	0.998	1.000	1.000	1.000
5	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000
6	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000

Note: 10,000 replications were used in each case.

Table 5. Rejection frequencies of the different versions of the tests of Robinson (1994), with $T = 500$.

k/θ	−1.00	−0.75	−0.50	−0.25	0.25	0.50	0.75	1.00
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: 10,000 replications were used in each case.

In brief, the versions of the tests of Robinson (1994) considered in this paper appear to perform well in finite samples, especially if the sample size is relatively large (i.e., $T \geq 300$).

4. Two Empirical Applications

Two financial time series are analyzed in this section. The first is the Eurodollar rate, daily (from Monday to Friday), for the time period January 9, 1995 to April 23, 2004.[3] The second is the Dow Jones stock price index, daily from January 7th, 2002 to May 7th, 2004. In both cases, if there is no value for a given day, the arithmetic mean using the previous and the following observation was computed. We have chosen to analyze these two series because their statistical properties are those typically found in most daily financial time series. Other financial series with similar features could also have been employed.

[3] The week corresponding to the September 11 attacks in 2001 was removed from the analysis.

The analysis is carried out with the original data, although identical results were obtained when using the log transformations. Note, however, that these are of interest only in the case of integer differentiation, which gives a return series. Deterministic components such as an intercept or an intercept and a linear time trend were also included in the models specified below, but these coefficients were found to be insignificantly different from zero in all cases. Thus, the analysis was carried out on the basis of Equation (5) for different ρ functions and different types of $I(0)$ disturbances.

4.1. *The Eurodollar rate*

The Eurodollar rate is the bid side of the Eurodollar quote in London. It is collected between 7 and 9 a.m. Eastern time (approximately late morning London time), and it has been obtained from the Federal Reserve Bank of St. Louis database.

Figure 2 contains plots of the original series, its first and 5-period differences along with their corresponding correlograms and periodograms. We see that the original series seems to be stationary for the first part of the sample, up to approximately December 2000. Then, it starts decreasing sharply and becomes relatively stable towards the end of 2003 and the beginning of 2004. Both the correlogram and the periodogram indicate nonstationarity, with the autocorrelation values decaying very slowly and with a large peak in the periodogram at the smallest frequency. The plots corresponding to the first differences suggest that the differenced series may be stationary, although there are significant values in the correlogram at some lags far away from zero, and the same happens with the 5-period differences. Finally, the periodogram of the 5-period differences has values close to 0 at some frequencies, indicating that the series may be overdifferenced with respect to them.

Denoting the series by x_t, we start by estimating the model given by Equations (5) and (6) with $k = 5$, i.e., under H_o (Equation (7)), we test

$$(1 - L^5)^d x_t = u_t, \quad t = 1, 2, \ldots,$$

with $d = 0, (0.10), 2$, and modeling u_t first as a white noise process, and then allowing for $I(0)$ autocorrelation. In the latter case, we first assumed AR(1) and AR(2) processes, and the null hypothesis was rejected in all cases except when $d = 0$, thus implying short memory. However, in these cases the coefficients corresponding to the AR parameters were extremely close

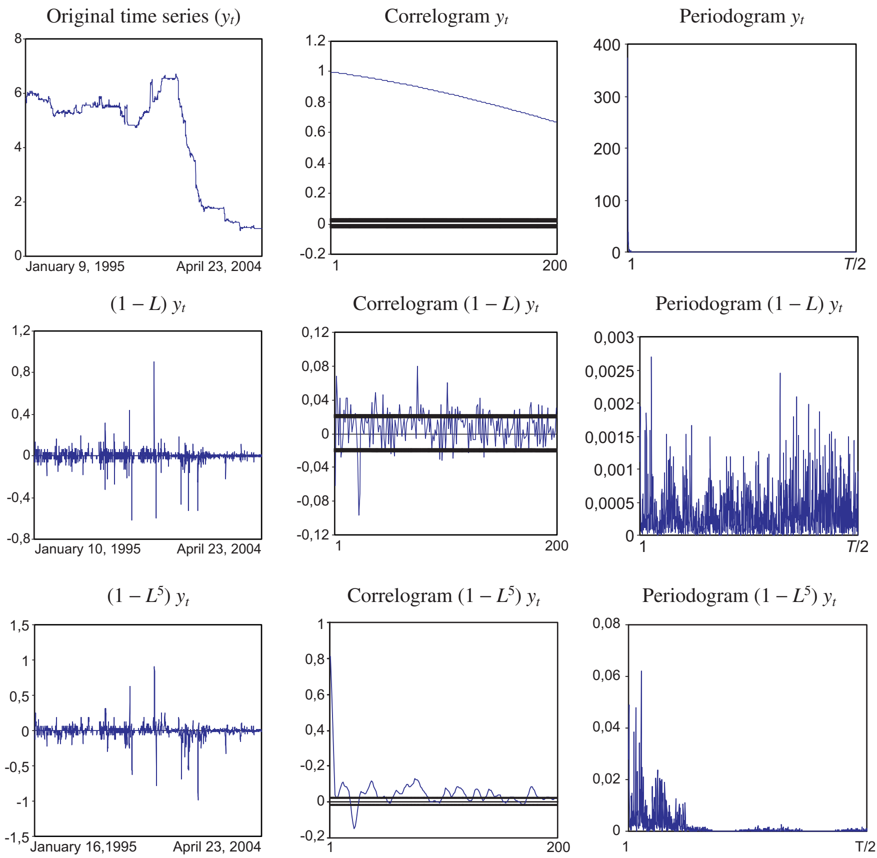


Figure 2. Eurodollar rate and its differences with their corresponding correlograms and periodograms. The large sample standard errors under the null hypothesis of no autocorrelation is $1/\sqrt{T}$ or roughly 0.02.

to the unit root circle, suggesting that these parameters are competing with the fractional differencing one in describing the nonstationarity of the series. Thus, we tried other less conventional forms for the $I(0)$ disturbances, which are very convenient in the context of the present tests. In particular, we used a model due to Bloomfield (1973), where the short-run components are defined exclusively in terms of the spectral density function, which is given by

$$f(\lambda; \tau) = \frac{\sigma^2}{2\pi} \exp\left(2 \sum_{r=1}^{m} \tau_r \cos(\lambda r)\right), \tag{12}$$

where m indicates the number of parameters required to describe the short-run dynamics. The intuition behind this model is the following. Suppose that u_t is an ARMA process of the form

$$u_t = \sum_{r=1}^{p} \phi_r u_{t-r} + \varepsilon_t - \sum_{r=1}^{q} \theta_r \varepsilon_{t-r},$$

where ε_t is a white noise process and all zeros of $\phi(L) = (1 - \phi_1 L - \cdots - \phi_p L^p)$ lie outside the unit circle and all zeros of $\theta(L) = (1 - \theta_1 L - \cdots - \theta_q L^q)$ lie outside or on the unit circle. Clearly, the spectral density function of this process is then

$$f(\lambda; \tau) = \frac{\sigma^2}{2\pi} \left| \frac{1 - \sum_{r=1}^{q} \theta_r \mathrm{e}^{\mathrm{i}r\lambda}}{1 - \sum_{r=1}^{p} \phi_r \mathrm{e}^{\mathrm{i}r\lambda}} \right|^2, \tag{13}$$

where τ now corresponds to all the AR and MA coefficients and σ^2 is the variance of ε_t. Bloomfield (1973) showed that the logarithm of an estimated spectral density function is often a fairly well-behaved function and, therefore, can be approximated by a truncated Fourier series. He showed that Equation (12) approximates Equation (13) well, with p and q being small values, which usually happens in economics. Like the stationary AR(p) model, the Bloomfield (1973) model has exponentially decaying autocorrelations, and thus can be used for u_t in Equation (5). It is a member of a large family of spectral density functions of which the most famous example is the Fourier transformation providing a spectral density for a given process (see Wong, 1971), and, while there exist alternative spectral density representations, we have chosen to use the Bloomfield (1973) specification in this paper because it is particularly suited to the functional form of the test statistics we employ. The formulae for Newton-type iterations for estimating the τ_r are very simple (involving no matrix inversion), updating formulae when m is increased are also simple, and we can replace $\hat{A}$ in Equation (9) by the population quantity

$$\sum_{l=m+1}^{\infty} l^{-2} = \frac{\pi^2}{6} - \sum_{l=1}^{m} l^{-2},$$

which indeed is constant with respect to the τ_r (in contrast to the AR case). The Bloomfield (1973) model, combined with fractional integration, has not been

Table 6. Testing the order of integration in $(1 - L^5)^d x_t = u_t$, in the Eurodollar rate.

d	White noise	Bloomfield ($m = 1$)	Bloomfield ($m = 2$)
0.00	154.26	65.04	41.09
0.10	146.90	64.23	38.37
0.20	138.16	63.91	34.56
0.30	126.42	61.05	34.63
0.40	110.07	53.70	33.05
0.50	88.65	43.72	29.39
0.60	64.27	32.66	23.63
0.70	41.14	23.48	22.48
0.80	22.68	15.70	14.23
0.90	9.59	8.35	9.09
1.00	0.79*	0.91*	1.02*
1.10	−5.10	−4.67	−3.35
1.20	−9.17	−8.63	−6.78
1.30	−12.07	−11.49	−9.04
1.40	−14.23	−13.66	−11.22
1.50	−15.88	−15.32	−13.42
1.60	−17.18	−16.67	−14.98
1.70	−18.22	−17.26	−16.67
1.80	−19.09	−18.65	−16.98
1.90	−19.81	−19.41	−17.09
2.00	−20.42	−20.07	−17.99

* Nonrejection values at the 5% significance level.

used very much in previous econometric models (although the model itself is well known in other disciplines — see, e.g., Beran, 1993). Our analysis shows that it is a credible alternative to the fractional ARIMA specifications, which have become conventional in the parametric modeling of long memory.[4]

The test statistics reported in Table 6 is the one-sided statistics given by $\hat{r}$ in Equation (8) for the three types of disturbances. A noteworthy feature emerging from this table is that it decreases monotonically with d. This is to be expected given the fact that it is a one-sided statistic. Thus, for example, it is desirable that if H_o (Equation (7)) is rejected with $d = 0.75$ in favor of alternatives of form $d > 0.75$, an even more significant rejection should occur when $d = 0.50$ or 0.25 are tested. It can be seen that the only value of d for which the null hypothesis cannot be rejected corresponds to $d = 1$, in all three cases of white noise and Bloomfield (with $m = 1$ and 2) disturbances.

[4]Empirical applications using the model of Bloomfield (1973) with $I(d)$ processes can be found in Velasco and Robinson (2000) and Gil-Alana (2001b).

Table 7. 95% confidence intervals in the Eurodollar rate.

Disturbances	Confidence intervals
White noise	[0.99 (1.01) 1.03]
Bloomfield ($m = 1$)	[0.99 (1.01) 1.04]
Bloomfield ($m = 2$)	[0.97 (1.00) 1.03]

Table 7 displays the 95% confidence intervals of the values of d for which H_o cannot be rejected. These intervals were constructed as follows: We recomputed the tests sequentially for d_o values $= 0, (0.01), 2$, choosing the values of d_o for which H_o cannot be rejected at the 5% significance level. Thus, the value corresponding to the lowest statistics in absolute value (which is reported in the table in parentheses within the square brackets) will be an approximation to the maximum likelihood estimator. We see that the intervals are very narrow, and the lowest statistics in absolute value corresponds to $d = 1$ or 1.01.

Next, we performed the test assuming that the process contains only one root at the long run or zero frequency. In other words, we tested for the presence of unit (or fractional) roots in a model given by

$$(1 - L)^d y_t = u_t, \quad t = 1, 2, \ldots,$$

for the same values of d and the same type of disturbances as in the previous case. The results are reported in Tables 8 and 9. Starting with the case of a white noise u_t, we see that the unit root cannot be rejected, implying that a simple random walk model may be a plausible alternative for this series, and similar results are obtained when autocorrelated disturbances are employed. Again, the values here are centered around the unit root, and the coefficients for the autocorrelation case were significantly close to 0 in all cases.

The results presented so far suggest that the daily Eurodollar rate can be described either as a pure random walk process, i.e.,

$$y_t = y_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots, \tag{14}$$

or as 5-period differences, such that

$$y_t = y_{t-5} + \varepsilon_t, \quad t = 1, 2, \ldots, \tag{15}$$

Table 8. Testing the order of integration in $(1 - L)^d x_t = u_t$, in the Eurodollar rate.

d	White noise	Bloomfield ($m = 1$)	Bloomfield ($m = 2$)
0.00	211.50	135.90	101.62
0.10	203.87	128.47	94.80
0.20	193.97	116.85	79.30
0.30	178.86	103.78	68.84
0.40	154.69	87.18	58.28
0.50	119.71	65.13	50.96
0.60	79.70	45.06	31.30
0.70	45.58	27.32	19.78
0.80	22.52	14.72	10.73
0.90	8.53	5.97	5.07
1.00	0.05*	0.74*	0.12*
1.10	−5.35	−3.22	−2.50
1.20	−9.04	−5.63	−3.81
1.30	−11.69	−7.15	−5.59
1.40	−13.69	−8.54	−8.19
1.50	−15.24	−9.91	−8.34
1.60	−16.48	−10.58	−9.05
1.70	−17.44	−11.41	−12.12
1.80	−18.33	−12.05	−13.37
1.90	−19.04	−12.56	−13.87
2.00	−19.64	−13.27	−15.03

* Nonrejection values at the 5% significance level.

Table 9. 95% confidence intervals in the Eurodollar rate.

Disturbances	Confidence intervals
White noise	[0.98 (1.00) 1.02]
Bloomfield ($m = 1$)	[0.99 (1.02) 1.06]
Bloomfield ($m = 2$)	[0.96 (1.00) 1.04]

both models implying that the data are completely unpredictable.[5] Note that the similarities for the cases of white noise and autocorrelated (Bloomfield) disturbances may be explained by the fact that the short-run dynamics are not important when modeling this series. In fact, the coefficients corresponding to

[5]Similar conclusions were obtained with the log-transformed data implying that the returns are also unpredictable.

the Bloomfield model (for the unit root case—not reported) were very close to 0, indicating that u_t can be specified as a white noise process. Finally, the plot of the series in Figure 2 also indicates the possible presence of a structural break in the data. An appealing feature of Robinson's (1994) testing procedure described in Section 2 is that it allows one to include deterministic components to take into account the break with no effect on its standard limit distribution. In contrast, a drawback of this procedure is that the type and the time of the break have to be specified. Visual inspection of Figure 2 suggests that the break is likely to have occurred in December 2000, and we choose as the breakpoint November 29, 2000, which corresponds to the highest value of the series, which then decreases sharply.[6] We experimented with different types of break, including slope and shift breaks. In all cases, the results were very similar to those reported, providing evidence of unit roots of the form given in Equation (14) or (15).

Therefore, it appears that the daily structure of the Eurodollar rate is completely unpredictable, supporting the efficiency market hypothesis, according to which, because of arbitrage, it should not be possible, using publicly available information, to make systematic profits over and above transaction costs and risk premia. Thus, evidence of mean reversion would be inconsistent with equilibrium asset pricing models. The fact that both a random walk and a 5-period difference model appear to be equally suitable to capture the stochastic behavior of the series can be interpreted as indicating that weekday effects do not play a crucial role in this series.[7]

To determine which of the two specifications better describe the data, we carry out a bootstrap simulation. Instead of standard normal increments, we consider the changes in the Eurodollar rate. For both the random walk (Equation (14)) and the 5-period random walk (Equation (15)), we simulate 500 times the future paths for $t = 250 - 2425$, drawing each time uniformly from the original changes. In Figure 3, we can see the Eurodollar rate and the averages of the simulated paths.

One can see that approximately up to observation 1645, the Eurodollar rate can be better modeled as a 5-period random walk, whereas for the whole

[6] Other breakpoints were also considered, obtaining very similar results.

[7] We also investigated the possible presence of day of the week effects by applying both versions of the tests to the data according to the day of the week. In all cases, we found evidence of unit roots and hence of no weekday effects.

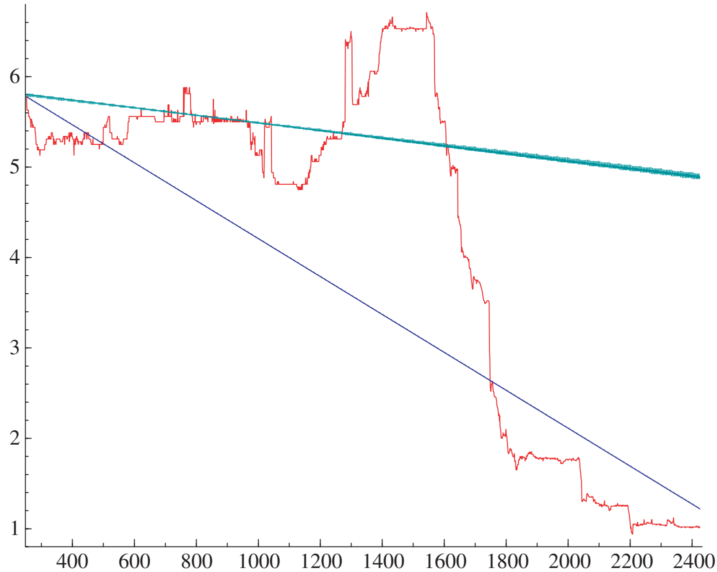


Figure 3. Eurodollar rate and the averages of its simulated paths for $t = 251 - 2425$. The upper line represents the average of 500 5-period random walks and the bottom line is the average of 500 random walks.

sample a random walk specification seems preferable. It also appears that there is a structural break around $t = 1645$ (a big drop on 19 April, 2001 from 4.97 to 4.44).

4.2. *The Dow Jones index*

The second empirical application is based on the Dow Jones (Equation (5)) index. This is a market index constructed as a subset of the Dow Jones Industrial Average. Of the 30 stocks in the Industrial Average, the five with the highest dividend yield during the 12-month period ending in December are selected as the Dow Jones (Equation (5)). The source of the data is http://www.djindexes.com.

Figure 4 is similar to Figure 2 but refers to the new series. This is also clearly nonstationary. Its first differences may be stationary, whereas the 5-period differences suggest overdifferencing with respect to some of the frequencies.

We proceed as in the previous case. Thus, we start by performing the tests for the case of Equations (5) and (6) with $k = 5$, with the same values of d as in the other cases (see Table 10). An interesting result we find is that the unit

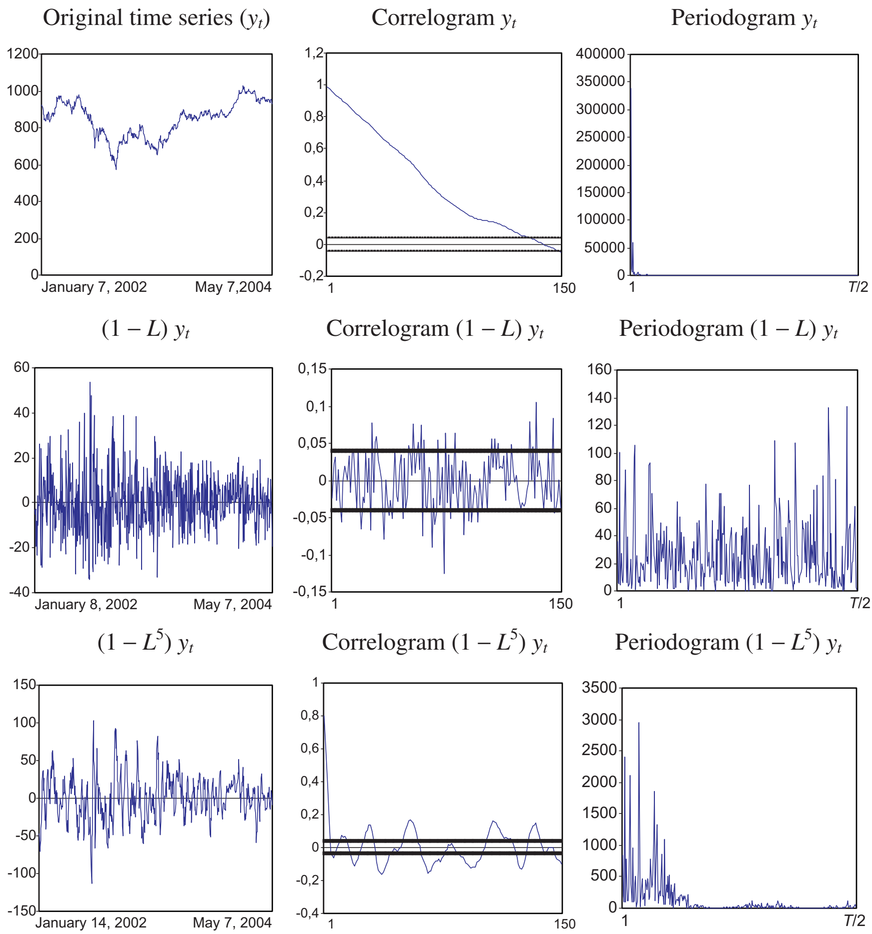


Figure 4. Dow Jones index and its differences with their corresponding correlograms and periodograms. The large sample standard errors under the null hypothesis of no autocorrelation is $1/\sqrt{T}$ or roughly 0.04.

root null (i.e., $d = 1$) is rejected for the three types of disturbances in favor of smaller orders of integration. If u_t is white noise, H_o cannot be rejected at $d = 0.9$, and, assuming autocorrelation in the case of Bloomfield (1973) disturbances, the nonrejection values occur at $d = 0.60$ with $m = 1$ and at $d = 0.50$ with $m = 2$. Thus, in the three cases, we find evidence of mean reversion. Table 11 displays the 95% confidence intervals. If u_t is white noise, the values of d where H_o cannot be rejected range between 0.89 and 0.98, and

Table 10. Testing the order of integration in $(1 - L^5)^d x_t = u_t$, in the Dow Jones index.

d	White noise	Bloomfield ($m = 1$)	Bloomfield ($m = 2$)
0.00	55.54	24.51	8.30
0.10	49.74	19.00	6.94
0.20	41.38	13.89	5.01
0.30	34.55	10.13	3.24
0.40	28.13	6.89	1.90
0.50	21.73	3.74	0.20*
0.60	15.51	0.73*	−1.57
0.70	9.80	−1.98	−2.83
0.80	4.90	−4.27	−4.09
0.90	0.96*	−6.13	−4.60
1.00	−2.07	−7.51	−4.78
1.10	−4.36	−8.60	−4.83
1.20	−6.06	−9.41	−5.06
1.30	−7.35	−10.07	−5.76
1.40	−8.34	−10.58	−6.04
1.50	−9.12	−10.99	−7.11
1.60	−9.74	−11.33	−8.06
1.70	−10.25	−11.62	−8.87
1.80	−10.76	−11.86	−9.21
1.90	−11.03	−12.08	−9.56
2.00	−11.33	−12.26	−10.04

* Nonrejection values at the 5% significance level.

Table 11. 95% confidence intervals in the Dow Jones index.

Disturbances	Confidence intervals
White noise	[0.89 (0.93) 0.98]
Bloomfield ($m = 1$)	[0.57 (0.62) 0.68]
Bloomfield ($m = 2$)	[0.44 (0.50) 0.56]

in the case of Bloomfield (1973) disturbances, they range between 0.57 and 0.68 (with $m = 1$) and between 0.44 and 0.56 (with $m = 2$).

In Table 12, we assume that the correct model is given by Equations (5) and (6) with $k = 1$, and the nonrejection values now occur at $d = 1$ (white noise u_t) along with $d = 0.9$ with autocorrelated disturbances. Thus, assuming a single pole (or singularity) at the zero frequency, the values of d for which the null cannot be rejected are much higher than in the previous case. The 95% confidence intervals are reported in Table 13. However, these values may

Table 12. Testing the order of integration in $(1 - L)^d x_t = u_t$, in the Dow Jones index.

d	White noise	Bloomfield ($m = 1$)	Bloomfield ($m = 2$)
0.00	82.73	49.67	32.69
0.10	75.28	42.02	28.12
0.20	64.63	33.77	22.14
0.30	54.84	28.46	16.80
0.40	44.50	21.99	16.30
0.50	33.59	17.16	9.63
0.60	23.08	11.85	9.51
0.70	14.12	7.32	4.83
0.80	7.28	3.75	2.51
0.90	2.42	1.06*	1.55*
1.00	−0.95*	−0.63*	0.002*
1.10	−3.30	−2.22	−2.89
1.20	−4.99	−3.25	−3.27
1.30	−6.25	−4.16	−3.36
1.40	−7.21	−4.79	−3.80
1.50	−7.96	−5.44	−5.20
1.60	−8.57	−5.75	−5.87
1.70	−9.07	−5.13	−6.34
1.80	−9.48	−5.43	−7.09
1.90	−9.83	−6.82	−7.88
2.00	−10.13	−6.99	−9.11

* Nonrejection values at the 5% significance level.

be biased. Several studies conducted in a hydrological context (Montanari, Rosso, and Taqqu, 1995, 1996, 1997) showed that the presence of periodicities might influence the reliability of the estimators of the long-memory parameter. Analyzing the series of the monthly flows of the Nile River at Aswan, these authors found that many heuristic estimators gave a positive value for d, indicating long memory where none was present.[8]

As a final step, we investigated whether the day of the week has any influence on the order of integration of the series. Therefore, we separated the data according to the day of the week, and performed again the tests of Robinson (1994) based on the model given by Equations (5) and (6) with

[8]In another recent paper, Montanari *et al.* (1999) performed an extensive Monte Carlo investigation in order to find out how reliable the estimators of long memory are in the presence of periodicities, and they concluded that the best results are those obtained using likelihood-type methods.

Table 13. 95% confidence intervals in the Dow Jones index.

Disturbances	Confidence intervals
White noise	[0.92 (0.97) 1.02]
Bloomfield ($m = 1$)	[0.88 (0.96) 1.04]
Bloomfield ($m = 2$)	[0.90 (1.00) 1.04]

Table 14. 95% confidence intervals for the values of d, at the zero frequency, for each day of the week.

	White noise	Bloomfield ($m = 1$)	Bloomfield ($m = 2$)
Monday	[0.82 (0.92) 1.05]	[0.74 (0.90) 1.15]	[0.71 (0.92) 1.16]
Tuesday	[0.83 (0.92) 1.06]	[0.74 (0.92) 1.17]	[0.72 (0.92) 1.16]
Wednesday	[0.84 (0.93) 1.06]	[0.77 (0.96) 1.20]	[0.76 (0.95) 1.22]
Thursday	[0.84 (0.93) 1.07]	[0.78 (0.95) 1.22]	[0.77 (0.97) 1.24]
Friday	[0.84 (0.94) 1.07]	[0.77 (0.96) 1.21]	[0.77 (0.97) 1.24]

$k = 1$, testing the degree of dependence at the long run or zero frequency. We find that the unit root null hypothesis cannot be rejected for any series and any type of disturbances, although the lowest statistics are in all cases smaller than 1 (Table 14). Another interesting feature emerging from Table 14 is that the degree of dependence increases with the day of the week. On Mondays, d is around 0.92. It is slightly higher on Tuesdays and Wednesdays, and a bit higher on Thursdays and Fridays. These differences, although small, give further support to the model given in Equations (5) and (6) with $k = 5$ as an adequate specification for this series, and, given the fact that d is smaller than 1, future values of the series are predictable to some extent.

In the case of the Dow Jones index, therefore, we find evidence of nonstationarity, but also of shocks dying away in the long run, which would imply a degree of predictability apparently inconsistent with market efficiency. We also show that this is a function of the day of the week being considered: it appears that there are significant weekday effects, resulting in predictable values throughout the past history of the series.

5. Conclusions

This paper has considered a version of the tests of Robinson (1994) that enables one to test models of the form $(1 - L^k)^d x_t = u_t$, where k is an integer value, d can be any real number, and u_t is $I(0)$. The most common cases are those with $k = 1$ (unit or fractional roots) and $k = 4$ and 12 (seasonal unit or fractional models). However, we extend the analysis to cover situations such as $(1 - L^5)^d x_t = u_t$, which might be relevant, for example, in the context of daily financial data. Our Monte Carlo experiments show that these tests perform well against fractional alternatives if the sample size is relatively large (e.g., $T \geq 300$).

Two empirical applications were carried out to shed light on day of the week effects in financial series. This is an important issue, as the existence of such predictable patterns might enable investors to devise trading strategies which result in excess returns, thereby violating market efficiency.[9] First, we examined the Eurodollar rate, and found no evidence of fractional integration either at the long run or zero frequency, or in the more elaborated version based on $(1 - L^5)^d$. In fact, the most adequate specifications for this series were a pure random walk model ($x_t = x_{t-1} + \varepsilon_t$) or its 5-period difference ($x_t = x_{t-5} + \varepsilon_t$), implying that the series is unpredictable.

The second application focused on the Dow Jones (Equation (5)) daily index. Here, using a model with a single pole at the zero frequency, the unit root cannot be rejected. However, using the version based on $(1 - L^5)^d$, the hypothesis of a unit root (i.e., $d = 1$) was decisively rejected in favor of smaller degrees of integration, implying mean-reverting behavior. The value of d is found to range between 0.50 and 1, indicating nonstationarity but with shocks disappearing in the long run. Finally, it was also found that the degree of dependence between the observations is higher at the end of the week. In the presence of such mean-reverting behavior, i.e., if asset prices over time move back to some "fundamental" value, their changes are highly predictable, implying that there are unexploited profit opportunities, which might indicate that investors are not fully rational, and the market is not efficient.

For further research, it may be of interest to apply the same type of model as the one employed here to other financial daily time series data. Also, it would be interesting to develop procedures to estimate the fractional differencing parameter in the context of the models presented here. In the seasonal case ($k = 4$ or 12) some attempts have been made by some authors. Ooms

[9] Note, however, that, because of transaction costs, profits might still not be gained. Furthermore, in addition to returns, stock price volatility and the risk profile of investors will also determine their buying–selling decisions.

(1995) proposes Wald tests based on the same model as in Robinson (1994), but requiring efficient estimates of the fractional differencing parameters (he uses a modified periodogram regression estimation procedure due to Hassler, 1994). Also, Hosoya (1997) establishes the limit theory for long-memory processes with the singularities not restricted at the zero frequency and proposes a set of quasi-log-likelihood statistics to be applied in raw time series. Arteche and Robinson (2000) and Arteche (2002) propose a model for the cyclical component in raw time series. Specifically, they estimate d at any frequency in the spectrum, thereby including seasonal or cyclical structures. Unlike previous methods, Robinson's (1994) tests do not require estimation of the long-memory parameter, as the differenced series have short memory under the null. More recently, Giraitis *et al.* (2001) extend the estimation to the frequency parameter, i.e., assuming that the pole occurs at an unknown frequency. The robustness of our results to using such methods will be the object of future research.

Acknowledgments

Comments of an anonymous referee are gratefully acknowledged. Gil-Alana gratefully acknowledges financial support from the Minsterio de Ciencia y Tecnologia (SEC2002-01839, Spain).

References

Abraham, A. and D. Ikenberry, "The Individual Investor and the Weekend Effect." *Journal of Financial and Quantitative Analysis* 29, 263–277 (1994).

Agrawal, A. and K. Tandon, "Anomalies or Illusions? Evidence from Stock Markets in Eighteen Countries." *Journal of International Money and Finance* 13, 83–106 (1994).

Arteche, J., "Semiparametric Robust Tests on Seasonal and Cyclical Long Memory Series." *Journal of Time Series Analysis* 23, 1–35 (2002).

Arteche, J. and P. M. Robinson, "Semiparametric Inference in Seasonal and Cyclical Long Memory Processes." *Journal of Time Series Analysis* 21, 1–25 (2000).

Baillie, R. T., "Long Memory Processes and Fractional Integration in Econometrics." *Journal of Econometrics* 73, 5–59 (1996).

Beaulieu, J. J. and J. A. Miron, "A Cross Country Comparison of Seasonal Cycles and Business Cycles." *The Economic Journal* 102, 702–788 (1992).

Beran, J., "Fitting Long Memory Models by Generalized Linear Regressions." *Biometrika* 80, 817–822 (1993).

Bhargava, A., "On the Theory of Testing for Unit Roots in Observed Time Series." *Review of Economic Studies* 53, 369–384 (1986).

Bloomfield, P., "An Exponential Model in the Spectrum of a Scalar Time Series." *Biometrika* 60, 217–226 (1973).

Box, G. E. P. and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day (1970).

Cornell, B., "The Weekly Patterns in Stock Returns Cash Versus Futures: A Note." *Journal of Finance* 40, 583–588 (1985).

Cross, F., "The Behavior of Stock Prices on Fridays and Mondays." *Financial Analyst Journal* 29, 67–69 (1973).

Dickey, D. A. and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74, 427–431 (1979).

Dickey, D. A., D. P. Hasza and W. A. Fuller, "Testing for Unit Roots in Seasonal Time Series." *Journal of the American Statistical Association* 79, 355–367 (1984).

Diebold, F. X. and G. D. Rudebusch, "Long Memory and Persistence in the Aggregate Output." *Journal of Monetary Economics* 24, 189–209 (1989).

Dolado, J. J., J. Gonzalo and L. Mayoral, "The Fractional Dickey–Fuller Unit Root Tests." *Econometrica* 70, 1963–2006 (2001).

Dyl, E. and E. D. Maberly, "The Anomaly that isn't there: A Comment on Friday the Thirteenth." *Journal of Finance* 43, 1285–1286 (1986).

Fama, E. and K. French, "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47, 427–465 (1992).

Fortune, P., *"Weekends can be Rough: Revisiting the Weekend Effect in Stock Prices."* Working Paper No. 98-6, Federal Reserve Bank of Boston (1998).

French, K., "Stock Returns and the Weekend Effect." *Journal of Financial Economics* 8, 55–69 (1980).

Fuller, W. A., *Introduction to Statistical Time Series*, New York: Wiley (1976).

Gibbons, M. and P. Hess, "Day of the Week Effects and Asset Returns." *Journal of Business* 54, 579–596 (1981).

Gil-Alana, L. A., "Testing Fractional Integration with Monthly Data." *Economic Modelling* 16, 613–629 (1999).

Gil-Alana, L. A., "A Fractionally Integrated Exponentially Spectral Model for the UK Unemployment." *Journal of Forecasting* 20, 329–340 (2001b).

Gil-Alana, L. A. and P. M. Robinson, "Testing of Unit Roots and Other Nonstationary Hypotheses in Macroeconomic Time Series." *Journal of Econometrics* 80, 241–268 (1997).

Gil-Alana, L. A. and P. M. Robinson, "Testing Seasonal Fractional Integration in the UK and Japanese Consumption and Income." *Journal of Applied Econometrics* 16, 95–114 (2001).

Giraitis, L., J. Hidalgo and P. M. Robinson, "Gaussian Estimation of Parametric Spectral Density with Unknown Pole." *Annals of Statistics* 29, 987–1023 (2001).

Hassler, U., "Misspecification of Long Memory Seasonal Time Series." *Journal of Time Series Analysis* 15, 19–30 (1994).

Hosoya, Y., "A Limit Theory for Long-Range Dependence and Statistical Inference on Related Models." *Annals of Statistics* 25, 105–137 (1997).

Hsieh, D. A., "The Statistical Properties of Daily Foreign Exchange Markets: 1974–1983." *Journal of International Economics* 24, 129–145 (1988).

Hylleberg, S., R. F. Engle, C. W. J. Granger and B. S. Yoo, "Seasonal Integration and Cointegration." *Journal of Econometrics* 44, 215–238 (1990).

Jaffe, J. F. and R. Westerfield, "The Week-end Effect in Common Stock Returns: The International Evidence." *Journal of Finance* 40, 433–454 (1985).

Jaffe, J. F., R. Westerfield and C. Ma, "A Twist on the Monday Effect in Stock Prices: Evidence from the US and Foreign Stock Markets." *Journal of Banking and Finance* 13, 641–650 (1989).

Keim, D. B. and F. Stambaugh, "A Further Investigation of Weekend Effects in Stock Returns." *Journal of Finance* 39, 819–840 (1984).

Kiymaz, H. and H. Berument, "The Day of the Week Effect on Stock Market Volatility and Volume: International Evidence." *Review of Financial Economics* 12(4), 363–380 (2003).

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin, "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root." *Journal of Econometrics* 54, 159–178 (1992).

Lakonishok, J. and E. Maberly, "The Weekend Effect: Trading Patterns of Individual and Institutional Investors." *Journal of Finance* 45(1), 231–244 (1990).

Miller, E., "Why a Weekend Effect?." *Journal of Portfolio Management* 14, 42–48 (1988).

Montanari, A., R. Rosso and M. S. Taqqu, "A Seasonal Fractional Differenced ARIMA Model: An Application to the River Nile Monthly Flows at Aswan." Preprint (1995).

Montanari, A., R. Rosso and M. S. Taqqu, "Some Long-Run Properties of Rainfall Records in Italy." *Journal of Geophysics and Research in Atmosphere* 101, 431–438 (1996).

Montanari, A., R. Rosso and M. S. Taqqu, "Fractionally Differenced ARIMA Models Applied to Hydrologic Time Series: Identification, Estimation and Simulation." *Water Resources Research* 33l, 1035–1044 (1997).

Montanari, A., M. S. Taqqu and V. Teverowsky, "Estimating Long Range Dependence in the Presence of Periodicity: An Empirical Study." *Mathematical and Computer Modelling* 29, 217–238 (1999).

Nelson, C. R. and C. I. Plosser, "Trends and Random Walks in Macroeconomic Time Series." *Journal of Monetary Economics* 10, 139–162 (1982).

Ooms, M., "*Flexible Seasonal Long Memory and Economic Time Series*." Discussion Paper 9951/A. Presented at the World Congress Econometric Society 1995, Tokyo, Econometric Institute, Erasmus University Rotterdam (1995).

Osborne, M. F. M., "Periodic Structure in the Brownian Motion of the Stock Market." *Operations Research* 10, 345–379 (1962).

Phillips, P. C. B., "Time Series Regression with a Unit Root." *Econometrica* 55, 277–301 (1987).

Phillips, P. C. B. and P. Perron, "Testing for a Unit Root in a Time Series Regression." *Biometrika* 75, 335–346 (1988).

Porter-Hudak, S., "An Application of the Seasonal Fractionally Differenced Model to the Monetary Aggregates." *Journal of the American Statistical Association* 85(410), 338–344 (1990).

Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Wetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge: Cambridge University Press (1986).

Ray, B. K., "Long Range Forecasting of IBM Product Revenues using a Seasonal Fractionally Differenced ARMA Model." *International Journal of Forecasting* 9, 255–269 (1993).

Robinson, P. M., "Efficient Tests of Nonstationary Hypotheses." *Journal of the American Statistical Association* 89, 1420–1437 (1994).

Robinson, P. M., "Log-Periodogram Regression of Time Series with Long Range Dependence." *Annals of Statistics* 23, 1048–1072 (1995a).

Robinson, P. M., "Gaussian Semiparametric Estimation of Long Range Dependence." *Annals of Statistics* 23, 1630–1661 (1995b).

Schmidt, P. and P. C. B. Phillips, "LM Tests for a Unit Root in the Presence of Deterministic Trends." *Oxford Bulletin of Economics and Statistics* 54, 257–287 (1992).

Sowell, F., "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models." *Journal of Econometrics* 53, 165–188 (1992).

Sutcliffe, A., "Time Series Forecasting using Fractional Differencing." *Journal of Forecasting* 13, 383–393 (1994).

Tanaka, K., "The Nonstationary Fractional Unit Root." *Econometric Theory* 15, 549–582 (1999).

Velasco, C. and P. M. Robinson, "Whittle Pseudo-Maximum Likelihood Estimation for Nonstationary Time Series." *Journal of the American Statistical Association* 95, 1229–1243 (2000).

Wong, E., *Stochastic Processes in Information and Dynamic Systems*, New York, Mc-Graw Hill (1971).

Chapter 3

Equity Restructuring via Tracking Stocks: Is there any Value Added?

Beni Lauterbach*
Bar-Ilan University, Israel

Joseph Vu
Depaul University, USA

In a tracking stock restructuring, the parent company issues a stock that tracks the earning performance of one of its divisions or subsidiaries. We study the effect of such an equity restructuring on the parent stock value. Parent stock response is insignificant in the short and long run. Thus, unlike equity carve-outs and spin-offs, issuing tracking stock does not create value, on average. This explains the complete cessation of tracking stock issuing since 2000. Our evidence also suggests that parent firms may have exploited tracking stock shareholders, which further explains the disappearance of tracking stocks.

Keywords: Tracking stocks; long-term returns; exploiting shareholders.

1. Introduction

Tracking stocks, also called targeted stocks, are a class of the parent company stock that tracks the earning performance of a division or a subsidiary of the parent company. Although the first tracking stock was issued back in 1984 by General Motors, tracking stocks have not been popular until the booming stock market of the 1990s. From 1984 to 1993, there were only seven tracking stock announcements, and since then there were 47 more, the peak year being 1999 with 19 tracking stock announcements. At the end of 1999, the total market value of outstanding tracking stocks exceeded $100 billions.

Tracking stocks are different from other forms of equity restructuring such as carve-outs and spin-offs (see Chemmanur and Paeglis, 2001). In both carve-outs and spin-offs, a new corporation is created with a new and separate board

*Corresponding author.

of directors, and the division (or subsidiary) assets are transferred to the new corporation. In a carve-out, the parent corporation uses an IPO to sell a stake in the division or subsidiary, yet keeps a majority interest in the issued firm. In a spin-off, the parent corporation distributes all subsidiary/division shares to shareholders as dividends.

In contrast, the issuance of tracking stock does not create a new corporation. The tracked subsidiary or division does not have a separate board—it is controlled and managed by the parent company. In addition, the tracking stock assets remain an integral part of the parent company, as there is no physical separation between the parent company and the targeted division. Tracking stock shareholders receive dividends from the earnings of the tracked division, which are reported separately from the earnings of the parent corporation. Tracking stock shareholders also receive voting rights in the parent corporation. Thus, tracking stocks are the mildest form of equity restructuring, with minimal business and operational changes.

Tracking stocks have not fared well after their issuance. Billett and Vijh (2004) show that tracking stocks underperform various benchmarks by 15–20% (20–40%) in the 2 (3) years after their issuance. This evidence contrasts with the postissue excess returns of spin-offs, which are known to be positive, and of carve-outs, which are known to be insignificant. Billett and Vijh (2004) also find, in a rather small sample, that parent stocks have insignificant excess returns after the tracking stock issuance.

We focus on the parent stock performance, trying to understand what (if at all) they gained from the tracking stock restructuring. First, we extend the sample period till the end of 2000, which increases the parent stock samples from 19 in Billett and Vijh (2004) to 32. It is possible that our larger and most updated samples would facilitate more reliable inferences on the short- and long-run excess returns of parent stocks.

Second, we extend the sample in the direction of firms that announced but did not eventually issue tracking stocks. These 22 firms are a natural control group for our 32 firms that issued tracking stocks. We find that firms that cancelled a planned tracking stock issue severely underperform in the 2 years after the tracking stock announcement. In contrast, parent firms that issued tracking stocks achieve a “normal” stock performance in the 2 years after. Thus, firms that issued tracking stocks might have avoided severe performance shortfalls in the subsequent years.

2. Why Issue Tracking Stocks?

The finance literature has suggested several possible answers to the question of: how can tracking stock issuance generate value for parent-stock shareholders?

2.1. *Information explanations*

Tracking stock issuance can alleviate the problems generated by the asymmetric information between the firm and its shareholders. After the issuance, investors receive information on both the parent firm and its tracked division. Hence, they know more on what happens inside the firm and can more accurately assess firm value.

Empirical tests examine the asymmetric information argument by looking at analyst coverage and earning forecast accuracy. Zuta (2000), Gilson *et al.*, (1998), and Chemmanur and Paeglis (2001) find that the number of analysts following the firm increased after the tracking stock issue, but D' Souza and Jacob (2000) show that the change in number of analysts is statistically insignificant. An increase in number of analysts could improve public available information about the firm. For example, it could increase the accuracy of future earning forecasts. Unfortunately, direct tests such as Chemmanur and Paeglis (2001) and Billett and Vijh (2004) do not find any improvement in earning forecast accuracy after the tracking stock issuance. Thus, it is unclear how much of the information asymmetry can be solved by tracking stock issuance, and this motivation appears weak.

A second information-based motive is that the tracking stock issue unveils the firm's true value. Many corporations argue that they are undervalued, and issue tracking stocks to show the market their undervalued asset. By doing so, these firms hope to gain by unlocking their "hidden value." Evidence on the hidden value proposition is mixed. In support of the hidden value proposition, it is found that parent stocks respond positively to an announcement of a tracking stock issue. The positive announcement excess return is about 2–3% (see Logue *et al.*, 1996; Billett and Mauer, 2000; Elder and Westra, 2000; and Harper and Madura, 2002). Moreover, Clayton and Qian (2004) also find a significantly positive ex-date excess return. However, the longer-term perspective is gloomy, as the postissue performance of tracking stocks is negative, and the postissue performance of parent stocks is neutral. Hence, it is unclear whether or not there was any hidden value that was unlocked.

2.2. *The diversification discount motive*

Berger and Ofek (1995) document a diversification discount of about 15% for US conglomerates. Zuta (2000) argues that issuing tracking stocks can solve some of the diversification-induced problems. Thus, tracking stock issues may create value by reducing the diversification discount. Billett and Mauer (2000) examine the diversification motive, and conclude that it cannot explain the positive revaluation (positive excess return) on tracking stock announcement. Hence, the diversification-discount motivation remains unsupported.

2.3. *Investor clientele*

The tracked division is sometimes from a different industry than the parent company. For example, the tracked division might be a growth company, whereas the parent firm is a more traditional (slowly growing) "value" company. In such a case, the tracked stock may attract some new investors, who value it most, leading to an increase in the conglomerate overall market value.

There is evidence that the tracking stock attracted new investors. For example, a year after US West issued its Media Group tracking stock, new investors owned more than 86% of the Media Group stock. Thus, the new clientele argument is pertinent. The clientele effect is also consistent with the positive response to tracking stock announcements. However, it cannot explain the negative postissue performance of tracking stocks.

2.4. *Agency perspectives*

The tracking stock discloses the division performance, affording incentive (pay for performance) plans for the division executives. This should improve managerial input in the division and increase division value. In contrast, Hass (1996) and Harper and Madura (2002) argue that the tracking stock may be a source of friction because the parent's Board of Directors, which also controls the tracked division, may sacrifice some of the division's value for the sake of maximizing the parent's value. Billett and Vijh (2004) present newspaper reports on severe conflicts between tracking and parent stock shareholders, which led the authors to conclude that in some cases tracking stocks create more problems than they solve.

The agency approach is consistent with the accumulated evidence. The positive announcement response may be due to the initial hopes for improved managerial input, whereas the later negative excess returns may reflect the new agency problem that emerged—conflicts of interest between parent and tracking stock shareholders.

Harper and Madura (2002) test the agency explanation. They find that the announcement response is more positive when the parent company is larger, less leveraged, and underperforming. All these firm characteristics are indicators for relatively heavy agency problems. Hence, firms that are more prone to agency problems appear to benefit more upon announcing a tracking stock issue, which leads Harper and Madura (2002) to conclude that the agency explanation is supported.

Interesting recent evidence is reported in Elder *et al.* (2005). They find that: (1) relative to similar (control) stocks, parent stock liquidity fell after the tracking stock issuance, (2) the adverse-selection component of parent stock bid-ask spread increased after the issuing, and (3) tracking stocks have lower liquidity than comparable firm stocks. All these findings suggest that investors concluded that firms issuing tracking stocks are more prone to agency misbehavior (such as insider trading). Another possible interpretation of Elder *et al.* (2005) evidence is that information asymmetry increased. However, previous research (reviewed above) finds an increase in number of analysts and information accuracy following tracking stock issuing. Thus, we maintain that the issuing has stained investor perception of the decency of parent firms.

2.5. *Other motivations*

Tracking stocks were also issued as a "currency" for acquisitions. In some cases, acquisitions were accomplished only after target shareholders were offered the choice between the acquirer's stock and a cash payment plus a tracking stock that follows the target's performance.

Finally, some tracking stocks were probably issued because of parent firms' "fad-following" behavior. Some parent firms simply joined the tracking stock bandwagon. Issuing tracking stocks during the hot market of the late 1990s (i.e., at peak prices) was definitely a clever strategy that served well parent-stock shareholders.

In this context, we realize that the negative postissue performance of tracking stocks can also be explained as a consequence of their "peak price" issue, after which came the inevitable rough landing. Postissuance negative excess returns are observed in other equity issues as well. If the postissue underperformance of tracking stocks is a typical equity issue phenomenon, then some of our previous explanations (the investor clientele effect, and the unlocking of hidden value argument) regain credibility, and remain plausible alongside the agency explanation.

3. Market Response to Tracking Stock Announcements

Tracking stock announcements are collected from the *Wall Street Journal* Index and the Dow Jones Newswire. Daily and monthly stock returns are downloaded from the CRSP database, and accounting information is from the Disclosure CD-ROM, the National Automated Accounting Research System (available on Lexis/Nexis), and 10k reports. The final sample comprises 54 tracking stock announcements from 1984 in 2000.

Table 1 presents the parent-stock response on announcement of a tracking stock issue. Like several previous studies (e.g., Harper and Madura, 2002), we observe statistically significant excess returns on days −1, 0, and 1 relative to the announcement. Thus, we use days −1 to 1 to estimate the announcement response.

The average raw return in days −1 to 1 is 1.6% in the overall sample, 2.0% in the sample of firms that had no confounding news (in the week before and after the announcement), and 1.6% in the sample of firms that eventually issued tracking stocks.

Excess returns are estimated in three ways: (1) net of market method, i.e., as $R_i - R_M$, where R_i is the return on the stock and R_M the return on the value-weighted index of NYSE-AMEX-NASDAQ stocks; (2) market model method, using the standard event study methodology, with the value-weighted market index and a parameter estimation period from day −315 to day −61 relative to the announcement; (3) net of matched-firm method, i.e., as $R_i - R_{Match}$, where R_i is the return on the stock and R_{Match} is the return on the firm in the same industry (4-digit SIC code) that is closest in total equity capitalization to the announcing firm.

The announcement excess returns in Table 1 are statistically significant. The net of market and market model methodology estimate a positive

Table 1. Stock response to tracking stock announcements, 1984–2000.

	All announcements ($N = 54$)	Announcements without confounding news ($N = 31$)	Announcements of firms that eventually issued tracking stocks ($N = 32$)
Days −1 to 1 around the announcement			
Average raw return	1.6% (1.4)	2.0% (1.5)	1.6% (1.1)
Average net of market return	1.2% [3.0]	1.6% [3.2]	1.0% [2.5]
Average market model excess return	1.3% [3.2]	1.7% [3.6]	1.2% [2.9]
Average net of matched-firm return	2.8% (1.7)	5.1% (3.1)	4.2% (1.9)
Days −5 to 5 around the announcement			
Average raw return	0.5% (0.4)	−0.6% (−0.3)	0.4% (0.2)
Average net of market return	−0.7% [−0.2]	−1.3% [−0.2]	−0.9% [−0.2]
Average market model excess return	−0.2% [0.6]	−0.6% [0.5]	−0.1% [0.7]
Average net of matched-firm return	1.6% (0.9)	2.4% (1.3)	3.0% (1.1)

Notes: t-statistics in parentheses and Z-statistics in brackets. The net of market return is computed by subtracting from the return on the stock, the return on the value-weighted index of NYSE-AMEX-NASDAQ stocks. The market model excess return methodology uses the value-weighted index and a parameter estimation period from day −315 to day −61 relative to the announcement. Net of matched-firm return is calculated as the announcing firm stock return minus the return on the stock of the firm in the same industry (4-digit SIC code) that is closest in size (total capitalization of equity) to the announcing firm.

revaluation of 1–1.7%, whereas the net of matched firm method assessed a positive response of 3–5%. These findings are consistent with previous evidence on the announcement response. Also noteworthy that the positive revaluation result is robust to the exclusion of firms that did not eventually issue tracking stocks (see the last column in Table 1), and to the exclusion of firms that had confounding news in the week before or after the tracking stock announcement (see the middle column in Table 1).

Table 1 also presents stock returns and excess returns in days −5 to 5 relative to the tracking stock announcement. This window is chosen because we note some short-term stock price drifts before and after the announcement. The net of market and market model methodology assess a slightly negative, yet statistically insignificant, response in days −5 to 5, whereas the matched-firm methodology estimates an insignificant positive response in that interval. In any case, the (−5, 5) window results serve as a caution. It is possible that the announcement response is close to zero and insignificant. Given our doubts about the "true" announcement response and given the small magnitude of the days $(-1, 1)$ response (about 1–2% only), we conclude that the parent stocks' "true" announcement response is (economically) negligible.

4. The Long-Term Response of Parent Stocks

Table 2 summarizes the long-term performance of parent stocks. First, we will examine the overall (all announcements) results reported in column 2 of the table. In the 2 years before the announcement, parent stocks performed poorly. The average market model excess return in the 2 years before the announcement is about -27%, with a t-statistics of -3.5. However, net of market and net of matched-firm preannouncement excess returns are about -9% and statistically insignificant.

After the announcement, our overall sample parent-stock performance continues to be dismal (see column 2 in the middle third of Table 2). The average 2-year postannouncement excess return is about -29% (t-statistics $= -2.3$) according to the market model, -16% (t-statistics $= -1.7$) according to the net of market methodology, and -8% (t-statistics $= -0.6$) according to the net of matched-firm technique. The difference between the pre- and postannouncement periods, reported in the lower third of Table 2, is statistically insignificant in the "all announcements" sample. Thus, based on the overall sample, the tracking stock announcement appears like an insignificant event in the long run because, on average, parent stocks continue to underperform at the same rate before and after the tracking stock announcement.

If tracking stock announcements do not help the parent firms, then tracking stock issuance may be redundant. That is, if parent-firm shareholders do not gain in the short or long run, why issue tracking stocks? The situation is even more complex because Billett and Vijh (2004) report that the issued tracking stocks severely underperform in the years after their issuance. (Clayton and Qian (2004) claim that the underperformance of tracking stocks is statistically insignificant.) It appears that equity restructuring via tracking stock may be a value-decreasing endeavor.

It is possible to offer some defense for tracking stock issuance. Table 2 also compares the 32 firms that went on to issue tracking stocks with the 22 firms that cancelled (indefinitely postponed) the planned issue. In the pre-announcement period, there is no statistically significant difference between "cancelled" and "issued" parent stocks (see the last column of Table 2), as all

Table 2. Long-term stock performance around tracking stock announcements, 1984–2000.

	All announcements ($N = 54$)	Announcements by firms that cancelled the planned issue ($N = 22$)	Announcements of firms that issued tracking stocks ($N = 32$)	t-of difference between issued and cancelled
Months −24 to −1 before the announcement				
Average net of market return	−8.3% (−1.5)	−5.9% (−0.7)	−9.9% (−1.4)	−0.4
Average market model excess return	−26.9% (−3.5)	−37.5% (−2.8)	−19.9% (−2.2)	1.1
Average net of matched-firm returns	−9.1% (−1.0)	3.4% (0.2)	−17.9% (−2.0)	−1.1
Months 1 to 24 after the announcement				
Average net of market return	−15.7% (−1.7)	−41.6% (−2.5)	1.6% (0.2)	2.4
Average market-model excess return	−28.7% (−2.3)	−71.5% (−2.8)	−0.1% (−0.0)	2.6
Average net of matched-firm returns	−7.5% (−0.6)	−27.0% (−2.3)	6.1% (0.3)	1.5
Return improvement between pre- and postannouncement periods				
Average difference in net of market return	−7.4% (−0.7)	−35.7% (−1.8)	11.5% (0.9)	2.1
Average difference in market model excess return	−1.8% (−0.2)	−34.0% (−1.8)	19.8% (1.5)	2.3
Average difference in net of matched-firm returns	1.6% (0.1)	−30.4% (−1.3)	24.0% (1.0)	1.6

Notes: t-statistics in parentheses. The net of market return is computed by subtracting from the return on the stock, the return on the value-weighted index of NYSE-AMEX-NASDAQ stocks. The market model excess return methodology uses the value-weighted index and a parameter estimation period from month −84 to month −25 relative to the announcement. Net of matched-firm return is calculated as the announcing firm stock return minus the return on the stock of the firm in the same industry (4-digit SIC code) that is closest in size (total capitalization of equity) to the announcing firm.

stocks perform poorly. However, in the postannouncement period, these two groups differ substantially. The 2-year postannouncement stock performance of firms that issued tracking stocks is neutral (slightly positive, yet statistically insignificant), whereas the 2-year postannouncement stock performance of firms that cancelled the tracking stock issue is significantly negative.

To complement the picture, Table 2 also presents "return improvement" statistics. Stocks of firms that cancelled the planned issue worsened their performance, in fact accelerated their downhill slide, in the 2 years after the announcement. In contrast, stocks of firms that issued tracking stocks improved their performance relative to the 2-year preannouncement period. The difference in "return improvement" between firms that issued tracking stocks and firms that did not (i.e., cancelled) is statistically significant at the 10% level at least (see the last column of Table 2). It appears that firms that issued tracking stocks managed to recover (regain normal performance), whereas firms that cancelled the planned issue continued to deteriorate.

The evidence in Table 2 suggests that parent firms that issued tracking stocks have benefited from it. Without issuing the tracking stocks, their fortune might have been similar to that of the firms that cancelled the planned issue. Issuing tracking stocks helped parent firms to stop their decay, possibly via some exploitation of tracking stock shareholders.

5. Summary and Conclusions

Equity restructuring via tracking stock issuance does not generate any significant excess return opportunities for the parent-firm stocks in the short or long run. This contrasts with the evidence on spin-offs and carve-outs. In spin-off and carve-out restructuring, parent stocks gain in the short run and have neutral performances in the long run (see Desai and Jain, 1999; Vijh, 1999).

Despite the insignificant excess returns, parent firms benefited from the tracking stock issuance. Before the issuance, parent stocks manifested poor performance, and after the issuance their performance turned into "average" or "normal." This recovery or stabilization may be attributed to the tracking stock issuance, also because of our finding that firms that announced but did not eventually issue tracking stocks demonstrate poor stock performance *before and after* the tracking stock announcement.

The likely source of parent-stock recovery is some exploitation of poor tracking stock investors. Tracking stocks severely underperformed in the years following their issuance (see Billett and Vijh, 2004).

For the sake of clarity, we do not claim that tracking stock issuance is necessarily a scheme to transfer wealth from poor tracking stock investors to parent firms. Nevertheless, the complete cessation of tracking stock issuance since year 2000 (see Krantz, 2004) appears consistent with the view that investors have realized the "true" nature of tracking stocks, thus refusing to participate in any more tracking stock offerings.

Krantz (2004) also reports that most (all but five) of the tracking stock issues were reabsorbed by their parent companies by the end of 2004. This quick move of the firms to wipe out these failing securities reinforces our impression that tracking stock issuance is an inferior (and perhaps even a value-destroying) equity restructuring solution.

Acknowledgments

We are grateful to two reviewers of *Advances in Quantitative Analysis of Finance and Accounting* for their constructive comments. All remaining errors are our own. The research was supported by a grant from the Investment Analysts Society of Chicago and by the Research Foundation of the Association for Investment Management and Research (AIMR).

References

Berger, P. and E. Ofek, "Diversification Effect on Firm Value." *Journal of Financial Economics* 37, 39–65 (1995).

Billett, M. and D. Mauer, "Diversification and the Value of Internal Capital Markets: The Case of Tracking Stock." *Journal of Banking and Finance* 24, 1457–1490 (2000).

Billett, M. and A. Vijh, "Long Term Returns from Tracking Stocks." *Journal of Financial Research* 27, 559–583 (2004).

Chemmanur, T. and I. Paeglis, "Why Issue Tracking Stocks? Insights from a Comparison with Spin-offs and Carve-outs." *Journal of Applied Corporate Finance* 14, 102–114 (2001).

Clayton, M. and Y. Qian, "Wealth Gains from Tracking Stocks: Long-Run Performance and Ex-Date Returns." *Financial Management* 33(3), 83–106 (2004).

D' Souza, J. and J. Jacob, "Why Firms Issue Targeted Stock." *Journal of Financial Economics* 56, 459–483 (2000).

Desai, H. and P. Jain, "Firm Performance and Focus: Long-run Stock Market Performance Following Spinoffs." *Journal of Financial Economics* 54, 75–101 (1999).

Elder, J. and P. Westra, "The Reaction of Security Prices to Tracking Stock Announcements." *Journal of Economics and Finance* 24(1), 36–55 (2000).

Elder, J., P. Jain and J. C. Kim, "Do Tracking Stocks Reduce Information Asymmetries? An Analysis of Liquidity and Adverse Selection." *Journal of Financial Research* 28(2), 197–213 (2005).

Gilson, S., P. Healy, C. Noe and K. Palepu, "*Information Effects of Spin-Offs, Equity Curve-Outs, and Targeted Stock Offerings*." Working Paper, Harvard Business School (1998).

Harper, J. and J. Madura, "Sources of Hidden Value and Risk within Tracking Stock." *Financial Management* 31(3), 91–109 (2002).

Hass, J., "Directorial Fiduciary Duties in a Tracking Stock Equity Structure: The Need for a Duty of Fairness." *Michigan Law Review* 94, 2089–2177 (1996).

Krantz, M., "Remember Tracking Stocks? Most are History." *09/20/2004 Edition of USA Today* (2004).

Logue, D., J. Seward and J. Walsh, "Rearranging Residual Claims: A Case for Targeted Stock." *Financial Management* 25, 43–61 (1996).

Vijh, A., "Long-term Returns from Equity Carveouts." *Journal of Financial Economics* 51, 273–308 (1999).

Zuta, S., "*Diversification Discount and Targeted Stock: Theory and Empirical Evidence*." Working Paper, University of Maryland (2000).

Chapter 4

Stock Option Exercises and Discretionary Disclosure

Wei Zhang*
University at Albany, USA

Steven F. Cahan
University of Auckland Business School, New Zealand

In this paper, we empirically examine recent claims by politicians and business journalists that stock option exercises give managers incentives to adopt certain disclosure policies. We predict that when firms have bad news, large stock option exercises will have a negative effect on discretionary disclosures. Using Association for Investment Management and Research ratings to measure discretionary disclosures, for a sample of 359 firm-year observations from 18 industries, we find evidence supporting our prediction after we control for stock option grants. Although we confirm previous findings that option awards create incentives for more disclosures, our results also suggest that stock option exercises could have an adverse effect on accounting disclosures when firms are experiencing a period of negative returns in stock markets.

Keywords: Discretionary disclosure; stock options; option exercises; managerial incentives.

1. Introduction

Recent accounting scandals, e.g., at Enron, WorldCom, Xerox, and Global Crossing, have brought unprecedented attention to accounting reports and the auditing profession, and have led to corporate reform in the US (i.e., the Sarbanes-Oxley Act of 2002).[1] In this paper, we focus on two elements of the ongoing debate regarding corporate accountability, i.e., the level of accounting disclosures and the incentives created by stock options. Specifically, we

---

*Corresponding author.

[1]The Sarbanes-Oxley Act of 2002 was passed in July 2002. Several of its provisions affect auditors, e.g., it bans nine types of nonaudit services, requires mandatory rotation of audit partners every 5 years, and requires the auditor to report "critical" accounting policies to the audit committee. Other provisions are directed at corporate management, e.g., the CEO and CFO must certify the appropriateness and accuracy of the financial statements.

consider whether the level of disclosure and managers' decision to exercise stock options are related.

Accounting researchers have been interested in disclosures at both theoretical (e.g., Diamond and Verrecchia, 1991) and empirical levels (e.g., Botosan and Plumlee, 2002). However, the recent accounting scandals have brought the issue of transparency and full disclosure into the public's eye. For example, *The Economist* (Clambering back up, 2002, p. 55) wrote "shareholders are demanding more 'honest' numbers, designed to illuminate, rather than disguise, the profitability of businesses." Also, the Sarbanes-Oxley Act of 2002 has numerous provisions related to disclosures. For instance, section 401(a) requires disclosures on off-balance sheet transactions and unconsolidated entities.

Likewise, although stock-based compensation has long been an issue of academic interest (see Core *et al.*, 2003, for a review), only recently have politicians and the business press begun to take a more critical look at stock options. For example, some criticize stock options for creating incentives for risk-taking and for short-term decision-making (e.g., Clambering back up, 2002).[2] Regarding the possible influence of stock options on accounting, *The Economist* (Clambering back up, 2002, p. 55) suggests that by "fiddling with their accounts, company bosses could hope to drive up the share price, cash in their options, and set sail in their yachts."[3] Critics in politics and in the business press claim that accounting may be used to increase profits from option exercises (e.g., Use and abuse, 2002). However, whether stock option exercises are related to lower disclosure levels is an empirical question.

[2]Core *et al.* (2003) note that there has been a huge increase in the amount of stock-based compensation in use. For example, in 1980, Hall and Liebman (1998) note that 57% of CEO had stock options, whereas this was 90% in 1994. Not surprisingly, Core and Guay (2001) find cross-industry differences in use of options. More specifically, Ittner, Lambert, and Larcker (2003) find that stock options are used more frequently in high-technology firms than in "old economy" manufacturing firms.

[3]Also, section 305 of the Sarbanes-Oxley Act requires managers to reimburse the firm for any equity-based compensation received within 12 months of issuing statements that require restatement due to "material noncompliance." However, in this study, we do not specifically examine fraudulent reporting. Rather, we concentrate on variations in disclosure levels within the parameters of generally accepted accounting principles.

Recently, Nagar *et al.* (2003) find a positive association between accounting disclosures and stock option grants. Our research differs from theirs because we focus on option exercises rather than option grants. We focus on option exercises because of the long-term nature of stock options. Specifically, there is a vesting period (usually a number of years) before options can be exercised, and the realized value of options will depend on the share price at the time of the exercise. As a result, the manager's incentives may also be affected by stock options around the time of their exercise.

In fact, Bartov and Mohanram (2004) provide empirical evidence that managers manage earnings using more positive discretionary accruals in years when option exercises are abnormally large. Their findings are consistent with the argument that top-level managers inflate earnings to increase the cash payout from options exercises. Thus, even though stock options can be long-term efficient (as they align the interests of managers and owners), Bartov and Mohanram's (2004) results suggest that the exercise of stock options can induce short-term opportunism.[4] In this study, we complement Bartov and Mohanram (2004) by examining the effect of stock option exercises on accounting disclosures (rather than on earnings), and we complement Nagar *et al.* (2003) by examining disclosures around the time of the exercise (rather than the time of the grant).

Penman (1982) and others propose that managers exploit their informational monopolies in their stock trading. To retain or enhance the benefits associated with private information, the manager may rationally choose to provide disclosures that are unclear, incomplete, or even misleading. We hypothesize that there is a negative association between option exercises and the level of disclosure if the firm is having bad news. If the firm has good news, managers will maintain or expand disclosures to deliver the good news to investors and to boost the share price before the options are exercised. However, if the firm has bad news that negatively affects the market's perception, the manager has incentives to withhold information. In this way, the manager may be able to prevent a decrease in the share value and, consequently, a decrease in the value of the firm's options. As a result, in periods when there are abnormally large

[4] Bartov and Mohanram's (2004) results are consistent with earlier positive accounting research that finds that although debt contacts and accounting-based bonus plans are efficient contracting mechanisms in the long run, they can still induce opportunistic behavior ex post.

option exercises, we expect a negative relationship between option exercises and disclosure levels for the firms with bad news.

Since Nagar *et al.* (2003) find that equity-based compensation can improve disclosures in the long-run, we control for the level of stock option grants using the ratio of stock option grants to total compensation as an important control variable. Thus, our tests examine whether option exercises have an incremental effect on disclosures over and above that of option grants. We define abnormally large option exercises the same way as Bartov and Mohanram (2004), where the ratio of stock option exercises to total compensation increases by more than 50% over the average for prior 2 years. Using a sample of 359 firm-year observations from 18 industries, we find a negative and significant relationship between stock option exercises and disclosure levels when there is bad news. When there is good news, we do not find such a relation.

We contribute to the literature by extending Nagar *et al.* (2003) and Bartov and Mohanram (2004). Our results shed light on the relationship between stock option exercises and accounting disclosures. Specifically, although we confirm the allegations of some observers that managers will reduce disclosures to make their options more valuable (e.g., Clambering back up, 2002), we show that this only holds where the firm has bad news. Further, we show that, for bad news firms, this effect is incremental to any disclosure effects created by the equity grants. Thus, even though stock options can be an efficient incentive mechanism for firms, our evidence suggests that they can still induce ex post opportunism in some settings.

The remainder of this paper is divided as follows. Section 2 reviews the relevant literature. Section 3 develops the hypothesis. Section 4 describes the research method and data. Section 5 provides the results and Section 6 is a conclusion.

2. Prior Literature

2.1. *Disclosure*

As Healy and Palepu (2001) point out, in capital market economies corporate disclosure is needed to ensure efficient allocation of resources. In particular, disclosure helps to address problems associated with information asymmetry and problems arising from agency relationships within the firm. Diamond and Verrecchia (1991) show that disclosures can increase firm value by reducing the information asymmetry, resulting in increased liquidity (Welker, 1995) and a lowering of the cost of capital (Botosan, 1997) and the cost of debt (Sengupta, 1998).

Disclosures can also reduce agency costs by improving monitoring by shareholders and debtholders (Watts and Zimmerman, 1986). For example, disclosures can help identify shirking and perquisite consumption. Leftwich *et al*. (1981) find that firms with high agency costs were more likely to provide interim reports before they were legally required. Managers may also increase disclosures to reduce the likelihood of being the target of a takeover bid (Brennan, 1999) and to provide signals to the market about their managerial talent (Trueman, 1986).

Coincidentally, there are costs associated with increased disclosures. For example, there are proprietary costs, e.g., disclosures can reveal information to competitors (Verrecchia, 1983; Darrough and Stoughton, 1990; Hayes and Lundholm, 1996; Piotroski, 1999). Information production costs can also be significant (Watts and Zimmerman, 1978), and litigation costs may discourage disclosures of certain types of information (Francis *et al*., 1994; Skinner, 1994, 1997; Miller and Piotroski, 2000).

2.2. *Disclosure and option grants*

Prior research examining the association between stock options and disclosures has developed along two lines, one focusing on option grants and the other on option exercises. Studies on the relationship between option grants and accounting disclosures include (but are not limited to) Yermack (1997), Aboody and Kasznik (2000), and Nagar *et al.* (2003).

Yermack (1997) examines whether stock option grants are issued prior to announcements that increase the firm's share price. He notes that managers play an active role in structuring their own compensation packages in many companies, and also hypothesizes that managers will use their influence to manipulate the timing of their option awards to maximize the compensation they receive. Consistent with this view, Yermack (1997) finds significant and positive abnormal stock returns after the option grant date.

Aboody and Kasznik (2000) focus on the manipulation of disclosures rather than manipulation of the grant date as in Yermack (1997). They hypothesize that managers would accelerate announcing bad news and delay announcing good news to make the options more valuable. Using a sample of 2039

CEO options that were awarded on a fixed schedule, they find that the analyst earnings forecast error is more negative in award months relative to non-award months. They also find that abnormal returns are insignificant and negative prior to the award date but are significant and positive after the award date.

Because it is difficult to write contracts that specify both the quantity and quality of disclosures, Nagar *et al.* (2003) hypothesize that stock-based compensation should help align the incentives of the manager and the shareholders, leading to more and better disclosures. Nagar *et al.* (2003) calculate the ratio of stock-based compensation (the sum of total value of stock option grants plus the value of restricted stock grants) to total direct compensation. Their results indicate that this ratio is positively related to both the frequency of earnings forecasts provided by management and to the firm's Association for Investment Management and Research (AIMR) rating.

The aforementioned studies focus on the effect of option grants on managerial behavior. However, the realized value of the options will also depend on the stock price at the time of the exercise.[5] Will managers behave opportunistically and adjust overall disclosures to increase the share price in the exercise period? Bartov and Mohanram (2004) examine the question, but their study concentrates on how earnings and discretionary accruals are related to option exercises. In our study, we investigate the relationship between overall accounting disclosures and option exercises, and we take a more detailed look at how this relationship may vary under different circumstances.

2.3. *Disclosures, option exercises, and privation information*

Carpenter and Remmers (2001) examine whether managers use private information in timing the exercise of stock options.[6] They expect to find negative

[5]As many including Huddart and Lang (1996) point out, the Black–Scholes valuation model ignores exercise decisions of the option holder. Although the Black–Scholes model assumes that exercise will occur at the expiration date, Huddart (1994) shows that risk-adverse employees will generally be better off if they exercise the options before the expiration date.

[6]We focus just on Carpenter and Remmers' (2001) theory and results for stock option exercises after May 1991 when managers were able to sell the shares from call option exercises immediately. They also examine stock option exercises prior to May 1991 when managers faced an SEC-required 6-month holding period before they could sell their shares. However, as our data are from 1994 and 1995, we do not review their pre-May 1991 findings.

postexercise abnormal returns if the managers exercise options based on private information, whereas they expect to find no postexercise abnormal returns if the managers exercise for noninformation-related reasons such as liquidity or diversification. Carpenter and Remmers' (2001) empirical results are more consistent with the argument that managers exercise options mainly to diversify risks or to increase liquidity.

Limiting option exercises to abnormally large ones, Bartov and Mohanram (2004) document a significant decline in stock prices in the postexercise period. Thus, different from Carpenter and Remmers (2001), Bartov and Mohanram (2004) find that top management uses private information to time large option exercises. Bartov and Mohanram (2004) also find that managers inflate earnings by recording higher discretionary accruals in the period of large option exercises. These discretionary accruals reverse subsequently, causing disappointing earnings in the postexercise period. Examining a proprietary data set covering stock option exercises of over 50,000 employees at seven corporations, Huddart and Lang (2003) are also able to document employees' option exercises predict subsequent returns. These findings suggest that employees base their option exercise decisions, in part, on private information.

Research on managers' exercise of stock options fits into the strand of research on insider trading. There are studies that look at disclosures and insider trading more generally.[7] For example, Penman (1982) examines whether insider trading and the release of insiders' forecasts are related. He expects that insider sales should precede bad news forecasts, whereas insider purchases should precede good news forecasts. His results are consistent with his expectations, and later studies such as Hirschey and Zaima (1989), Seyhun (1990), Karpoff and Lee (1991), and John and Lang (1991) provide similar findings.

In contrast, other studies find that managers do not seem to utilize insider information to their benefit. Sivakumar and Waymire's (1994) results suggest that managers concentrate their insider transactions after material news events when information asymmetry is low. Noe (1999) looks at insider trading around management earnings forecasts and finds that there are fewer insider

[7] See Beny (2004) for a comprehensive review of the insider trading literature.

transactions before the forecasts than after the forecasts. He suggests that insiders may not exploit their short-term information monopolies.

Thus, although the research is not conclusive, there is evidence that some managers behave opportunistically when buying or selling their firm's shares. Given that managers may take advantage of undisclosed nonpublic information with regard to insider trades, a related question is whether managers adapt their disclosure policies to increase the value of options that they exercise.

3. Hypothesis

Before formally stating our hypothesis, we note that although managers can sell their shares on exercise or hold them, Huddart and Lang (1996) report that most employees in their sample sold the shares immediately.[8] This makes exercises of call options equivalent to insider sales.[9] To sell their shares at a higher price, if there is bad news, managers may refrain from more disclosures to avoid price declines; but if there is good news, managers may want to maintain the original disclosure level or even expand disclosure to further drive up the share price. Thus, the extent to which the disclosure level is decreased to facilitate option exercises may not be identical across firms with good news and bad news. As managers will have incentives to decrease disclosure levels when the news is bad, our hypothesis is formally stated as follows:

H1:Stock option exercises are negatively related to disclosure levels when there is bad news.

Of course, insider trading rules impose penalties for nondisclosures, but whether managers who exercise stock options reduce disclosures to increase the value of their shares in the bad news setting is an empirical question. Additionally, different from Bartov and Mohanram (2004) who report that managers record more positive discretionary accruals in the year of large

[8]In some cases, firms may have share ownership targets that require managers to maintain a certain share holding. Core and Larcker (2002) examine performance changes in response to target ownership plans.

[9]The SEC regulation pertaining to stock options underwent a major change in May 1991 when the SEC eliminated the requirement that the shares acquired through option exercises be held for 6 months before they could be resold. After the new regulation, managers can sell the shares immediately after acquiring them, making the exercise of call options equivalent to insider sales.

option exercises to inflate earnings, our study focuses on the overall accounting disclosure.

4. Method

4.1. *Measurements of main variables*

The empirical proxy for disclosure levels is taken from the Corporate Information Committee Report (CICR) prepared by the Financial Analysts Federation, a branch of the Association for Investment Management and Research (AIMR). When preparing the report, analysts for a specific industry form a subcommittee to evaluate the adequacy of a firm's disclosures along three dimensions: annual and required published information, quarterly and other nonrequired published information, and investor relations. For each firm, the industry committee assigns a score to each dimension and, based on a list of important industry-specific factors, aggregates the three component scores to assign a composite score. Following Bushee and Noe (2000) and Bens and Monahan (2004), this study uses composite disclosure scores, and we label this as DISCLEV.

Several studies have investigated the validity and neutrality of this database.[10] Compared with the information on a firm's particular disclosure choices, the AIMR ratings represent a more comprehensive measure "that includes both quantifiable and nonquantifiable aspects of disclosures" (Lang and Lundholm, 1993, p. 247). Thus, the AIMR data facilitate investigations of a firm's overall disclosure choices. In addition, the AIMR data are free from the editorial bias that may confound the findings obtained from using press announcements because "certain types of firms or announcements may be more likely to be covered in the financial press" (Lang and Lundholm, 1993, p. 247). Disadvantages of the AIMR data are that they may be biased by analysts' perceptions, the firms covered by the ratings may not be representative, and analysts doing the ratings might not take the ratings seriously (Healy and Palepu, 2001).

[10]For a complete list of criteria analysts use to evaluate a firm's disclosure practice and the procedures analysts follow to ensure a fair evaluation, see Lang and Lundholm (1993, 1996), Welker (1995), and Sengupta (1998). More recent research has also used the AIMR disclosure scores as a proxy for disclosure levels (Bushee and Noe 2000; Botosan and Plumlee 2002).

The option exercise variable is calculated in the same manner as Bartov and Mohanram (2004). Specifically, we obtain the ratio of stock option exercises to total compensation[11] for each of the five most highly compensated executives, average this ratio across the executives, and compare the ratio with the average from the past 2 years.[12] If the current year's ratio is up by more than 50%, then there are abnormally large option exercises during the year. We code an indicator variable, OPEX, as one if there are abnormally large option exercises.

To test H1, we need to define good news firms and bad news firms. Like Nagar *et al.* (2003), bad news firms are those experiencing a negative annual return, computed as the market-adjusted compounded monthly return. If a firm experiences a positive annual return, then the firm is in the good news category. Further, because we are interested in the incremental effects on disclosures above and beyond that caused by the granting of stock options, we control the level of option grants. We compute the weight of option grants in total compensation as the ratio of stock options' Black–Scholes values to total compensation, which is then averaged across the five most highly paid executives. We label the variable as GRANT.

4.2. *Model specification*

We use the following primary regression model to test the effect of the option exercises on the disclosure level:

$$\begin{aligned}\mathrm{DISCLEV}_{it} = {} & \alpha_0 + \alpha_1\,\mathrm{OPEX}_{it} + \alpha_2\,\mathrm{GRANT}_{it} + \alpha_3\,\mathrm{SIZE}_{it} + \alpha_4\,\mathrm{CORR}_{it} \\ & + \alpha_5\,\mathrm{STDEV}_{it} + \alpha_6\,\mathrm{PERF}_{it} + \alpha_7\,\mathrm{ANA}_{it} + \epsilon, \end{aligned} \tag{1}$$

[11] Total compensation is calculated as the sum of salary, bonus, other annual pay, values of restricted stock granted, values of stock options granted, long-term incentive pay, and all other payouts. The final year for which AIMR scores are available is 1995. Thus, we are unable to expand the sample period beyond 1995.

[12] Bartov and Mohanram (2004) use a 3-year average as the benchmark. However, as we can only use the compensation data from 1992 to 1995 (see footnote 11), comparison with a 3-year benchmark will only enable us to examine 1995 disclosures in relation to 1995 option exercises. To examine the relation in a multiyear setting, we use the option exercises during the previous 2 years as a comparison basis. We also calculated the correlation between the large option exercise variable obtained from using the 2-year benchmark and that from using the 3-year benchmark. The Pearson correlation is 0.645 ($p < 0.0001$).

where as discussed above DISCLEV is the AIMR rating, OPEX the indicator variable for abnormally large option exercises, and GRANT the weight of stock options in total compensation.[13]

Lang and Lundholm (1993) find that analysts' ratings of corporate disclosure practices are positively related to firm size and firm performance and also positively related to the level of information asymmetry (which is measured as the correlation between earnings and returns and the standard deviation of market-adjusted returns). Nagar *et al.* (2003) further find that disclosure level is positively related to the number of analysts following the firm.

Therefore, we include the following variables as control variables in Equation (1): the log of the market value (price times number of common shares outstanding) at the beginning of the fiscal year (SIZE), market-adjusted annual returns (PERF), the correlation (CORR) between market-adjusted annual returns and annual EPS (before extraordinary items and discontinued operations), and the standard deviation of the market-adjusted annual returns (STDEV). Both CORR and STDEV are computed over the 10-year period $t - 10$ to $t - 1$, where t is the current AIMR rating year. Like Nagar *et al.* (2003), the variable for analyst following (ANA) is obtained by taking the log form of (1 + number of analysts).

When conducting the regression analysis, we rank the original AIMR scores within each industry-year and convert the ranks into percentiles through the transformation: (rank – 1) / (number of firms – 1) (Lang and Lundholm, 1993). Botosan and Plumlee (2002) argue that, although it is appropriate to use the ranks of ordinal measures (such as analyst ratings) in a regression model, using the ranks of cardinal measures reduces the informativeness of the variables and makes it difficult to interpret the magnitude of the coefficients. Therefore, we include all independent variables (which are all cardinals) with their original values. Additionally, to correct for time-series and cross-sectional correlation, we report the Newey–West t-statistics in square brackets.

[13]Bartov and Mohanram (2004) document that discretionary accruals are significantly higher in the year of abnormally high exercise of stock options. In the preceding years, however, the discretionary accruals are not significantly different from those of a control sample. This suggests that option exercises may provide managers incentives to decrease the level of disclosure for the current period. Consequently, we hypothesize and test for a concurrent relationship between option exercises and disclosures.

5. Results

5.1. *Sample and descriptive statistics*

The sample firms are drawn from the following sources: (1) 1994 and 1995 AIMR reports for disclosure level measures, (2) the ExecuComp database for compensation data from 1992 to 1995,[14] and (3) COMPUSTAT, CRSP, and IBES databases for all other variables. All variables are measured on the fiscal year basis. Table 1 summarizes the sample selection procedures, the application of which yields a total of 359 firm-year observations spanning over 18 industries. The largest reduction of the sample occurs when we require that firms covered by the AIMR reports also be covered by the ExecuComp database; 148 observations are eliminated as a result of the requirement.

Table 2 classifies firm-year observations based on industries defined by the AIMR reports. The observations are mostly from the food, beverage, and tobacco industry, insurance industry, paper and forest industry, and retail trade industry; the combined observations account for 45.68% (164 out of 359) of the sample. Observations from industries such as airline, environmental control, homebuilding, and natural gas-distribution are relatively scant.

Table 1. Sample screening process.

Firm-years appearing in the 1994 and 1995 AIMR reports		541
Less:	Observations for which compensation data are not available in the ExecuComp database from 1992 to 1995	148
Less:	Observations for which the abnormal option exercise variable cannot be computed for 1994 or 1995 from the ExecuComp database	16
Less:	Observations for which there is inadequate information in the COMPUSTAT, CRSP, or IBES data sets for computing control variables	18
Final sample		359

Notes: The abnormal option exercise variable is calculated in the same manner as Bartov and Mohanram (2004). Specifically, we obtain the ratio of stock option exercises to total compensation for each of the five most highly compensated executives, average this ratio across the executives, and compare the ratio with the average from the past 2 years. If the current year's ratio is up by more than 50%, then there are abnormally large option exercises during the year.

[14]Executive compensation data for 1992 and 1993 are necessary to calculate the option exercise measure (OPEX).

Table 2. Sample composition based on industry-years.

Industry	1994	1995
Airline	6	6
Automotive and related products	—	10
Chemical	8	—
Electrical equipment	11	10
Environmental control	5	4
Food, beverage, and tobacco	23	26
Health care	14	13
Homebuilding	—	6
Insurance	19	19
Machinery	12	—
Media	14	12
Natural Gas—distribution	5	7
Natural Gas—pipeline	7	—
Paper and forest products	18	16
Petroleum—domestic	9	9
Precious metals mining	7	7
Railroad	7	6
Retail trade	21	22
Total	186	173

Table 3A reports descriptive statistics for the sample. The disclosure scores are in percentages,[15] averaging at 73.2% with a median of 74.2%. For the firm-year observations in our sample, the proportion of executive compensation derived from option exercises is 9.7%. The abnormal option exercise variable averages at 0.382, indicating that 38.2% of the firm-year observations saw a 50% jump in the ratio of option exercise benefits to total compensation. In general, the Black–Scholes values of stock option grants account for 24.0% of top executives' compensation.

The sample has a mean market value of $7087 million. In comparison, the average market value of a larger group of firms, covered by the ExecuComp database and having the stock option exercise variable and other variables available in 1994 and 1995, is $2645 million (untabulated). These statistics point to the large firm bias of the AIMR coverage. The average correlation between earnings and firms' market performance computed over the preceding 10 years is 0.038. The sampled observations also experienced a negative return

[15]When nonpercentage scores are reported, the scores are converted into percentages by dividing the firm's actual score by the total available score.

Table 3A. Sample summary statistics ($N = 359$).

Measures	Mean	Standard Deviation	Minimum	Median	Maximum
AIMR ratings (%)	73.224	13.752	36.5	74.2	95.7
Exercise ratios	0.097	0.121	0	0.051	0.491
Abnormally large option exercises (a 0/1 variable)	0.382	0.486	0	0	1
Equity grant ratios	0.240	0.169	0	0.215	0.717
Market value (in billions)	7.087	11.414	0.188	3.347	60.911
Correlation between earnings and returns	0.038	0.373	−0.711	−0.004	0.872
Standard deviation of the returns	0.270	0.197	0.087	0.218	0.820
Market-adjusted returns	−0.025	0.266	−0.678	−0.025	0.744
Number of analysts following	8.969	6.720	0	9	25

Notes: AIMR ratings are in percentages. When nonpercentage scores are reported, the scores are converted into percentages by dividing the firm's actual score by the total available score.

Exercise ratios are the portion of compensation derived from stock option exercises, averaged across the five most highly compensated executives.

Abnormally large option exercises are as defined in Table 1. The variable is coded as 1 if the observation has abnormally large option exercises during the year.

Equity grant ratios are the ratio of values of stock options (based on the Black–Scholes formula) to total compensation, averaged across the five most highly paid executives.

Market value is the price multiplied by number of common shares outstanding at the beginning of the fiscal year.

Correlation between earnings and returns is computed between the firm's annual returns (adjusted for value-weighted market returns) and annual EPS (before extraordinary items and discontinued operations) over the previous 10 years.

Standard deviation of the returns is the standard deviation of the firm's annual returns (adjusted for value-weighted market returns), computed over the previous 10 years.

Market-adjusted returns are compounded monthly returns, adjusted for value-weighted market returns.

Number of analysts following is obtained from the IBES data.

(after adjusting for market returns), −2.5%, which is significantly below zero. This is probably caused by the sample's large firm bias (as large firms tend to earn lower returns). Finally, on average, the number of analysts following the firms has a mean of 8.9 and a median of 9.

Table 3B tabulates the correlation matrix for the variables appearing in the primary regression. There is a significant negative correlation between disclosure scores DISCLEV and large stock option exercises OPEX ($r = -0.097$),

Table 3B. Pearson correlations for variables appearing in the regression model ($N = 359$).

	OPEX	GRANT	SIZE	CORR	STDEV	PERF	ANA
DISCLEV	−0.097*	0.074	0.297***	−0.026	−0.095*	−0.016	0.032
OPEX		−0.027	−0.139***	0.027	0.015	0.171***	−0.024
GRANT			0.175***	−0.105**	0.115**	0.025	−0.008
SIZE				−0.311***	−0.282***	−0.004	0.198***
CORR					0.040	−0.061	−0.117**
STDEV						0.038	−0.039
PERF							−0.068

Notes: DISCLEV is the disclosure scores. The reported scores are transformed to industry-year percentiles through the following: (rank−1) / (number of firms−1).

OPEX is coded as 1 if the observation has abnormally large option exercises during the year. See Table 1 for the definition of abnormally large option exercises.

GRANT is the ratio of the Black–Scholes values of stock option grants to total compensation, averaged across the five most highly paid executives.

SIZE is the log of firms' market value (price multiplied by number of common shares outstanding) at the beginning of the fiscal year.

CORR is the correlation between earnings and returns, computed between firms' annual returns (adjusted for value-weighted market returns) and annual EPS (before extraordinary items and discontinued operations) over the previous 10 years.

STDEV is the standard deviation of firms' annual returns (adjusted for value-weighted market returns) over the previous 10 years.

PERF is the market-adjusted returns, obtained by compounding monthly returns and then adjusting for value-weighted market returns.

ANA is the log form of (1 + number of analysts).

*Significant at the 0.10 level (two-tailed).

**Significant at the 0.05 level (two-tailed).

***Significant at the 0.01 level (two-tailed).

consistent with our hypothesis for the bad news subsample. Consistent with previous research, DISCLEV is positively associated with the size variable SIZE ($r = 0.297$) and negatively associated with one of the information asymmetry variables STDEV ($r = -0.095$). The option grant variable GRANT is positively correlated with SIZE ($r = 0.175$), suggesting that larger firms are more likely to use options as a form of compensation. Furthermore, GRANT and CORR have a negative correlation ($r = -0.105$) but GRANT and STDEV have a positive correlation ($r = 0.115$). As lower CORR and higher STDEV imply higher information asymmetry, these two correlation coefficients suggest that the use of stock option grants is related with greater information asymmetry. The correlation with the highest absolute value is between CORR and SIZE ($r = -0.311$). We believe that this poses no serious multicollinearity problem for the model.[16]

5.2. *Regression results from partitioned samples*

We divide the sample based on news type. This process yields a bad news subsample containing 166 firm-years and a good news subsample containing 193 firm-years. Table 4 contains results of model estimations for this subdivision. The first column is for the good news subsample, whereas the second column for the bad news. Both models are significant with adjusted R^2 of 6.10% and 15.18%, respectively. The variable of interest is OPEX. The results show that it is significantly negative (Newey–West t-statistics $= -2.246$) in the bad news regression yet is insignificant in the good news regression. GRANT is significant and positive in the bad news regression (t-statistics $= 1.719$),

Table 4. Regression models for the sample partitioned by good and bad news.

	Good news	Bad news
Constant	−0.330 [−1.095]	−0.935*** [−3.306]
OPEX	0.046 [0.893]	−0.094** [−2.246]
GRANT	−0.885 [−0.670]	0.238* [1.719]
SIZE	0.141*** [3.159]	0.219*** [4.852]
CORR	0.057 [0.870]	0.077 [1.452]
STDEV	0.007 [0.053]	−0.041 [−0.260]
PERF	−0.259*** [−2.831]	0.212* [1.749]
ANA	−0.055 [−1.147]	0.045 [1.066]
N	166	193
F-statistics	2.53 ($p < 0.05$)	5.91 ($p < 0.00$)
Adjusted R^2	6.10%	15.18%

Notes: Bad news firms are those experiencing a negative annual return, computed as the market-adjusted compounded monthly return. If a firm experiences a positive annual return, then the firm is in the good news category. All variables are as defined in Table 3B. Newey–West t-statistics are reported in square brackets.
*Significant at the 0.10 level (two-tailed).
**Significant at the 0.05 level (two-tailed).
***Significant at the 0.01 level (two-tailed).

[16]We tested the model's sensitivity to the correlation between SIZE and other independent variables by dropping the variable SIZE from the regression. The results pertaining to OPEX remained unchanged.

confirming Nagar *et al.*'s (2003) finding that stock-based compensation mitigates the managerial disclosure agency problem by providing positive incentives for disclosures. Thus, after we control the positive disclosure incentives brought about by the grants of stock options, we obtain evidence supporting our hypothesis that when there is bad news, executives' decision to exercise options will decrease the disclosure level but this may not be the case when there is good news.

As for other control variables, SIZE is highly significant in both regressions. In the bad news regression, there is evidence that PERF is positively associated with DISCLEV (t-statistics $= 1.749$); but in the good news regression, the performance variable is unexpectedly negative (t-statistics $= -2.831$).[17]

5.3. *Results from regressions with interactions*

We also form the interaction between option exercises and news type, and estimate the following equation for the entire sample:

$$\begin{aligned}\mathrm{DISCLEV}_{it} = {} & \alpha_0 + \alpha_1\,\mathrm{OPEX}_{it} + \alpha_2\,\mathrm{OPEX}_{it} \times \mathrm{GOODNEWS}_{it} \\ & + \alpha_3\,\mathrm{GRANT}_{it} + \alpha_4\,\mathrm{SIZE}_{it} + \alpha_5\,\mathrm{CORR}_{it} + \alpha_6\,\mathrm{STDEV}_{it} \\ & + \alpha_7\,\mathrm{PERF}_{it} + \alpha_8\,\mathrm{ANA}_{it} + \eta, \end{aligned} \tag{2}$$

where the variable GOODNEWS (a 0/1 variable) is as previously defined. We expect that the coefficient on OPEX $\times$ GOODNEWS to be positive. The reason is that, relative to bad news firms, managers in good news firms will have weaker incentives to conceal or distort real information when exercising their options. Table 5 contains the results.

In the first column where the interaction term is absent from the model, OPEX is not significant. In the second column, after adding the interaction OPEX $\times$ GOODNEWS, OPEX becomes significantly negative with a Newey–West t-statistics -2.024 ($p < 0.05$, two-tailed). Also, consistent with our expectation, OPEX $\times$ GOODNEWS has a significantly positive coefficient (t-statistics $= 1.748$). Thus, the impact of option exercises on disclosure levels is conditioned on the news type because OPEX becomes significantly negative only after we control the news type by adding the interaction term. Of the

[17] The difference in the findings between our paper and prior research (e.g., Lang and Lundholm, 1993; Nagar *et al.*, 2003) could be due to different sample size and sample period examined.

Table 5. Regression model including the interaction term as independent variable.

	Without interaction	With interaction
Constant	−0.646*** [−3.165]	−0.665*** [−3.280]
OPEX	−0.036 [−1.070]	−0.088** [−2.024]
OPEX × GOODNEWS		0.113* [1.748]
GRANT	0.051 [0.534]	0.040 [0.424]
SIZE	0.178*** [5.717]	0.181*** [5.843]
CORR	0.061 [1.448]	0.054 [1.299]
STDEV	−0.023 [−0.186]	−0.021 [−0.185]
PERF	−0.003 [−0.052]	−0.074 [−1.192]
ANA	−0.014 [−0.425]	−0.014 [−0.437]
N	359	359
F-statistics	5.42 ($p < 0.01$)	5.20 ($p < 0.01$)
Adjusted R^2	7.95%	8.57%

Notes: Variable GOODNEWS (having positive market-adjusted annual returns; a 0/1 variable) is as defined in Table 4. All other variables are as defined in Table 3B. Newey–West t-statistics are reported in square brackets.
*Significant at the 0.10 level (two-tailed).
**Significant at the 0.05 level (two-tailed).
***Significant at the 0.01 level (two-tailed).

control variables, SIZE is significant with a positive sign in both models at the 0.01 level. This indicates that larger firms tend to have higher AIMR ratings, consistent with Lang and Lundholm (1993).

In summary, we find consistent evidence supporting H1. For bad news subsample, the coefficient on OPEX is negative and significant; for good news subsample, the negative relationship between option exercises and disclosures is absent. We obtain similar results by analyzing the interaction term using the whole sample. Thus, our finding suggests that it is important that we take different settings into consideration when we attempt to assess the impact of option exercises on disclosures.

6. Discussion and Conclusion

The purpose of this paper is to examine whether stock option exercises and discretionary disclosures are related. In doing so, we also test the assertions of some business commentators (e.g., Clambering back up, 2002) that stock option exercises create incentives for managers to manipulate disclosures opportunistically. Our evidence shows that option exercises are negatively related to the disclosure level for a subset of firms that are experiencing negative market-adjusted returns. Thus, our study finds limited support for the observation made by certain politicians and by members of the business press (e.g., Use and abuse, 2002).

We contribute to the literature by extending two recent studies that examine the managerial incentives created by stock options. We extend Nagar *et al.* (2003) by showing that stock options not only can affect disclosures when the options are granted, but also can affect disclosures in the period when they are exercised if the amount of options exercised is large and if the firm has bad news. Thus, we show that stock options can induce ex post opportunism in certain settings, and this effect is incremental to the effect documented by Nagar *et al.* (2003). Further, we extend Bartov and Mohanram (2004) by showing that the opportunistic effect that they find in periods of abnormally high option exercises can be generalized beyond discretionary accruals. Specifically, we show that managers in bad news firms can also behave opportunistically with respect to other types of accounting disclosures.

Like all research, our study has certain limitations. Our study does not pair the compensation received by an executive with the option exercising activities carried out by that same individual. Instead, the compensation and option exercise information is aggregated across individual executives to form firm-level measurements. Measuring option exercises and compensation at the firm level corresponds to the level of disclosure measured at the firm level. However, in doing so, we implicitly assume that the same group of managers both make disclosure decisions and exercise options. To the extent that this assumption does not hold in practice, measuring variables at the firm level introduces noise into our empirical tests.

Also, we note that managers' opportunistic behavior in the context of executive compensation and private information might not be suboptimal. For example, allowing managers to exploit inside information may allow firms to adjust compensation levels when recontracting costs are nonzero (Manne, 1966; Carlton and Fischel, 1983).

We suggest several areas for future research. First, a new proxy for disclosure levels is necessary to overcome the large firm bias in AIMR ratings and to extend the time period beyond what is examined by this study. Second, other possible research venues (e.g., survey or experimental economics research) could be explored to investigate the impact of stock option exercises on the level of disclosure in a more controlled setting. Finally, the endogeneity of option exercises and the simultaneity between option exercises and disclosure levels should be addressed. For this purpose, the determinants of option exercises should be identified at both theoretical and empirical levels. We leave this task for future research.

References

Aboody, D. and R. Kasznik, "CEO Stock Option Awards and the Timing of Corporate Voluntary Disclosure." *Journal of Accounting and Economics* 29(1), 73–100 (2000).

Bartov, E. and P. Mohanram, "Private Information, Earnings Manipulations, and Executive Stock-option Exercises." *Accounting Review* 79(4), 889–920 (2004).

Bens, D. and S. Monahan, "Disclosure Quality and the Excess Value of Diversification." *Journal of Accounting Research* 42(4), 691–730 (2004).

Beny, L., "*A Comparative Empirical Investigation of the Agency and Market Theories of Insider Trading*." Working Paper Series #04–004, University of Michigan Law School (2004).

Botosan, C., "Disclosure Level and the Cost of Equity Capital." *The Accounting Review* 72(3), 323–349 (1997).

Botosan, C. and M. Plumlee, "A Re-examination of Disclosure Level and the Expected Cost of Capital." *Journal of Accounting Research* 40(1), 21–40 (2002).

Brennan, N., "Voluntary Disclosure of Profit Forecasts by Target Companies in Takeover Bid." *Journal of Business Finance and Accounting* 26(7/8), 883–918 (1999).

Bushee, B. and C. Noe, "Corporate Disclosure Practices, Institutional Investors, and Stock Return Volatility. *Journal of Accounting Research* 38 (Supplement), 171–202 (2000).

Carlton, D. and D. Fischel, "The Regulation of Insider Trading." *Stanford Law Review* 35(5), 857–895 (1983).

Carpenter, B. and B. Remmers, "Executive Stock Option Exercises and Inside Information." *Journal of Business* 74(4), 513–534 (2001).

Clambering back up. *The Economist* 364 (8282), 55 (2002).

Core, J. and W. Guay, "Stock Option Plans for Non–executive Employees." *Journal of Financial Economics* 61(2), 253–287 (2001).

Core, J. and D. Larcker, "Performance Consequences of Mandatory Increases in Executive Stock Options." *Journal of Financial Economics* 64(3), 317–340 (2002).

Core, J., W. Guay and D. Larcker, "Executive Equity Compensation and Incentives: A Survey." *Economic Policy Review—Federal Reserve Bank of New York* 9(1), 27–50 (2003).

Darrough, M. and N. Stoughton, "Financial Disclosure Policy in an Entry Game." *Journal of Accounting and Economics* 12(1–3), 219–244 (1990).

Diamond, D. and R. Verrecchia, "Disclosure, Liquidity, and the Cost of Capital." *Journal of Finance* 46(4), 1325–1355 (1991).

Francis, J., D. Philbrick and K. Schipper, "Shareholder Litigation and Corporate Disclosures." *Journal of Accounting Research* 32(2), 137–165 (1994).

Hall, B. J. and J. B. Liebman, "Are CEOs Really Paid like Bureaucrats?" *Quarterly Journal of Economics* 113(3), 653–691 (1998).

Hayes, R. M. and R. Lundholm, "Segment Reporting to the Capital Market in the Presence of a Competitor." *Journal of Accounting Research* 34(2), 261–280 (1996).

Healy, P. and K. Palepu, "Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Literature." *Journal of Accounting and Economics* 31(1–3), 405–440 (2001).

Hirschey, M. and J. K. Zaima, "Insider Trading, Ownership Structure, and the Market Assessment of Corporate Sell-offs." *Journal of Finance* 44(4), 971–980 (1989).

Huddart, S., "Employee Stock Options." *Journal of Accounting and Economics* 18(2), 207–231 (1994).

Huddart, S. and M. Lang, "Employee Stock Option Exercises: An Empirical Analysis." *Journal of Accounting and Economics* 21(1), 5–43 (1996).

Huddart, S. and M. Lang, "Information Distribution Within Firms: Evidence from Stock Option Exercises." *Journal of Accounting and Economics* 34(1–3), 3–31 (2003).

Ittner, C., R. Lambert and D. Larcker, "The Structure and Performance Consequences of Equity Grants to Employees of New Economy Firms." *Journal of Accounting and Economics* 34(1–3), 89–127 (2003).

John, K. and L. Lang, "Insider Trading around Dividend Announcements: Theory and Evidence." *Journal of Finance* 46(4), 1361–1389 (1991).

Karpoff, J. M. and D. Lee, "Insider Trading before New Issue Announcements." *Financial Management* 20(1), 18–26 (1991).

Lang, M. H. and R. J. Lundholm, "Cross-sectional Determinants of Analyst Ratings of Corporate Disclosures." *Journal of Accounting Research* 31(2), 246–271 (1993).

Lang, M. H. and R. J. Lundholm, "Corporate Disclosure Policy and Analyst Behavior. *The Accounting Review* 71(4), 467–492 (1996).

Leftwich, R., R. Watts and J. Zimmerman, "Voluntary Corporate Disclosure: The Case of Interim Reporting." *Journal of Accounting Research* 19 (Supplement) 50–77 (1981).

Manne, H., *Insider Trading and the Stock Market*. New York: Free Press (1966).

Miller, G. and J. Piotroski, "*The Role of Disclosure for High Book-to-market Firms.*" Unpublished Working Paper, Harvard University (2000).

Nagar, V., D. Nanda and P. Wysocki, "Discretionary Disclosure and Stock-based Incentives." *Journal of Accounting and Economics* 34(1–3), 283-309 (2003).

Noe, C., "Voluntary Disclosure and Insider Trading." *Journal of Accounting and Economics* 27(3), 305–326 (1999).

Penman, S., "Insider Trading and Dissemination of Firms' Forecast Information." *Journal of Business* 55(4), 479–503 (1982).

Piotroski, J., "*Discretionary Segment Reporting Decisions and the Precision of Investor Beliefs*." Working Paper, University of Chicago (1999).

Sengupta, P., "Corporate Disclosure Quality and the Cost of Debt." *The Accounting Review* 73(4), 459–474 (1998).

Seyhun, N., "Overreaction to Fundamentals: Some Lessons from Insiders' Response to the Market Crash of 1987." *Journal of Finance* 45(5), 1363–1388 (1990).

Sivakumar, K. and G. Waymire, "Insider Trading Following Material News Events: Evidence from Earnings." *Financial Management* 23(1), 23–32 (1994).

Skinner, D., "Why Firms Voluntarily Disclose Bad News." *Journal of Accounting Research* 32(1), 38–61 (1994).

Skinner, D., "Earnings Disclosures and Stockholder Lawsuits." *Journal of Accounting and Economics* 23(3), 249–282 (1997).

Trueman, B., "Why do Managers Voluntarily Release Earnings Forecasts?" *Journal of Accounting and Economics* 8(1), 53–72 (1986).

Use and abuse. *The Economist* 364 (8282), 14 (2002).

Verrecchia, R., "Discretionary Disclosure." *Journal of Accounting and Economics* 5(3), 179–194 (1983).

Watts, R. and J. Zimmerman, "Toward a Positive Theory of the Determination of Accounting Standards." *The Accounting Review* 53(1), 273–305 (1978).

Watts, R. and J. Zimmerman, *Positive Accounting Theory.* Englewood Cliffs, NJ: Prentice-Hall (1986).

Welker, M., "Disclosure Policy, Information Asymmetry, and Liquidity in Equity Markets." *Contemporary Accounting Research* 11 (2), 801–827 (1995).

Yermack, D., "Good Timing: CEO Stock Option Awards and Company News Announcements." *Journal of Finance* 52 (2), 449–476 (1997).

Chapter 5

Do Profit Warnings Convey Information About the Industry?

Dave Jackson
University of Texas–Pan American, USA

Jeff Madura*
Florida Atlantic University, USA

Judith Swisher
Western Michigan University, USA

We find that profit warnings result in negative industry effects, which indicate that profit warnings convey negative industry information rather than favorable information for industry rivals about the competition. Multivariate analyses show that profit warnings carry stronger industry signals when the warning firm has a greater stock price adjustment, are larger, and have more analyst coverage. The adverse industry effects in response to profit warnings are attenuated since the inception of Regulation Fair Disclosure (RFD). However, the sensitivity of industry effects to the warning firm's stock price adjustment and size has increased since RFD. Adverse industry effects are also attenuated when market sentiment is more favorable.

Keywords: Profit warnings; industry effects; regulation; fair disclosure.

1. Introduction

The market response to earning information is well documented. Earning announcement effects have been detected in numerous studies, such as Brown (1978), Watts (1978), and Rendleman *et al.* (1982). More recently, researchers have studied the market reaction to profit warnings issued by firms, which indicate lowered profit expectations on the part of management. Profit warning announcement effects have been found by studies such as Libby and Tan (1999), Heflin *et al*. (2003), and Jackson and Madura (2003). Less is known about whether the market perceives a profit warning to be firm-specific or to have industry-wide implications. If a firm's profit warning results from underlying conditions that are industry-wide, the impact should spread beyond

*Corresponding author.

the individual firm to other industry firms as the market revalues nonwarning firms in response to the information.

Our first objective is to determine whether profit warnings convey information only about the warning firms or whether they also carry implications for the corresponding industries. Second, we analyze the conditions under which warnings carry information about the corresponding industries. As part of this analysis, we examine the effect of Regulation Fair Disclosure (RFD) on the industry response to profit warnings. If firms increase the frequency with which they release information, the information content of each announcement could be reduced. Similarly, if more firms within an industry release information, the industry prices will more quickly adjust to include industry-wide information, and the information content of an individual announcement for industry firms could be reduced.

Overall, we find evidence of significant negative industry effects in response to profit warnings. In addition, our multivariate analyses indicate that industry effects are significantly related to firm, industry, and market characteristics.

2. Related Literature and Hypotheses

2.1. *Related literature*

There is substantial evidence that a firm's profit warning announcements contain valuable information about the firm. Evidence also shows that some firm-specific announcements convey industry-wide information. Akhigbe *et al*. (1997) find evidence that a bond rating downgrade inflicts damage on the corresponding industry of the downgraded firm. Lang and Stulz (1992) detect negative industry effects due to bankruptcy announcements. Other studies show that more subtle forms of information can alter the valuations of firms in the same industry. Foster (1981) and Clinch and Sinclair (1987) find industry effects in response to earning announcements, whereas Firth (1996), and Laux *et al*. (1998) detect industry effects in response to dividend announcements. However, Hertzel (1991) finds that although firm-initiated announcements of stock repurchases affect the values of the repurchasing firms, they have negligible effects on other firms in the same industry.

Our research focuses on profit *warnings*, not actual earning announcements or earning forecasts by analysts. Profit warnings are unique in that they disclose the firm's perspectives that the market expectations of its profits are too high. Although a warning does not represent a report of recent quarterly earnings, it could be seen as indicative of negative future performance that the management has voluntarily disclosed. As profit warnings may result from the firm's inside information, they should contain more reliable information than bond rating downgrades issued by independent agencies.[1]

The underlying cause for the warning could be firm-specific or the result of conditions that affect the corresponding industry. If the underlying conditions that precipitate a firm's profit warnings are unique to the firm, a profit warning may affect only the firm. However, if the underlying conditions reflect industry-wide factors, the effects of the warning may spill over to industry rivals. Even though this information is firm-specific, it could be attributed to industry-wide weakness, and should be a cause for industry concern.

Alternatively, bad news about the firm may create opportunities for rival firms. Such competitive effects have been tested and detected for specific types of bankruptcy events by Lang and Stulz (1992) and for specific types of dividend reductions by Laux *et al.* (1998).

2.2. *Factors that could influence industry effects*

The impact of a firm's profit warnings on other firms could be attributed to the following factors.

Regulation fair disclosure. RFD requires that all material information disclosed by a firm is simultaneously released to all investors, not only to analysts, institutional investors, or any other special-interest group. The SEC implemented this rule because of widespread evidence of selective disclosure to analysts. Heflin *et al*. (2001) assess stock price volatility surrounding earning announcements before and after RFD and conclude that RFD does not reduce the amount of information to the public before an earning announcement. Similarly, Bailey *et al.* (2003) find no increase in volatility around earning announcements following RFD. Jackson and Madura (2006) find that the negative valuation effects of profit warnings are reduced following the introduction of RFD, and that the information leakage is reduced. These results

[1]A small proportion of profit warnings are issued because the expected earnings are more favorable than market expectations. This type of profit warning is not considered in this study.

suggest that the market is more informed about firm-specific information that could shock the share price.

Bailey *et al.* (2003) find that firms more frequently provide voluntary disclosures following RFD. To the extent that the public has access to more information following RFD, industry conditions could already be reflected in stock prices before a firm issues a profit warning. Investors should not need to rely as heavily on profit warning information about a firm to derive valuations of other firms in the industry. Thus, industry effects from profit warnings should be reduced since the inception of RFD.

It is also possible that the contrasting market conditions of the two time periods affect the industry response. From an ex post perspective, a bull market existed during the pre-RFD period, whereas a bear market existed during the post-RFD sample. However, the market participants would not necessarily have recognized these differences on an ex ante basis. We control market sentiment at the time of each profit warning with a market sentiment variable. To the extent that investors consider macroenvironment when assessing the market, they may punish related industry stocks to a smaller degree when market sentiment is more favorable.

Size of the surprise contained in the profit warning. A larger surprise (revised estimate of earnings relative to expected earnings) represented by the profit warning should emit a larger signal and cause more pronounced industry effects.

Revaluation of the warning firm. If a firm experiences no stock price adjustment from issuing a profit warning, the market must have already processed the information or does not view the information to be relevant. In this case, the effect on the industry should be negligible. Conversely, if the profit warning contains new information, as evidenced by the size of the revaluation of the firm that issues the warning, and the warning contains industry information, then we expect the market to inflict more damage on the industry.

Size of the warning firm. The profits of larger firms are more likely to reflect industry conditions, whereas the profits of smaller firms could be attributed to situations regarding a single supplier, a single customer, or other firm-specific conditions. Hong *et al.* (2000) suggest that bad news of a well-known firm that has a greater presence in the market is expected to emit a stronger signal about industry conditions. As large firms control a larger market share, their financial conditions are more likely to indicate information about the industry.

We, therefore, hypothesize that the profit warnings of larger firms will emit more pronounced industry signals.

Analyst coverage of the warning firm. Ivkovic and Jegadeesh (2004) assess the information content of analyst forecasts and note that market participants clearly value analysts' forecasts. To the extent that analysts rely on profit warnings as an indicator of the industry, profit warnings by firms with a greater degree of analyst coverage may convey more information. In addition, those firms that receive more coverage tend to receive more media attention, and their profit warnings may reach a wider set of investors, who could act on the signal emitted. The relatively larger collective action will result in more pronounced industry effects. Bhushan (1989) points out that analyst coverage is strongly correlated with firm size. Thus, we also expect warnings by firms with larger numbers of analysts to have greater industry effects.

Sequence of the profit warnings. A profit warning that is the first in the industry within a given quarter could be expected to signal more valuable information about the industry than subsequent profit warnings in the same industry in that same quarter. The subsequent warnings may have been partially anticipated, and the market response attenuated, if the market detected the industry signal from the first warning in the quarter for the corresponding industry. However, a counterargument is that subsequent announcements could signal underlying industry problems, as more than one firm in the industry has issued a warning.

3. Sample Selection

We obtain our sample of profit warnings from First Call. The First Call database is a comprehensive record of profit forecasts by firms along with the concurrent analyst consensus forecasts for each firm. The CRSP database is our source of stock price information and firm market values. Compustat is our source for SIC codes used for industry matching and for sales data used in construction of the Herfindahl index.[2]

First Call surveys analysts and compiles a consensus forecast, which is made available to the public via the electronic media. The individual analyst forecasts are not made public by First Call, but in some cases are publicized

[2]The Herfindahl index is a measure of the degree of industry concentration (see, e.g., Lang and Stultz, 1992).

by the analysts themselves. There is no set timing of a profit warning by a firm in relation to First Call's consensus estimate. In some cases, the firm responds to First Call's estimate with a public profit warning and in other cases the timing of the firm's profit warning is based on other factors. In the event that both the firm and First Call provide profit estimates, the data are presented in First Call on a similar basis (i.e., extraordinary income is indicated where relevant).

Our initial data set includes all profit forecasts made by firms between October 1, 1998 and September 30, 2001 as recorded in the First Call database. After applying several screens that are described next, our final sample consists of 1421 profit warnings. Table 1 summarizes the sample determination process. During our sample time period, First Call has information on 7766 announcements by firms regarding profit expectations. First Call uses a coding system to classify announcements based on the nature of the information. We drop 2860 announcements that confirm analysts' forecasts (First Call code

Table 1. Sample determination.

	Number of announcements
Total announcements in First Call during sample period	7766
Confirm analyst forecasts (First Call code M)	2860
Remaining	4906
Warning within the prior 2 weeks	99
Remaining	4807
Missing data (CRSP or Compustat)	1004
Remaining	3803
REITs	58
Remaining	3745
Convey positive news (First Call code E)	646
Remaining	3099
Missing analyst forecasts or company guidelines	1203
Number of analysts not available	365
Remaining	1531
Misclassified by First Call	93
Remaining	1438
No industry firms with same 4-digit SIC code	17
Final sample	1421

Notes: This table reports the process by which the sample is determined. The final sample consists of 1421 profit warnings issued by firms between October 1, 1998 and September 30, 2001.

M), leaving 4906 announcements that provide new information to investors. As firms often make multiple forecasts in a single announcement, these 4906 announcements include forecasts for 6204 quarters or years. We limit our analysis to one forecast per announcement and keep the forecast for the period end that is closest to the announcement date. If a firm warns about annual and quarterly profits with the same period end, we use the annual warning. Of these 4906 forecasts, we eliminate 99 forecasts by firms that warned within the prior 2 weeks.

The requirement that returns are available in CRSP and that SIC codes and sales data are available from Compustat eliminates 1004 announcements. As the inclusion of REITs and mutual funds has been shown to distort the results for other types of firms (see, e.g., Ling and Ryngaert, 1997), we also drop 58 forecasts by REITS or mutual funds, reducing our sample to 3745 announcements. A small portion of these announcements conveys favorable information about future profits (First Call code E). As our focus is on profit warnings, we drop the 646 announcements with positive news.

We are left with the announcements that First Call defines as profit warnings (First Call code D): announcements that indicate that the firm will fall short of the analyst consensus estimates, or announcements indicating that the firm's performance will be at the lower end of a range of analysts' forecasts. In the absence of consensus estimates, First Call classifies firm forecasts as warnings if the new forecast indicates erosion in profits relative to the previously announced projections. We drop 1203 announcements that are missing either analyst forecasts or company-issued guidelines in the First Call data. This requirement reduces our sample to 1896 profit warnings. Requiring data on the number of analysts eliminates 365 firms, reducing the sample to 1531 warnings. We delete an additional 93 announcements that are classified as profit warnings (First Call code D), but have analyst forecasts that are lower than the company guidelines. We also drop 17 firms for which no industry firms (using 4-digit SIC codes) are available. Our final sample consists of 1421 announcements by firms that profits will fall short of prior estimates.

4. Descriptive Statistics

Descriptive statistics for the sample are displayed in Table 2 for the whole sample and also partitioned for periods before and after RFD. Results are also

Table 2. Descriptive statistics for a sample of firms that issue profit warnings.

	Whole sample		Surprise %			
	Mean	Median	≤10%	−10% to −5%	−5% to −1%	−1% to 0%
Analysts forecast ($)						
Whole sample	0.461	0.295	0.115	0.085	0.264	0.543
Pre-RFD	0.470	0.310	0.053	0.074	0.278	0.555
Post-RFD	0.454	0.280	0.171	0.095	0.254	0.534
Profit warnings ($)						
Whole sample	0.458	0.295	0.098	0.078	0.258	0.543
Pre-RFD	0.468	0.310	0.045	0.067	0.271	0.555
Post-RFD	0.451	0.280	0.146	0.089	0.248	0.533
Surprise ($)						
Whole sample	−0.002	0.000	−0.017	−0.007	−0.006	−0.001
Pre-RFD	−0.002	0.000	−0.008	−0.007	−0.006	−0.001
Post-RFD	−0.003	0.000	−0.025	−0.007	−0.006	−0.001
Surprise						
Whole sample	−0.013	0.000	−0.253	−0.066	−0.022	−0.001
Pre-RFD	−0.013	0.000	−0.219	−0.066	−0.020	−0.001
Post-RFD	−0.014	0.000	−0.283	−0.066	−0.023	−0.001
Number of analysts						
Whole sample	8.4	7.0	7.2	6.7	8.2	8.6
Pre-RFD	7.7	6.0	6.5	7.4	7.3	7.9
Post-RFD	9.0	8.0	7.8	6.0	8.9	9.2
Firm market value ($ millions)						
Whole sample	6599.9	1055.7	1426.8	1907.3	7594.2	6599.9
Pre-RFD	5319.6	845.6	724.9	1522.0	6901.6	5319.6
Post-RFD	7560.2	1217.9	2050.7	2292.6	8066.5	7560.2
Number of observations						
Whole sample	1421	1421	34	44	296	1047
Pre-RFD	609	609	16	22	120	451
Post-RFD	812	812	18	22	176	596
Percent of observations						
Whole sample (%)	100.0	100.0	2.4	3.1	20.8	73.7
Pre-RFD (%)	100.0	100.0	2.6	3.6	19.7	74.1
Post-RFD (%)	100.0	100.0	2.2	2.7	21.7	73.4

Note: The descriptive statistics are for a sample of 1421 profit warnings issued by firms between October 1, 1998 and September 30, 2001. Statistics are also reported separately for the 609 warnings during the time prior to October 23, 2000, when RFD became effective and for the 812 warnings following the implementation of RFD. Analyst forecasts and profit warnings are obtained from First Call. Surprise ($) is calculated as (profit warning–analyst forecast). Surprise is calculated as (surprise ($)/absolute value of analyst forecast). Firm market value is determined 1 month (20 trading days) prior to the announcement.

broken out by the size of the surprise, where the surprise is the percentage deviation of the firm's warning from the absolute value of the mean analyst forecast. Table 2 shows that the mean (median) analyst forecast for the whole sample is \$0.461 (\$0.295), whereas the mean (median) firms' profit warning is \$0.458 (\$0.290). For firms with the worst surprises (−10% or lower), the mean analyst forecast is \$0.115. The mean analyst forecasts for the surprise ranges from −10% to −5%, from −5% to −1%, and from −1% to 0% are, respectively, \$0.085, \$0.264, and \$0.543. The corresponding profit warnings for the four surprise ranges are \$0.098, \$0.078, \$0.258, and \$0.543.

During the pre- (post-) RFD period, the mean analyst's forecast is \$0.470 (\$0.454). The mean profit warning is \$0.468 (\$0.451). The mean analyst forecasts for the surprise ranges of −10% or lower, from −10% to −5%, from −5% to −1%, and from −1% to 0% during the pre- (post-) RFD period are, respectively, \$0.053, \$0.074, \$0.278, and \$0.555 (\$0.171, \$0.095, \$0.254, and \$0.534). The corresponding profit warnings for the four surprise ranges during the pre- (post-) RFD period are \$0.045, \$0.067, \$0.271, and \$0.555 (\$0.146, \$0.089, \$0.248, and \$0.533).

The surprise in dollars is calculated as the difference between the analyst forecast and the firm's profit warning. The surprise is relatively small in dollar terms. For the whole sample, the mean percentage surprise is −1.3%, and for the pre- (post-) RFD period, it is −1.3% (−1.4%). These descriptive statistics demonstrate that profit warnings are frequently issued even if the firm's expectations differ only slightly from the consensus forecast, as a small adjustment in earning expectations can have a large effect on future cash flow expectations and therefore on stock valuations.

Sample firms have, on average, 8.4 analysts following them. During the pre- (post-) RFD period, the average number of analysts is 7.7 (9.0). For the four surprise ranges, the mean numbers of analysts following the firms during the pre- (post-) RFD period are 6.5, 7.4, 7.3, and 7.9 (7.8, 6.0, 8.9, and 9.2). On average, the market value of sample firms in millions is \$6,599.9, with an average pre- (post-) RFD market value of \$5319.6 (\$7560.2). The mean market values in millions by surprise range for the pre- (post-) RFD periods are \$724.9, \$1522.0, \$6901.6, and \$5319.6 (\$2050.7, \$2292.6, \$8066.5, and \$7560.2). Our sample period spans the time from October 1, 1998 to September 30, 2001, with roughly 2 years falling prior to the October 23, 2000 effective date of RFD and 1 year after that date. Less than half (609) of our profit warnings occur during the time prior to effective date for RFD, whereas the remainder (812) occur during the time following the effective date of RFD.

5. Industry Effects

Panel A of Table 3 reports announcement returns for sample firms and their corresponding industry effects for days −5 to +5 relative to the announcement date. Because we expect larger firms to better reflect industry conditions, we calculate return averages as value-weighted averages. Sample firm market values as of 20 days prior to the profit warnings serve as the weights. Industry effects are determined using portfolios of other firms with the same 4-digit SIC code. The industry portfolio return is calculated as the value-weighted[3] return for industry firms. Daily abnormal returns are derived using the market model based on a 200-day estimation period for sample firms and their corresponding industries. The CAR for days −20 to −1 measures the leakage prior to the announcements.

Examination of the sample firm returns shows that market participants clearly consider the profit warnings to be informative, which is consistent with the findings of Jackson and Madura (2003). Sample firm returns are negative and significant on several days near the announcement, with a −6.69% mean (−5.38% median) CAR for days 0 to +1. In addition, industry effects are negative and significant on several days around the announcement for the sample as a whole, with a −1.23% mean (−0.25% median) CAR for days 0 to +1. Figure 1 provides a comparison of abnormal returns of the firms issuing profit warnings to those of their corresponding industry portfolios.

5.1. *Industry effects partitioned by pre- and post-RFD*

The implementation of RFD requires major changes in the manner in which firms release information. Consequently, Panels B and C of Table 3 examine the impact of those changes on the industry information contained in profit warnings. Specifically, we examine industry effects partitioned by pre- and post-RFD. The industry effects are negative and significant on days 0 and +1

[3]Industry portfolio returns are computed by weighting each firm's return by its market capitalization.

Table 3. Industry effects in response to profit warning announcements.

Day	Sample firms		Industry portfolios	
	Mean	Median	Mean	Median
Panel A. Whole sample				
−5	0.357***	−0.205**	0.058	−0.101
−4	−0.360***	−0.321***	0.324***	−0.025
−3	−0.029	−0.194***	0.112*	−0.108
−2	−0.548***	−0.315***	−0.461***	−0.107*
−1	−1.295***	−0.440***	−0.438***	−0.084*
0	−3.190***	−1.794***	−0.943***	−0.074*
+1	−3.502***	−1.599***	−0.287***	−0.227***
+2	−0.404***	−0.266*	−0.083	−0.097
+3	−0.283***	−0.158	−0.288***	0.134***
+4	0.282***	−0.182*	0.172***	−0.118
+5	0.562***	−0.062	0.085	−0.045
0 to +1	−6.692***	−5.383***	−1.230***	−0.251***
−20 to −1	−4.337***	−3.162***	−1.410***	−0.147
Panel B. Pre-RFD				
−5	0.105	−0.342**	0.162***	−0.031
−4	−0.744***	−0.448***	0.085	−0.002
−3	−0.148	−0.436***	0.069	−0.177*
−2	−0.622***	−0.370***	0.008	−0.113
−1	−0.950***	−0.421***	−0.375***	−0.161***
0	−7.608***	−3.123***	−1.054***	−0.098
+1	−3.567***	−2.559***	−0.226***	−0.178***
+2	0.299	−0.301	−0.333***	−0.134
+3	−0.779***	−0.182	−0.488***	0.134***
+4	0.226	−0.208	0.424***	−0.153*
+5	−0.035	−0.214	−0.057	−0.155
0 to +1	−11.175***	−9.447***	−1.280***	−0.255**
−20 to −1	−5.649***	−5.123***	−1.024***	−1.171**
Panel C. Post-RFD				
−5	0.490***	−0.134	0.003	−0.123
−4	−0.157	−0.199**	0.450***	−0.026
−3	0.033	−0.115**	0.135	−0.030
−2	−0.509***	−0.233***	−0.709***	−0.106*
−1	−1.478***	−0.481***	−0.471***	0.013
0	−0.858***	−1.158***	−0.884***	−0.050
+1	−3.468***	−1.203***	−0.319***	−0.252***
+2	−0.776***	−0.232	0.049	−0.034
+3	−0.022	−0.128	−0.182**	0.135***

(*Continued*)

Table 3. (*Continued*)

Day	Sample firms		Industry portfolios	
	Mean	Median	Mean	Median
+4	0.311***	−0.150	0.040	−0.051
+5	0.877***	0.036	0.160**	0.002
0 to +1	−4.326***	−3.165***	−1.203***	−0.247***
−20 to −1	−3.644***	−1.914***	−1.614***	0.498

Notes: This table reports abnormal returns around the announcements of 1421 profit warnings. Mean returns are value-weighted, using sample firm market values as of 20 days prior to the announcements as weights. Industry portfolio returns are the value-weighted returns of portfolios of other firms with the same 4-digit SIC codes as sample firms. The CAR for days 0 to +1 (−20 to −1) is the sum of the returns for days 0 and +1 (days −20 to −1).
*Statistically significant at 10% level.
**Statistically significant at 5% level.
***Statistically significant at 1% level.

during both the pre- and post-RFD periods, but are more pronounced in the pre-RFD period.

5.2. *Industry effects partitioned by SIC classification*

Table 4 shows the breakdown by SIC classification for the whole sample and for the pre- and post-RFD periods. Manufacturing firms dominate the sample, comprising 60% of all profit warnings. During the pre- (post-) RFD period, manufacturing firms account for 56.96% (62.19%) of firm warnings. Retail trade and the service sector follow manufacturing in the number of warnings with retail sector accounting for 12.10% of the whole sample, and 13.58% and 10.96% for the pre- and post-RFD periods, respectively, whereas the service sector accounts for 11.33% (whole sample), 11.95% (pre-RFD), and 10.84% (post-RFD).

The announcement returns for sample firms for nearly all the SIC classifications are negative and significantly different from zero. Agriculture, wholesale trade, finance, insurance, and real estate, services, and public administration are SIC classifications in which the industry effects for the whole sample are not significantly different from zero. Two of those industries (agriculture and public administration) have less than 10 profit warning announcements in the whole sample. Overall, the results suggest that the industry effects detected in our analysis are not confined to a particular industry

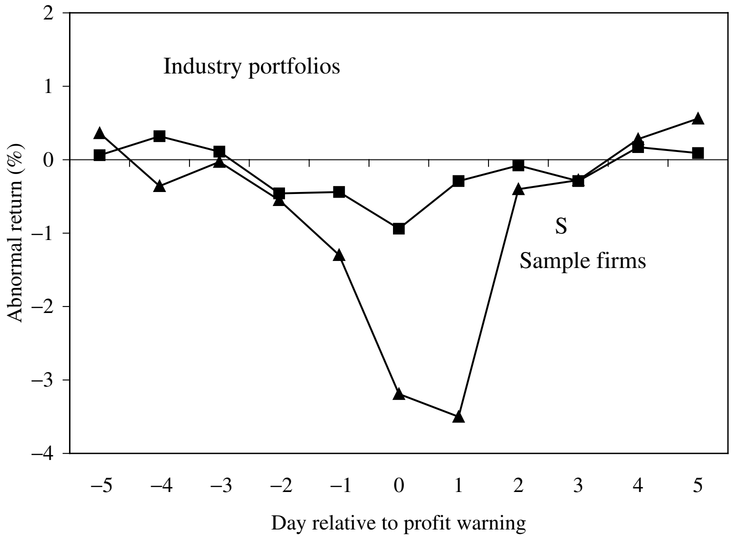


Figure 1. Comparison of abnormal returns of sample firms and industry portfolios. This figure shows value-weighted mean abnormal returns around the announcements of 1421 profit warnings between October 1, 1998 and September 30, 2001.

or sector.[4] The following cross-sectional analyses offer some insight on the impact of certain firm, industry, and market characteristics on the industry effects.

5.3. *Industry effects partitioned by size of the surprise*

Table 5 examines returns partitioned by the size of the surprise. Results show that for sample firms, the mean CARs for days 0 to +1 are generally more pronounced for larger negative surprises within the whole sample, and both the pre- and post-RFD periods. However, the mean returns do not decrease monotonically with the size of the surprise. The industry effects appear to be unrelated to the size of the surprise, and are significantly different from zero primarily for smaller surprises. Similarly, sample firm leakage does not decrease monotonically with the size of the surprise, and the industry response

[4] Although analysis of variance tests find significant differences among the industry effects of the various industry classifications, Tukey pairwise tests find few significant differences. Tukey results show that the only classification that differs somewhat consistently from other classifications is public administration, which has only two observations.

Table 4. Industry effects by SIC Industry Classification.

Standard industrial classification	SIC code	Sample firm CAR		Industry port CAR		# of Obs.	% of Obs.
		Mean	Median	Mean	Median		
Panel A: Whole sample							
Agriculture, forestry, and fishing	0000–0999	−4.465	−0.009	0.515	0.009	8	0.56
Mining	1000–1499	−6.501***	−0.035***	−1.403*	−0.010	36	2.53
Construction	1500–1999	−16.989***	−0.134***	−3.507**	−0.032	8	0.56
Manufacturing	2000–3999	−7.061***	−0.051***	−0.856***	−0.003**	853	60.03
Trans., commun., elec, gas, and san. svcs.	4000–4999	−6.975***	−0.038***	−1.053***	−0.003**	84	5.91
Wholesale trade	5000–5199	−8.057***	−0.067***	−0.553	0.005	54	3.80
Retail trade	5200–5999	−9.834***	−0.067***	−2.063***	−0.004	172	12.10
Finance, insurance, and real estate	6000–6999	−7.245***	−0.041***	−0.472	−0.001	43	3.03
Services	7000–8999	−7.218***	−0.077***	−0.019	−0.003*	161	11.33
Public administration	9000–9999	−0.749	−0.011	−5.100	−0.049	2	0.14
Panel B: Pre-RFD							
Agriculture, forestry, and fishing	0000–0999	−3.161	−0.009	0.422	0.007	6	0.98
Mining	1000–1499	−10.312***	−0.035***	−3.725***	−0.012	17	2.78
Construction	1500–1999	−8.347	−0.057	0.823	0.000	2	0.33
Manufacturing	2000–3999	−9.978***	−0.100***	−1.028***	−0.003	348	56.96
Trans., commun., elec, gas, and san. svcs.	4000–4999	−12.130***	−0.059***	−1.509***	−0.006**	34	5.56
Wholesale trade	5000–5199	−10.273***	−0.081***	1.832***	0.005	18	2.95
Retail trade	5200–5999	−16.029***	−0.108***	−2.676***	−0.002	83	13.58
Finance, insurance, and real estate	6000–6999	−9.845***	−0.072***	−1.215***	−0.002	28	4.58
Services	7000–8999	−15.823***	−0.130***	−1.182***	−0.005*	73	11.95
Public administration	9000-9999	—	—	—	—	0	0.00

(*Continued*)

Table 4. (*Continued*)

Standard industrial classification	SIC code	Sample firm CAR		Industry port CAR		# of Obs.	% of Obs.
		Mean	Median	Mean	Median		
Panel C: Post-RFD							
Agriculture, forestry, and fishing	0000–0999	−15.908	−0.069	1.324**	0.014	2	0.25
Mining	1000–1499	−2.207*	−0.029***	1.213	−0.001	19	2.34
Construction	1500–1999	−18.198**	−0.166**	−4.112**	−0.044*	6	0.74
Manufacturing	2000–3999	−4.914***	−0.031***	−0.729***	−0.003**	505	62.19
Trans., commun., elec, gas, and san. svcs.	4000–4999	−4.752***	−0.019***	−0.856**	0.000	50	6.16
Wholesale trade	5000–5199	−6.815***	−0.053***	−1.890**	0.005	36	4.43
Retail trade	5200–5999	−3.713***	−0.036***	−1.457***	−0.004	89	10.96
Finance, insurance, and real estate	6000–6999	−5.519**	−0.022	0.021	0.002	15	1.85
Services	7000–8999	−5.333***	−0.058***	0.236	0.000	88	10.84
Public administration	9000–9999	−0.749	−0.011	−5.100	−0.049	2	0.25

Notes:This table reports announcement CARs for days 0 to +1 for sample firms and industry effects by SIC code for a sample of 1421 profit warnings between October 1, 1998 and September 30, 2001. Mean returns are value-weighted, using sample firm market values as of 20 days prior to the announcements as weights. Abnormal returns are calculated using the market model. Industry portfolio returns are the value-weighted returns of portfolios of other firms with the same 4-digit SIC codes as sample firms. Results are reported for the whole sample and partitioned by pre- and post-RFD. p-Values for t-tests of the differences of the returns from zero are reported in parentheses.

*Statistically significant at the 10% level.

**Statistically singnificant at the 5% level.

***Statistically singnificant at the 1% level.

Table 5. Industry effects partitioned by size of the surprises.

	Whole sample		Surprise %			
	Mean	Median	≤−10%	−10% to −5%	−5% to −1%	−1% to 0%
Sample firm CAR (days 0 to +1)						
Whole sample	−6.69***	−5.38***	−8.85***	−8.35***	−6.40***	−6.75***
Pre-RFD	−11.18***	−9.45***	−20.06***	−22.18***	−9.52***	−11.56***
Post-RFD	−4.33***	−3.17***	−5.33**	0.84	−4.58***	−4.30***
Industry portfolio CAR (days 0 to +1)						
Whole sample	−1.23***	−0.25***	−0.96	0.83	−0.45**	−1.51***
Pre-RFD	−1.28***	−0.26**	0.08	−1.22	−0.64***	−1.51***
Post-RFD	−1.20***	−0.25***	−1.29	2.19**	−0.34	−1.50***
Sample firm leakage (days −20 to −1)						
Whole sample	−4.34***	−3.16***	−1.89	−8.89***	−3.86***	−4.45***
Pre-RFD	−5.65***	−5.12***	−6.40**	−10.22*	−5.00***	−5.81***
Post-RFD	−3.64***	−1.91***	−0.47	−8.01***	−3.20***	−3.76***
Industry portfolio response to sample firm leakage (days −20 to −1)						
Whole sample	−1.41***	−0.15	1.20	−4.14*	−2.31***	−1.11***
Pre-RFD	−1.02***	−1.17**	−3.48***	2.24	−1.36**	−0.94**
Post-RFD	−1.61***	0.50	2.67	−8.38**	−2.86***	−1.19***

(*Continued*)

Table 5. (*Continued*)

	Whole sample		Surprise %			
	Mean	Median	≤ −10%	−10% to −5%	−5% to −1%	−1% to 0%
Number of observations						
Whole sample	1421	1421	34	44	296	1047
Pre-RFD	609	609	16	22	120	451
Post-RFD	812	812	18	22	176	596

Note: This table reports abnormal announcement CARs for days 0 to +1 and leakage for a sample of 1421 profit warnings between October 1, 1998 and September 30, 2001. Leakage and industry response are measured as the CAR for days −20 to −1. Results are reported for the whole sample, partitioned by pre- and post-RFD, and by percentage surprise ranges. Industry portfolio returns are the value-weighted returns of portfolios of other firms with the same 4-digit SIC codes as sample firms. Abnormal returns are calculated using the market model. Mean returns are value-weighted, with sample firm market values 20 days prior to the announcements as the weights. Asterisks indicate results of tests of differences from zero (mean: t-test; median: signed rank test).

*Statistically significant at the 10% level.

**Statistically significant at the 5% level.

***Statistically significant at the 1% level.

to sample firm leakage is significantly different from zero primarily for smaller surprises.

5.4. *Industry effects partitioned by the revaluation of the warning firm*

Table 6 shows industry effects partitioned by various characteristics. Panel A shows that industry effects are more pronounced when the firm's stock price response is more pronounced. That is, a profit warning that elicits a larger drop in a firm's stock price signals more negative information about the corresponding industry. Results of tests of differences in the value-weighted means and the medians, reported in the last row, show significant differences between the industry effects of quartiles 1 and 4.

5.5. *Industry effects partitioned by size of the warning firm*

The relation between the sample firm size and industry effects is provided in panel B of Table 6. In general, the industry effects are more pronounced for large firms (quartile 4). Tests of differences show that the median returns, but not the value-weighted means, differ significantly between quartiles 1 and 4.

5.6. *Industry effects partitioned by analyst coverage of the warning firm*

The industry effects are assessed according to the number of analysts following the firm that issued the profit warning in Panel C of Table 6. The effects are more pronounced when the firms that issue the profit warnings are covered by a large number of analysts (quartile 4).

Overall, our univariate results suggest that industry effects are more pronounced when the firm issuing the profit warning is larger, has more analyst coverage, and experiences a larger revaluation at the time of the warning. To more comprehensively assess the conditions under which profit warnings contain industry information, we turn now to multivariate analysis.

6. Multivariate Analysis

We employ cross-sectional regressions of the whole sample and of the sample partitioned by pre- and post-RFD periods to more comprehensively examine

Table 6. Industry effects partitioned by various characteristics.

Quartile		Sample firm CARs (%)		Industry effects (%)	
		Mean	Median	Mean	Median
Panel A: Industry effects partitioned by revaluation of warning firms					
Sample firm CAR					
Quartile 1		−25.28***	−21.94***	−2.87***	−0.62***
Quartile 2		−9.45***	−9.17***	−1.02***	−0.53***
Quartile 3		−2.68***	−2.61***	−0.78***	−0.33**
Quartile 4		5.04***	2.91***	−0.83***	0.53***
p-value (quartile 1–quartile 4)				0.00	0.00
Panel B: industry effects partitioned by size of the warning firms					
	Sample firm size ($ MM)				
Quartile 1	182.90	−9.90***	−7.99***	0.51***	0.30***
Quartile 2	656.02	−8.77***	−6.17***	−0.38*	−0.37**
Quartile 3	2103.74	−6.18***	−4.19***	−0.54***	−0.26**
Quartile 4	23475.09	−6.66***	−3.68***	−1.33***	−0.59***
p-value (quartile 1–quartile 4)				0.27	0.00
Panel C: industry effects partitioned by analyst coverage of the warning firms					
Analyst coverage	Number of analysts				
Quartile 1	1.63	−10.29***	−9.97***	0.11	0.14
Quartile 2	5.21	−7.05***	−4.80***	−0.46***	−0.30
Quartile 3	9.63	−9.49***	−5.49***	−0.73***	−0.39***
Quartile 4	17.99	−5.89***	−3.50***	−1.44***	−0.45***
p-value (quartile 1–quartile 4)				0.09	0.00

Note: Abnormal returns are calculated using the market model. Mean returns are value-weighted, using firm market values 20 days prior to announcement as the weights. *p*-Values are for tests of differences in value-weighted means (t-test) and medians (Wilcoxon rank sum test) between quartiles 1 and 4. Asterisks indicate results of tests of differences from zero (mean: t-test; median: signed rank test).
* Statistically significant at the 10% level.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.

the variables that influence the industry effects of profit warnings. In addition to the characteristics described earlier, we control other characteristics that could affect the industry effects in response to profit warnings. Since Jackson and Madura (2003) find statistically significant leakage for firms that issue profit warnings, a leakage variable is included to capture the effect of any leakage prior to the profit warning. We control a firm's growth prospects, as a profit warning may reflect a more severe signal to a high-growth industry. We include an industry concentration variable as Kohers (1999) shows that dividend cuts carry stronger industry signals for less concentrated (more competitive) industries. Thus, adverse industry effects may be attenuated for firms in highly concentrated industries. We also control stock market sentiment, as the industry effects could be influenced by market sentiment at the time of the profit warning.

As discussed earlier, industry effects could be influenced by whether the warning is the first quarterly warning or a subsequent warning for an industry. We also expect adverse industry effects to be more pronounced in an industry characterized as having a low degree of concentration, and for warnings issued before the implementation of RFD. Thus, we include these variables in our regressions to assess their influence on industry effects. We use an RFD dummy variable to test the hypothesis of differences in industry effects between the pre- and post-RFD periods. In addition, we test whether the effects of some variables differ between the pre- and post-RFD periods through the use of interactive variables.

We also examine the influence of firm size and analyst coverage on industry effects. Small firms are more susceptible to firm-specific problems such as those associated with a single supplier or customer. Thus, a profit warning of a large firm is more likely to reflect industry information. Industry effects may also be conditioned on analyst coverage, because firms with greater analyst coverage have greater information flow to the market.

6.1. *Multivariate model*

To investigate the determinants of industry effects, we estimate the following cross-sectional model:

$$\begin{aligned}\mathrm{IND} = {} & \beta_0 + \beta_1\,\mathrm{RFD} + \beta_2\,\mathrm{SURPRISE} + \beta_3\,\mathrm{CAR} + \beta_4\,\mathrm{FSIZE} \\ & + \beta_5\,\mathrm{ANALYSTS} + \beta_6\,\mathrm{FIRST} + \beta_7\,\mathrm{LEAKAGE} \\ & + \beta_8\,\mathrm{EARN/PRICE} + \beta_9\,\mathrm{HINDEX} + \beta_{10}\,\mathrm{SENTIMENT} \\ & + \beta_{11}\,\mathrm{RFD} \times \mathrm{CAR} + \beta_{12}\,\mathrm{RFD} \times \mathrm{LEAKAGE} \\ & + \beta_{13}\,\mathrm{RFD} \times \mathrm{FSIZE} + \varepsilon \qquad (1)\end{aligned}$$

where IND is the 2-day (days 0 to +1) announcement cumulative abnormal value-weighted return for a portfolio of industry rivals (constructed using

4-digit SIC codes) of firms making profit warnings; RFD a dummy variable assigned a value of 1 for profit warnings issued after the implementation of RFD and 0 otherwise; SURPRISE the difference between the company's profit warning and the consensus estimate, scaled by the absolute value of the consensus estimate; CAR the 2-day (days 0 to +1) announcement cumulative abnormal return for the firm issuing the profit warning; FSIZE the size (natural log of market capitalization) of the firm issuing the profit warning on day −20 relative to the profit warning; ANALYSTS the number of analysts covering the firm issuing the profit warning; FIRST a dummy variable set equal to 1 for profit warnings that are the first for an industry within the quarter and 0 otherwise; LEAKAGE the firm's CAR for the month (days −20 to −1) before the profit warning;[5] EARN-PRICE the earnings–price ratio of the firm that issued the profit warning, and serves as an indicator of growth (higher earnings–price ratio implies lower growth prospects); HINDEX the natural log of one plus the Herfindahl index value for the respective industry. Following Lang and Stulz (1992), we use the Herfindahl index as a proxy for the degree of industry concentration. The Herfindahl index for each industry is computed as:

$$\mathrm{HI} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left(\frac{\mathrm{Sales_i}}{\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{Sales_i}} \right)^2 \tag{2}$$

where N is the number of firms in the industry.

Industries with relatively high (low) values of the Herfindahl index are characterized as having a high degree of concentration (being more competitive). SENTIMENT is an indicator of market sentiment according to the recent runup in the market during the 20-day period prior to the profit warning, measured by the holding period return on the CRSP value-weighted index.

In addition, three interaction terms are considered when assessing the whole sample in order to capture the potential change in the sensitivity of industry effects to characteristics of the firm that issued the profit warning.

RFD × CAR is an interaction term used to test for a change in the sensitivity of industry effects to announcement CARs for warning firms since RFD.

[5]To mitigate the effects of outliers and improve the distributional properties of IND, CAR, LEAKAGE, SURPRISE, RFD × CAR, and RFD × LEAKAGE, we winsorize the distributions at the 5th and 95th percentiles.

RFD × LEAKAGE is an interaction term used to test for a change in the sensitivity of industry effects to the LEAKAGE for warning firms since RFD.

RFD × FSIZE is an interaction term used to test for a change in the sensitivity of industry effects to the size of the firm that issued a profit warning since RFD.

We provide results for three different models; as firm size and the number of analysts exhibit a high degree of colinearity, they are not included in the same model. Consequently, we include a variable for firm size in Model 1 and exclude the analyst coverage variable. In Model 2, the firm size variable is dropped and the analyst coverage variable is included. Model 3 deletes both variables and includes a variable to indicate if the event occurred in the pre- or post-RFD period, as well as several RFD interaction variables.

6.2. *Results of multivariate analysis*

Results from applying a multivariate cross-sectional analysis are provided in Table 7. All the three models show that industry effects are positively and significantly related to the announcement abnormal return (CAR) of the firm that issued the profit warning. Thus, the adverse industry signal of a profit warning is more pronounced when the market punishes the issuing firm more severely. In Model 1, the firm size variable is negative and significant, which is consistent with our earlier finding of a more pronounced industry effect in response to profit warnings by larger firms.

In Model 2, the analyst variable replaces the firm size variable. The coefficient is negative and significant, which supports our earlier finding of a more pronounced industry effect in response to profit warnings by firms with a greater analyst following. The negative coefficient of HINDEX indicates that industry effects are attenuated for more competitive industries. As the industry portfolio returns are value-weighted, this unexpected result could be due to a competitive advantage that accrues to larger industry rivals when a firm issues a profit warning. The positive coefficient of the SENTIMENT variable shows that when recent market returns have been positive, profit warnings elicit less pronounced industry effects.

Model 3 contains variables that isolate the impact of RFD on industry effects. The RFD variable is positive and significant, confirming that industry effects are less pronounced in the post-RFD period even after controlling

Table 7. Multivariate analysis of industry effects.

Dependent variable: IND	Model 1	Model 2	Model 3
Intercept	0.038***	0.008***	0.002
RFD			0.030***
SURPRISE	−0.030	−0.045	−0.042
CAR	0.051***	0.049***	0.034***
FSIZE	−0.003***		
ANALYSTS $\times 10^{-2}$		−0.048***	
FIRST $\times 10^{-2}$	0.192	0.207	0.193
LEAKAGE	−0.003	−0.001	0.007
EARN-PRICE	−0.016	−0.015	−0.007
HINDEX	−0.006	−0.011*	−0.009
SENTIMENT	0.021	0.023*	0.018
RFD × CAR			0.040**
RFD × LEAKAGE			−0.016
RFD × FSIZE			−0.002***
Adjusted R^2	0.054	0.044	0.042
F-statistics	11.11***	9.14***	6.70***

Note: This table shows OLS regression results for industry effects on sample firm announcement returns and other explanatory variables. IND is the value-weighted return for days 0 to +1 of industry portfolios, which are comprised other firms with the same 4-digit SIC codes as warning firms. RFD is a dummy variable assigned a value of 1 for profit warnings issued after the implementation of RFD and 0 otherwise. SURPRISE is the difference between the company's profit warning and the consensus estimate, scaled by the absolute value of the consensus estimate. CAR is the announcement CAR for days 0 to +1 for the warning firm. FSIZE is the natural log of market capitalization of the warning firm on day −20 relative to the profit warning. ANALYSTS is the number of analysts covering the warning firm. FIRST is a dummy variable set equal to 1 for profit warnings that are the first for an industry within the quarter, 0 otherwise. LEAKAGE is the warning firm's CAR for the month (days −20 to −1) before the profit warning. EARN-PRICE is the earnings–price ratio of the firm that issued the profit warning. HINDEX is the natural log of one plus the Herfindahl index. SENTIMENT is an indicator of market sentiment, measured as the holding period return on the CRSP value-weighted index during the 20-day period prior to the profit warning. The sample consists of 1421 profit warnings issued between October 1, 1998 and September 30, 2001. The following variables are winsorized at the 5th and 95th percentiles to mitigate the effect of outliers: sample and industry portfolio CARS, LEAKAGE, SURPRISE, RFD×CAR, and RFD×LEAKAGE. p-Values are calculated using White-corrected standard errors.

*Statistically significant at the 10% level.

**Statistically significant at the 5% level.

***Statistically significant at the 1% level.

other factors. The coefficient of the interaction term RFD×CAR is positive and significant, suggesting that the positive relation between the sample firm CAR and industry effects is more pronounced since the inception of RFD. In addition, the coefficient of the interaction term RFD×SIZE is negative and significant, suggesting that the inverse relation between firm size and industry effects is more pronounced since the inception of RFD. The other variables are not significant determinants of industry effects.

6.3. *Results of the multivariate analysis applied to pre- and post-RFD periods*

Multivariate models are also applied separately to subsamples of profit warnings in the pre- and post-RFD periods. One obvious difference in the models is that the RFD dummy variable and interaction terms are no longer applicable, because each subsample represents either the pre- or post-RFD period. Results are disclosed in Table 8. Again, multiple models are used to examine effects of variables that exhibit a high degree of colinearity.

As shown in Table 8, the sample firm CAR variable remains positive and significant, and the firm size variable remains negative and significant for each subsample. The variable representing number of analysts remains negative and significant for each subsample. The dummy variable representing the first quarterly profit warning is positive and significant for the post-RFD in all the three models, which is consistent with earlier findings of less pronounced industry effects in response to the first profit warning for a particular industry in a given quarter.

Although the market sentiment variable is not significant when applied to the pre-RFD period, it is positive and significant within the post-RFD period. The general market conditions are weak in this period, but the market sentiment variable captures the recent market movements over the previous 20-day period. The results suggest that the industry effects are attenuated when the market has recently experienced a runup. That is, the market punishes related industry stocks to a smaller degree when market sentiment is relatively favorable in the post-RFD period. The coefficient of the industry concentration variable, HINDEX, is negative and significant in one of the three models, suggesting attenuated industry effects for more competitive industries in the post-RFD period.

Table 8. Multivariate analysis of industry effects in the pre- and post-RFD periods.

Dependent variable: IND	Model 1	Model 2	Model 3
Panel A: Pre-RFD			
Intercept	0.051***	0.012***	0.004
SURPRISE	0.030	0.007	−0.008
CAR	0.037***	0.035***	0.034***
FSIZE	−0.003***		
ANALYSTS $\times 10^{-2}$		−0.075***	
FIRST$\times 10^{-2}$	−0.079	−0.067	−0.029
LEAKAGE	0.007	0.012	0.010
EARN-PRICE	−0.035	−0.037	−0.012
HINDEX	0.003	−0.006	−0.004
SENTIMENT	−0.036	−0.032	−0.033
Adjusted R^2	0.054	0.040	0.013
F-statistic	5.37***	4.19***	2.15**
Panel B: Post-RFD			
Intercept	0.031***	0.006**	0.001
SURPRISE	−0.067	−0.079	−0.094
CAR	0.066***	0.064***	0.061***
FSIZE	−0.002***		
ANALYSTS $\times 10^{-2}$		−0.032**	
FIRST $\times 10^{-2}$	0.395*	0.408*	0.451**
LEAKAGE	−0.009	−0.008	−0.008
EARN-PRICE	−0.003	−0.001	0.011
HINDEX	−0.013	−0.015*	−0.012
SENTIMENT	0.039**	0.040***	0.040***
Adjusted R^2	0.069	0.061	0.057
F-statistic	8.45***	7.62***	7.97***

Note: This table shows OLS regression results for industry effects on sample firm announcement returns and other explanatory variables, which are defined in the previous table. p-Values are calculated using White-corrected standard errors.
*Statistically significant at the 10% level.
**Statistically significant at the 5% level.
***Statistically significant at the 1% level.

7. Conclusion

Research shows that profit warnings have a substantial impact on firm valuations, but has not examined whether there is a corresponding impact on the rest of the industry. We investigate this issue, because profit warnings could signal problems that extend beyond warning firms and have industry-wide implications. Our objective in this research is to determine the conditions under which profit warnings convey information about the corresponding industries.

We find that profit warnings result in negative industry effects. This evidence confirms that the negative information contained in a firm's profit warning conveys negative information about both the firm and its industry, rather than favorable information for industry rivals about the competition.

Multivariate analysis reveals that profit warnings emit stronger industry signals when the warning firm CARs are more pronounced. In addition, the adverse industry effects are worse when the firm that issues the warning is larger and has more analyst coverage. The adverse industry effects are attenuated since the inception of RFD, which we attribute to increased information flow. However, the sensitivity of the industry effects to the warning firm's stock price adjustment and size has increased since RFD. In addition, the first profit warning within a given quarter has a less pronounced effect on the industry than subsequent warnings in the same quarter since RFD. Furthermore, the industry effects in response to a profit warning are less pronounced when the recent market sentiment is more favorable.

Acknowledgments

The authors thank the anonymous reviewers, Joel Harper, and Xenophon Koufteros for their helpful suggestions and First Call/Thomson Financial for providing the profit warnings data.

References

Akhigbe, A., J. Madura and A. Whyte, "Intra-Industry Effects of Bond Rating Announcements." *Journal of Financial Research* 20, 547–563 (1997).

Bailey, W., H. Li, C. X. Mao and R. Zhong, "Regulation Fair Disclosure and Earnings Information: Market, Analyst, and Corporate Responses." *The Journal of Finance* 33, 2487–2514 (2003).

Bhushan, R., "Firm Characteristics and Analyst Following." *Journal of Accounting and Economics* 11, 255–274 (1989).

Brown, S. L., "Earnings Changes, Stock Prices, and Market Efficiency." *The Journal of Finance* 33, 17–28 (1978).

Clinch, G. J. and N. A. Sinclair, "Intra-Industry Information Releases: A Recursive Systems Approach." *Journal of Accounting and Economics* 9, 89–106 (1987).

Firth, M., "Dividend Changes, Abnormal Returns, and Intra-Industry Firm Valuations." *Journal of Financial and Quantitative Analysis* 31, 189–211 (1996).

Foster, G., "Intra-Industry Information Transfers Associated with Earnings Releases." *Journal of Accounting and Economics* 3, 201–232 (1981).

Heflin, F., K. R. Subramanyam and Y. Zhang, "*Stock Return Volatility before and after Regulation FD*." Working paper (2001).

Heflin, F., K. R. Subramanyam and Y. Zhang, "Regulation FD and the Financial Information Environment." *Accounting Review* 78, 1–37 (2003).

Hertzel, M. G., "The Effects of Stock Repurchases on Rival Firms." *Journal of Finance* 46, 707–716 (1991).

Hong, H., T. Lim and J. Stein, "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies." *The Journal of Finance* 55, 265–295 (2000).

Ivkovic, Z. and N. Jegadeesh, "The Timing and Value of Forecast and Recommendation Revisions." *Journal of Financial Economics* 73, 433–463 (2004).

Jackson, D. and J. Madura, "Profit Warnings and Timing." *Financial Review* 38, 497–513 (2003).

Jackson, D. and J. Madura, "Impact of Regulation Fair Disclosure on the Information Flow Associated with Profit Warnings." *Journal of Economics and Finance*, In press.

Kohers, N., "The Industry-Wide Implications of Dividend Omission and Initiation Announcements and the Determinants of Information Transfer." *Financial Review* 34, 137–158 (1999).

Lang, L. H. P. and R. M. Stulz, "Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements." *Journal of Financial Economics* 32, 45–60 (1992).

Laux, P. L., L. T. Starks and P. S. Yoon, "The Relative Importance of Competition and Contagion in Intraindustry Information Transfers: An Investigation of Dividend Announcements." *Financial Management* 27, 5–16 (1998).

Ling, D. C. and M. Ryngaert, "Valuation Uncertainty, Institutional Involvement, and the Underpricing of IPOs: The Case of REITs." *Journal of Financial Economics* 43, 433–457 (1997).

Libby, R. and H. Tan, "Analysts' Reactions to Warning of Negative Earning Surprises." *Journal of Accounting Research* 37, 415–435 (1999).

Rendleman, R., C. P. Jones and H. Latane, "Empirical Anomalies Based on Unexpected Earnings and the Importance of Risk Adjustments." *Journal of Financial Economics* 10, 269–288 (1982).

Watts, R. L., "Systematic 'Abnormal' Returns after Quarterly Earnings Announcements." *Journal of Financial Economics* 6, 127–157 (1978).

This page intentionally left blank

Chapter 6

Are Whisper Forecasts More Informative than Consensus Analysts' Forecasts?

Erik Devos
Ohio University, USA

Yiuman Tse*
University of Texas at San Antonio, USA

In this paper, we compare whisper forecasts with consensus forecasts of quarterly earnings. Analysts' private whisper forecasts are optimistic, whereas their public consensus forecasts are pessimistic. We find that whispers are more accurate for high-tech and high-growth firms. The introduction of Regulation FD does not alter these findings. In a multivariate setting, we find that the whisper and consensus forecast errors are related to the abnormal returns around the earnings announcement, suggesting that whisper and consensus forecasts are value-relevant.

Keywords: Whisper forecasts; analyst forecasts; earnings per share.

1. Introduction

In the 1990s, the investment community witnessed the birth of whisper forecasts of earnings. These forecasts, privately released by analysts, have since become an important focus of investors. Associated Press recently reported that "… It didn't matter that earnings from companies like Yahoo and Intel surpassed analysts' expectations earlier this month. Their stocks still fell, partly because their results missed the so-called 'whisper numbers'…" (Lexis-Nexis Newswires, 01/26/2004). This suggests that whispers indeed provide information to the market and investors use whisper forecasts as the basis for trading. Whether whisper forecasts provide information over what is already publicly known is a question that has been sparsely addressed in the academic literature. A notable exception is the work of Bagnoli *et al*. (1999), who look at a small relatively small sample of mainly high-tech firms in the mid-1990s and suggest that whispers are informative.

*Corresponding author.

The sparseness of research on whisper forecasts is surprising, given the fact that the academic literature has extensively documented that public information of analysts in the form of analysts' reports (Asquith, Mikhail, and Au, 2005), recommendations (Womack, 1996), forecasts, and revisions (Francis and Soffer, 1997) is valuable to investors. In this paper, we investigate the value of whisper forecasts, by comparing them to public consensus forecasts, using a large sample of whisper forecasts in the period 1999–2001. By using a sample of 9,764 whisper forecasts that includes a large number of high-tech and non high-tech firms, we are able to test whether the earlier findings of Bagnoli *et al.* (1999) are sample and/or time specific. Their paper is confined to a sample of 943 whisper forecasts for 127 firms and consists of about 90% of high-tech firms, all collected in the mid-1990s. In addition, our sample period spans a time frame that is characterized by heavy internet use by investors, and spans the introduction of Regulation FD, all of which could affect the accuracy and usefulness of whisper forecasts.

Private whisper forecasts are provided anonymously by analysts, which implies that if they turn out to be inaccurate the analysts providing them would not lose a significant amount of reputation. The lack of loss in reputation and the nature/source of the whisper may lead whispers to be open to manipulation and possible inaccuracy. Conversely, consensus forecasts are not anonymous and therefore analysts providing them have an incentive to be accurate, in order to maintain their reputation. Prior evidence, e.g., Mikhail *et al.* (1999), suggests that analysts who are less accurate are more likely to lose their jobs, whereas Stickel (1992) finds that in order to be included in the All-American Research Team an analyst needs to be more accurate than his peers. In addition, inaccurate forecasting could lead to litigation by investors.

In contrast, there are several potential explanations why whispers could be more accurate then consensus forecasts. First, reputational concerns of analysts and investment banks may not be sufficient to force analysts to disclose all their information publicly, as evidenced by the settlement between 10 large investment banks and the SEC in 2003. This settlement resolves the claim that the analysts knowingly issued inaccurate recommendations. Second, analysts may not want to disclose all their available information through public consensus forecasts because they do not want to increase market earnings expectations and therefore the likelihood of a negative earnings surprise.

A third reason could be avoidance of disclosing negative information and therefore alienating management. This could also explain the scarcity of sell recommendations as found by Womack (1996). Analysts could choose to provide more accurate information through private whisper forecasts. A fourth phenomenon that could lead whisper forecasts to be more accurate than consensus forecasts is the likelihood that analysts may wish to protect their valuable information and disclose it only to their preferred customers through whisper forecasts. This would be the case especially for information (and forecasts) pertaining to fast-growing, innovative firms. For these firms, regular financial reports are less informative (Lev and Zarowin, 1999), and, as Amir *et al*. (1999) report for this type of firm, financial analyst forecasts are most valuable. Hence, we expect whisper forecasts to be more accurate for these firms.

We investigate these claims by comparing whisper and consensus forecasts. Our main findings shed light on the importance of whisper forecasts relative to consensus forecasts and the way investors react to them. We find that the majority of whisper forecasts are optimistic, whereas a majority of consensus forecasts are pessimistic. In only 49.1% of cases do we find that whisper forecasts are more accurate than consensus forecasts. We do find that whisper forecasts are more accurate for high-tech firms and for firms with high book-to-market equity ratios. However, whisper forecasts are only better in these subsamples for those firms reporting positive earnings (profit firms).

We also investigate whether regulation FD had any effect on the relative accuracy of whisper and consensus forecasts. We do not find any evidence that our observed patterns change pre- and post-introduction. Hence, whispers remain useful even after regulation FD was introduced. This suggests that the relative superiority of whisper forecasts is not due to selective disclosure. Finally, in univariate and multivariate tests, we find that whisper and consensus forecast errors are associated with abnormal stock returns around earnings announcements. This suggests that whisper forecasts have incremental information content over and above what is available from public consensus forecasts.

The results in our paper confirm that the main finding of Bagnoli *et al.* (1999) is robust. When analysts face pressures from management to bias their public forecasts or when information is especially valuable (i.e., when financial statements are less informative), analysts provide private whisper forecasts that are more informative than public consensus forecasts. An important implication of our findings is that for high-tech and high-growth firms the consensus forecast may not be the best measure of market earnings expectations.

The rest of this paper is organized as follows: Section 2 contains our hypotheses. Section 3 describes the sample collection process and the data. In Section 4, we present our main results containing the whisper and consensus forecasts. Section 5 documents the relationship between abnormal returns around the earnings announcements and the whisper and consensus forecast errors. Section 6 concludes the paper.

2. Hypotheses

As mentioned in the previous section, Bagnoli *et al.* (1999) find that whisper forecasts are more accurate than public, First Call forecasts. They also document that a portfolio strategy based on whisper forecast errors yields positive returns, suggesting that whisper forecasts are informative and that public forecasts may not accurately represent the market's earnings expectations. In this section, we form the four main hypotheses that we test in this paper.

2.1. *Accuracy of whisper forecasts*

Because analysts provide whisper forecasts anonymously, they do not lose reputation in case they are inaccurate. This may leave whisper forecasts open for manipulation and/or inaccuracy. On the other hand, consensus forecasts are provided publicly, which may make inaccuracy very costly. See, for example, Mikhail *et al.* (1999) who show that inaccurate analysts are more likely to loose their jobs. Also, accuracy is an important element to be included in the All-American Research team (Stickel, 1992), and inaccurate analysts could be subject to litigation. This would suggest that whisper forecasts are not more accurate than consensus forecasts. The main findings of Bagnoli *et al.* (1999) provide evidence that is inconsistent with this hypothesis.

However, there are several reasons that do suggest that whispers are superior. First, reputation may not be sufficient to force analysts to disclose all their information in public forecasts. This claim can be supported by the settlement of between 10 large investment banks and the SEC in 2003. This settlement resolves the claim that the analysts knowingly issued inaccurate recommendations. Second, analysts may not want to disclose all their available information through public consensus forecasts because they do not want to increase market earnings expectations and therefore the likelihood of a negative earnings surprise. Third, the fear of alienating management could prevent analysts from disclosing negative information publicly. Womack (1996) finds a paucity of sell recommendation, which may be caused by the same fear.

In these cases, analysts could choose to provide more accurate information through private whisper forecasts. A final reason why whisper forecasts could be more accurate than consensus forecasts is the likelihood that analysts may wish to protect their valuable information and disclose it only to their preferred customers through whisper forecasts. Specifically, this could be the case for high-growth and high-tech firms, as we will hypothesize next.

2.2. *Accuracy of whisper forecasts for high-growth firms*

Lev and Zarowin (1999) suggest that for rapidly growing firms in innovative industries, audited financial reports and other public financial information are less informative. These firms have relatively large amounts of intangible assets that are not represented in financial reports (e.g., human capital) and are characterized by high investment in R&D and relatively quick product obsolescence. The results in Amir *et al.* (1999) suggest that the incremental contribution of financial analysts is largest in such firms. Additionally, they show that the analysts' contribution to valuation in firms with substantial capital invested in R&D is larger than in firms with low R&D investment.

Das *et al*. (1998) show that the bias in analysts' forecasts is more pronounced for firms whose earnings are less predictable. As many of these firms are in emerging industries whose potential for growth is unknown, we expect the market to be less severe in punishing analysts who provide imprecise public forecasts. Besides, investors are more likely to highly value any private information for such firms. In such cases, analysts may protect their private information by disseminating their information via whisper forecasts. Hence, we expect whisper forecasts to be more accurate than public consensus analysts' forecasts for rapidly growing firms in innovative industries. In our empirical tests, such firms are those that we classify as either high-tech or high-growth firms.

2.3. *Accuracy of whisper forecasts for high-tech firms*

The superiority of whisper forecasts could also be due to other reasons. For example, stocks where a large fraction of the value is due to future growth are likely to exhibit a larger negative response to negative earnings surprises (e.g., Skinner and Sloan, 2002). Managers of these firms may exert pressure on analysts to bias their public earnings forecasts downwards and reduce the likelihood of the firm announcing a negative earnings surprise (Matsumoto, 2002). Managers may selectively reward compliant analysts with improved access to their private information that analysts could then reveal through whisper forecasts.

If the earnings for these firms are less predictable when based upon publicly available information, analysts may also have an incentive to bias their public forecasts and gain access to managers' private information as an additional, accurate source of information. Analysts not wanting to alienate management may respond to such managerial persuasion and report lower public forecasts to avoid negative earnings surprises. Public forecasts for these firms may thus be inaccurate relative to whisper forecasts. In our empirical analysis, we expect these issues to be more relevant for high-tech and high-growth firms.

2.4. *Effects of regulation FD on whisper forecast accuracy*

Our analysis also provides an opportunity to investigate whether the greater accuracy of whisper forecasts is due to selective disclosure. The U.S. SEC was concerned about the impact of selective disclosure on the efficiency of U.S. capital markets. In a written statement about the impact of Reg. FD filed before a U.S. House of Representatives subcommittee on May 17, 2001 (http://www.sec.gov/news/testimony/051701wssec.htm), the SEC stated,

> The primary issue is the basic unfairness of providing a select few with a significant informational advantage over the rest of the market. This unfairness damages investor confidence in the integrity of our capital markets … Further, if selective disclosure is permitted, corporate management can treat material information as a commodity to be used to gain or maintain favor with particular analysts or investors. This practice could undermine analyst objectivity, in that analysts will feel pressured to report favorably about a company or slant their analysis to maintain access to selectively disclosed information. Thus, selective disclosure may tend to reduce serious, independent analysis.

The introduction of Reg. FD allows us to test whether this hypothesis explains the relative superiority of whisper forecasts. Reg. FD, which took effect on October 23, 2000, prohibits managers from releasing material information privately to analysts. If whisper forecasts are the outcome of selective disclosure of information by managers, we would expect the usefulness of whisper forecasts to decline after the introduction of Reg. FD.

3. Data

3.1. *Sample*

We collect whisper forecasts from the internet site *earningswhispers.com*. The initial sample consists of all firms that had whisper forecasts of quarterly earnings per share that were reported from October 14, 1999 through September 12, 2001. We employ several screens to ensure the integrity of our data. First, we minimize the possibility of back filling by ensuring that the whisper forecasts are collected before the actual earnings are released. We also verify the accuracy of the data by comparing a random sample of whisper forecasts from our source with whisper forecasts reported in the financial press (such as the *Wall Street Journal*) and in a competing source www.whispernumbers.com. Overall, we find little evidence of systematic differences between the whisper forecasts that we use and those available from these other sources. Finally, during our data collection, we verified from the *earningswhispers.com* website that they use inputs from analysts as well as from other sources to compile the whisper forecasts.

Data on returns are from the Center for Research in Security Prices (CRSP), and financial data are from Compustat. The analysts' consensus forecasts are from the Institutional Brokers Estimate System (IBES). We also collect the announcement dates and reported earnings data from IBES.[1]

The initial sample of whisper forecasts consists of 17,360 distinct firm-quarters. Information from IBES is missing for 2,707 firm-quarters. We delete observations that we could not match on Compustat and retain firms only if they are based in the U.S. (excluding observations if the share code from CRSP is not equal to 10 or 11). This eliminates an additional 811 firm-quarters. We exclude 229 observations where the share price at the beginning of the quarter is not available, or if either the whisper/consensus forecast or actual

[1]In order to prevent errors associated with using the split-adjusted IBES files we use the unadjusted files, as suggested by Payne and Thomas (2003).

earnings are missing. We further exclude 327 firm-quarters where the whisper/consensus forecast or the actual earnings are zero. As our primary tests involve a comparison of whisper and consensus analysts' forecasts, we eliminate 2,225 observations where the whisper and consensus forecasts are the same. To minimize the effect of possible data errors, we eliminate 1,297 firm-quarters where the absolute value of the whisper forecast error or the consensus analyst forecast error (scaled by the absolute value of reported EPS) is greater than 1.

The final sample consists of 9,764 whisper forecasts of quarterly earnings per share made during a two-year period (October 1999 to September 2001) that spans the introduction of Reg. FD. This is significantly larger than the sample size in Bagnoli *et al.* (1999), who study 943 whisper forecasts.

3.2. *Descriptive statistics*

In Table 1 (panel A), we present the summary statistics for our sample. The mean (median) market capitalization of the sample firms is $5.2 billion ($672 million). The book-to-market equity ratio is computed as the ratio of equity book value (data item # 60 from Compustat) divided by equity market value (#24 * #25) at the beginning of the calendar year of the forecast. The average (median) book-to-market equity ratio is 0.57 (0.43). The average (median) number of analysts following a firm is 6.58 (5.00). The inter-quartile range suggests that there is wide variation in these variables in our sample.

Table 1 (Panel B) presents the distribution across industries (as defined by the two-digit SIC code) for the sample firms. Firm in the business services industry (SIC 73, 14% of the sample) and depository institutions (SIC 60, 11% of the sample) are present in large numbers. Chemicals (SIC 28), industrial machinery (SIC 35), and the electronic equipment (SIC 36) firms each represent about 5–8% of the sample. A majority of the sample (55%) is in the "others" category, which includes industries with less than 5% representation. We classify firms in the following SIC codes as high-tech firms: 2834–36, 3570–79, 3660–69, 3670–79, 3690, 3694–95, 3820–29, 3840–49, 3861, and 7370–79.[2] About 70% of our sample are non-high-tech firms. In contrast, Bagnoli *et al.* (1999) note that over 90% of the firms in their sample are high-tech firms, which may limit the generality of their results.

[2]Chandra *et al*. (1999) and Amir *et al.* (1999) use a similar classification.

Table 1. Summary statistics and industry distribution of the sample.

	Median	Mean	Q1	Q3	# Obs.
Panel A: Summary statistics					
Market value of equity ($ millions)	671.92	5, 206.12	202.26	2, 274.50	9,764
Equity book-to-market ratio	0.43	0.57	0.21	0.72	9,764
Number of analysts' following	5.00	6.58	2.00	9.00	9,764
Panel B: Industry (two digit SIC code)				Number (%)	
Chemicals (28)				635 (6.50%)	
Industrial machinery (35)				556 (5.69%)	
Electronic equipment (36)				749 (7.67%)	
Depository institutions (60)				1,079 (11.05%)	
Business services (73)				1,366 (13.99%)	
Others				5,379 (55.09%)	
High-tech industries				2,880 (29.50%)	
Non-high-tech industries				6,884 (70.50%)	

Notes: The sample consists of 9,764 firm-quarters that had available whisper and consensus forecasts of current quarter earnings from October 14, 1999 to September 12, 2001. The market value of equity and the book-to-market ratio of equity are computed from Compustat at the calendar year-end prior to the earnings announcement. The number of analysts' following is prior to the earnings announcement date and comes from I/B/E/S. Panel A lists the mean, median and the first and third quartile for these variables. Panel B lists the industry concentration for the major industries represented in the sample (those with at least 5% of the sample) and the actual number of firm-quarters from that industry. High-tech firms are firms in industries with the following four-digit SIC codes: 2834–36, 3570–79, 3660–69, 3670–79, 3690, 3694–95, 3820–29, 3840–49, 3861, and 7370–79.

4. Accuracy of Whisper and Consensus Forecasts

We begin our empirical analysis by comparing the accuracy of whisper and consensus analyst forecasts. The whisper forecast error (consensus analyst forecast error) is defined as the difference between the actual quarterly earnings per share and the whisper forecast (consensus analysts' forecast). We normalize the forecast error by either the absolute value of the actual quarterly earnings per share or by the share price at the beginning of the quarter of the earnings announcement (from CRSP). We compare forecast errors for firms with positive earnings versus those for firms incurring losses, as previous research, such as Dowen (1996), shows significant differences in the optimistic bias of analysts for the two groups. Further, we report the percentage of instances where whisper forecasts represent an improvement in accuracy over the consensus analysts' forecasts (i.e., smaller absolute forecast error).

4.1. *Full sample results*

We present the results for the full sample in Table 2. For all firms, the median whisper forecast error (WFE) is 0.00% for both the absolute EPS and the share price.[3] Only 37.8% of whisper forecast errors are positive, i.e., they are pessimistic and underestimate earnings. Thus, about 62% of whisper forecasts are optimistic. There is wide dispersion in whisper forecast errors, as evidenced by the inter-quartile range (Q3–Q1) of 17% (as a percentage of the absolute value of the EPS).[4] In contrast, the median consensus analyst forecast error (CFE) is 2.47% of the reported EPS and 0.03% of the share price, and 56.1% are positive. Thus, consensus analyst forecasts are *pessimistic,* as opposed to whisper forecasts which are *optimistic*.

We illustrate the relative accuracy of whisper forecasts and consensus forecasts by dividing the sample into five categories. Category 1 represents a whisper forecast that is between the consensus and the actual EPS without overshooting the earnings (see Figure 1). Category 2 represents the case in which the whisper overshoots the actual EPS but the magnitude of the forecast error is less than that for the consensus forecast. Category 3 represents cases in which the whisper overshoots the actual EPS and the forecast error is not better than the consensus forecast error. Category 4 represents cases in which the whisper forecast is in the opposite direction of the actual EPS and the consensus forecast, and category 5 represents cases where the consensus forecast error is zero while the whisper forecast is non-zero. Thus, categories 1 and 2 represent a better whisper forecast than consensus forecast while the other categories represent cases in which the whisper forecast is worse than the consensus forecast.

Only in 49.1% of the cases does the whisper forecast represent an improvement over the consensus forecast. This proportion is statistically significantly different from 50% with a p-value of 0.04. When we analyze the forecast errors separately for the profit firms, we find that the median whisper forecast error is 0%. The median consensus forecast error is 2.70% of the absolute value of EPS. The median whisper forecast error is significantly negative for loss firms (−2.86%), and the median consensus forecast error is zero. Whisper forecasts improve the consensus forecast in 49.4% of the cases (p-value 0.59).

[3]The Wilcoxon rank sum test is used throughout for the medians.

[4]Since our results when we deflate using the absolute value of the EPS or share price are similar, we only briefly discuss the latter results.

Table 2. Whisper forecast error and consensus forecast error.

	All firms (− = 9,764)		Profit firms (− = 7,779)		Loss firms (− = 1,985)	
	WFE (% > 0)	CFE (% > 0)	WFE (% > 0)	CFE (% > 0)	WFE (% > 0)	CFE (% > 0)
Forecast error as % of abs (EPS)	0.00***	2.47***	0.00***	2.70***	−2.86***	0.00**
	(37.8)	(56.1)	(37.2)	(58.3)	(40.3)	(47.4)
Forecast error as % of price	0.00***	0.03***	0.00***	0.03***	−0.05***	0.00***
	(37.8)	(56.1)	(37.2)	(58.3)	(40.3)	(47.4)
% Whispers improve consensus (p-value)	49.1**		49.0**		49.4	
	(0.04)		(0.04)		(0.59)	

Notes: WFE (CFE) is the whisper (consensus) forecast error, computed as actual earnings per share minus whisper (consensus) forecast per share, deflated by either the absolute value of the reported earnings per share or the share price. The table reports the median forecast errors and the percentage positive (in parentheses). The significance levels are from the Wilcoxon rank sum test for the median. The third row lists the percentage of observations where the whisper forecasts are more accurate than the consensus and the associated p-value (testing for difference from 50%, excluding instances where whisper and consensus are of equal accuracy). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

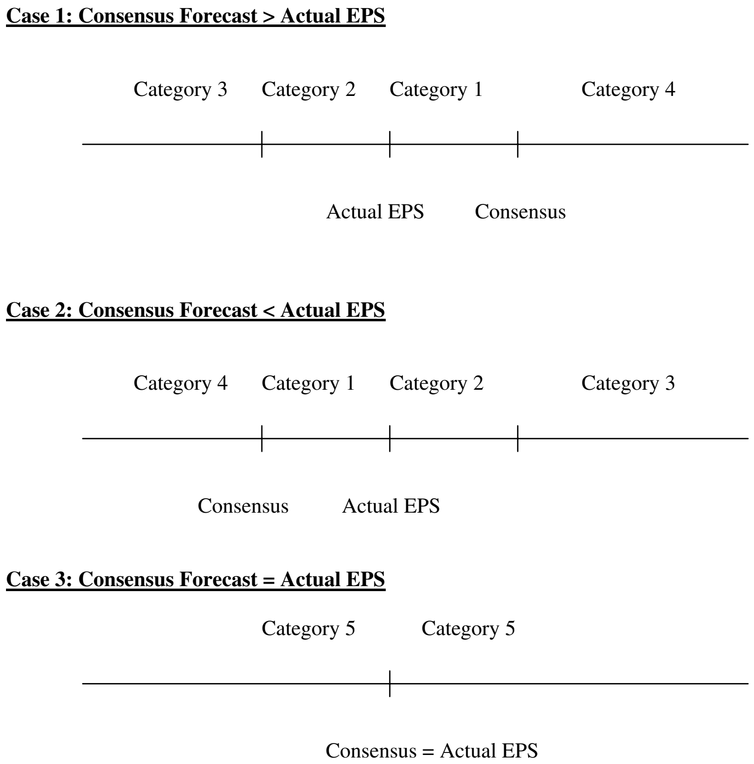


Figure 1. Relative accuracy of whisper and consensus forecasts.

Thus, unlike Bagnoli *et al.* (1999), our results show that whisper forecasts are not more accurate for all firms than consensus analysts' forecasts.

4.2. *Accuracy of whisper and consensus forecast in different types of firms*

While our overall sample results differ from Bagnoli *et al.* (1999), it is likely that the differences are due to differences in the sample composition in the two studies. About 30% of our sample is composed of high-tech firms, whereas about 90% of their sample is from high-tech firms. Also, as discussed in Section 1, analysts may not be able to (or may not wish to) provide accurate public forecasts for rapidly evolving, high-tech, high-growth firms. We now examine whether the accuracy of private and public forecasts is indeed different in these cases.

4.2.1. *High-tech versus non-nigh-tech firms*

We divide our sample into two broad categories: high-tech firms and non-high-tech firms. In Table 3 (panel A), we report the results for high-tech firms. Consistent with our hypothesis, the median whisper forecast error for high-tech firms is 0%, when normalized by the absolute value of EPS. About 40% of the whisper forecast errors are positive, i.e., the whisper forecast underestimates the reported EPS. In contrast, the median consensus forecast error is 4.76%, and 59.3% of the forecasts are pessimistic. Further, whispers improve the consensus forecast 53.5% of the time, and this proportion is statistically significantly different from 50% at the 1% significance level.

When the high-tech firms are further split into profit and loss firms, we find that the results hold only for profitable firms. Though the median whisper forecast error is zero both for profit and loss firms, the median consensus forecast (5.56%) is significantly positive only for the profit firms. The median consensus forecast error is not significantly different from zero for the loss firms. Moreover, for the profit group, whispers improve the consensus forecast 54.8% of the time (p-value 0.01), whereas for the loss firms whispers improve the consensus only 51.4% of the time (p-value 0.38). Thus, for high-tech firms that make profits, our results are consistent with the findings in Bagnoli *et al.* (1999). However, whisper forecasts are not better than consensus forecasts for high-tech firms that report losses.

The results for non-high-tech firms presented in panel B are different from those for the high-tech firms. The median whisper forecast error is −1.21% and is significantly different from zero at the 1% level. Whisper forecast errors are pessimistic in 37% of the observations. However, the median consensus forecast error is 1.96% and 54.8% of the forecasts are pessimistic. Moreover, whispers represent an improvement over the consensus in only 47.2% (p-value 0.01) of the cases. The results are similar both for profit and loss firms.

These results indicate that whisper forecasts are more accurate than consensus analysts' forecasts for profitable high-tech firms, but not for non-high-tech firms. Thus, analysts may want to retain their private information when it is more valuable. However, consensus forecasts may be as good as or better than whisper forecasts for other firms with lower uncertainty of earnings.

Table 3. Whisper forecast error and consensus forecast error across industry categories.

	All firms			Profit firms			Loss firms		
	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*
Panel A: High-tech firms									
Forecast error as % of abs (EPS)	0.00*** (39.7)	4.76*** (59.3)	2,880	0.00** (38.2)	5.56*** (65.5)	1,724	0.00*** (41.9)	0.68 (50.0)	1,156
Forecast error as % of price	0.00*** (39.7)	0.03*** (59.3)	2,880	0.00** (38.2)	0.03*** (65.5)	1,724	0.00*** (41.9)	0.01 (50.0)	1,156
% Whispers improve consensus (*p*-value)	53.5*** (0.01)			54.8*** (0.01)			51.4 (0.38)		
Panel B: Non-high-tech firms									
Forecast error as % of abs (EPS)	−1.21*** (37.0)	1.96*** (54.8)	6,884	−0.94*** (36.8)	2.17*** (56.3)	6,055	−4.55*** (38.1)	0.00*** (43.8)	829
Forecast error as % of price	−0.02*** (37.0)	0.03*** (54.8)	6,884	−0.02*** (36.8)	0.04*** (56.3)	6,055	−0.10*** (38.1)	0.00*** (43.8)	829
% Whispers improve consensus (*p*-value)	47.2*** (0.01)			47.3*** (0.01)			46.6* (0.06)		

Notes: WFE (CFE) is the whisper (consensus) forecast error, computed as actual earnings per share minus whisper (consensus) forecast per share, deflated by either the absolute value of the reported earnings per share or the share price. Firms are classified as "high-tech" (Panel A) and "non-high tech" (Panel B), based on Compustat two-digit SIC code. Each panel in the table reports the median forecast errors and the percentage positive (in parentheses). The significance levels are from the Wilcoxon rank sum test for the median. The third row in each panel lists the percentage of observations where the whisper forecasts are more accurate than the consensus and the associated *p*-value (testing for difference from 50%, excluding instances where whisper and consensus are of equal accuracy). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

4.2.2. *Whisper and analyst forecast errors by book to market ratio quartiles*

We now turn our attention to the pattern of forecast errors for firms classified by the book-to-market ratio (BM). We examine whether whisper and consensus forecast errors differ across quartiles based on the equity BM, as predicted. In Table 4 (panel A), we document that the results for firms in the lowest BM quartile (high-growth stocks) are similar to those for high-tech firms. The median whisper forecast error is zero, whereas the median consensus forecast error is significantly positive. The whisper (consensus) forecast error is pessimistic in 37% (62.9%) of the cases. The whisper forecast improves the consensus forecast in 54% of the observations (p-value 0.01).

The consensus forecast pessimism declines as we move from the low BM to the high BM quartiles. Only 50.9% of the consensus forecasts are pessimistic in the highest BM quartile, compared with 62.9% in the lowest BM quartile. In contrast, there is little change by quartile in the whisper forecasts. The frequency of pessimistic forecasts remains between 36% and 39% in all four BM quartiles. A similar pattern emerges for the profit firms, but not for the loss firms. However, for firms in the three highest BM quartiles, whisper forecasts are more accurate than consensus forecasts only about 47% of the time. Overall, this evidence suggests that consistent with our expectations, whisper forecasts are more accurate than consensus forecasts only for high-tech and high-growth firms. This is not so for other firms where the earnings are more likely to be predictable. Our findings suggest that the conclusions of Bagnoli *et al.* (1999) continue to hold during this later time period.

The classification of firms on the basis of BM ratio could be correlated with our classification based on industry groups. We find that the BM is negatively correlated to our high-tech dummy (−0.147 with a p-value of 0.01). Specifically, high-tech firms may consist primarily of low-BM growth stocks. Further, we find that for both high-tech and high-growth firms, whisper forecasts are better than consensus forecasts only for profit firms.

In order to verify whether the consensus forecast pessimism and the improved accuracy of whisper forecasts are limited to high-tech firms or whether they are characteristics of growth stocks in general, we proceed as follows. We divide the non-high-tech firms into quartiles based upon the BM. For the firms in the lowest book-to-market quartile (non-high-tech, growth firms), the median consensus forecast error is 2.67% of EPS or 0.03% of the

Table 4. Whisper forecast error and consensus forecast error across book-to-market categories.

	All firms			Profit firms			Loss firms		
	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*
Panel A: BM quartile 1(low)									
Forecast error as % of abs (EPS)	0.00*** (37.0)	4.26*** (62.9)	2,441	0.00*** (34.1)	4.55*** (66.9)	1,717	0.00* (43.6)	3.01 (53.5)	724
Forecast error as % of price	0.00*** (37.0)	0.03*** (62.9)	2,441	0.00*** (34.1)	0.03*** (66.9)	1,717	0.00*** (43.6)	0.03 (53.5)	724
% Whispers improve consensus (*p*-value)	54.0*** (0.01)			54.9*** (0.01)			51.9 (0.38)		
Panel B: BM quartile 2									
Forecast error as % of abs (EPS)	0.00*** (38.5)	2.50*** (58.1)	2,443	0.00*** (38.1)	2.63*** (59.7)	2,031	−2.58** (40.3)	0.54 (50.0)	412
Forecast error as % of price	0.00*** (38.5)	0.04*** (58.1)	2,443	0.00*** (38.1)	0.04*** (59.7)	2,031	−0.04** (40.3)	0.00 (50.0)	412
% Whispers improve consensus (*p*-value)	48.0** (0.05)			47.7** (0.04)			49.4 (0.85)		
Panel C: BM quartile 3									
Forecast error as % of abs (EPS)	−1.38*** (36.7)	1.45*** (52.5)	2,438	−1.07*** (36.7)	1.56*** (53.8)	2,166	−5.56*** (37.1)	0.00*** (41.9)	272
Forecast error as % of price	−0.03*** (36.7)	0.03*** (52.5)	2,438	−0.02*** (36.7)	0.03*** (53.8)	2,166	−0.10*** (37.1)	0.00*** (41.9)	272
% Whispers improve consensus (*p*-value)	47.4*** (0.01)			47.5** (0.02)			46.3 (0.25)		

(*Continued*)

Table 4. (*Continued*).

	All firms			Profit firms			Loss firms		
	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*
Panel D: BM quartile 4 (high)									
Forecast error as % of abs (EPS)	−1.72*** (39.0)	1.55** (50.9)	2,442	−1.06*** (39.5)	2.50*** (54.1)	1,865	−5.66*** (37.6)	−0.55*** (40.6)	577
Forecast error as % of price	−0.04*** (39.0)	0.04*** (50.9)	2,442	−0.03** (39.5)	0.06*** (54.1)	1,865	−0.22*** (37.6)	−0.04*** (40.6)	577
% Whispers improve consensus (p-value)	46.8*** (0.01)			46.5*** (0.01)			47.8 (0.32)		

Note: WFE (CFE) is the whisper (consensus) forecast error, computed as actual earnings per share minus whisper (consensus) forecast per share, deflated by either the absolute value of the reported earnings per share or the share price. Firms are classified into quartiles based on the equity book-to-market ratio. Each panel in the table reports the median forecast errors and the percentage positive (in parentheses). The significance levels are from the Wilcoxon rank sum test for the median. The third row in each panel lists the percentage of observations where the whisper forecasts are more accurate than the consensus and the associated p-value (testing for difference from 50%, excluding instances where whisper and consensus are of equal accuracy). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

share price.[5] These are statistically different from zero at the 1% level. The consensus forecast is pessimistic in 60.5% of the cases. The median whisper forecast error is 0.00% of EPS and 0.00% of the share price (significant at the 1% level). For value stocks in the fourth quartile, the median consensus forecast error is 1.54% of EPS and 0.04% of the share price (both significant at 5% or better). The median whisper forecast error is −1.92% of EPS and –0.06% of the share price (both significant at the 1% level).

However, the whisper forecasts are not better than the consensus forecast in any category. In fact, for the high-growth loss firms, the whisper forecasts are better than the consensus in 49.9% of the cases. This evidence suggests that the results we find for growth firms in general are driven mainly by the high-tech firms. Whisper forecasts are not better than consensus forecasts for non-high-tech, high-growth firms.

4.3. *The impact of regulation FD on whisper and consensus forecasts*

We now investigate whether the superiority of whisper forecasts documented here arises as a result of selective disclosure by management. Specifically, the results in Skinner and Sloan (2002) and Matsumoto (2002) suggest that managers of high-tech and high-growth firms may be more willing to exert pressure on analysts and guide their public forecasts downward. They could also reward compliant analysts with information that the analysts could disclose privately. Reg. FD prohibits firms from disclosing information to a few, selected analysts. If the source of whisper forecasts is selective disclosure by management, we expect that whisper forecasts would no longer be more accurate than consensus forecasts after introduction of Reg. FD. On the other hand, if whisper forecasts are the result of analysts disseminating their own information privately while making pessimistic, public forecasts, we would not expect to see any change after Reg. FD.

Since the introduction of Regulation FD, extant research has studied the effects of Reg. FD on analyst forecasts. Heflin *et al.* (2003), and Bailey *et al.* (2003) do not find any change in forecast errors. Agrawal and Chadha (2004) find that analyst forecasts become less accurate. While these

[5]These results are not reported in a table, but are available from the authors upon request.

studies analyze public analyst forecasts, our focus is on private, whisper forecasts.

Table 5 shows that the results we document in Table 2 for the full sample continue to hold both before (panel A) and after (panel B) the introduction of Reg. FD. Whisper forecasts do not improve upon consensus forecasts for the full sample. Consistent with the assertion that in the pre-Reg.-FD period, analysts publicly issued pessimistic forecasts to curry favor with management, we find that the median consensus forecast error declines from 3.51% in the pre-FD period to 1.85% after Reg. FD was introduced. Furthermore, the proportion of pessimistic forecasts drops from 61.7% to 53%. Thus, after Reg. FD made it difficult for managers to use selective disclosure to reward compliant analysts, forecast pessimism declined.

In Table 6, we compare the forecast errors for high-tech and non-high tech firms (panel A) and firms in the lowest and highest BM quartile (panel B). We find that our earlier results are unchanged. Specifically, for high-tech firms, we find that whisper forecasts continue to be more accurate than consensus forecasts in about 53% of the cases, both before and after Reg. FD. A similar result is obtained for the low BM quartile firms. Also, for non-high-tech and high-BM firms, consensus forecasts continue to be more accurate than whisper forecasts both before and after Reg. FD. Hence, even though Reg. FD has reduced consensus forecast pessimism, whisper forecasts continue to be more accurate than consensus forecasts for high-tech and high-growth firms. The hypothesis that whisper forecasts are the outcome of selective disclosure is not supported by our results.

To further investigate whether Regulation FD has affected the information environment in which analysts work, and thus may have had an effect on the accuracy of whisper and consensus forecasts, we use a set of measures introduced by Barron *et al.* (1998). Their framework allows for measuring the quality (or precision) of private and common information by analysts. Although our earlier findings suggest that the continued superiority of the whisper forecast is not the outcome of selective disclosure, we want to investigate whether the information environment has changed as was the goal of Regulation FD (i.e., analysts no longer are allowed to get information privately from firms and use this information in their forecasts). Because we find that after the introduction of regulation FD whisper forecasts are still more accurate for certain groups of firms (high-tech and high-growth), we want to ensure that regulation FD did indeed have an effect on the environment

Table 5. Comparison of whisper and consensus forecast error around the introduction of Reg. FD.

	All firms			Profit firms			Loss firms		
	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*	WFE (% > 0)	CFE (% > 0)	*N*
Panel A: Pre-FD									
Forecast error as % of abs (EPS)	−1.02*** (37.4)	3.51*** (61.7)	3,471	0.00*** (36.9)	3.70*** (63.8)	2,948	−5.00*** (40.0)	0.00 (49.9)	523
Forecast error as % of Price	−0.01*** (37.4)	0.05*** (61.7)	3,471	0.00*** (36.9)	0.05*** (63.8)	2,948	−0.07*** (40.0)	0.00* (49.9)	523
% Whispers improve consensus (p-value)	49.3 (0.31)			49.8 (0.74)			45.9* (0.08)		
Panel B: Post-FD									
Forecast error as % of abs (EPS)	0.00*** (38.0)	1.85*** (53.0)	6,293	0.00*** (37.3)	2.13*** (55.0)	4,831	−2.24*** (40.4)	0.00** (46.5)	1,462
Forecast error as % of price	0.00*** (38.0)	0.02*** (53.0)	6,293	0.00*** (37.3)	0.03*** (55.0)	4,831	−0.04*** (40.4)	0.00*** (46.5)	1,462
% Whispers improve consensus (p-value)	48.9* (0.07)			48.4** (0.02)			50.7 (0.68)		

WFE (CFE) is the whisper (consensus) forecast error, computed as actual earnings per share minus whisper (consensus) forecast per share, deflated by either the absolute value of the reported earnings per share or the share price. The table lists the median forecast errors and the percentage positive (in parentheses) for forecasts made before Reg. FD (before October 2000, panel A) and after Reg. FD (panel B). The significance levels are from the Wilcoxon rank sum test for the median. The third row in each panel lists the percentage of observations where the whisper forecasts are more accurate than the consensus and the associated p-value (testing for difference from 50%, excluding instances where whisper and consensus are of equal accuracy). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 6. Whisper and consensus forecast error around Reg. FD across industry and book-to market categories.

	High-tech				Non-high-tech			
	Pre-FD (– = 1,003)		Post-FD (– = 1,877)		Pre-FD (– = 2,468)		Post-FD (– = 4,416)	
	WFE (% > 0)	CFE (% > 0)	WFE (% > 0)	CFE (% > 0)	WFE (% > 0)	CFE (% > 0)	WFE (% > 0)	CFE (% > 0)
Panel A: High-tech versus non-high-tech firms								
Forecast error as % of abs (EPS)	0.00***	7.02***	0.00***	3.33***	−1.21***	2.63***	−1.20***	1.52***
	(39.2)	(67.7)	(40.0)	(54.8)	(36.6)	(59.2)	(37.2)	(52.3)
Forecast error as % of price	0.00***	0.04***	0.00***	0.02***	−0.02***	0.05***	−0.02***	0.02***
	(39.2)	(67.7)	(40.0)	(54.8)	(36.6)	(59.2)	(37.2)	(52.3)
% Whispers improve consensus (p-value)	53.45*		53.46***		47.55**		46.99***	
	(0.06)		(0.01)		(0.02)		(0.01)	
	BM quartile 1 (low)				BM quartile 4 (high)			
	Pre-FD (– = 980)		Post-FD (– = 1,466)		Pre-FD (– = 769)		Post-FD (– = 1,673)	
Panel B: Low versus high BM quartile firms								
Forecast error as % of abs (EPS)	0.00***	6.02***	0.00***	3.10***	−1.79***	2.76***	−1.44***	0.00***
	(36.6)	(71.3)	(37.7)	(57.6)	(38.0)	(56.2)	(40.0)	(49.1)
Forecast error as % of price	0.00***	0.04***	0.00***	0.02***	−0.04***	0.07***	−0.04***	0.00***
	(36.6)	(71.3)	(37.7)	(57.6)	(38.0)	(56.2)	(40.0)	(49.1)
% Whispers improve consensus (p-value)	54.44**		53.57***		46.18**		47.41**	
	(0.03)		(0.01)		(0.03)		(0.04)	

Notes: WFE (CFE) is the whisper (consensus) forecast error, computed as actual earnings per share minus whisper (consensus) forecast per share, deflated by either the absolute value of the reported earnings per share or the share price. The table lists the median forecast errors and the percentage positive (in parentheses) for forecasts made before and after Reg. FD for high-tech firms versus others (panel A) and for low BM versus high BM firms (panel B). The significance levels are from the Wilcoxon rank sum test for the median. The third row in each panel lists the percentage of observations where the whisper forecasts are more accurate than the consensus and the associated p-value (testing for difference from 50%, excluding instances where whisper and consensus are of equal accuracy). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

to which an analyst is subjected when generating forecasts. The method of Barron *et al.* (1998) allows for such a test (for examples of others who have used this method see Venkataraman 2001; Barron *et al.*, 2002; and Botosan *et al.*, 2004).

Precision of public information is measured by $\underline{h}$ and is defined as:

$$h = \frac{\mathrm{SE} - D/N}{((1 - 2/N)D + \mathrm{SE})^2} \tag{1}$$

where SE is the squared error of the consensus mean forecast, $\underline{\mathrm{D}}$ is the dispersion among forecasts, and $\underline{\mathrm{N}}$ is the number of analysts. Private information is measured as $\underline{\mathrm{s}}$ and is defined as follows:

$$\underline{\mathrm{s}} = \frac{\underline{\mathrm{D}}}{((1 - 1/\underline{\mathrm{N}})\underline{\mathrm{D}} + \mathrm{SE})^2} \tag{2}$$

The extent to which public information is used in analysts' forecasts is measured by $\underline{\mathrm{h}}$, whereas $\underline{\mathrm{s}}$ measures the amount of idiosyncratic information in analysts' forecasts. When we calculate $\underline{\mathrm{h}}$ and $\underline{\mathrm{s}}$ for all the available forecasts we find that there is no significant difference in the amount of private information in the consensus forecasts. The median is 57.34 in the pre-period and is 58.57 in the post-period. At the same time, we do find that the median $\underline{\mathrm{h}}$ decreased from 65.39 to 48.22. (Results are not reported but are available upon request.) These findings lend credence to our finding that the private information environment has changed as a result of reg. FD.

4.4. *What determines relative accuracy?*

In the previous sections we found that whisper forecasts were more accurate for profitable high-tech firms and profitable firms with a low book-to-market. We did not find that these patterns changed after the introduction of Reg. FD. In Table 7 we perform a simple logistic regression to see whether these findings hold in a multivariate setting. The dependent variable takes a value of 1 if the whisper forecast represents an improvement over the consensus forecast. In all four models we find that being in the high-tech sector is significantly positively related to the dependent variable. Also consistent with our univariate findings we find that having a low book-to-market is significantly positively related to the whispers being more accurate.

There is no evidence that Reg. FD has any material effect on the chances of whispers being more accurate as indicated by the Pre-FD dummy (equals

Table 7. Determinants of relative forecast accuracy.

	Model 1	Model 2	Model 3	Model 4
Intercept	0.001	0.026	0.063	0.078
High-tech	0.229***	0.218***	0.226***	0.210***
BM	−0.074**	−0.074**	−0.103***	−0.102***
Pre-FD	−0.004	−0.010	−0.010	−0.011
Pos. EPS		−0.001	0.057*	0.036
Pos. EPS* High-tech		0.050		0.035
Pos. EPS* BM			−0.057*	−0.054*
Log(MVE)	0.017	−0.021*	−0.027**	−0.027**
N	9,764	9,764	9,764	9,764

Notes: The dependent variable is a dummy equal to 1 if the whisper forecast is more accurate then the consensus forecast. The market value of equity (MVE) and the book-to-market ratio of equity (BM) are computed from Compustat at the calendar year-end prior to the earnings announcement. High-tech is a dummy equal to 1 if firms are firms in industries with the following four-digit SIC codes: 2834–36, 3570–79, 3660–69, 3670–79, 3690, 3694–95, 3820–29, 3840–49, 3861, and 7370–79. Pre-FD is a dummy equal to 1 if forecasts are made before the introduction of Reg. FD. Pos. EPS is dummy equal to 1 if the actual earnings are positive. The table lists the logistic regression coefficients and the number of observations. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

1 if the forecast is before the introduction of Reg. FD). We do not find that the dummy for positive actual earnings per share (equal to 1 if actual EPS > 0) is significant. When we interact this dummy with book-to-market we do find significant negative coefficients, suggesting that low book-to-market firms with positive earnings are more likely to have more accurate whispers. In three of our four models we find the log of market-value-of-equity to be significantly negatively related to our dependent variable. Overall these findings confirm that high-tech firms and low book-to-market firms are more likely to have whispers that are more accurate than consensus forecasts.

5. Abnormal Returns Around Earnings Announcements

While the results documented so far indicate that whisper forecasts may be more accurate than consensus forecasts for high-tech and high-growth firms, we have not investigated whether they are related to returns around earnings announcements. If, as at least some market participants claim, whisper forecasts represent the market's true earnings expectations, then we should observe that abnormal returns around earnings announcements should be closely related to the whisper forecast error. On the other hand, if the credibility of whisper forecasts is poor because of their anonymity, or if market participants adjust the consensus expectations for any predictable bias, then whisper forecast errors would not be related to abnormal returns. We compute abnormal returns (CAR) as market-adjusted returns over days (−1 to +1), where day 0 is the earnings announcement day. We use the returns from the CRSP equally weighted index as our proxy for the market return. The results of this analysis are presented in Table 8.

We find that the price reaction to the whisper forecast relates to the signs of the whisper forecast errors and consensus forecast errors (Table 8, panel A). When the actual earnings are less than the whisper forecast, the median CAR is −1.76%, −0.61%, and −0.04% for negative consensus forecasts, consensus forecasts of zero, and positive consensus forecast errors, respectively (the first two significantly different from zero at 1%, and the latter not significantly different from zero).

When the whisper forecast equals the actual EPS, the abnormal return is −0.86% for those cases where the actual is smaller than the consensus forecast, which is significantly different from zero at the 5% level, and 0.63% (significant at the 1% level as well) for those instances where the actual is larger than the consensus forecast.

For firms where the whisper forecast underestimates actual EPS, the abnormal return again depends on the sign of the consensus forecast error. When the consensus forecast is larger than the actual the abnormal return is –0.43% and not significant. When the consensus forecast error is zero the abnormal return is not significantly different from zero at –0.36%. If the actual is larger than the consensus forecast error the abnormal return is 1.59% (significant at 1%). Most striking is that when the actual is lower or equal to the consensus forecast the abnormal return is negative (and significantly different from zero). This may arise if the market expects the analysts to provide conservative forecasts and hence the market's true expectations would be higher. In this case, earnings that just meet the consensus forecast may be a negative signal. Our results do not materially change when we divide the sample into profit and loss firms (panels B and C).

Regression analysis shows that the earnings response coefficients for the whisper forecast error and consensus forecast error are similar. Also, the adjusted R-squares are similar. This suggests that, for the entire sample, whisper forecasts do not convey information beyond what is conveyed by

Table 8. Forecast errors and abnormal returns around earnings announcements.

	Actual < Consensus		Actual = Consensus		Actual > Consensus	
	Median CAR	#Obs	Median CAR	# Obs	Median CAR	# Obs
Panel A: All firms						
Actual < Whisper forecast	−1.76***	2,101	−0.61***	1,087	−0.04	1,670
Actual = Whisper forecast	−0.86**	299	n.a.		0.63***	910
Actual > Whisper forecast	−0.43	429	−0.36	366	1.59***	2,893
Panel B: Profit firms						
Actual < Whisper forecast	−1.40***	1,413	−0.56***	987	0.14***	1,430
Actual = Whisper forecast	−0.83*	249	n.a.		0.78***	808
Actual > Whisper forecast	−0.65	294	−0.06	299	1.66***	2,297
Panel C: Loss firms						
Actual < Whisper forecast	−3.03***	688	−1.86	100	−1.73***	240
Actual = Whisper forecast	−1.62	50	n.a.		−1.87	102
Actual > Whisper forecast	0.41	135	−2.68***	67	1.08*	596

Notes: We compute the market-adjusted returns around the earnings announcements using the return on the CRSP equally weighted index as our proxy for the market return. Daily abnormal returns (percent) are cumulated over days −1 to 1 (CAR (−1, 1). The earnings announcement date is day zero. We report the median percentage CAR (−1, 1) separately for cases where the actual earnings is less than, equal to or greater than the whisper forecast or the consensus analysts' forecast. We report the results for the entire sample in panel A, for firms reporting profits (losses) in panel B (panel C). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

consensus forecasts. When we compare earnings response coefficients for the sample of forecasts that are likely to have biased consensus forecasts our results do not change. Hence, we conclude that even if whisper forecasts improve the consensus forecasts, the market may recognize any predictable bias on the part of analysts and react accordingly. If so, trading strategies based on analysts' whisper forecasts may not earn returns over and above those based on the consensus forecasts. This evidence is contrary to the findings of Bagnoli *et al.* (1999).

6. Conclusions

In this paper, we investigate private whisper and public consensus forecasts along two dimensions: relative accuracy and value-relevance. We find that whisper forecasts tend to be optimistic, whereas consensus forecasts tend to be pessimistic in nature. Only in cases where analysts may have a tendency to bias their public forecasts (high-tech firms) or when analysts want to protect their private information (high growth firms) do we find that whisper forecasts are more accurate than consensus forecasts. For other firms, we find that the consensus tends to be more accurate. This can be explained by the fact that for these other firms earnings will be more predictable and hence there will be a larger reputational penalty for providing imprecise public forecasts. Also, we find that both whisper and consensus forecast errors are related to stock returns around earnings announcements, suggesting that both forecasts convey valuable information.

Finally, we find that after the introduction of regulation FD the existing accuracy patterns do not materially change. Both pre-and post-regulation FD whisper forecasts remain more accurate than consensus forecasts. This suggests that selective disclosure is not the reason for the relative accuracy of whisper forecasts.

In a more general context, our results suggest that for high-tech and high-growth firms, published consensus forecasts may not be the best estimate of the market's earnings expectations. For other firms, consensus forecasts may well serve as a proxy for market earnings expectations.

Acknowledgments

Erik Devos acknowledges financial support from the Gardner Fellowship, Ohio University. We thank Thomson Financial for providing analyst forecast

data, available through the Institutional Brokers Estimate System. This paper was started when Devos was at Binghamton University — SUNY.

References

Agrawal, A. and S. Chadha, *Who is Afraid of Reg FD? The Behavior and Performance of Sell-Side Analysts following the SEC's Fair Disclosure Rules*. Working Paper, University of Alabama, Alabama, 2004.

Amir, E., B. Lev and T. Sougiannis, *What Value Analysts?* Working Paper, New York University, New York, 1999.

Asquith, P., M. Mikhail and A. Au, "Information Content of Equity Analyst Reports." *Journal of Financial Economics* 75, 245–282 (2005).

Bagnoli, M., M. Beneish and S. Watts, "Whisper Forecasts of Quarterly Earnings per Share." *Journal of Accounting and Economics* 28, 27–50 (1999).

Bailey, W., H. Li, C. Mao and R. Zhong, "Regulation Fair Disclosure and Earnings Information: Market, Analyst, and Corporate Responses." *Journal of Finance* 58, 2487–2514 (2003).

Barron, O., D. Byard, C. Kile and E. Riedl, "High-Technology Intangibles and Analysts' Forecasts." *Journal of Accounting Research* 40, 289–312 (2002).

Barron, O., O. Kim, S. Lim and D. Stevens, "Using Analysts's Forecasts to Measure Properties of Analyst's Information Environment." *Accounting Review* 73 421–434 (1998).

Botosan, C., M. Plumlee and Y. Xie, "The Role of Information Precision in Determining the Cost of Equity Capital." *Review of Accounting Studies* 9, 233–259 (2004).

Chandra, U., C. Wasley and G. Waymire, *Share Price Responses to Technology Firms' Negative Quarterly Earnings Surprises*. Working Paper, University of Rochester, Rochester, 1999.

Das, S., C. Levine and K. Sivaramakrishnan, "Earnings Predictability and Bias in Analysts' Earnings Forecasts." *Accounting Review* 73, 277–294 (1998).

Dowen, R., "Analyst Reaction to Negative Earnings for Large Well-Known Firms." *Journal of Portfolio Management* 23, 49–55 (1996).

Francis, J. and L. Soffer, "The Relative Informativeness of Analysts' Stock Recommendations and Earnings Forecast Revisions." *Journal of Accounting Research* 35, 193–211 (1997).

Heflin, F., K. Subramanyam and Y. Zhang, "Regulation FD and the Financial Information Environment: Early Evidence." *Accounting Review* 78, 1–37 (2003).

Lev, B. and P. Zarowin, "The Boundaries of Financial Reporting and how to Extend Them." *Journal of Accounting Research* 37, 353–385 (1999).

Matsumoto, D., "Management's Incentives to Avoid Negative Earnings Surprises." *Accounting Review* 77, 483–514 (2002).

Mikhail, M., M. Walther and R. Willis, "Does Forecast Accuracy Matter to Security Analysts?" *Accounting Review* 74, 185–200 (1999).

Payne, J. and W. Thomas, "The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research." *Accounting Review* 78, 1049–1067 (2003).

Skinner, D. and R. Sloan, "Earnings Surprises, Growth Expectations, and Stock Returns or Don't Let an Earnings Torpedo Sink your Portfolio." *Review of Accounting Studies* 7, 289–312 (2002).

Stickel, S., "Reputation and Performance among Security Analysts." *Journal of Finance* 47, 1811–1836 (1992).

Venkataraman, R., *The Impact of SFAS 131 on Financial Analysts' Information Environment*. Working paper, Pennsylvania, State University, Pennsylvania, 2001.

Womack, K.L., "Do Brokerage Analysts' Recommendations have Investment Value?" *Journal of Finance* 51, 137–167 (1996).

Chapter 7

Earning Forecast-Based Return Predictions: Risk Proxies in Disguise?

Le (Emily) Xu
University of New Hampshire, USA

This paper explores a new approach to evaluate the likelihood that omitted risk factors or market mispricing underlies the anomalous security returns to price-scaled analysts' forecasts of earnings investigated by Elgers, Lo, and Pfeiffer (2001). This approach is to reduce measurement errors by incorporating predictable errors in analysts' earnings forecasts and used the adjusted forecasts to evaluate the underlying phenomenon. Under the explanation of market mispricing, the price-scaled adjusted analysts' forecasts are expected to generate greater hedge portfolio returns as the adjusted forecasts with less measure errors reduce the noise embedded in the relation of price-scaled analyst forecasts of earnings to subsequent abnormal returns. The results show that the adjusted forecasts do *not* enable improvements in the abnormal security returns resulting from the trading strategy. The inability to show improvements is inconsistent with the interpretation of market mispricing. Therefore, it provides indirect support for the interpretation of omitted risk factors.

Keywords: Analysts' forecasts of earnings; market mispricing; omitted risk proxies; predictable error pattern.

1. Introduction

A recent study (Elgers *et al.*, 2001, hereafter ELP) shows that subsequent abnormal security returns are associated with the ratio of early-in-the-year analysts' forecasts of earnings to market value of common equity. Given the results, the authors infer that the delayed security returns are due to investors' misunderstanding of the value-relevant information in publicly available analysts' estimates. However, a salient feature of the study is the use of market values as the method of scaling the forecast variable. Market value scaling introduces the possibility that a given firm-specific attribute, scaled by market value, serves as a "yield surrogate," i.e., a proxy for some pertinent risk factors that have not been controlled in the researcher's measures of abnormal security returns. Therefore, the market-value-scaled attributes would also generate significant subsequent abnormal security returns. Those apparent abnormal returns would actually represent equilibrium returns to the omitted risk factor, instead of biases in investors' interpretation of analysts' expected earnings.

The possibility that omitted risk factors may account for observed "abnormal" security returns has been widely acknowledged in the extant literature. Numerous studies have recognized omitted risk proxies as alternative explanations for market anomalies documented in the literature, e.g., earnings yield (Basu, 1977), accounting accruals (Sloan, 1996), analysts' forecasts (ELP), fundamental accounting ratios (Abarbanell and Bushee, 1997), and a firm's fundamental value (Frankel and Lee, 1998).

ELP have used conventional approaches to mitigate the problem. First, they use the mean returns for all the sample firms in the same market capitalization decile as the specification for expected market returns. In addition, an exhaustive list of previously documented proxies that may potentially be risk factors is included in the analyses. Unfortunately, however, both the approaches require the *ad hoc* intuition by the researcher in specifying the alternative security return metrics and/or in identifying a suitable (and sufficient) set of control variables. For this reason, a misspecified measure of equilibrium security returns continues to be a potential underlying explanation for the delayed security price adjustment documented by ELP.

This study provides a new approach to evaluating the likelihood that omitted risk factors or market mispricing underlies the anomalous security returns reported in ELP. My approach is motivated by a body of prior research that demonstrates predictable errors in analysts' earnings forecasts and shows how the predictable errors in the analysts' forecasts of earnings may be employed to address the ambiguity in the literature.

Recent literature offers the evidence that a variety of firm-specific variables are related to errors in early-in-the-year analysts' earning forecasts. Prior studies show that analysts fail to fully appreciate earning-relevant information imbedded in a variety of attributes: prior-year analysts' forecast errors, size-adjusted returns, actual earning changes, accruals, analysts' long-term growth rate forecasts, and other firm-specific characteristics (Abarbanell, 1991; Abarbanell and Bushee, 1997; Ali *et al*., 1992; Bradshaw *et al*., 2001; Elgers and Lo, 1994; Elgers and Murray, 1992; Frankel and Lee, 1998; Klein, 1990; La Porta, 1996; Lobo and Nair, 1990).

To the extent that analysts' earning forecasts reflect these predictable errors, those forecasts may add systematic classification "noise" to the relations tested in ELP. Given that the adjusted analyst earning forecasts are likely to be more accurate earnings predictors, it may expectted even greater returns to forecast-based hedge portfolios under the hypothesis of market mispricing. This argument is based on an assumption that the more accurate forecasts contain less measurement errors for the value-relevant component of forecasts. This approach aims to reduce measurement errors and reinvestigate the market anomaly using the attributes with less measurement errors.

2. Literature Review

2.1. *Discussion of ELP*

ELP report a profitable trading strategy based on price-scaled early-in-the-year analysts' forecasts of earnings. That is a positive significant association between price-scaled analysts' forecasts and subsequent securities returns. In addition, the authors partition their sample into two groups according to the analyst coverage. The analyst coverage is assumed a proxy for the quality of an information environment (Barth and Hutton, 2004; Barth *et al.*, 2001; Bhushan, 1994; Hong *et al.* 2000; Walther, 1997). Analysts are perceived as information intermediaries and to increase information flow to the security market. They document that the hedge portfolio strategy based on price-scaled analysts' forecasts is more profitable for firms with lower analyst coverage than for those with higher analyst coverage. As such, their interpretation of the results is that analysts can predict value-relevant earnings more accurately than investors can. In this case, analysts are viewed as being more informed than investors, and their publicly available forecasts are not efficiently impounded in securities prices, especially for firms with less analyst following. Furthermore, their results appear robust, with control for a list of the previously documented risk factors and anomalies. However, as argued before, the list of variables that are included to control for risk requires the *ad hoc* intuition of researchers. In addition, analyst following may be argued as a proxy for risk. That is firms with less analyst following are more risky than those with more analyst following. Therefore, it is still conceivable that the profitability of the trading strategy is due to omitted risk proxies rather than to market mispricing.[1]

[1] ELP also report 1.35% security returns, significant at 5%, in earning announcement months and 0.51%, not significant at conventional levels, in nonannouncement months and they conclude that the disproportionate concentration of returns in earning announcement months supports

2.2. *Predictable error patterns in analysts' forecasts of earnings*

A stream of studies has documented predictable error patterns embedded in analysts' forecasts of earnings, especially early-in-the-year forecasts. Abarbanell (1991) and Klein (1990) show that prior price changes predict errors in subsequent quarterly analysts' forecasts. Furthermore, analysts' inability or unwillingness to incorporate information in prior security returns is greater for firms with bad prior security return performance. Similarly, Elgers and Murray (1992) show that price-based forecasts, based on historical security returns and price-to-earnings ratios, can provide additional information for forecasting subsequent actual earnings beyond analysts' earning forecasts. Both studies show that the accuracy of analysts' forecasts can be improved by incorporating prior stock price changes, especially for firms with prior negative securities returns.

In addition, Elgers and Lo (1994) report that both prior earning changes and prior security returns are incrementally associated with analysts' forecast errors. Firms with prior poor return/earnings performance are more inclined to have more optimistic analysts' forecasts. A number of other studies have found similar results. According to Lobo and Nair (1990), the accuracy of the most accurate individual analyst forecasts can be improved by incorporating the time-series properties of earnings. Ali *et al.* (1992) show a positive serial relation in analysts' forecast errors. They interpret the evidence as analysts underestimating the persistence of past earnings forecast errors. They show that analysts tend to omit the earning-relevant information in prior forecast errors for firms with relatively permanent earnings and that this tendency is diminished for the firms with relatively transitory earnings. In sum, all the studies cited above conclude that the accuracy of analysts' forecasts can be improved by impounding prior earnings, in addition to prior securities returns.

Abarbanell and Bushee (1997) show a significant relationship between analysts' forecast errors and *ex ante* firm-specific fundamental signals, defined earlier in Lev and Thiagarajan (1993). Their results suggest that analysts

---

the interpretation of market mispricing. However, they have not statistically tested whether the security returns in earning announcement months are different from those in nonannouncement months. In addition, the magnitudes of the security returns (1.35% and 0.51%) and the difference in the security returns (0.84%) are economically insignificant. Therefore, this evidence does not sufficiently support the interpretation of market mispricing.

are not fully aware of the earning-relevant information embedded in the fundamental signals and as such respond with a delay. Bradshaw *et al.* (2002) suggest that analysts overestimate the persistence of accruals, consistent with investors' behavior, documented by Sloan (1996). La Porta (1996) shows that analysts are overly optimistic for firms with extremely high growth rate forecasts, and therefore leading to a negative association between growth rate forecasts and subsequent forecast errors.

Overall, errors in analysts' early-in-the-year earning forecasts are partly predictable based on (1) positive serial dependencies among successive forecast errors; (2) a tendency among analysts to overestimate reversals of prior-year poor earning performance; (3) a failure of analysts fully to incorporate the earning implications of poor security returns preceding the earnings year; (4) the inability of analysts to incorporate the differential persistence of the accruals and cash flow components of prior-year earnings; (5) a tendency for tail-area earning growth rate forecasts to be too extreme; and (6) the relations between subsequent-year earnings and other firm-specific characteristics that are associated with subsequent security price adjustments. Incorporating these predictable forecast errors may be able to help us to differentiate the two competing explanations for the delayed security price adjustments to price-scaled analyst forecasts of earnings.

3. Hypothesis Development

The approach that this paper uses to differentiate the two competing explanations is to evaluate the relation of adjusted analyst forecasts of earnings, by incorporating predictable analyst forecast errors, to subsequent securities. The arguments of which the adjustment of analyst forecasts may help to differentiate the two competing explanations are articulated below.

Under market mispricing, the ratio, FY1/P, may be expressed as ($F^T/P + n/P$), where F^T/P is the "true" anomaly variable and n/P is noise in the expectation proxy. Therefore, the measurement error in FY1/P due to n/P would bias hedge portfolio returns based on F^T/P toward zero. If the measurement error can be reduced, then the adjusted price-scaled analysts' forecasts should generate higher hedge portfolio returns. If the adjustments of analysts' forecasts increase the accuracy of analysts' forecasts, as more accurate analysts' forecasts contain less measurement errors, n/P, the adjusted analysts' forecasts are expected to generate more significant hedge portfolio returns under market mispricing. The hypothesis is stated as the null form:

$H1_0$: If the adjusted analyst forecasts are more accurate than the unadjusted analyst forecasts, the subsequent hedge portfolio returns based on the adjusted analyst forecasts are expected to be same under the hypothesis of market mispricing.

4. Data and Variable Definitions

The analysis uses all December fiscal-year firms available from the intersection of Compustat (including both the Active and Research firms), CRSP, and I/B/E/S, with sufficient information to measure the following variables (firm-specific subscripts are omitted). There are 20,702 firm-years over 1983–2003.

FE_{t+1}	Actual earnings (#123) of year $t+1$ minus I/B/E/S mean consensus forecast of earnings for year $t+1$, reported in May of year $t+1$, scaled by market value of equity at the end of year t.[2]
FE_t	Actual earnings (#123) of year t minus I/B/E/S mean consensus forecast of earnings for year t, reported in May of year t, scaled by market value of equity at the end of year t.
$EPSCHG_t$	Earnings (#123) of year t minus earnings (#123) of year t_1, scaled by market value of equity at the end of year t.
SAR_t	One-year size-adjusted security returns, measured as the raw security returns from CRSP cumulated over 12 months beginning June 1 of t, less the corresponding mean returns for all the sample firms in the same market capitalization decile at the start of year t.
$INVENT_t$[3,4]	$\Delta\text{Inventory}_t - \Delta\text{Sales}_t$

[2]ELP use analysts' forecasts of earnings made in March in their main results. However, footnote 3 indicates that their results are robust to the use of I/B/E/S consensus forecasts reported in April or May. Our main results are robust to the use of March analyst forecasts of earnings. In addition, using May analyst forecasts help to enlarge the sample size. Lastly, the main results remain robust using I/B/E/S median consensus forecast of earnings.

[3]According to Abarbanell and Bushee (1997), the Δ operator represents a percentage change in the variable based on a 2-year average expectation model, e.g., $\Delta\text{Sales} = [\text{Sales}_t - E(\text{Sales}_t)]/E(\text{Sales}_t)$, where $E(\text{Sales}_t) = (\text{Sales}_{t-1} + \text{Sales}_{t-2})/2$.

[4]Only the significant variables from Abarbanell and Bushee (1997) are included in this study.

GM_t^2	$\Delta Sales_t - \Delta GrossMargin_t$
$EQ1_t$	One if FIFO and zero otherwise.
$EQ2_t$	One if LIFO and zero otherwise.
LF_t	$(Sales_{t-1}/\#Employees_{t-1} - Sales_t/\#Employees_t)/(Sales_{t-1}/\#Employees_{t-1})$
LTG_t	I/B/E/S mean consensus forecast of long-term growth rate, reported in May of year t.[5]
$DLTG_t$	A scaled-decile variable with the range [0,1] annually converted from LTG_t.
$WCACC_t$	Working capital accruals of period t, deflated by market value of equity at the end of period t. Working capital accruals are defined as follows using the Statement of Cash Flows: Increase in Accounts Receivable (#302) + Increase in Inventory (#303) + Decrease in Accounts Payable and Accrued Liabilities (#304) + Decrease in Accrued Income Taxes (#305) + Increase (Decrease) in Assets (Liabilities) − Other (#307). When the components of working capital accruals are not reported in the Statement of Cash Flows, working capital accruals are defined using the Balance Sheet.[6]
$DWCACC_t$	A scaled-decile variable with the range [0,1] annually converted from $WCACC_t$.
$(FY1/P)_{t+1}$	I/B/E/S mean consensus forecast of earnings per share for year $t+1$, reported in May of year $t+1$, scaled by share price at the end of year t.
SAR_{t+1}	One-year size-adjusted security returns, measured as the raw security returns from CRSP cumulated over 12 months beginning June 1 of $t+1$, less the corresponding mean returns for all the sample firms in the same market capitalization decile at the start of year $t+1$.

[5]When LTG is missing, the implicit growth rate imbedded in FY1 and FY2, (FY2/FY1-1), is used instead. The main inferences remain robust without this replacement. I/B/E/S long-term growth rate forecasts are available from 1981 to the most current year.

[6]$Accrual_t$ is defined as working capital accruals based on the finding from Bradshaw *et al.* (2001) that errors in investors' expectations of mean reversion in accruals mainly derive from working capital accruals rather than from total accruals.

To mitigate the potential impact of outliers, the 1% top and bottom cases of the pooled distribution of price-scaled earnings per share and 1-year size-adjusted returns are deleted. All the observations with the values of price-scaled analysts' forecast errors or changes in actual earnings exceeding 1 are deleted. Observations with share price of under \$1 as of the end of June in year $t + 1$ are deleted. Lastly, as extremely low values of inventory, sales, and gross margin may lead to unreasonably high values of the fundamental signals: INVENT_t and GM_t, observation with the values of INVENT_t and GM_t greater than 50 are deleted. These outlier screens result in 18,782 firm-years from 1983 to 2003.

5. Methodology and Empirical Results

Table 1 provides descriptive statistics for the main variables. This table presents the means, the medians, and the standard deviations (SD) of the pooled distribution for 1983–2003. Consistent with prior literature, analysts' forecast errors (actual less forecasted earnings), for both period $t+1$ and period t, are negative, reflecting analysts' overall optimistic bias. Table 1 also includes the Spearman correlations between the main variables. All the associations of subsequent analysts' forecast errors with the attributes are directly consistent

Table 1a. Descriptive statistics for analysis variables ($n = 18,782$ firm-years, 1983–2003).

	Mean	Median	SD
FE_{t+1}	−0.025	−0.004	0.110
FE_t	−0.034	−0.003	0.180
EPSCHG_t	−0.012	0.006	0.191
SAR_t	0.077	−0.004	0.664
INVENT_t	−0.091	−0.069	1.862
GM_t	−0.140	−0.003	14.647
EQ1_t	0.347	0.000	0.476
EQ2_t	0.178	0.000	0.382
LF_t	−0.164	−0.048	3.364
LTG_t	0.159	0.147	0.867
DLTG_t	0.482	0.500	0.284
WCACC_t	0.000	0.000	0.000
DWCACC_t	0.454	0.500	0.276
$(\text{FY1/P})_{t+1}$	0.068	0.065	0.237
SAR_{t+1}	0.002	−0.016	0.374

Table 1b. Spearman correlation.

	FE_{t+1}	FE_t	$EPSCHG_t$	SAR_t	$INVENT_t$	GM_t	$EQ1_t$	$EQ2_t$	LF_t	LTG_t	$DLTG_t$	$WCACC_t$	$DWCACC_t$	$(FY1/P)_{t+1}$	SAR_{t+1}
FE_{t+1}	1.00														
FE_t	0.31	1.00													
$EPSCHG_t$	0.15	0.64	1.00												
SAR_t	0.37	0.25	0.17	1.00											
$INVENT_t$	−0.03	−0.10	−0.13	−0.04	1.00										
GM_t	−0.09	−0.29	−0.32	−0.10	$0.01^{\#}$	1.00									
$EQ1_t$	−0.07	−0.06	$0.00^{\#}$	−0.02	0.12	$0.01^{\#}$	1.00								
$EQ2_t$	$0.01^{\#}$	$0.01^{\#}$	$0.01^{\#}$	$0.01^{\#}$	0.11	$-0.01^{\#}$	−0.34	1.00							
LF_t	−0.03	−0.14	−0.24	−0.02	0.29	$0.00^{\#}$	$0.01^{\#}$	0.03	1.00						
LTG_t	−0.07	$-0.01^{\#}$	0.07	$0.01^{\#}$	−0.14	−0.05	0.21	−0.19	−0.06	1.00					
$DLTG_t$	−0.06	$-0.01^{\#}$	0.08	$0.01^{\#}$	−0.12	−0.04	0.21	−0.16	−0.07	0.97	1.00				
$WCACC_t$	−0.07	0.07	0.09	−0.06	0.15	$-0.01^{\#}$	0.05	0.02	−0.02	0.07	0.08	1.00			
$DWCACC_t$	−0.07	0.06	0.08	−0.07	0.15	$-0.00^{\#}$	0.06	$0.02^{\#}$	$-0.01^{\#}$	0.08	0.08	0.98	1.00		
$(FY1/P)_{t+1}$	0.11	0.20	0.20	0.13	0.08	−0.07	−0.09	0.14	$-0.01^{\#}$	−0.29	−0.26	0.11	0.09	1.00	
SAR_{t+1}	0.26	0.02	0.02	0.04	0.02	$-0.01^{\#}$	−0.04	0.03	0.02	−0.09	−0.09	−0.04	−0.04	0.11	1.00

(*Continued*)

Table 1b. (*Continued*)

Notes:

FE_{t+1}: Actual earnings (#123) of year $t+1$ minus I/B/E/S mean consensus forecast of earnings for year $t+1$, reported in May of year $t+1$, scaled by market value of equity at the end of year t.

FE_t: Actual earnings (#123) of year t minus I/B/E/S mean consensus forecast of earnings for year t, reported in May of year t, scaled by market value of equity at the end of year t.

$EPSCHG_t$: Earnings (#123) of year t minus earnings (#123) of year t_1, scaled by market value of equity at the end of year t.

SAR_t: One-year size-adjusted security returns, measured as the raw security returns from CRSP cumulated over 12 months beginning June 1 of t, less the corresponding mean returns for all the sample firms in the same market capitalization decile at the start of year t.

$INVENT^{\&}_t$: $\Delta \text{Inventory}_t - \Delta \text{Sales}_t$.

$GM^{\&}_t$ $\Delta \text{Sales}_t - \Delta \text{Gross Margin}_t$.

$EQ1_t$: One if FIFO and zero otherwise.

$EQ2_t$: One if LIFO and zero otherwise.

LF_t: $(\text{Sales}_{t-1}/\#\text{Employees}_{t-1} - \text{Sales}_t/\#\text{Employees}_t)/(\text{Sales}_{t-1}/\#\text{Employees}_{t-1})$.

LTG_t: I/B/E/S mean consensus forecast of long-term growth rate, reported in May of year t.

$DLTG_t$: A scaled-decile variable with the range [0,1] annually converted from LTG_t.

$WCACC_t$ Working capital accruals of period t, deflated by market value of equity at the end of period t. Working capital accruals are defined as follows using the Statement of Cash Flows: Increase in Accounts Receivable (#302) + Increase in Inventory (#303) + Decrease in Accounts Payable and Accrued Liabilities (#304) + Decrease in Accrued Income Taxes (#305) + Increase (Decrease) in Assets (Liabilities) − Other (#307). When the components of working capital accruals are not reported in the Statement of Cash Flows, working capital accruals are defined using the Balance Sheet.

$DWCACC_t$: A scaled-decile variable with the range [0,1] annually converted from $WCACC_t$.

$(FY1/P)_{t+1}$: I/B/E/S mean consensus forecast of earnings per share for year $t+1$, reported in May of year $t+1$, scaled by share price at the end of year t.

SAR_{t+1}: One-year size-adjusted security returns, measured as the raw security returns from CRSP cumulated over 12 months beginning June 1 of $t+1$, less the corresponding mean returns for all the sample firms in the same market capitalization decile at the start of year $t+1$.

[&] According to Abarbanell and Bushee (1997), the Δ operator represents a percentage change in the variable based on a 2-year average expectation model, e.g.,
$\Delta \text{Sales} = [\text{Sales}_t - E(\text{Sales}_t)]/E(\text{Sales}_t)$, where $E(\text{Sales}_t) = (\text{Sales}_{t-1} + \text{Sales}_{t-2})/2$.

[#] These Spearman correlations are not significant at 1% level. All the others are significant at 1% level.

with the literature. For example, subsequent analysts' forecast errors are positively related to previous analysts' forecast errors, consistent with Ali *et al.* (1992). Analysts' forecast errors are negatively associated with prior period long-term growth rate forecasts, consistent with La Porta (1996). The correlations between SAR_{t+1} and the attributes are also consistent with the prior studies. For example, the Spearman correlation between SAR_{t+1} and SAR_t is 0.04, consistent with Jegadeesh and Titman's (1993) finding that short-term returns tend to persist in the successive years. The negative association between SAR_{t+1} and LTG_t is consistent with La Porta (1996). As expected, FE_t and EPSCHG_t are highly correlated. The Spearman correlation is 0.64.

5.1. *Adjustments of analysts' forecasts*

This section describes the adjustment of analysts' forecasts by previous error patterns, and evaluates the accuracy of the adjusted analysts' forecasts. To adjust the analysts' forecasts by previous error patterns, the historical relations between the documented attributes and the subsequent forecast errors need to be examined. Based on prior literature (e.g., Abarbanell, 1991; Elgers and Lo, 1994; Klein, 1990), the inability or unwillingness of analysts to incorporate earning-relevant information from prior returns or earnings information is asymmetric. The biases are stronger for firms with prior poor performance than for those with prior good performance. Therefore, the regression analysis will differentiate the firms with good performance from those with bad performance. Actual analysts' forecast errors are regressed annually on the pertinent variables:

$$\text{FE}_t = \alpha_0 + \sum_{i=1}^{12} \alpha_i X_{i,t-1} \tag{1}$$

where the $X_{i,t}$ includes the variables that prior studies have shown to be systematically associated with forecast errors, as discussed above and described in the footnote of Table 2.

The coefficients, α_0 and α_i, are estimated by ordinary least-square (OLS). The statistical tests are based on intertemporal tests, using the mean and the standard deviation of the 21 annual coefficients.

Panel A of Table 2 replicates the pertinent results of prior studies. Consistent with Ali *et al.* (1992), there is a positive serial correlation between errors in analysts' forecasts of earnings. The coefficient of FE_t is 0.291 with

Table 2. Predictable error patterns in analysts' forecasts ($n = 18,782$ firm-years, 1983–2003).

Coefficients	Intercept	FE_t	$EPSCHG\_P_t$	$EPSCHG\_G_t$	SAR_P_t	SAR_G_t	$INVENT_t$	GM_t	$EQ1_t$	$EQ2_t$	LF_t	$DLTG_t$	$DAccrual_t$	Adj. R^2
Panel A: Replication														
Predicted sign		+	+	+	+	+	−	−	−	−		−	−	
Ali *et al.* (1992)														
Mean	−0.014	0.291												0.240
t-statistics	−4.91***	6.72***												
Elgers and Lo (1994)														
Mean	−0.006		0.149	0.101	0.123	0.013								0.263
t-statistics	−3.20***		3.62***	3.48***	13.26***	4.04***								
La Porta (1996)														
Mean	−0.018											−0.008		0.001
t-statistics	−6.73***											−2.18**		
Bradshaw *et al.* (2001)														
Mean	−0.019												−0.008	0.013
t-statistics	−4.77***												−1.05	
Abarbanell and Bushee (1997)														
Mean	−0.018						−0.004	0.005	−0.012	0.000	−0.002			0.024
t-statistics	−4.53***						−1.37*	0.51	−3.55***	0.11	−0.68			
Panel B: Equation (1) $FE_t = \alpha_0 + \sum_{i=1}^{12} \alpha_i X_{i,t-1}$														
Mean	0.007	0.314	−0.043	−0.065	0.103	0.014	0.000	0.005	−0.005	−0.001	0.000	−0.003	−0.015	0.379
t-statistics	2.78**	5.13***	−0.92	−2.21**	11.23***	4.04***	−0.38	1.17	−1.84**	−0.52	−0.07	−0.87	−3.44***	

(*Continued*)

Table 2. (*Continued*)

Notes:

FE_{t+1}: Actual earnings (#123) of year $t+1$ minus I/B/E/S mean consensus forecast of earnings for year $t+1$, reported in May of year $t+1$, scaled by market value of equity at the end of year t.

FE_t: Actual earnings (#123) of year t minus I/B/E/S mean consensus forecast of earnings for year t, reported in May of year t, scaled by market value of equity at the end of year t.

$EPSCHG\_P_t$: $EPSCHG_t$ for firms with poor performance and zero otherwise.

EPSCHG_G: $EPSCHG_t$ for firms with good performance and zero otherwise, where the negative (positive) prior returns are defined poor (good) prior performance.

$EPSCHG_t$: Earnings (#123) of year t minus earnings (#123) of year t_1, scaled by market value of equity at the end of year t.

SAR_P_t: SAR_t for firms with poor performance and zero otherwise.

SAR_G_t: SAR_t for firms with good performance and zero otherwise, where the negative (positive) prior returns are defined poor (good) prior performance.

SAR_t: One-year size-adjusted security returns, measured as the raw security returns from CRSP cumulated over 12 months beginning June 1 of t, less the corresponding mean returns for all the sample firms in the same market capitalization decile at the start of year t.

$INVENT_t^*$: $\Delta Inventory_t - \Delta Sales_t$.

GM_t^*: $\Delta Sales_t - \Delta Gross\ Margin_t$.

$EQ1_t$: One if FIFO and zero otherwise.

$EQ2_t$: One if LIFO and zero otherwise.

LF_t: $(Sales_{t-1}/\#Employees_{t-1} - Sales_t/\#Employees_t)/(Sales_{t-1}/\#Employees_{t-1})$.

LTG_t: I/B/E/S mean consensus forecast of long-term growth rate, reported in May of year t.

$DLTG_t$: A scaled-decile variable with the range [0,1] annually converted from LTG_t.

$WCACC_t$: Working capital accruals of period t, deflated by market value of equity at the end of period t. Working capital accruals are defined as follows using the Statement of Cash Flows: Increase in Accounts Receivable (#302) + Increase in Inventory (#303) + Decrease in Accounts Payable and Accrued Liabilities (#304) + Decrease in Accrued Income Taxes (#305) + Increase (Decrease) in Assets (Liabilities) − Other (#307). When the components of working capital accruals are not reported in the Statement of Cash Flows, working capital accruals are defined using the Balance Sheet.

$DWCACC_t$: A scaled-decile variable with the range [0,1] annually converted from $WCACC_t$.

If we have directional expectations, the statistical tests are one-tailed. Otherwise, the statistical tests are two-tailed.

* Significant at a probablity below 0.10 based on intertemporal tests.

** Significant at a probablity below 0.05 based on intertemporal tests.

*** Significant at a probability below 0.01 based on intertemporal tests. Each annual measure is treated as a single observation, and statistical tests are based on the means and standard deviations of the annual observations (Bernard, 1987).

t-statistics of 6.72. Consistent with Elgers and Lo (1994), changes in prior earnings and prior returns are positively associated with analysts' forecast errors and analysts' ability or willingness to impound prior returns and earnings in their earning forecasts is asymmetric. The negative coefficient on the growth variable reflects analysts' bias in their long-term growth rate forecasts (La Porta, 1996). The estimated coefficient for accruals descriptively replicates the finding of prior studies that analysts do not appear to fully understand the high reversion property of accruals over successive years (Bradshaw *et al.*, 2001). Lastly, Panel A of Table 2 also shows that the coefficient estimates for the fundamental signals in Abarbanell and Bushee (1997) are negative, except for GM_t.[7]

Panel B of Table 2 presents the multivariate analysis including all the attributes that are correlated to FE_{t+1}. The inferences regarding the coefficient estimates remain the same except the following. The coefficient estimates of $EPSCHG\_P_t$ and $EPSCHG\_G_t$ are no longer consistent with the expectation due to the high correlation between FE_t and $EPSCHG_t$. A noteworthy evidence in Panel B of Table 2 is a significant increase in the adjusted R^2. When all the attributes are included in the regression as explanatory variables, the adjusted R^2 is significantly increased to 37.9%, compared to the highest adjusted R^2 from Panel A of Table 2, 26.3%. The cross-sectional variation in the subsequent analyst forecast errors can be explained by the list of the attributes by 37.9%. This shows that the adjustments of forecasts by all the attributes are superior to the adjustments based on a subset of the attributes used in prior studies. It highlights the importance of including all the attributes to adjust for analyst' forecast errors.

The next step in the analysis is to adjust the analysts' forecasts using the average of the coefficient estimates estimated by Equation (1) from 3 years prior to the analysis year. For example, the forecasts for the year 1995 are adjusted based on the average of the annual estimates of Equation (1) from 1992 to 1994. Because the adjustment of forecasts requires estimates of the parameters from 3 prior years, the analysis period covers 1986–2003. Adjusted

[7] Abarbanell and Bushee (1997) defines EQ_t to be 1 if FIFO and 0 if LIFO. However, this requirement significantly reduces the sample size due to the fact that a lot of firms use other inventory methods. Therefore, two variables $EQ1_t$ and $EQ2_t$ are defined. $EQ1_t$ equals to 1 if FIFO and 0 otherwise and $EQ2_t$ equals to 1 if LIFO and 0 otherwise. The difference in the coefficients of these two variables ($-0.012 - 0.000 = -0.012$) can be interpreted as the difference in the impact of using FIFO vs. LIFO on errors in analyst forecast earnings.

analysts' forecasts are computed in the following fashion:

$$\mathrm{FY1\_adj}_{t+1} = \mathrm{FY1}_{t+1} + \hat{\alpha}_0 + \sum_{i=1}^{12} \hat{\alpha}_i X_{i,t-1} \tag{2}$$

where $\mathrm{FY1\_adj}_{t+1}$ is the *ex ante* adjusted forecasts in year $t+1$; F_{t+1} is mean analysts' forecasts issued in May of year $t+1$; $\hat{\alpha}_i$ is the average of the annual coefficient estimates of α_i from Equation (1) from 3 years prior to the analysis year; and the $X_{i,t}$ are the variables included in Equation (1).

5.2. *Accuracy of the adjusted analysts' forecasts*

Although analysts' forecasts are adjusted by predictable error patterns, the impact upon the accuracy of analysts' forecasts is yet to be determined. This study uses two measures to evaluate the improvement in the accuracy of adjusted analysts' forecasts of earnings: bias and mean square errors (MSE). Bias is defined within each year as actual earnings minus analysts' forecasts of earnings of year $t+1$ scaled by the market value of equity at the end of year t. MSE is defined within each year as the squared difference between actual earnings and analysts' forecasts of earnings of year $t+1$ scaled by the market value of equity at the end of year t. The statistical tests are based on intertemporal tests, using the mean and the standard deviation of the 18 annual coefficients.

Table 3 reports the bias and MSE for the unadjusted and adjusted analysts' forecasts. MSE is multiplied by 100 to suppress the leading zeros. The results show that the forecast adjustments provide a significant reduction in the bias as well as the MSE. The negative bias before adjustments reflects analysts' overall optimistic bias. Yet, after the adjustments, analysts' forecasts of earnings no longer show this overall optimistic bias. MSE is decreased to 0.942 from 1.110. The error reductions are significant at 1% level. In addition, the adjusted analysts' forecasts of earnings produce lower biases for all the sample years and lower MSE for 15 years out of 18 sample years.

5.3. *Differentiation of market mispricing vs. omitted risk factor for ELP*

ELP document a significantly positive relationship between price-scaled analysts' forecasts, FY1/P, and subsequent security returns. They interpret their

Table 3. Accuracy of analysts' forecasts ($n = 18,626$ firm-years, 1986–2003).

	$(\text{FY1/P})_{t+1}$		$(\text{FY1\_adj/P})_{t+1}$		Difference (%)
	Bias	MSE × 10^2	Bias	MSE × 10^2	
Intertemporal mean	−0.023	1.110	0.001	0.942	−239.1%
Error reduction			0.024	0.168	11.3%
t-statistics for error reduction			9.370***	3.410***	
Number of years with lower error			18/18***	15/18***	

Notes:

$(\text{FY1/P})_{t+1}$: I/B/E/S mean consensus forecast of earnings per share for year $t+1$, reported in May of year $t+1$, scaled by share price at the end of year t.

$(\text{FY1\_adj/P})_{t+1}$: Adjusted I/B/E/S mean consensus forecast of earnings per share for year $t+1$, reported in May of year $t+1$, scaled by share price at the end of year t.

Adjusted analysts' forecasts are computed in the following fashion:

$$\text{FY1\_adj}_{t+1} = \text{FY1}_{t+1} + \hat{\alpha}_0 + \sum_{i=1}^{12} \hat{\alpha}_i X_{i,t-1}, \tag{2}$$

where $\text{FY1_adj}_{t+1}$is the ex ante-adjusted forecasts in year $t+1$; F_{t+1} is mean analysts' forecasts issued in May of year $t+1$; $\hat{\alpha}_i$ is the average of the annual coefficient estimates of α_i from Equation (1) from 3 years prior to the analysis year; and the $X_{i,t}$ are the variables included in Equation (1).

Bias: Actual earnings minus analysts' forecasts of earnings of year $t+1$ scaled by the market value of equity at the end of year t within each year.

MSE: The squared difference between actual earnings and analysts' forecasts of earnings of year $t+1$ scaled by the market value of equity at the end of year t within each year.

* Significant at a probablity below 0.10 based on two-tailed intertemporal tests.

** Significant at a probablity below 0.05 based on two-tailed intertemporal tests.

*** Significant at a probability below 0.01 based on two-tailed intertemporal tests. Each annual measure is treated as a single observation, and statistical tests are based on the means and standard deviations of the annual observations (Bernard, 1987).

results as investors underestimating the value-relevant analysts' forecasts. As stated above, the measurement error in FY1/P would bias hedge portfolio returns toward zero. If the measurement error can be reduced, then the adjusted price-scaled analysts' forecasts should generate higher hedge portfolio returns than FY1/P. The adjustments of analysts' forecasts based on Equation (2) increase the accuracy of analysts' forecasts as shown in Table 3. Because more accurate analysts' forecasts contain less measurement errors, the adjusted analysts' forecasts are expected to generate more significant hedge portfolio returns under the explanation of market mispricing.

Table 4 reports the hedge portfolio returns based on price-scaled analysts' forecasts before and after adjustments. It followed ELP's partitioning of the

Table 4. Differentiation of market mispricing vs. omitted risk factor for ELP ($n = 18,626$ firm-years, 1986–2003).

	$(\mathrm{FY1/P})_{t+1}$	$(\mathrm{FY1\_adj/P})_{t+1}$
Panel A: Lower analyst coverage ($n = 8686$)		
Bottom quintile	−0.071	−0.059
Top quintile	0.055	0.054
Hedge t-statistics	0.126	0.112
	(3.16)***	(3.87)***
Years positive	16/18***	15/18***
Difference in hedge t-statistics	−0.013	
	(−0.74)	
Panel B: Higher analyst coverage ($n = 9940$)		
Bottom quintile	−0.033	0.002
Top quintile	0.028	0.022
Hedge (t-statistics)	0.060	0.020
	(1.04)	(0.50)
Years positive	9/18	9/18
Difference in hedge (t-statistics)	−0.040	
	(−1.34)	

Notes:

$(\mathrm{FY1/P})_{t+1}$: I/B/E/S mean consensus forecast of earnings per share for year $t+1$, reported in May of year $t+1$, scaled by share price at the end of year t.

$(\mathrm{FY1\_adj/P})_{t+1}$: Adjusted I/B/E/S mean consensus forecast of earnings per share for year $t+1$, reported in May of year $t+1$, scaled by share price at the end of year t. The adjustment of analysts' forecasts of earnings is explained in detail in Section 4.1.

Adjusted analysts' forecasts are computed in the following fashion:

$$\mathrm{FY1\_adj}_{t+1} = \mathrm{FY1}_{t+1} + \hat{\alpha}_0 + \sum_{i=1}^{12} \hat{\alpha}_i X_{i,t-1}, \tag{2}$$

where $\mathrm{FY1\_adj}_{t+1}$ is the ex ante-adjusted forecasts in year $t+1$; F_{t+1} is mean analysts' forecasts issued in May of year $t+1$; $\hat{\alpha}_i$ is the average of the annual coefficient estimates of α_i from Equation (1) from 3 years prior to the analysis year; and the $X_{i,t}$ are the variables included in Equation (1).

* Significant at a probablity below 0.10 based on two-tailed intertemporal tests.

** Significant at a probablity below 0.05 based on two-tailed intertemporal tests.

*** Significant at a probability below 0.01 based on two-tailed intertemporal tests. Each annual measure is treated as a single observation, and statistical tests are based on the means and standard deviations of the annual observations (Bernard, 1987).

sample into two groups, based on the extent of analysts' coverage (number of analysts with May annual earnings forecasts for a given firm). The cases below (above) the yearly cross-sectional medians of the number of analysts' forecasts are classified as the lower (higher) analyst coverage group. The hedge portfolio returns to analysts' forecasts are expected to be significantly greater for markets with a less-rich information environment (the lower analyst coverage group) than for markets with a stronger information environment (the higher analyst coverage group). To compute the hedge portfolio returns, all the observations are sorted annually into quintiles based upon the magnitudes of the price-scaled analysts' forecasts. The hedge portfolio returns equal to the annual means of size-adjusted returns for the firms in the top quintile minus those in the bottom quintile.

Consistent with ELP, Table 4 shows that for the lower analyst coverage group, the size-adjusted hedge portfolio return is 12.6% ($t-$statistics $= 3.16$), whereas for the higher analyst coverage group, the hedge portfolio return is only 6% ($t -$ statistics $= 1.04$). Moreover, Table 4 shows the hedge portfolio returns based on the adjusted analysts' forecasts of earnings. For the lower (higher) analyst coverage group, the returns based on adjusted analysts' forecasts are 11.2% (2%). The change in hedge portfolio returns is -1.3% (-4%) for the lower (higher) analyst coverage. These changes are both statistically and economically insignificant, that is the null hypothesis $\mathrm{H1}_0$ cannot be rejected. These results show that the hedge portfolio returns are the essentially same, no matter whether the less or more accurate analysts' forecasts are used. The inability of generating greater hedge portfolio returns using the more accurate forecasts with less measurement errors does not support the explanation of market mispricing. Therefore, it provides indirect support for the explanation of omitted risk factors.

6. Conclusion

This paper reexamines evidence of delayed security returns associated with early-in-the-year analysts' forecasts of annual earnings investigated in ELP. A salient feature of this study is the use of market values as the method of scaling the forecast variables. Market value scaling, however, introduces the possibility that a given firm-specific attribute, scaled by market value, serves as a "yield surrogate," i.e., a proxy for some unidentified pertinent risk factors.

This study develops and applies a new approach to evaluating the likelihood that omitted risk factors or market mispricing underlies the anomalous security returns reported in recent research by ELP. This approach is motivated by a body of prior research that demonstrates that errors in analysts' earning forecasts are to some extent predictable based on firm-specific characteristics, and indicates that the predictable errors may be employed to adjust and improve the accuracy of the analysts' earning forecasts. Under the hypothesis of market mispricing, greater hedge portfolio returns to be expected when the analysts' forecasts are adjusted by the predictable error patterns. The argument is based on the assumption that the adjusted analysts' forecasts contain less measurement errors and thus reduce the noise embedded in the relation tested in ELP.

The results of this study indicate that the main results in ELP are replicated in the sample, which extends over a later time period. In addition, the adjustments that applied here to extant analysts' forecasts improve the predictive accuracy of the analysts' forecasts. These adjusted forecasts, however, do *not* enable improvements in the abnormal security returns resulting from the trading strategies employed in ELP. The inability to show improvements is inconsistent with the interpretation that the predictable security returns documented in the study are due to market mispricing. Therefore, it provides indirect support for the alternative interpretation of omitted risk proxies.

Similar method that aims to reduce the measurement errors in the variable of interest may be applied to other papers in the anomaly area in order to differentiate the two competing explanations—omitted risk proxies vs. market mispricing. For example, Frankel and Lee (1998) have shown a significant association between price-scaled intrinsic firm value and subsequent security returns, where the intrinsic value estimates rely upon several variables as inputs to the residual income valuation model. Researchers may be able to differentiate the two competing explanations for the results reported in Frankel and Lee (1998) using variables that contain less measurement errors as inputs to the residual income model.

Acknowledgments

The author appreciates the invaluable comments from Ben Branch, Bernie Morzuch, and Sue Porter as well as the workshop participants at the 2004

Northeast American Accounting Association meeting. The author thanks Pieter Elgers for his substantive suggestions. She thanks the University of New Hampshire for providing research support. The author acknowledges I/B/E/S International, Inc. for providing earnings per share forecast data available through the Institutional Brokers Estimate System.

References

Abarbanell, J., "Do Analysts' Earnings Forecasts Incorporate Information in Prior Stock Price Changes?" *Journal of Accounting and Economics* 14, 147–65 (1991).

Abarbanell, J. and B. Bushee, "Fundamental Analysis, Future Earnings, and Stock Prices." *Journal of Accounting Research* 35, 1–24 (1997).

Ali, A., A. Klein and J. Rosenfeld, "'Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS." *The Accounting Review* 67, 183–198 (1992).

Barth, M. and A. Hutton, "Information Intermediaries and the Pricing of Accruals." *Review of Accounting Studies* 9, 55–96 (2004).

Barth, M., R. Kasznik and M. McNichols, "Analyst Coverage and Intagible Assets." *Journal of Accounting Research* 39, 1–34 (2001).

Basu, S., "Investment Performance of Common Stocks in Relation to Price-Earnings Ratios." *Journal of Finance* 32, 663–682 (1977).

Bernard, V., "Cross-Sectional Dependence and Problems in Inference in Market-Based Accounting Research." *Journal of Accounting Research* 25, 1–48 (1987).

Bhushan, R., "An Information Efficiency Perspective on the Post-Earnings-Announcement Drift." *Journal of Accounting and Economics* 18, 45–65 (1994).

Bradshaw, M., S. Richardson and R. Sloan, "Do Analysts and Auditors Use Information in Accruals?" *Journal of Accounting Research* 39, 45–74 (2001).

Elgers, P. and M. Lo, "Reductions in Analysts' Annual Earnings Forecast Errors Using Information in Prior Earnings and Security Returns." *Journal of Accounting Research* 32, 290–303 (1994).

Elgers, P. and D. Murray, "The Relative and Complementary Performance of Analysts and Security-Price-Based Measures of Expected Earnings." *Journal of Accounting and Economics* 15, 303–316 (1992).

Elgers, P. T., M. H. Lo and R. Pfeiffer, "Delayed Security Price Adjustments to Financial Analysts' Forecasts of Annual Earnings." *The Accounting Review* 76, 613–632 (2001).

Frankel, R. and C. M. C. Lee, "Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns." *Journal of Accounting and Economics* 25, 283–319 (1998).

Hong, H., T. Lim and J. Stein, "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies." *Journal of Finance* 55, 265–295 (2000).

Jegadeesh, N. and S. Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48, 65–91 (1993).

Klein, A., "A Direct Test of the Cognitive Bias Theory of Share Price Reversals." *Journal of Accounting and Economics* 13, 155–166 (1990).

La Porta, R., "Expectations and the Cross-Section of Stock Returns." *Journal of Finance* 51, 1715–1742 (1996).

Lev, B. and S. R. Thiagarajan, "Fundamental Information Analysis." *Journal of Accounting Research* 31, 190–215 (1993).

Lobo, G. and R. Nair, "Combining Judgmental and Statistical Forecasts: An Application to Earnings Forecasts." *Decision Sciences* 21, 446–460 (1990).

Sloan, R., "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?" *The Accounting Review* 71, 289–315 (1996).

Walther, B., "Investor Sophistication and Market Earnings Expectations." *Journal of Accounting Research* 35, 157–117 (1997).

This page intentionally left blank

Chapter 8

On Simple Binomial Approximations for Two Variable Functions in Finance Applications

Hemantha S. B. Herath*
Brock University, Canada

Pranesh Kumar
University of Northern British Columbia, Canada

We extend the volatility stabilization transformation technique to two correlated Brownian motions. This technique allows to construct a computationally simple binomial tree and to obtain the probabilities for the up and down movements. We derive the expressions for correlated geometric Brownian motions by considering two variable functions. We discuss particular functions of two variables, which are commonly employed in finance. Further, we simulate results for the numerical accuracy of the approximations using an exchange option.

Keywords: Option pricing; Correlated assets; binomial lattice; volatility transform; binomial

1. Introduction

Nelson and Ramaswamy (1990) used an elegant instantaneous volatility stabilization transformation to approximate diffusions commonly used in finance such as the Ornstein–Uhlenbeck (OU or mean reversion) process and the constant elasticity of variance (CEV) to a computationally simple binomial lattice. Although binomial approximations for these types of diffusions may exist, the binomial tree structures may not necessarily recombine. Such binomial tree structures are computationally complex because the number of nodes in the tree doubles at each time step. The idea is to obtain a computationally simple binomial tree structure where an up move followed by a down move causes a displacement which is equal to a displacement caused by a down move that is followed by an up move. This objective is achieved by employing a transformation that makes the heteroscedastic process a homoscedastic process, that is, employing a transformation that makes the instantaneous volatility of the transformed process constant.

*Corresponding author.

In this paper, we extend the volatility stabilization transformation technique for two variable functions. There are numerous situations where two variable functions are commonly encountered when pricing options (Kamrad and Ritchken, 1991; Boyle, 1988; Boyle *et al.*, 1989; Johnson, 1987; Stulz, 1982). These models are useful for valuing real options having multiple sources of uncertainty (Cortazar and Schwartz, 1993). We derive general expressions for correlated geometric Brownian motions. Then we consider some cases, which are commonly employed in finance applications. The paper is organized as follows. Section 2 includes the transformation technique applied by Nelson and Ramaswamy (1990) to a single asset that follows a diffusion process. We extend the transformation technique to two correlated Brownian motions in Section 3. Log-transformed variables are presented in Section 4. Section 5 discuses the numerical accuracy of the approximations using an exchange option. A summary of findings and conclusions is included in Section 6.

2. Nelson–Ramaswamy Instantaneous Volatility Stabilization Transformation

The basic intuition of the instantaneous volatility stabilization transformation is as follows. Consider the stochastic differential equation

$$\mathrm{d}y = \mu(y, t)\mathrm{d}t + \sigma(y, t)\mathrm{d}W \tag{1}$$

where W is the standard Brownian motion, $\mu(y, t)$, $\sigma(y, t) \geq 0$, are the instantaneous drift and standard deviation of y at time t and the initial value y_0 is a constant. The time interval $[0, T]$ is divided into n equal time steps of size $\Delta t = T/n$. The objective is to find a sequence of binomial processes that converge in probability to the process (1) on $[0, T]$.

Nelson and Ramaswamy (1990) consider a transform $X(y, t)$ which is twice differentiable in y and once in t. By Ito's lemma,

$$\begin{aligned}\mathrm{d}X(y, t) = {} & \left(\mu(y, t)\frac{\partial X(y, t)}{\partial y} + \frac{1}{2}\sigma^2(y, t)\frac{\partial^2 X(y, t)}{\partial^2 y} + \frac{\partial X(y, t)}{\partial t}\right)\mathrm{d}t \\ & + \left(\sigma(y, t)\frac{\partial X(y, t)}{\partial y}\right)\mathrm{d}W. \end{aligned} \tag{2}$$

Now make the term

$$\frac{\partial X(y,t)}{\partial y}\sigma(y,t)\,\mathrm{d}W = \mathrm{d}W,$$

in (2) so that the instantaneous volatility of the transformed process $x = X(y,t)$ is constant by taking

$$\sigma(y,t)\frac{\partial X(y,t)}{\partial y} = 1.$$

Then by integrating the above term

$$\int \partial X(y,t) = \int \frac{\partial y}{\sigma\,(y,t)},$$

and substituting y by z, we get

$$X(y,t) = \int^{y} \frac{\mathrm{d}Z}{\sigma(z,t)}, \tag{3}$$

on the support of y. The above transformation allows one to construct a computationally simple binomial tree for the transformed process x where the variance of local change in x is constant at each node. The binomial lattice for the X process can be obtained by defining $X_0 = X(y_0)$ and drawing the X tree as shown in Figure 1.

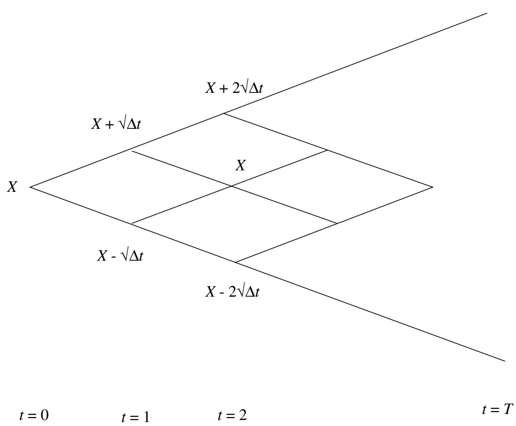


Figure 1. Simple binomial tree for X.

In order to arrive at the binomial process for y, one has to transform from x back to y.Using an inverse transformation defined as

$$Y(x,t) = \{y : X(y,t) = x\} \tag{4}$$

does this. Substituting Equation (4) into Equation (3), we get

$$x = \int^{Y} \frac{\mathrm{d}Z}{\sigma(z,t)}$$

and taking the partial derivative we obtain $\partial y/\partial x = \sigma(y,t)$ which implies that $Y(x,t)$ is weakly monotonic in x for a fixed value of t. The inverse transform in Equation (4) can be used to construct the lattice for y such that the up movement $Y^{+}(x,t)$ and a down movement $Y^{-}(x,t)$ are given by

$$Y^{+}(x,t) = Y\left(x + \sqrt{\Delta t}, t + \Delta t\right), \tag{5}$$

$$Y^{-}(x,t) = Y\left(x - \sqrt{\Delta t}, t + \Delta t\right), \tag{6}$$

and the up movement probability

$$p = \frac{\Delta t \mu(Y(x,t),t) + Y(x,t) - Y^{-}(x,t)}{Y^{+}(x,t) - Y^{-}(x,t)}. \tag{7}$$

The use of the transform, inverse transform, and a feasible probability enables one to construct computationally simple binomial approximation for y. The binomial tree for y is shown in Figure 2.

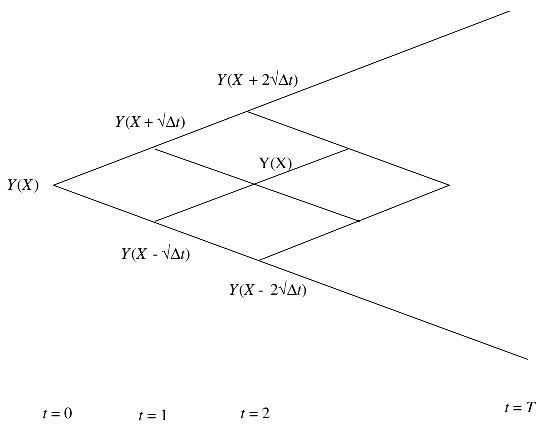


Figure 2. A simple binomial tree for $y = Y(X)$.

3. Transform for Two Variables

We extend the transformation technique for two correlated Brownian motions. Consider a function of two variables S_1 and S_2 each following a geometric Brownian motion where

$$\mathrm{d}S_1 = \mu_1 S_1 \,\mathrm{d}t + \sigma_1 S_1 \,\mathrm{d}W_1, \tag{8a}$$

$$\mathrm{d}S_2 = \mu_2 S_2 \,\mathrm{d}t + \sigma_2 S_2 \,\mathrm{d}W_2, \tag{8b}$$

with $\varepsilon[\mathrm{d}W_1 \mathrm{d}W_2] = \rho \mathrm{d}t$, the correlation between S_1 and S_2.

Consider a general functional form as a power function of S_1 and S_2 given by $F(S_1, S_2, t) = S_1^a S_2^b$, where constants a and b are real number. Since $\partial F/\partial S_1 = aS_1^{a-1}S_2^b$, $\partial F/\partial S_2 = bS_1^a S_2^{b-1}$, $\partial F/\partial t = 0$, $\partial^2 F/\partial S_1^2 = a(a-1)S_1^{a-2}S_2^b$, $\partial^2 F/\partial S_2^2 = b(b-1)S_1^a S_2^{b-2}$, $\partial^2 F/\partial S_1 \partial S_2 = abS_1^{a-1}S_2^{b-1}$, from the Ito lemma

$$\begin{aligned} \mathrm{d}F &= aS_1^{a-1}S_2^b \mathrm{d}S_1 + bS_1^a S_2^{b-1}\mathrm{d}S_2 + \frac{1}{2}a(a-1)S_1^{a-2}S_2^b(\mathrm{d}S_1)^2 \\ &\quad + \frac{1}{2}b(b-1)S_1^a S_2^{b-2}(\mathrm{d}S_2)^2 + abS_1^{a-1}S_2^{b-1}\mathrm{d}S_1\mathrm{d}S_2. \end{aligned} \tag{8c}$$

Substituting for $\mathrm{d}S_1$ and $\mathrm{d}S_2$ and rearranging terms, we obtain

$$\begin{aligned} \mathrm{d}F &= \left[(a\mu_1 + b\mu_2) + \frac{1}{2}\left(a(a-1)\sigma_1^2 + b(b-1)\sigma_2^2 + \rho ab\sigma_1\sigma_2 \right) \right] F \,\mathrm{d}t \\ &\quad + (a\sigma_1 \mathrm{d}W_1 + b\sigma_2 \mathrm{d}W_2)\, F. \end{aligned} \tag{8d}$$

Notice that F follows a geometric Brownian motion with

$$\mu(F,t) = \left[(a\mu_1 + b\mu_2) + \frac{1}{2}\left(a(a-1)\sigma_1^2 + b(b-1)\sigma_2^2 + \rho ab\sigma_1\sigma_2 \right) \right],$$

and the noise term given by $(a\sigma_1 \mathrm{d}W_1 + b\sigma_2 \mathrm{d}W_2)$ where W_i are Wiener processes. Now define $\mathrm{d}W = a\sigma_1 \mathrm{d}W_1 + b\sigma_2 \mathrm{d}W_2$, where $\mathrm{d}W_i \sim N(0, \mathrm{d}t)$ and σ_i are constant, for $i = 1, 2$. Considering the Wiener processes per unit of time $\mathrm{d}W_i \sim N(0, 1)$ and we have

$$\mathrm{d}W \sim N\left(0, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2\right). \tag{8e}$$

The per unit standardized value is

$$\mathrm{d}W_z = \frac{\mathrm{d}W}{\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}}. \tag{8f}$$

Substituting for $\mathrm{d}W$ in Equation (8d), we have

$$\mathrm{d}F = \left[(a\mu_1 + b\mu_2) + \frac{1}{2}\left(a(a-1)\sigma_1^2 + b(b-1)\sigma_2^2 + \rho ab\sigma_1\sigma_2\right)\right] F\,\mathrm{d}t + F\left(\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}\right)\mathrm{d}W_z, \tag{8g}$$

where $dW_z \sim N(0, dt)$.

In order to obtain a computationally simple binomial approximation, we need to make the volatility term constant in Equation (8g). The transform is

$$H(F,t) = \int^F \frac{\mathrm{d}Z}{\sigma(Z,t)} = \frac{\ln F}{\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma}}. \tag{8h}$$

The inverse transformation provides $F(H) = e^{H\sqrt{a^2\sigma_1^2+b^2\sigma_2^2+2ab\rho\sigma_1\sigma}}$ and defining $H(F_0) = \ln F_0/\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma}$, we obtain the H-tree and F-tree as previously. From the F-tree in Figure 3 and Equation (7), we can obtain the expressions for up and down movements and the probability on an up movement.

In what follows now, we present some commonly used functions in finance applications. The multiplicative function form can be used when an option on an underlying asset value has two sources of uncertainty. For example, when

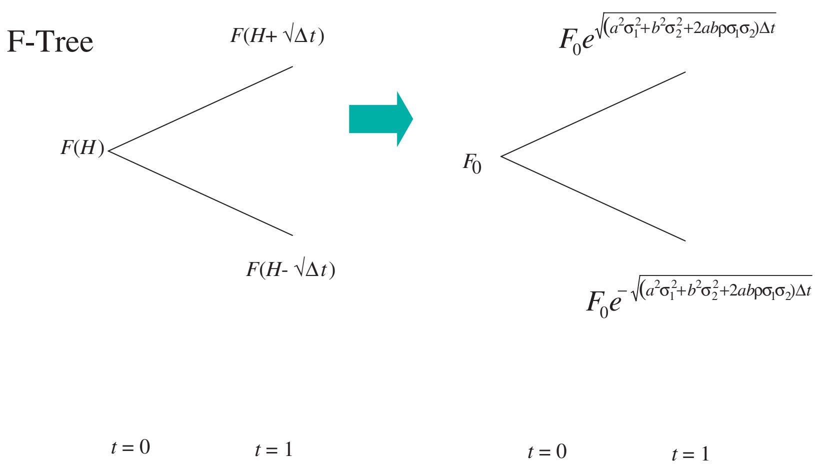


Figure 3. One period binomial tree $F = S_1^a S_2^b$.

valuing a forest concession, where the value of standing timber is a function of price and inventories, each following a diffusion process.

3.1. *Product of the underlying variables: a = b = 1, i.e., F = F(S₁, S₂, t) = S₁S₂*

From Equation (8d), we have

$$\mathrm{d}F = (\mu_1 + \mu_2 + \rho\sigma_1\sigma_2)F\,\mathrm{d}t + (\sigma_1\,\mathrm{d}W_1 + \sigma_2\,\mathrm{d}W_2)F. \tag{9a}$$

Hence F follows a geometric Brownian motion with $\mu(F,t) = (\mu_1 + \mu_2 + \rho\sigma_1\sigma_2)$, the noise term given by $(\sigma_1\,\mathrm{d}W_1 + \sigma_2\,\mathrm{d}W_2)$ and W_i are Wiener processes. Define $\mathrm{d}W = \sigma_1\,\mathrm{d}W_1 + \sigma_2\,\mathrm{d}W_2$, then since the per unit time processes are $\mathrm{d}W_i \sim N(0,1)$ and σ_i are constant for $i = 1, 2$ we have

$$\mathrm{d}W \sim N(0, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2). \tag{9b}$$

The per unit of time standardized value is

$$\mathrm{d}W_z = \frac{\mathrm{d}W}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}} \tag{9c}$$

Substituting for $\mathrm{d}W$ in Equation (9a), we have

$$\mathrm{d}F = (\mu_1 + \mu_2 + \rho\sigma_1\sigma_2)F\,\mathrm{d}t + F\left(\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\right)\mathrm{d}W_z, \tag{9d}$$

where $\mathrm{d}W_z \sim N(0, \mathrm{d}t)$.

To make the volatility term constant in Equation (9d), the transformation is

$$H(F,t) = \int^{F} \frac{\mathrm{d}Z}{\sigma(Z,t)} = \frac{\ln F}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}. \tag{9e}$$

The inverse transformation provides $F(H) = e^{H\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}}$ and defining $H(F_0) = \ln F_0/\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}$, we obtain the H-tree and F-tree as previously.

Now we consider the ratio functional form, which is typically encountered among others in exchange options and real options to abandon a project for its salvage value. For example, an opportunity to exchange one company's securities for those of another within a stated time period (Margrabe, 1978).

3.2. *Relative value of the underlying variables a = 1, b = −1, i.e., F = F(S₁, S₂, t) = S₁/S₂*

From Equation (8d), we have

$$dF = (\mu_1 - \mu_2 + \sigma_2^2 - \rho\sigma_1\sigma_2)F\,dt + (\sigma_1\,dW_1 - \sigma_2\,dW_2)F. \tag{10a}$$

Therefore, F follows a geometric Brownian motion with $\mu(F, t) = (\mu_1 - \mu_2 + \sigma_2^2 - \rho\sigma_1\sigma_2)$, and the noise term given by $(\sigma_1\,dW_1 - \sigma_2\,dW_2)$. Define $dW = \sigma_1\,dW_1 - \sigma_2\,dW_2$, then since the per unit time processes are $dW_i \sim N(0, 1)$ and σ_i are constant for $i = 1, 2$ we have

$$dW \sim N\left(0, \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2\right). \tag{10b}$$

The per unit of time standardized value is

$$dW_z = \frac{dW}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}. \tag{10c}$$

Substituting for dW in Equation (10a), we have

$$dF = (\mu_1 - \mu_2 + \sigma_2^2 - \rho\sigma_1\sigma_2)F\,dt + F\left(\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\right)dW_z, \tag{10d}$$

where $dW_z \sim N(0, dt)$.

Making the volatility term constant in Equation (10d) gives us a computationally simple binomial approximation. The transform is

$$H(F, t) = \int^{F} \frac{dZ}{\sigma(Z, t)} = \frac{\ln F}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}} \tag{10e}$$

The inverse transformation provides $F(H) = e^{H\sqrt{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}}$ and defining $H(F_0) = \ln F_0/\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$, we obtain the H-tree and F-tree as in the previous cases.

The case discussed in Section 3.3 has applications, for example, in the valuation of basket options (Rubinstein, 1992) where the distribution of the weighted forward price of all assets in the basket is approximated by the geometric average.

3.3. *Geometric average of underlying variables a = b = 0.5, i.e., F = F(S₁, S₂, t) = (S₁S₂)^(1/2)*

Substituting $a = b = 0.5$ in Equation (8d), we have

$$\begin{aligned} \mathrm{d}F &= \frac{1}{2}\left((\mu_1+\mu_2) - \frac{1}{4}(\sigma_1^2+\sigma_2^2-\rho\sigma_1\sigma_2)\right)F\,\mathrm{d}t \\ &\quad + \frac{1}{2}\left(\sigma_1\,\mathrm{d}W_1 + \sigma_2\,\mathrm{d}W_2\right)F. \end{aligned} \tag{11a}$$

In the above equation, F follows a geometric Brownian motion with $\mu(F,t) = \frac{1}{2}((\mu_1+\mu_2) - \frac{1}{4}(\sigma_1^2+\sigma_2^2-\rho\sigma_1\sigma_2))$, and the noise term given by $(\sigma_1\mathrm{d}W_1 + \sigma_2\mathrm{d}W_2) \geq 0$. Define $\mathrm{d}W = \frac{1}{2}(\sigma_1\mathrm{d}W_1+\sigma_2\mathrm{d}W_2)$, then since the per unit time processes are $\mathrm{d}W_i \sim N(0,1)$ and σ_i are constant for $i = 1, 2$, we have

$$\mathrm{d}W \sim N\left(0, \frac{1}{4}(\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)\right). \tag{11b}$$

The per unit of time standardized value is

$$\mathrm{d}W_z = \frac{2\,\mathrm{d}W}{\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}}. \tag{11c}$$

Substituting for $\mathrm{d}W$ in Equation (11a), we have

$$\begin{aligned} \mathrm{d}F &= \frac{1}{2}\left((\mu_1+\mu_2) - \frac{1}{4}(\sigma_1^2+\sigma_2^2-\rho\sigma_1\sigma_2)\right)F\,\mathrm{d}t \\ &\quad + F\left(\frac{1}{2}\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}\right)\mathrm{d}W_z, \end{aligned} \tag{11d}$$

where $\mathrm{d}W_z \sim N(0, \mathrm{d}t)$.

Making the volatility term constant in Equation (11d) gives us a computationally simple binomial approximation. The transform is

$$H(F,t) = \int^F \frac{\mathrm{d}Z}{\sigma(Z,t)} = \frac{2\ln F}{\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}} \tag{11e}$$

The inverse transformation provides $F(H) = e^{\frac{1}{2}H\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}}$ and defining $H(F_0) = 2\ln F_0/\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}$, we obtain the H-tree and F-tree as in the previous cases. We summarize the parameters, mean, and variance of the processes for two state variables in Table 1.

Table 1. Mean and variances of the commonly used functions.

(a, b)	$F(S_1, S_2, t)$	Mean	Variance
(1, 1)	$S_1 S_2$	$\mu_1 + \mu_2 + \rho\sigma_1\sigma_2$	$\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2$
(1, −1)	S_1/S_2	$\mu_1 - \mu_2 + \sigma_2^2 - \rho\sigma_1\sigma_2$	$\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$
(0.5, 0.5)	$(S_1 S_2)^{1/2}$	$\frac{1}{2}\left((\mu_1 + \mu_2) - \frac{1}{4}(\sigma_1^2 + \sigma_2^2 - \rho\sigma_1\sigma_2)\right)$	$\frac{1}{4}(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$

In the following section, we consider $F(S_1, S_2, t)$ as a function of log-transformed variables. The log-transformed variables are useful in valuing complex investments with multiple interactive options, options with nonproportional dividends, and compound options (with a series of exercise prices) (Trigeorgis, 1991).

4. Log-Transformed Variables

In general, let $F(S_1, S_2, t) = \ln(S_1^a S_2^b)$ where constants a and b are real numbers and ln is the natural logarithm. As $\partial F/\partial S_1 = a/S_1$, $\partial F/\partial S_2 = b/S_2$, $\partial F/\partial t = 0$, $\partial^2 F/\partial S_1^2 = -a/S_1^2$ $\partial^2 F/\partial S_2^2 = -b/S_2^2$, $\partial^2 F/\partial S_1 \partial S_2 = 0$, from the Ito lemma

$$\mathrm{d}F = \frac{a}{S_1}\mathrm{d}S_1 + \frac{b}{S_2}\mathrm{d}S_2 + \frac{1}{2}\left(\frac{-a}{S_1^2}\right)(\mathrm{d}S_1)^2 + \frac{1}{2}\left(\frac{-b}{S_2^2}\right)(\mathrm{d}S_2)^2. \tag{12a}$$

Substituting for $\mathrm{d}S_1$ and $\mathrm{d}S_2$ and rearranging terms, we obtain

$$\mathrm{d}F = \left[(a\mu_1 + b\mu_2) - \frac{a\sigma_1^2}{2} - \frac{b\sigma_2^2}{2}\right]\mathrm{d}t + (a\sigma_1\,\mathrm{d}W_1 + b\sigma_2\,\mathrm{d}W_2). \tag{12b}$$

Here, F follows a geometric Brownian motion with

$$\mu(F, t) = \left[(a\mu_1 + b\mu_2) - \frac{a\sigma_1^2}{2} - \frac{b\sigma_2^2}{2}\right],$$

and the noise term given by $(a\sigma_1\,\mathrm{d}W_1 + b\sigma_2\,\mathrm{d}W_2)$ where W_i are Wiener processes. Define $\mathrm{d}W = a\sigma_1\,\mathrm{d}W_1 + b\sigma_2\,\mathrm{d}W_2$, then since the per unit time processes are $\mathrm{d}W_i \sim N(0, 1)$ and σ_i are constant for $i = 1, 2$ we have

$$\mathrm{d}W \sim N(0, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2). \tag{12c}$$

The per unit of time standardized value is

$$dW_z = \frac{dW}{\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}}. \tag{12d}$$

Substituting for dW in Equation (12b), we have

$$\begin{aligned} dF &= \left[(a\mu_1 + b\mu_2) - \frac{a\sigma_1^2}{2} - \frac{b\sigma_2^2}{2}\right] F\, dt \\ &\quad + F\left(\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}\right) dW_z, \end{aligned} \tag{12e}$$

where $dW_z \sim N(0, dt)$.

In order to obtain a computationally simple binomial approximation for the log-variables, we need to make the volatility term constant in Equation (12e). The transform is

$$H(F,t) = \int^F \frac{dZ}{\sigma(Z,t)} = \frac{\ln F}{\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}}. \tag{12f}$$

The inverse transformation provides $F(H) = e^{H\sqrt{a^2\sigma_1^2+b^2\sigma_2^2+2ab\rho\sigma_1\sigma_2}}$ and defining $H(F_0) = \ln F_0/\sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}$, we obtain the H-tree and F-tree as previously. We discuss special processes, which include the sum, and difference of two log-transformed variables.

4.1. *Sum of log-transformed variables:* $a = b = 1$*, i.e.,* $F = F(S_1, S_2, t) = \ln(S_1 S_2) = \ln(S_1) + \ln(S_2)$

Substituting $a = b = 1$ in Equation (12e), we have

$$dF = \left(\mu_1 + \mu_2 - \frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right) F\, dt + (\sigma_1\, dW_1 + \sigma_2\, dW_2)F, \tag{13a}$$

where F follows a geometric Brownian motion with

$$\mu(F,t) = (\mu_1 + \mu_2 - \frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}),$$

a noise term given by $(\sigma_1\, dW_1 + \sigma_2\, dW_2)$ and W_i Wiener processes. Define $dW = \sigma_1\, dW_1 + \sigma_2\, dW_2$, then since the per unit time processes are $dW_i \sim$

$N(0, 1)$ and σ_i are constant for $i = 1, 2$ we have

$$\mathrm{d}W \sim N(0, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2). \tag{13b}$$

The per unit of time standardized value is given by

$$\mathrm{d}W_z = \frac{\mathrm{d}W}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}. \tag{13c}$$

Substituting for $\mathrm{d}W$ in Equation (13a), we have

$$\mathrm{d}F = \left(\mu_1 + \mu_2 - \frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right) F\,\mathrm{d}t + F\left(\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\right)\mathrm{d}W_z, \tag{13d}$$

where $\mathrm{d}W_z \sim N(0, \mathrm{d}t)$.

In order to obtain a computationally simple binomial approximation, we need to make the volatility term constant in Equation (13d). The transform is

$$H(F, t) = \int^{F} \frac{\mathrm{d}Z}{\sigma(Z, t)} = \frac{\ln F}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}. \tag{13e}$$

The inverse transformation provides $F(H) = e^{H\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}}$ and defining $H(F_0) = \ln F_0/\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}$, we obtain the H-tree and F-tree as previously.

4.2. *Difference of log-transformed variables: $a = 1, b = -1$, i.e., $F = F(S_1, S_2, t) = ln(S_1/S2) = ln(S_1) - ln(S_2)$*

Substituting $a = 1, b = -1$ in Equation (12e), we have

$$\mathrm{d}F = \left(\mu_1 - \mu_2 - \frac{\sigma_1^2}{2} + \frac{\sigma_2^2}{2}\right) F\,\mathrm{d}t + (\sigma_1\,\mathrm{d}W_1 - \sigma_2\,\mathrm{d}W_2)F. \tag{14a}$$

Hence F follows a geometric Brownian motion with

$$\mu(F, t) = \left(\mu_1 - \mu_2 - \frac{\sigma_1^2}{2} + \frac{\sigma_2^2}{2}\right),$$

a noise term given by $(\sigma_1\,\mathrm{d}W_1 - \sigma_2\,\mathrm{d}W_2) \geq 0$ and W_i Wiener processes. Define $\mathrm{d}W = \sigma_1\,\mathrm{d}W_1 - \sigma_2\,\mathrm{d}W_2$, then since the per unit time processes are

Table 2. Mean and variances of the commonly used functions.

(a, b)	$F(S_1, S_2, t)$	Mean	Variance
$(1, 1)$	$\ln(S_1) + \ln(S_2)$	$\mu_1 + \mu_2 - \frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}$	$\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2$
$(1, -1)$	$\ln(S_1) - \ln(S_2)$	$\mu_1 - \mu_2 - \frac{\sigma_1^2}{2} + \frac{\sigma_2^2}{2}$	$\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$
$(1, 0)$	$\ln(S_1)$	$\mu_1 - \frac{\sigma_1^2}{2}$	σ_1^2

$\mathrm{d}W_i \sim N(0, 1)$ and σ_i are constant for $i = 1, 2$ we have

$$\mathrm{d}W \sim N\left(0, \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2\right). \tag{14b}$$

The per unit of time standardized value is

$$\mathrm{d}W_z = \frac{\mathrm{d}W}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}. \tag{14c}$$

Substituting for $\mathrm{d}W$ in Equation (14a), we have

$$\mathrm{d}F = \left(\mu_1 - \mu_2 - \frac{\sigma_1^2}{2} + \frac{\sigma_2^2}{2}\right) F\ \mathrm{d}t + F\left(\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\right)\mathrm{d}W_z \tag{14d}$$

where $\mathrm{d}W_z \sim N(0, \mathrm{d}t)$.

In order to obtain a computationally simple binomial approximation, we need to make the volatility term constant in Equation (14d). The transform is

$$H(F, t) = \int^F \frac{\mathrm{d}Z}{\sigma(Z, t)} = \frac{\ln F}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}. \tag{14e}$$

The inverse transformation provides $F(H) = e^{H\sqrt{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}}$ and defining $H(F_0) = \ln F_0/\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$, we obtain the H-tree and F-tree as previously. The mean and variances of the processes with the log-transformed variables discussed above are given in Table 2.

5. Numerical Accuracy

In order to study the numerical accuracy of the binomial approximations with the volatility transformation for two variable functions, we consider the option to exchange one asset for another. For this purpose, we use the relative value of underlying assets discussed in Section 3.2. We compare the exchange option values obtained from a one-period and a two-period binomial approximations (Rubinstein, 1992) with Margrabe's (1978) continuous time exchange option model. We choose the following parameters: asset values (in \$) $S_1 = S_1 = 10, 20, 30, 40, 50$; volatility $\sigma_1 = \sigma_2 = 5\%, 20\%$; correlation $\rho = 0$; and time to expiration $T = 1$ and 10 weeks. The percentage relative error in estimates is calculated as

$$\% \text{ Relative error} = 100 \left(\frac{\text{Estimate – Margrabe}}{\text{Margrabe}} \right).$$

The results for exchange option values with respect to different periods and the percentage relative errors (in parenthesis) are given in Table 3. For the given parameters in Table 3, we have the following indicative observations:

(1) When the volatility is low, one-period binomial approximations overestimate the exchange option values and the relative error is 25.3%, whereas the two-period binomial approximations underestimate the exchange option values and the relative error is -9.52%.
(2) When the volatility is high, then both the one-period and two-period binomial approximations overestimate the exchange option values and the relative errors are 25.25% and 15.3%, respectively.
(3) For both time periods, the two-period binomial approximations provide better exchange option values (smaller relative error) than the one-period binomial approximations.

The one-period binomial approximations provide option values accurate to within (0.009 to 0.049) for $\sigma_1 = \sigma_2 = 5\%$, $\rho = 0$, $T = 1$ and (-0.003 to -0.018) for $\sigma_1 = \sigma_2 = 20\%$, $\rho = 0$, $T = 10$. For the two-period binomial approximations, estimates are accurate within (0.125 to 0.624) for $\sigma_1 = \sigma_2 = 5\%$, $\rho = 0$, $T = 1$ and (0.075 to 0.378) for $\sigma_1 = \sigma_2 = 20\%$, $\rho = 0$, $T = 10$. The binomial approximations deteriorate as the option life is lengthened which is consistent with Nelson and Ramaswamy (1990).

Next by varying values of parameters of the exchange option, we simulated the option values presented in Table 4. We observe the following from numerical results in Table 4:

(1) Given $S_1 = S_2 = 10, \sigma_1 = \sigma_2 = 5\%, 20\%$, and $S_1 = S_1 = 10, \sigma_1 = 5\%$, $\sigma_2 = 20\%$, and $T = 1$. With increasing ρ, it is observed that for the

Table 3. Exchange option values and percentage relative errors.

$S_1 = S_2$	$\sigma_1 = \sigma_2 = 5\%$, $\rho = 0$, $T = 1$			$\sigma_1 = \sigma_2 = 20\%$, $\rho = 0$, $T = 10$		
	Margrabe	One-period	Two-period	Margrabe	One-period	Two-period
10	0.0391196	0.0490286 (25.3%)	0.0353973 (−9.52%)	0.4945114	0.6193798 (25.25%)	0.5701518 (15.3%)
20	0.0782392	0.0980573 (25.3%)	0.0707947 (−9.52%)	0.9890228	1.2387596 (25.25%)	1.1403036 (15.3%)
30	0.1173588	0.1470859 (25.3%)	0.106192 (−9.52%)	1.4835342	1.8581394 (25.25%)	1.7104554 (15.3%)
40	0.1564783	0.1961146 (25.3%)	0.1415893 (−9.52%)	1.9780456	2.4775192 (25.25%)	2.2806072 (15.3%)
50	0.1955979	0.2451432 (25.3%)	0.1769866 (−9.52%)	2.4725569	3.096899 (25.25%)	2.850759 (15.3%)

Table 4. Percentage relative errors.

S_1	S_2	T	σ_1	σ_2	ρ	Margrabe	% Relative error	
							One-period	Two-period
10	10	0.0192	0.05	0.05	−1	0.0553	25.33	−8.73
10	10	0.0192	0.05	0.05	0	0.0391	25.33	−9.52
10	10	0.0192	0.05	0.05	0.5	0.0277	25.33	−10.06
10	10	0.0192	0.05	0.05	0.95	0.0087	25.33	−10.96
10	10	0.0192	0.2	0.2	−1	0.2213	25.31	−0.31
10	10	0.0192	0.2	0.2	0	0.1565	25.32	−3.69
10	10	0.0192	0.2	0.2	0.5	0.1106	25.33	−6.01
10	10	0.0192	0.2	0.2	0.95	0.0350	25.33	−9.71
10	10	0.0192	0.05	0.2	−1	0.1383	25.32	−4.61
10	10	0.0192	0.05	0.2	0	0.1140	25.33	−5.84
10	10	0.0192	0.05	0.2	0.5	0.0997	25.33	−6.55
10	10	0.0192	0.05	0.2	0.95	0.0848	25.33	−7.29
20	10	0.0192	0.05	0.05	−1	10	0	0.07
20	10	0.0192	0.05	0.05	0	10	0	0.03
20	10	0.0192	0.05	0.05	0.5	10	0	0.02
20	10	0.0192	0.05	0.05	0.95	10	0	0
20	10	0.0192	0.2	0.2	−1	10	0	1.08
20	10	0.0192	0.2	0.2	0	10	0	0.54
20	10	0.0192	0.2	0.2	0.5	10	0	0.27
20	10	0.0192	0.2	0.2	0.95	10	0	0.03

one-period binomial approximation, the percent relative error remains constant, whereas relative error is reduced in the two-period binomial approximations.

(2) Given $S_1 = 20$, $S_2 = 10$, $\sigma_1 = \sigma_2 = 5\%$, 20%, and $T = 1$. When ρ is increased, the exchange option values for both the one- and two-period binomial approximations and Margrabe models are very close.

6. Conclusions

To construct a computationally simple binomial approximation for diffusions, we considered a family of two correlated variables and of two log-transformed variables. In particular, we showed how one could obtain the transforms for functions of two variables in multiplicative and ratio forms. We simulated exchange option values using one-period and two-period binomial approximations and compared with the Margrabe's model. Our numerical results indicate that the approximations work well for options with short maturity. Further, as also noted by Nelson and Ramaswamy (1990), the error in estimate increases when option life is lengthened.

Acknowledgments

We thank Gordon Sick, University of Calgary, for suggesting the research work by Nelson and Ramaswamy in the context of multinomial approximation models. We also thank the session participants at the 12th Annual Conference on Pacific Basin Finance Economics Accounting and Business for their suggestions.

References

Boyle, P. P., “A Lattice Framework for Option Pricing with Two State Variables.” *Journal of Financial and Quantitative Analysis* 23 (1), 1–12 (1988).

Boyle, P. P., J. Evnine and S. Gibbs, “Numerical Evaluation of Multivariate Contingent Claims.” *The Review of Financial Studies* 2 (2), 241–250 (1989).

Cortazar, G. and E. S. Schwartz, “A Compound Option Model of Production and Intermediate Inventories.” *Journal of Business* 66 (4), 17–540 (1993).

Johnson, H., “ Options on the Maximum or the Minimum of Several Assets.” *Journal of Financial and Quantitative Analysis* 22 (3), 277–283 (1987).

Kamrad, B. and P. Ritchken, “Multinomial Approximating Model for Options with k-State Variables.” *Management Science* 37 (12), 1640–1652 (1991).

Margrabe, W., “ The Value of an Option to Exchange One Asset for Another.” *The Journal of Finance* 33 (1), 177–186 (1978).

Nelson, D. B. and K. Ramaswamy, “ Simple Binomial Processes as Diffusion Approximations in Financial Models.” *The Review of Financial Studies* 3 (3), 393–430 (1990).

Rubinstein, M., “One for Another.” *RISK* (July–August) 191–194 (1992).

Stulz, R. M., “Options on the Minimum or the Maximum of Two Risk Assets: Analysis and Applications.” *Journal of Financial Economics* 10, 161–185 (1982).

Trigeorgis, L., “ A Log-Transformed Binomial Numerical Analysis Method for Valuing Complex Multi-Option Investments.” *Journal of Financial and Quantitative Analysis* 26 (3), 309–326 (1991).

This page intentionally left blank

Chapter 9

The Prime Rate–Deposit Rate Spread and Macroeconomic Shocks

Bradley T. Ewing*
Texas Tech University, USA

Jamie Brown Kruse
East Carolina University, USA

This paper examines the response of the prime rate–deposit rate spread to shocks in real output growth, inflation, and the stance of monetary policy. A simple model of the lending and deposit markets is introduced that provides insight as to how these macroeconomic factors might affect the spread. The paper employs the recently developed technique of *generalized* impulse response analysis proposed. This method does not impose *a priori* restrictions as to the relative importance each of the variables in the underlying vector autoregression may play in the transmission process. Thus, the results provide robust evidence as to the relationship between the prime rate–deposit rate spread and these macroeconomic factors. Specifically, the model suggests and the empirical results confirm that shocks to inflation widen the spread while unexpected changes in the federal funds rate and real output growth lead to a narrower spread.

Keywords: Macroeconomic shocks; VAR; interest rates.

1. Introduction

A number of researchers have concluded that announcements of prime lending rate changes are important determinants of equity prices, see Slovin *et al*. (1994), Johnson and Jensen (1994), for examples. It has also been found that the spread between the prime rate and the deposit rate is an important determinant of financial service sector stock returns, see Ewing *et al*. (1998) and, more recently, Gobbi and Lotti (2004) found that the spread may play an important role in a bank's decision to enter a market. To date, however, little research has been concerned with how the prime rate–deposit rate spread responds to macroeconomic shocks.[1] Given Chatrath *et al*. (1997) documented a trend

*Corresponding author.

[1] The effect of various exogenous variables on prime rate changes has been examined. For example, Mester and Saunders (1995) modeled the probability of prime rate changes as depending on market conditions.

toward globalization of banking activity and found that the US lending and deposit rates influence the rates of other countries, it is particularly interesting and important to examine what leads to movements in the prime rate–deposit rate spread in the United States. This paper identifies and examines the extent to which innovations in several key macroeconomic variables are transmitted to the prime rate–deposit rate spread. We focus on three fundamental macroeconomic variables that previous findings have identified as important state variables in stock and bond returns — namely, the stance of monetary policy, inflation, and real economic activity.

The relationship between the prime rate–deposit rate spread and the macroeconomic factors is examined by computing *generalized* impulse response functions derived from the estimation of a four-equation vector autoregression model.[2] This is an interesting and potentially informative endeavor in light of the work of Scholnick (1999) who, examining long-term lending and deposit markets, found that the corresponding individual rates do not, in general, respond symmetrically to changes in market rates. Thus, though the spread between the prime rate and the deposit rate may be stationary, the spread may not be constant period-to-period and may increase/decrease in response to shocks. An innovation to any of the macroeconomic variables may be interpreted as (unexpected) economic *news*.[3] The generalized response functions allow us to compare and contrast the effects of unanticipated changes in the macroeconomic factors on the interest rate spread without imposing *a priori* restrictions about the relative importance of the macroeconomic variables. Clearly, financial institutions and financial market participants may be affected by movements in the prime rate–deposit rate spread. Knowledge of what leads to movements in the spread and how long shocks may last is, therefore, important.

2. A Simple Model of the Lending and Deposit Markets

In order to understand movements in the spread between the prime rate and the deposit rate we begin by developing a model of the interest rate spread

[2]The paper employs the recently developed econometric technique of generalized impulse response analysis (Koop *et al.*, 1996; Pesaran and Shin, 1998).

[3]Fleming and Remolona (1999) find that the reactions of the bond markets depend on the unexpected component of a given macroeconomic announcement.

derived from general specifications of supply and demand in the lending and deposit markets.[4] We specify supply and demand in the lending market as follows:

$$L^{\mathrm{S}} = a \cdot \mathbf{X}^{\mathrm{LS}} + a_0 \cdot r^{\mathrm{L}}$$
$$L^{\mathrm{D}} = b \cdot \mathbf{X}^{\mathrm{LD}} - b_0 \cdot r^{\mathrm{L}}$$

where $\mathbf{X}^{\mathrm{LS}}$ and $\mathbf{X}^{\mathrm{LD}}$ are vectors of exogenous variables which affect the supply and demand for loans, respectively, and r^{L} is the lending rate.

Similarly, the supply and demand for deposits may be represented as:

$$D^{\mathrm{S}} = c \cdot \mathbf{X}^{\mathrm{DS}} + c_0 \cdot r^{\mathrm{D}}$$
$$D^{\mathrm{D}} = d \cdot \mathbf{X}^{\mathrm{DD}} - d_0 \cdot r^{\mathrm{D}}$$

where $\mathbf{X}^{\mathrm{DS}}$ and $\mathbf{X}^{\mathrm{DD}}$ are vectors of exogenous variables which affect the supply and demand for deposits and r^{D} is the deposit rate.

The underlying structural equations (i.e., demand and supply specifications) can be solved to obtain the reduced-form equation for the lending–deposit rate spread:

$$S = (r^{\mathrm{L}} - r^{\mathrm{D}}) = \alpha \cdot X^{\mathrm{LD}} - \beta \cdot X^{\mathrm{LS}} - \gamma \cdot X^{\mathrm{DD}} + \delta \cdot X^{\mathrm{DS}}$$

where $\alpha = [b/(a_0 + b_0)] > 0$, $\beta = [a/(a_0 + b_0)] > 0$, $\gamma = [d/(c_0 + d_0)] > 0$, and $\delta = [c/(c_0 + d_0)] > 0$.[5] A required reserve ratio of $0 < \sigma < 1$ implies that the equilibrium quantity of loans produced, L, will be related to the equilibrium quantity of deposits, D, such that $L = (1 - \sigma)D$ and $L < D$.

Changes in the exogenous variables that affect supply and demand in the lending and deposit markets may cause the spread to change with the response depending on the elasticities. The respective elasticities are

[4] We use terms lending and prime interchangeably throughout the remainder of the paper.

[5] The reduced-form variable allows us to focus on (aggregate or market-level) prime rate–deposit rate responses to macroeconomic variables without making restrictive assumptions about the separation of adjustment costs. Thus, costs are allowed to simultaneously exist on both the loan side and the deposit side of bank balance sheets. In fact, based on results obtained from a panel-data study of a number of individual banks. Elyasiani *et al*. (1995) conclude the restrictive assumption of asset–liability independence may not accurately characterize bank behavior.

given by:

$$
\begin{aligned}
\varepsilon^{\mathrm{LS}} &= (\mathrm{d}L^{S}\mathrm{d}r^{\mathrm{L}}) \cdot (r^{\mathrm{L}}/L^{\mathrm{S}}) = a_0 \cdot (r^{\mathrm{L}}L^{\mathrm{S}}) \\
\varepsilon^{\mathrm{LD}} &= (\mathrm{d}L^{\mathrm{D}}\mathrm{d}r^{\mathrm{L}}) \cdot (r^{\mathrm{L}}L^{\mathrm{D}}) = -b_0 \cdot (r^{\mathrm{L}}L^{\mathrm{D}}) \\
\varepsilon^{\mathrm{DS}} &= (\mathrm{d}D^{\mathrm{S}}\mathrm{d}r^{\mathrm{D}}) \cdot (r^{\mathrm{D}}D^{\mathrm{S}}) = c_0 \cdot (r^{\mathrm{D}}D^{\mathrm{S}}) \\
\varepsilon^{\mathrm{DD}} &= (\mathrm{d}D^{\mathrm{D}}\mathrm{d}r^{\mathrm{D}}) \cdot (r^{\mathrm{D}}D^{\mathrm{D}}) = -d_0 \cdot (r^{\mathrm{D}}D^{\mathrm{D}})
\end{aligned}
$$

where $\varepsilon^{\mathrm{LS}}$ ($\varepsilon^{\mathrm{DS}}$) is the elasticity of supply in the loan (deposit) market and $\varepsilon^{\mathrm{LD}}$ ($\varepsilon^{\mathrm{DD}}$) is the elasticity of demand in the loan (deposit) market. In order for lending institutions to remain in business the lending rate must be greater than the deposit rate (at least in the long run). Note that this requirement suggests that a long-run relationship between the lending rate and deposit rate will exist. Thus, if the spread becomes unusually high or low, competition in the marketplace will ensure that the spread adjusts to eliminate the disequilibrium, although this need not happen instantaneously.

From above we have that $(r^{\mathrm{D}}/D) < (r^{\mathrm{L}}/L)$. Further, if $[(r^{\mathrm{L}}/r^{\mathrm{D}})(D/L)] > (c_0/b_0)$ then $c_0 \cdot (r^{\mathrm{D}}/D) < b_0 \cdot (r^{\mathrm{L}}/L)$.[6] The meaning of this is that higher elasticity of demand in the lending market $[b_0(r^{\mathrm{L}}/L)]$ will generate a smaller response in r^{L} while the deposit market's less elastic supply $[c_0(r^{\mathrm{D}}/D)]$, relative to the lending market demand, will generate a larger response in r^{D} for given exogenous shocks (e.g., an unexpected change in an exogenous variable that affects lending demand/supply or deposit demand/supply).[7] Consequently, provided the above condition holds and consistent with the claims made in earlier empirical studies, the prime lending rate will be "sticky" and the spread should narrow when deposit rates are rising and widen when deposits rates are falling; for examples, see Forbes and Mayne (1989), Forbes and Paul (1992), Ewing, *et al.* (1998), and Thompson (in press). Similar results have been found in the personal loan market. For instance, Kahn *et al.* (2005) found that banks are slower to adjust personal loan rates downward when the spread between them and Treasury rates is widening. In the next section we suggest three macroeconomic factors that, by affecting the supply and demand in the lending and deposit markets, should affect the spread in a predictable manner.

[6]This holds if $c_0 < b_0$, $c_0 = b_0$, and also if c_0 is not too much larger than b_0.

[7]Note that this relationship between elasticities is consistent with a profit maximization problem in which deposits are an input of the production function for loans.

3. The Rate Spread and Macroeconomic Variables

In choosing the macroeconomic factors to include in our analysis, we borrow from the literature that has studied the relationship between stock and bond market returns and macroeconomic factors.[8] At the aggregate level, the stage of the business cycle — whether the economy is in a growth period or recession — affects the demand for and supply of loanable funds. Thus, interest rates will be affected by real activity and, if lending rates and deposit rates do not respond immediately and exactly in the same fashion, the lending–deposit rate spread can be expected to change. Economic growth increases the demand for loans as businesses foresee additional profit opportunities. According to Friedman and Kuttner (1998), corporate bankruptcy and default rise markedly during a downturn in the business cycle and fall during periods of economic growth. The supply of loans may increase with greater real economic activity as lenders perceive a lower risk of default. In the deposit market, lending institutions may increase their demand for deposits as more inputs are needed to produce a greater amount of loans. Consequently, changes in real output can be expected to affect the spread between the prime lending rate and the deposit rate.

Inflation affects the real cost of borrowing and the real return from lending. The model of Section 2 implies an increase in the demand for loans and a decrease in the supply of loans. The demand for deposits might fall if lenders cut back on loan production. Unanticipated inflation may create volatility and uncertainty regarding future price changes. Thus, it is possible that the supply of deposits may increase as savers desire safety. An inflation shock may have a fairly persistent effect on the spread by restricting lending and borrowing, as well as by altering (expected) future production activity. Stokes and Neuburger (1998) provide empirical evidence that inflation influences bond prices and returns. Thus, changes in inflation may affect the spread between the prime and deposit rates.

Mougoue and Wagster (1997) found that deposit rates are affected by monetary policy. Thorbecke (1997) and others argue that monetary policy can be measured by the change in the fed funds rate. In the model of Section 2,

[8] Chen *et al*. (1986) and He and Ng (1994) incorporate several macroeconomic variables when examining the relations among market fundamentals, economic forces, and stock returns. Ewing (2002) models the stock price returns of financial sector firms as depending on various macroeconomic factors.

the fed funds market can be viewed as a substitute for deposits. A monetary tightening (loosening) should raise (lower) the demand for deposits. Monetary policy can be expected to influence the supply of loans by influencing the operating costs of lenders. An increase in the fed funds rate raises the cost of borrowed reserves and the supply of loans should fall. Additionally, Ewing (2001) concluded that a monetary policy shock enters the economy mainly through its impact on the market risk premium. Thus, a monetary tightening may increase risk (through its effect on firm balance sheets) and further suppress the supply of loans.[9] Taken together, changes in the fed funds rate may influence the prime rate–deposit rate spread.

As the above literature attests to, the lending–deposit rate spread may be linked to macroeconomic factors. Based on the model of Section 2 where, under certain conditions, the prime lending rate is sticky, as well as the discussion above, it is expected that an unexpected increase in the federal funds rate will reduce the spread between the prime and deposit rates. Furthermore, it is expected that a positive shock to real output will lead to a narrower spread and inflation news will widen the spread. In the next sections, we describe the data used to examine and test these hypotheses.

4. The Data

The response of the prime rate–deposit rate spread to shocks in real output, the stance of monetary policy, and inflation are examined over the period from January 1981 to September 2000. The coincident index is chosen as a proxy for real output following Estrella and Mishkin (1997) and Ewing (2001). The choice of the federal funds rate as the stance of monetary policy is based on Thorbecke (1997), Park and Ratti, (2000), Bernanke and Kuttner (2005) and others. The consumer price index for all urban consumers (P) is the measure of the aggregate price level (Park and Ratti, 2000). Following Ewing, *et al.* (1998), the prime lending rate–deposit rate spread is computed as the difference between the prime rate charged on short-term business loans and the interest rate on 1-month (secondary market) certificates of deposit.[10]

[9] Jarrow and Turnbull (2000) discuss the inherent link between market risk and default risk.

[10] Interest rates are monthly averages of daily figures, annualized using a 360-day year or bank interest. The CD rate is based on the average of dealer offering rates on nationally traded certificates of deposit.

All variables except interest rates are entered in natural logarithms. Thus, the monthly data consist of the federal funds rate (FF), the index of coincident indicators (Y), the aggregate price level (P), and the spread between the prime rate and the CD rate (S). All data are from the Federal Reserve Bank of St. Louis.[11]

In preliminary analyses, it was determined that all variables except the interest rate spread required first-differencing to attain stationarity.[12] In what follows we use the first-differences of Y, FF, P, and the level of S. The first-difference operator is denoted by Δ. Table 1 presents descriptive statistics for the variables. The mean prime rate–deposit rate spread is 2.41 and varies from a low of 0.84 to a high of 4.39. Table 2 presents the estimated (contemporaneous) correlation matrix and reveals that the spread is negatively correlated with changes in real output, changes in the fed funds rate, and inflation.

Figure 1 presents a plot of the spread over the sample period. The spread appears relatively high and volatile at the beginning of the sample, a period that just follows the well-known monetary policy experiment of the Federal

Table 1. Descriptive Statistics (adjusted sample period is February 1981–September 2000).

	S	ΔP	ΔFF	ΔY
Maximum	4.3900	0.0114	2.800	0.0126
Minimum	0.8400	−0.0055	−3.1500	−0.0057
Mean	2.4067	0.0029	−0.0532	0.0021
Standard deviation	0.626	0.0022	0.4916	0.0028

Note: S = spread between prime rate and CD rate, ΔP = change in the price level, ΔFF = change in the fed funds rate, and ΔY = change in the coincident index.

[11]The sensitivity of the results were checked using industrial production for output, the CPI less food and energy for the price level, and nonborrowed reserves for monetary policy [see Strongin (1995) on using nonborrowed reserves to proxy monetary policy]. The findings and conclusions were robust to these changes.

[12]However, no evidence of cointegration was found. The results of unit root and Johansen–Juselius cointegration tests are available on request. Chatrath *et al.* (1997) determined that deposit rates and lending rates are first-difference stationary, while the spread between the lending rate and the deposit rate has been found to be stationary (see Ewing *et al.*, 1998).

Table 2. Estimated correlation matrix (adjusted sample period is February 1981–September 2000).

	S	ΔP	ΔFF	ΔY
S	1.0000	—	—	—
ΔP	−0.1647	1.0000	—	—
ΔFF	−0.3897	−0.0747	1.0000	—
ΔY	−0.2349	−0.1694	0.2945	1.0000

Note: S = spread between the prime rate and the CD rate, ΔP = change in the price level, ΔFF = change in the fed funds rate, and ΔY = change in the coincident index.

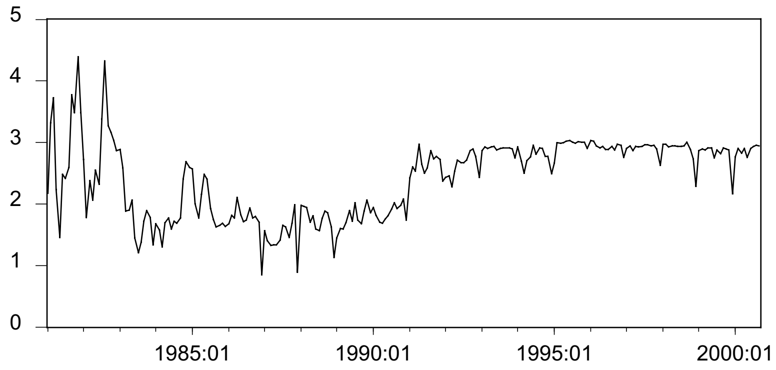


Figure 1. Spread between the prime rate and the CD rate (S). (January 1981–September 2000).

Reserve.[13] During much of the 1980s the spread appeared to be volatile with a number of noticeable increases and decreases. For example, the spread fell dramatically following the October 1987 stock market crash before rebounding quickly. The spread rose in a time period roughly corresponding to the 1990–1991 recession. During the 1990s, when the economy experienced an extended period of economic growth, the spread has remained relatively stable. However, the spread fell and then rose quickly in 1997 and late 1998 about the time that financial markets were rocked by the Asian and Russian crises, and the Long-Term Capital debacle. The spread also behaved erratically

[13]Beginning the study in 1981 has the advantage of avoiding the effects of the change in interest rate regime that accompanied the "monetary experiment." See Hays *et al.* (2000) for a discussion of this issue as it relates to interest rate behavior.

around the time of the Federal Reserve rate hikes of 1999–2000, a period in which the Fed's Alan Greenspan expressed concern over tight labor markets, excessive demand, and inflationary pressures. Certainly, no definitive statement about the response of the interest rate spread to changes in the stance of monetary policy, inflation, and output should be made based on a cursory review of the plot. However, given our previous discussion, we expect that: (1) a positive shock to the federal funds rate (i.e., an unexpected monetary tightening) will lead to a narrowing of the spread between the prime rate and the deposit rate; (2) an unanticipated increase (i.e., shock) in real economic growth will reduce the spread; and (3) a positive inflation shock will make the spread wider. In the next section we introduce a type of innovation accounting technique, based on the estimation of a vector autoregression, that allows us to examine the response of the spread to unanticipated changes in the macroeconomic variables.

5. Vector Autoregression and Generalized Impulse Response Analysis

We are interested in the response of the prime rate–deposit rate spread to shocks in each of the macroeconomic variables. Vector autoregressive (VAR) models and innovation accounting methods such as impulse response functions can be used for this type of dynamic analysis. The conventional impulse response method has been criticized because results are subject to the "orthogonality assumption" and may differ markedly depending on the ordering of the variables in the VAR (Lutkenpohl, 1991).[14] To overcome this problem, recent efforts have focused on developing impulse responses that are not sensitive to the ordering of the variables in the VAR. This paper employs the *generalized* impulse response function proposed by Pesaran and Shin (1998) and Koop *et al*. (1996).

Consider the following moving average representation of the VAR(m) model:

$$\mathbf{z}_t = \Psi(L)\mathbf{v}_t \tag{1}$$

[14]Recall that the traditional orthogonalized impulse response employs a Cholesky decomposition of the positive definite covariance matrix of the shocks (see Mills, 1999). The generalized version does not impose this restriction.

where $E(\mathbf{v}_t\mathbf{v}_t') = \Sigma_V$ such that shocks are contemporaneously correlated, and L is a polynomial in the lag operator. The generalized impulse response function of z_i to a unit (1 standard deviation) shock in z_j is given by:

$$\Psi_{ij,h} = (\sigma_{ii})^{-1/2}(\mathbf{e}_j'\Sigma_{\mathbf{v}}\mathbf{e}_i) \tag{2}$$

where σ_{ii} is the ith diagonal element of $\boldsymbol{\Sigma}_v$, $\mathbf{e}_i$ is a selection vector with the ith element equal to 1 and all other elements equal to zero, and h is the horizon.

A key feature of the generalized impulse response function is that the generalized responses are invariant to any re-ordering of the variables in the VAR.[15] Thus, the generalized impulse response function provides more robust results than the orthogonalized method. Furthermore, because orthogonality is not imposed, another key feature is that the generalized impulse response function allows for meaningful interpretation of the initial impact response of each variable to shocks to any of the other variables. In the study of financial markets where information is often quickly assimilated, the ability to capture these immediate responses of endogenous variables to shocks is appealing.

6. Discussion of Results

A vector autoregression was estimated where the four equations corresponded to ΔY, ΔFF, ΔP, and S. A constant term was included in each equation. The order of the VAR was determined to be three based on Akaike's information criterion, Schwartz Bayesian criterion, and likelihood ratio tests. If the shocks to the respective equations in a VAR are contemporaneously correlated, then orthogonalized and generalized impulse responses may be quite different. As discussed above, re-ordering the variables may lead to a number of vastly different conclusions based on orthogonalized responses. When shocks are not contemporaneously correlated the two types of responses would not be that different. Thus, before examining the dynamic responses of the lending–deposit rate spread to macroeconomic shocks, we performed a test to determine if innovations in the four individual equations in the VAR were

[15]According to Pesaran and Shin (1998) the "generalized impulse responses are unique and fully take account of the historical patterns of correlations observed amongst the different shocks" (p. 20). They caution against using orthogonalized responses as there is generally no clear guidance as to which of many possible parameterizations to employ. The generalized and orthogonalized impulse responses coincide only when the covariance matrix is diagonal.

contemporaneously correlated. The null hypothesis is that the off-diagonal elements in the covariance matrix equal zero and is tested against the alternative that none of the off-diagonal elements is equal to zero. A log-likelihood ratio test statistic is computed as $LR = 2(LL_u - LL_r)$ where LL_u and LL_r are the maximized values for the log-likelihood functions for the unrestricted and restricted models, respectively.[16] The LR statistic is distributed χ^2 with 4 degrees of freedom and equaled 144.4 which is significant at less than the 1% level. Thus, the null hypothesis is rejected and it is appropriate to examine generalized impulse response functions.

Figures 2–4 present the results of the generalized impulse response functions and are plotted out to the 24th month.[17] As can be seen in Figure 2, an unexpected positive change in the fed funds rate has a negative and significant initial impact effect (i.e., at horizon $h = 1$) on S. A monetary policy shock continues to squeeze the prime rate–deposit rate spread in the periods following the shock. The generalized impulse response is significant and negative

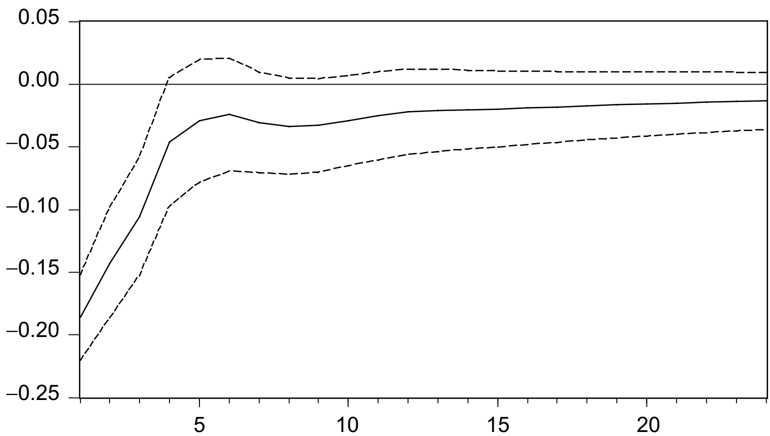


Figure 2. Generalized impulse response of S to a shock to stance of monetary policy (ΔFF).
Notes: The forecast horizon is measured in months and is given on the horizontal axis. The vertical axis measures the magnitude of the response. Confidence bands, used to determine statistical significance, are shown as dashed (----) lines and represent plus/minus 2 standard errors.

---

[16] LL_u is the system log-likelihood from the VAR and LL_r is computed as the sum of the log-likelihood values from the individual equations in the VAR.

[17] Following common practice, we present the generalized impulse response functions in graphical format. A table of the response information is available upon request.

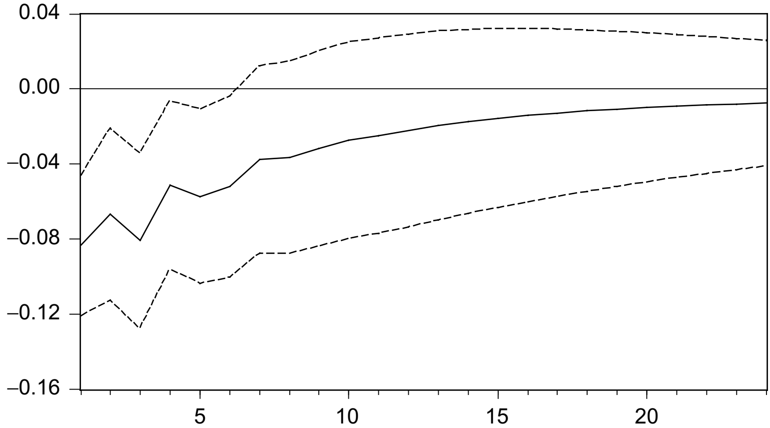


Figure 3. Generalized impulse response of S to a shock to change in real output (ΔY).
Notes: The forecast horizon is measured in months and is given on the horizontal axis. The vertical axis measures the magnitude of the response. Confidence bands, used to determine statistical significance, are shown as dashed (----) lines and represent plus/minus 2 standard errors.

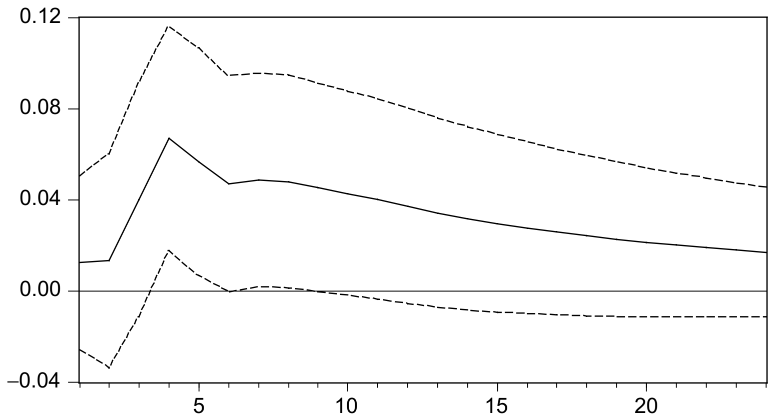


Figure 4. Generalized impulse response of S to a shock to inflation (ΔP).
Notes: The forecast horizon is measured in months and is given on the horizontal axis. The vertical axis measures the magnitude of the response. Confidence bands, used to determine statistical significance, are shown as dashed (----) lines and represent plus/minus 2 standard errors.

at $h = 1, 2, 3$. A sudden monetary tightening, as evidenced by an unanticipated rise in the fed funds rate, immediately lowers the spread and the spread remains below its baseline value for about 2 months after the shock. While rate hikes are typically viewed as an event that raises all rates, the CD rate may rise relatively faster or more than the lending rate if the prime rate is less responsive or sticky.[18] In this case, the spread remains lower than normal until the monetary policy shock mitigates, that is, until both rates adjust and the equilibrium spread is restored. After 2 months, the spread fully returns to its equilibrium value consistent with the notion that the prime rate rises more slowly than the CD rate. As discussed in Section 2, this would be the case if the elasticity of demand in the lending market was relatively greater than the elasticity of supply in the deposit market. This finding supports the claim of Forbes and Mayne (1989) and Forbes and Paul (1992) that when deposit rates are rising the spread becomes smaller, and when deposit rates are falling, sticky lending rates lead to a wider spread.

Figure 3 shows the generalized impulse response of the prime rate–deposit rate spread to a real output shock. The initial impact effect ($h = 1$) of an unanticipated increase in real output is negative and significant. Moreover, the impulse response is negative and significant through the fifth month following the shock. This finding suggests that unexpected real output growth lowers the interest rate spread. In the framework described in Sections 2 and 3, when the economy is performing better than expected several forces are at work that place pressure on these interest rates to change. The demand for loans may increase as the firms investment decision is affected by the higher than expected growth. Financial institutions must raise deposit rates to attract the additional funds required to facilitate the increase in loan activity (i.e., there is an increase in the demand for deposits). However, the supply of deposits may increase with additional wealth, particularly in the case of core deposits.[19] Provided that the elasticity of demand for loans is relatively more elastic than the elasticity of deposit supply, then a positive output shock works to narrow the spread. The deposit rate is more responsive to the macroeconomic event than is the "sticky" lending rate.

---

[18]In fact, Ewing *et al.* (1998) found the prime rate and the CD rate were cointegrated and that the normalized cointegrating vector was not significantly different from $[-1, 1]$. The fact that these two rates establish an equilibrium long run relationship suggests that the prime rate–deposit rate spread will return to an equilibrium following a shock. In terms of the VAR model, this means that the impact of macroeconomic shocks eventually die out as they have only temporary effects. It is in this context that we refer to a "baseline" or "normal" spread.

[19]It is possible that some suppliers of deposits may reallocate their portfolio and move funds from deposit accounts to other assets particularly if they expect these other assets to perform well given the unexpected economic growth. The standard view, however, is that the (net) aggregate supply of deposits will increase with greater wealth.

Figure 4 provides information as to how S responds to unexpected increases in inflation.[20] The initial impact on S is not significant. However, the response is positive and significant for $h = 4$–9, with the effect being only marginally significant at $h = 6$. An explanation for this lies in the analysis of the (expected) real interest rate. The inflation shock lowers the real return from lending which in turn provides an incentive to reduce the supply of loans, *ceteris paribus*. This produces upward pressure on the prime rate. The real cost of borrowing falls with the inflation shock and, *ceteris paribus*, the demand for loans should increase, thus placing additional upward pressure on the lending rate. On the deposit rate side, the former effect (declining supply of loans) coincides with lower demand for deposits and, *ceteris paribus*, this lowers deposit rates. Investors may perceive deposits to be relatively safe as compared to other investments making the supply of deposits increase when faced with possible price distortions and uncertainty resulting from an inflation shock. If agents in the economy are unable to perfectly discern relative price changes from overall price changes then the observed lag in the impulse response of the spread to an inflation shock is consistent with macroeconomic models with rigid output prices. Inflation shocks widen the spread between the prime and deposit rates, and this effect appears to be more persistent than the effects of the other shocks.

7. Concluding Remarks

This paper has examined and documented the response of the prime rate–deposit rate spread to shocks in three key macroeconomic variables using the newly developed technique of generalized impulse response analysis. The technique is robust in terms of the choice of ordering variables in the VAR, thus one can accurately examine and compare both the severity and extent of shocks to these variables on the spread.[21] A simple model of the lending and deposit markets was used to facilitate our understanding of how these

[20]The response to various measures of expected inflation produced similar results.

[21]A comparison to the traditional orthogonalized impulse responses was made and the results were quite different from those obtained from the robust generalized version. First, the orthogonalized version failed to detect an initial impact of a fed funds rate change and suggested a 3-month delayed impact. Second, the orthogonalized version failed to detect any prime–CD spread response from a real output shock. Finally, there was little difference between the orthogonalized and generalized responses to inflation shocks.

markets work. Under certain conditions pertaining to the elasticity of demand for loans and the supply elasticity of deposits, the lending rate will be relatively less responsive to changes in macroeconomic variables than the deposit rate. Thus, predictable short run changes in the spread will occur as a result of macroeconomic shocks.

The results of the paper can be summarized as follows. A monetary policy shock reduces the spread between the prime rate and the deposit rate and has a significant effect for around two months. Unanticipated changes in economic growth also lower the spread with the effects lasting about five months. Both output and monetary policy shocks have an immediate (i.e., initial) impact on the spread but the response of the spread to an output shock lasts longer than the response to an unexpected change in the stance of monetary policy. An inflation shock is associated with a widening of the spread between the prime rate and the certificate of deposit rate. The response of the spread to unexpected inflation occurs with a lag of about three months — and lasts for a relatively long time — over two quarters after the shock. The findings of this paper are consistent with the model and provide evidence in favor of the argument that the prime rate is sticky and less responsive to macroeconomic events than is the deposit rate. The results add to the literature on the relationship between the macroeconomy and lending–deposit markets. The paper has identified the magnitude and persistence of unanticipated changes in monetary policy, real output, and inflation on the spread between the prime rate and the certificate of deposit rate.

Acknowledgments

Earlier versions of this paper were presented in the Econometrics Workshop at Texas A&M University, and in seminars at Baylor University and University of Richmond. We especially thank Qi Li, Jian Yang, Steve Green, and David Van Hoose for valuable comments and suggestions. The usual disclaimer applies.

References

Bernanke, B. S. and K. N. Kuttner, “What Explains the Stock Market’s Reaction to Federal Reserve Policy.” *Journal of Finance* 60, 1221–1257 (2005).

Chatrath, A., S. Ramachander and F. Song, "International Linkages in Bank Lending and Borrowing Markets: Evidence from Six Industrialized Countries." *Applied Financial Economics* 7, 403–411 (1997).

Chen, N., R. Roll and S. Ross, "Economic Forces and the Stock Market." *Journal of Business* 59, 383–403 (1986).

Elyasiani, E., K. J. Kopecky and D. Van Hoose, "Costs of Adjustment, Portfolio Separation, and the Dynamic Behavior of Bank Loans." *Journal of Money, Credit, and Banking* 27, 955–974 (1995).

Estrella, A. and F. Mishkin, "Is there a Role for Monetary Aggregates in the Conduct of Monetary Policy?" *Journal of Monetary Economics* 40, 279–304 (1997).

Ewing, B. T., "Cross-effects of Fundamental State Variables." *Journal of Macroeconomics* 23, 633–645 (2001).

Ewing, B. T., "Macroeconomic News and the Returns of Financial Companies." *Managerial and Decision Economics* 23, 439–446 (2002).

Ewing, B. T., J. E. Payne and S. M. Forbes, "Co-movements of the Prime Rate, CD Rate, and the S&P Financial Stock Index." *Journal of Financial Research* 21, 469–482 (1998).

Forbes, S. M. and L. Mayne, "A Friction Model of the Prime." *Journal of Banking and Finance* 13, 127–135 (1989).

Forbes, S. M. and C. Paul II, "Modeling the Prime Rate: An Ordered-Response Approach." *International Review of Economics and Finance* 1, 147–157 (1992).

Fleming, M. J. and E. M. Remolona, "What Moves Bond Prices?" *Journal of Portfolio Management* 25, 28–38 (1999).

Friedman, B. and K. Kuttner, "Indicator Properties of the Paper-bill Spread: Lessons from Recent Experience." *Review of Economics and Statistics* 80, 34–44 (1998).

Gobbi, G. and F. Lotti, "Entry Decisions and Adverse Selection: an Empirical Analysis of Local Credit Markets." *Journal of Financial Services Research* 26, 225–244 (2004).

Hays, P., M. Schreiber, J. E. Payne, B. T. Ewing and M. J. Piette, "Are Net Discount Rates Stationary? Evidence of Mean-reversion and Persistence." *The Journal of Risk and Insurance* 67, 439–449 (2000).

He, J. and L. K. Ng, "Economic Forces, Fundamental Variables, and Equity Returns." *Journal of Business* 67, 599–609 (1994).

Jarrow, R. A. and S. M. Turnbull, "The Intersection of Market and Credit Risk." *Journal of Banking and Finance* 24, 271–299 (2000).

Johnson, R. R. and G. R. Jensen, "Prime Rate Changes and Returns to Industries: Announcement Period Evidence." *Quarterly Review of Economics and Finance* 34, 75–93 (1994).

Kahn, C., G. Pennacchi and B. Sopranzetti, "Bank Consolidation and the Dynamics of Consumer Loan Interest Rates." *Journal of Business* 78, 99–133 (2005).

Koop, G., M. H. Pesaran and S. M. Potter, "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics* 74, 119–147 (1998).

Lutkenpohl, H., *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin, 1991.

Mester, L. J. and A. Saunders, "When Does the Prime Rate Change?" *Journal of Banking and Finance* 19, 743–764 (1995).

Mills, T. C., *The Econometric Modelling of Financial Time Series*, 2nd ed. Cambridge University Press, Cambridge, UK, 1999.

Mougoue, M. and J. Wagster, "The Causality Effects of the Federal Reserve's Monetary Policy on US and Eurodollar Interest Rates." *Financial Review* 32, 821–844 (1997).

Park, K. and R. A. Ratti, "Real Activity, Inflation, Stock Returns, and Monetary Policy." *Financial Review* 35, 59–78 (2000).

Pesaran, M. H. and Y. Shin, "Generalized Impulse Response Analysis in Linear Multivariate Models." *Economics Letters* 58, 17–29 (1998).

Scholnick, B., "Interest Rate Asymmetries in Long-Term Loan and Deposit Markets." *Journal of Financial Services Research* 16, 5–26 (1999).

Slovin, M. B., M. E. Sushka and E. R. Waller, "Is there News in the Prime Rate?" *Journal of Financial and Quantitative Analysis* 29, 633–646 (1994).

Stokes, H. H. and H. M. Neuburger, *New Methods in Financial Modeling: Explorations and Applications*. Quorum Books, Westport, CT, 1998.

Strongin, S., "The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle." *Journal of Monetary Economics* 35, 463–497 (1995).

Thompson, M. A., "Asymmetric Adjustment in the Prime Lending — Deposit Rate Spread." *Review of Financial Economics* 15(4), 323–329, (2006).

Thorbecke, W., "On Stock Market Returns and Monetary Policy." *Journal of Finance* 52, 638–654 (1997).

This page intentionally left blank

Chapter 10

The Long-Run Performance of Firms that Issue Tracking Stocks

Charmen Loh
Rider University, USA

This paper examines the stock market response to companies that issue tracking stocks. We find that the issuance of tracking stock is preceded by a period of significant underperformance. But the market seems to welcome the decision to issuance tracking stock, at least in the short run, as our results show that announcement is associated with a significant price increase. Postissuance, the issuing companies continue to underperform their industry peers, although the underperformance is smaller when compared to the preissuance period. An analysis of the operating performance reveals that tracking stock issuers are significantly less profitable, especially over the immediate 12 months prior to the issuance date. Furthermore, their earnings per share do not grow as fast as that of the control group. Finally, there is no evidence that tracking stock issuers are undervalued, contradicting comments made by some executives.

Keywords: Tracking stock; trackers.

1. Introduction

The issuance of tracking stock — also known as targeted stock, letter stock, or alphabet stock — involves separating all revenues and expenses of the division from the parent's financial statements and having its shares trade independent of those of the parent. These shares are distributed to existing shareholders in the form of a nontaxable stock dividend. The share price of the tracking stock will "track" the performance of a specific division. This issuance is similar to a spin-off, but without the parent having to relinquish control of the division.

There are many reasons — put forward by company executives, Wall Street analysts, and finance scholars — to explain the motivation for undertaking tracking stock issuance. Unlocking hidden value, reducing diversification discount, and increased Wall Street interest are some of the often-cited reasons (see Clayton and Qian, 2004; D'Souza and Jacob, 2000).

This paper tests the opposite argument to the widely held notion of "if it ain't broke, don't fix it," by arguing that if a company fixes its corporate structure, then something must be "broke." Put in another way, we posit that companies that issue tracking stocks must have underperformed their peers in various measures, and the issuance is an attempt to prop up their sagging stock prices. This argument is broadly consistent with other studies examining the motivation for other corporate restructurings, such as spin-off, equity carve-outs, sell-offs, etc. (see Rosenfeld, 1984; Shleifer and Vishny, 1992; Ofek, 1993; and Lang *et al.*, 1995; Vijh, 1999; Billet and Vijh, 2004).

Unlike these restructuring measures, however, issuing tracking is unlikely to realize significant long-term benefits. There are several reasons for this. First, there is no legal separation of the division from the parent. In fact, the parent's board of directors is also the tracked division's board. Second, it is based on the questionable idea that the market is inefficient, in that it is unable to properly value a conglomerate simply because it is made up of many disparate divisions.

To test the central argument of "fix it if it is broke," this paper compiles a sample of all tracking stock issuance announcements in the 1984–1999 period. It finds that poor stock and operating performance is a significant factor influencing the company's decision to issue tracking stock. The result is broadly consistent with other studies on corporate restructuring that these measures are used to salvage, resurrect, reverse, or address a company's poor operating performance. Indeed, very successful companies have been documented to have undertaken restructurings. Additionally, this study finds that the market is not impressed with this effort at restructuring, as indicated by the postissuance negative long-term stock performance.

This paper examines the stock performance of companies that issue tracking stocks over a 6-year period surrounding the announcement date. It finds that a majority of issuers experience significant declines in stock value, dating back 36 months prior to the tracking stock announcement. Over a 3-day window surrounding the announcement of tracking stock issuance, we find that the issuers experience a significant average excess return of 1.18%. But this positive performance appears to be short-lived, as over the 3-year postissue period, the issuers continue the long-term trend of underperformance.

To find clues to this underperformance, we compare these tracking stock issuers to their industry counterparts. We find that the issuers are significantly less profitable, as measured by net profit margin and return on equity. Again, the gap in profitability is especially pronounced over the 1-year period just before the announcement of the issuance of the trackers.

Another factor that may account for the declining stock value is the lack of growth in key areas. Our analysis shows that growth — most notably earnings-per-share growth — has stalled for quite a while prior to announcing the tracking stock issuance. This slowdown becomes much more apparent for the 1-year period prior to the issuance announcement, when tracking stock issuers suffered a 42% decline in EPS growth, whereas the control group has a 17.5% growth.

We investigated if tracking stock issuers are undervalued in relation to their nearest industry counterparts. We found no evidence to support the notion. For these two groups, the price-to-book ratio has remained comparable throughout the various preissuance periods. This finding apparently contradicts statements made by a number of company executives, who explain the reason to the issuing tracking stock as one of unlocking hidden value.

This paper contributes to existing literature in the following ways. First, this paper expands the time frame for the study of the market reaction to the tracking stock issuance announcement by examining both the short-term and the long-term performance of tracking stock issuers. With the exception of Billett and Vijh (2002, 2004), researchers have focused on a relatively narrow window surrounding the event date (see Hass, 1996; Logue *et al.*, 1996; Elder and Westra, 2000; Billett and Mauer, 2000; D'Souza and Jacob, 2000). Short-term event studies provide a valuable gage of the market reaction to the tracking stock announcement. But a longer-term study complements these event studies by examining the long-term and more permanent impact.

Second, it adds tracking stock issuance to the list of corporate restructurings undertaken to counter and to reverse poor corporate performance. Researchers have concluded that some financially distressed firms use voluntary selloffs to raise cash (see Rosenfeld, 1984; Shleifer and Vishny, 1992; Ofek, 1993; Lang *et al.*, 1995). Others find that underperforming conglomerates often decide to spin off noncore businesses to increase their focus (see Daley *et al.*, 1997; Desai and Jain, 1999). Evidence uncovered in this paper points to long-term stock underperformance, deteriorating profitability, and declining growth as possible motivations for issuing tracking stock, a type of corporate restructuring.

Third, our paper adds declining performance to the list of reasons motivating companies to issue tracking stocks. Billett and Mauer (2000) and Zuta

(1997, 2002) examine the reduction of diversification discount as a possible motive. Prompting this research is their observation that diversified firms normally trade at a discount to firms with standalone businesses. They find that issuing tracking stock — akin to creating a quasi-pure play — seems to reverse this diversification discount. D'Souza and Jacob (2000) find the greater availability of information on the tracked business (a Securities and Exchange Commission requirement), an increase in the level of outside monitoring, and more focused managerial motivation to be the contributors to the decision to issue tracking stock.

The rest of the paper is organized as follows. In Section 2, we will discuss the data collection procedure and the description of the two samples. Section 3 contains a discussion of the research methodology. Results are covered in Section 4. Section 5 summarizes the main results and lays out the paper's main conclusions.

2. Data Selection and Description

We gather two samples of firms to conduct our analysis. The first sample consists of firms that announce the creation of a tracking stock in the period 1984–1999.[1] Data on the tracking stocks issued prior to 1998 are primarily taken from D'Souza and Jacob (2000), Billett and Mauer (2000), and Billett and Vijh (2002).[2] The remaining candidate firms are identified through the Dow Jones News Retrieval Service, magazines, 10-K statements, SEC filing statements, and the Internet. When using various search engines on the Internet, we input keywords such as "tracking stocks" and "trackers."[3] From these sources, the following information is gathered: (1) the names of firms that announce the issuance of a tracking stock; (2) the date of this announcement; and (3) the size, in terms of the number of shares and the offering price, of the tracking stock. The search yielded 42 firms.[4] Table 1 provides some baseline information on these firms, and Table 2 describes these 42 tracking

---

[1]Sample period for year 2000 ends in June.

[2]D'Souza and Jacob's (2000) sample spans from 1984 to 1998, and Billett and Mauer's (2000) sample covers the period 1984–1996.

[3]A number of companies announce multiple tracking stocks on the same day. For the purpose of this study, this will count as one data point.

[4]We search press releases associated with the announcement of tracking stock issuance to find out if the reason for the issuance is made public. Most companies state, with different degree

Table 1. The number of tracking stock issuance.

Year	Tracking stock issuers
1984	1
1991	1
1993	4
1994	2
1995	4
1996	4
1997	4
1998	6
1999	16
Total	42

Note: This table provides information about the year and the number of companies announcing tracking stock issuance.

stock issuers, their ticker symbols, the announcement dates, and the tracked divisions.

The second sample is comprised the industry peers of the tracking stock issuers. This sample is constructed by matching each tracking stock issuing firm with its industry peer, first using the 2-digit standard industry classification (SIC) code and then using the market capitalization, if there are more than one company with the same SIC code. Table 3 provides some summary statistics on the tracking stock sample and the control sample. Tracking stock issuers have total assets that are on average slightly larger than those of their matched counterparts. In annual sales, the matched firms have a higher average. Market value for both samples is comparable.

Table 4 provides information on the tracking stocks in the year of their issuance. It indicates that the average total assets is $5.4 billion, but the median is $1.4 billion, indicating that most of the tracked companies are relatively small. These tracking firms have an average sales of $3.9 billion and market value of $5.7 billion. The median figures for these two measures are significantly smaller than their respective means.

We use Datastream to access the stock return information. Datastream is a commercial provider of domestic and international financial information.

of detail, that their stocks are undervalued and that the issuance will help unlock the hidden value. A few firms mention that the issuance is part of the ongoing restructuring effort.

Table 2. Tracking stock issuers and announcement date.

Tracking stock issuer	TS issuer ticker	Announcement date	Tracked division(s)
AT&T	T	03/10/99	Liberty Media
Agouron Pharmaceuticals	AGPH	08/01/98	Oncology Division
American Health Properties	AHP	02/01/95	Psychiatric Group
Andrx	ADRX	12/22/99	Different groups
Cendant	CD	09/30/99	Move.com
Circuit City Stores	CC	02/04/97	Carmax
CMS Energy Corp.	CMS	07/21/95	CMS Gas Group
Connectiv Inc.	CIV	03/02/98	Connective Class A
Delmarva Power & Light	DPL	08/12/96	Class A stock
Disney (Walt)	DIS	11/18/99	Go.com
Donaldson Lufkin & Jenrette	DLJ	05/26/99	DLJ Direct
E.I. Du Pont (DD)	DD	03/10/99	Life Sciences
Electronic Arts	ERTS	11/23/99	Different groups
Epitope	EPTP	11/07/96	Agritope, Medical Products stock
Fletcher Challenge	FLC	12/13/93	Forrest
GE	GE	11/30/99	NBCi
General Motors	GM	10/19/84	Electronic Data Sys.
Genzyme General	GENZ	11/17/98	GenzymeMolecular 98
Georgia Pacific	GP	12/17/97	Timber Group
Hemespherx Biopharma	HB	12/07/98	Separating divisions
Inco Ltd.	NYT	09/09/96	Inco Ltd. — Class VBN Shares
Incyte	INCY	08/18/98	Noncore, core businesses
J.C. Penney	JCP	05/19/99	Eckerd Drugstore
K-Mart	KM	01/04/94	Specialty stores
MCI	WCOM	08/02/96	Three targeted stocks
Perkin-Elmer	PKI	05/06/99	Celera Genomics
Pittston Co.	PZB	07/06/93	Mineral
Quantum	HDD	08/04/99	Quantum HDDG
RJR Nabisco	RN	02/01/93	Nabisco Group & Reynolds Group
Ralston Purina	RAL	08/02/93	Continental Baking
Seagull Energy	SGO	03/11/94	Enstar Alaska
Snyder Communication	SNC	10/29/99	Circle.com

(*Continued*)

Table 2. (*Continued*).

Tracking stock issuer	TS issuer ticker	Announcement date	Tracked division(s)
Soveriegn Bancorp	SVRN	08/03/99	Banking core, nonbanking businesses
Sprint	FON	11/24/98	PCS Wireless
Staples	SPLS	09/16/99	Core business & divisions
Tele Communications Inc.	TCOMA	08/11/95	TCI — Liberty Media Group
Telephone and Data System	TDS	12/19/97	Separate divisions
US West Co.	USW	10/27/95	US West Media (3)
USX Marathon	MRO	05/07/91	Steel Group
VIA (MTV)	VIA	04/23/97	MTVi
Ziff-Davis	ZD	03/31/99	ZDNet
Knight-Ridder	KRI	11/11/99	Publishing, others

Note: The table details information on the 42 tracking stock issuers, their ticker symbols, the announcement date, and the proposed tracked division(s).

Table 3. Summary statistics of the tracking stock issuers and the matched firms.

Variable	Sample	Mean	Median	Max	Min
Total assets	Tracking stock issuers	33282	9559	355935	696
	Matched firms	29850	7954	353890	615
Sales	Tracking stock issuers	15411	5391	111234	646
	Matched firms	16571	5652	132391	587
Market value	Tracking stock issuers	61630	6330	462694	658
	Matched firms	62305	7019	420252	950

Notes: The table contains information on the sample of tracking stock issuers and the matched sample. The three variables used for the comparison are total assets, annual sales, and market value. All figures are obtained for the year of issuance and are expressed in millions of dollars.

Stock return information is retrieved by typing its stock ticker symbol on a pull-down menu. For the market proxy, Datastream's total-market index is used. It contains virtually all the stocks publicly traded in the US.

Compustat is used to gather accounting information, such as net profit margin, return on equity, sales growth, EPS growth, and price-to-book ratio.

Table 4. Summary statistics on the tracking stocks.

Variable	Mean	Median	Maximum	Minimum
Total assets	5428.3	1408.8	18261.0	344.7
Sales	3895.8	1302.4	21289.0	20.3
Market value	5665.3	502.0	28519.8	12.8

Notes: The table contains information on the 28 tracking stocks (14 out of the 42 tracking stock issuers that have previously announced their intention to issue tracking stock did not go through with the issuance). The three variables used for the comparison are total assets, annual sales, and market value. All figures are obtained for the year of issuance and are expressed in millions of dollars.

3. Research Methodology

The long-term performance of tracking stock issuers prior to the announcement is determined using the matching firm methodology (see Barber and Lyon, 1997; Spiess and Affleck-Graves, 1995; Desai and Jain, 1999). The matched firms are chosen according to two criteria. First, the matched firm should have the same 2-digit SIC code as that of the tracking stock issuer. Second, if there are more than one matched firm with the same SIC code, we will choose the one with market value closest to the tracking stock issuer as of the tracking stock announcement date.

The efficiency test for significance of the stock price performance prior to tracking stock issuance employs the test of whether the average cumulative abnormal return (ACAR) is significantly different from zero. The ACAR tests whether the average actual return is equal to the average expected return.

$$\begin{aligned}
\mathrm{ACAR}_T &= \frac{\sum_{i=1}^{N}\left(\prod_{t=1}^{T}(1+r_{it})\right)-1}{N} - \frac{\sum_{i=1}^{N}\left(\prod_{t=1}^{T}(1+E(r_{it}))-1\right)}{N} \\
&= \frac{\sum_{i=1}^{N}\mathrm{CAR}_i}{N},
\end{aligned}$$

where r_{it} is the actual monthly return for stock i, N the number of firms in the sample, $E(r_{it})$ the monthly return on the matched firm, and CAR_i the cumulative abnormal return for stock i. Abnormal returns over T period are calculated by subtracting the T period buy-and-hold return of tracking stock sample from the T period buy-and-hold return of the matched firm on day T.

The statistical significance of the average holding period abnormal return (ACAR_T) for any given holding period T is determined using the following t-statistics:

$$t = \frac{\mathrm{ACAR}_T}{\mathrm{SE}_T}$$

where SE_T is the cross-sectional standard error of ACAR_T. Barber and Lyon (1997) and Kothari and Warner (1997) find the significance levels and the t-statistics computed using this matching firm approach to be well specified in random samples.

4. Results

4.1. *Holding-period returns*

In this section, we examine the stock performance of tracking stock issuers, spanning over a 6-year period surrounding the announcement. The results, using the holding-period return methodology and event study methodology, are summarized in Table 5. Panel A of the table contains the preissuance results. For year -3 (i.e., the period between 24 and 36 months prior to the issuance announcement), the tracking stock sample has a mean raw buy-and-hold return of 23.02%, whereas the matched sample has a mean raw buy-and-hold return of 28.54%. The difference, calculated as average cumulative abnormal return (ACAR), is significant at the 0.05 level. For year -2, the tracking stock issuers have a mean raw buy-and-hold return of 18.13%, which is lower than the 20.94% returned by the matched firms. The difference is not statistically different. For year -1, the tracking stock sample has a return of 9.03%, which lags significantly (at the 0.01 level) behind that of the matched sample, which produces a return of 17.27%.

Panel B of Table 5 shows the findings from the event study that examines the effect of tracking stock issuance announcement on the equity value of the firms. Over a 3-day window ($t = -1$ to $+1$, where $t = 0$ is the announcement date), the tracking stock issuers register an average raw return of 2.01%, whereas the market produces a raw return of 0.83%. The abnormal return, which is the difference between these two raw returns, is 1.18% (significant at the 0.05 level). This result indicates that the market, in general, has a positive perception of the decision to issue tracking stock.

Table 5. Average cumulative abnormal returns prior to tracking stock issuance announcement.

Time period	N	RAWTS	RAWM	ACAR	t-stat	% negative
Panel A: Preissuance cumulative and buy-and-hold abnormal returns						
$t = -3$	43	0.2302	0.2854	−0.0552	2.54*	55.80
$t = -2$	43	0.1813	0.2094	−0.0281	1.65	46.50
$t = -1$	43	0.0903	0.1727	−0.0824	2.95**	72.10
Sample	**N**	**RAWTS**	**MKT**	**AR**	**t-stat**	**% negative**
Panel B: Announcement period abnormal returns						
All issuers	43	0.0201	0.0083	0.0118	2.17*	39.53
Time period	**N**	**RAWTS**	**RAWM**	**ACAR**	**t-stat**	**% negative**
Panel C: Postissuance cumulative and buy-and-hold abnormal returns						
$t = +3$	28	0.1873	0.3262	−0.1389	2.89*	74.41
$t = +2$	39	0.1942	0.2429	−0.0487	1.24	58.14
$t = +1$	43	0.0854	0.1332	−0.0478	0.83	55.81

Notes: Panels A and C contain the raw buy-and-hold returns for tracking stock companies (RAWTS), raw buy-and-hold returns for matched firm (RAWM), the average cumulative abnormal returns (ACAR), the t-statistics associated with the ACARs, and the percentage of negative abnormal returns. These statistics are calculated for the sample of 43 firms that announce tracking stock issuance in the period 1984–2000. The matched sample is formed by matching each tracking stock firm with a firm with similar 2-digit SIC code that has a market value nearest to the tracking stock issuer as of the announcement date. For Panel A, the study period ranges from 1 to 3 years *before* the date of tracking stock announcement, where $t = 0$ is the year of tracking-stock issuance. Panel B reports the 3-day announcement period (day $t = -1$ to $+1$, where $t = 0$ is the issuance announcement day) raw returns, corresponding returns for Datastream equal-weighted market index (MKT) and the abnormal return (AR) for the entire sample of tracking stock issuers. For Panel C, the study period ranges from 1 to 3 years *after* the date of tracking stock announcement, where $t = 0$ is the year of tracking-stock issuance.
* Difference between the two samples is significant at the 0.05 level.
** Difference between the two samples is significant at the 0.01 level.

Panel C of Table 5 contains the postissuance results. For the first 2 years following the issuance announcement, the issuing companies' stock performance is not significantly different from that of the matched sample. But for year +3, the ACAR of −13.89% (significant at the 0.05 level) indicates that the issuers begin to underperform their industry peers, again.[5]

Overall, Table 5 indicates that the issuers of tracking stock underperform their industry benchmark over a long period. The only exception is the positive

[5]There is a decline in sample size in the years $t = +2$ and $t = +3$, because of a lack of full-period stock return data for issuances in years 1999 and 2000.

announcement effect. These results are broadly consistent with Billet and Vijh (2002).

4.2. *Profitability, growth, and valuation comparison*

Table 6 contains a series of comparisons between the tracking stock issuers and their matched counterparts. The first measure — net profit margin (NPM) — is frequently used as a proxy for corporate profitability. For year −3, there is no significant difference between the two samples. For year −2, however, the tracking stock sample has a net profit margin of 5.7%, a figure that

Table 6. Profitability, growth, and valuation comparison between tracking stock issuers and their industry peers.

Variable	Sample	Fiscal year relative to the announcement of tracking stock issuance					
		−3	−2	−1	+1	+2	+3
Net profit margin	TS issuers	5.8	5.7*	7.3**	4.86	−0.31**	3.79**
	Matched	6.8	3.8	12.5	4.37	2.78	−5.45
Return on equity	TS issuers	11.6	14.5*	20.3**	15.8	7.88*	11.85*
	Matched	11.4	18.7	32.1	12.62	10.79	18.8
One-year sales growth	TS issuers	18.3**	36.5*	25.2	10.15*	0.81**	14.4*
	Matched	24.3	26.0	27.1	20.39	26.91	23.35
One-year EPS growth	TS issuers	11.2*	0.7**	−42.0**	−87.76**	37.01*	−66.15*
	Matched	14.8	13.3	19.3	156.57	24.41	−187.99
Price-to-book ratio	TS issuers	4.25	5.68	7.98	3.38	3.84*	3.70
	Matched	4.76	6.56	8.52	4.05	5.28	5.83

Notes: The table reports the comparison between tracking stock issuers and their industry peers. The comparison period ranges from 3 years to 1 year *before* tracking stock issuance announcement and 1 year to 3 years *after* the tracking stock issuance announcement. The median value of five variables are used for this comparison: (1) net profit margin is income of a company after all expenses but before provisions for common and/or preferred dividends, divided by net sales, (2) return on equity is income before extraordinary items and discontinued operations less preferred dividend requirements, but before adding savings due to common stock equivalents, divided by common shareholders' interest in the company, (3) 1-year sales growth is net sales divided by the previous year value of net sales minus 1, (4) 1-year EPS growth is EPS basic excluding extraordinary items minus the previous year's values of EPS basic excluding extraordinary items divided by the absolute value of the previous year's values of EPS basic excluding extraordinary items, and (5) price-to-earnings ratio is the monthly closing price divided by 12 months moving EPS. (The definition of variables is from Compustat.)

* Difference between the two samples is significant at the 0.05 level using the Wilcoxon signed rank test.

** Difference between the two samples is significant at the 0.01 level using the Wilcoxon signed rank test.

is significantly higher, at the 0.05 level, than the matched sample. But this NPM advantage disappears in year −1, when they produce a margin of 7.3%, whereas the matched sample turns in a margin of 12.5%, an advantage that is statistically significant at the 0.01 level.

When we examine the postissuance NPM, we come to two conclusions. First, the margin has decreased following the issuance. Second, TS issuers' average margin is significantly lower than that of the matched sample for the first 2 years after the issuance. The year +3, however, sees a reverse in the NPM comparison, with the TS issuers significantly outperforming their industry peers.

The second measure is the return on equity (ROE), another gage of company profitability. For year −3, both samples have figures that are comparable. But for years −2 and −1, the tracking stock sample shows significantly lower ROE. In year −1, for instance, the tracking stock sample averages 20.3%, whereas the matched sample produces a statistically greater return (at the 0.01 level) of 32.1%.

The postissuance ROE comparison indicates that matched firms are significantly more profitable than the TS firms for years +2 and +3. This comparison is comparable to the preissuance ROE figures.

When we look at the third measure, the 1-year sales growth, the difference between the two samples is less conclusive. For year −3, the tracking stock sample has a slower growth, but this trend is reversed for year −2, when sales growth for the issuers of tracking stock significantly outpaces that for the matched firms. For year −1, there is no appreciable difference between the two samples.

The 1-year sales growth figures for the postissuance periods show that the matched firms significantly outsell the TS issuers for years +1, +2, and +3. The TS firms' sales growth for year +1, for instance, of 10.15%, is less than half that of the 20.39% realized by the matched firms. And the comparison for years +2 and +3 is equally significant. These findings are consistent with the widely held belief that a firm tends to issue tracking stock on divisions that are growing faster than the firm as a whole, thus allowing the divisions to reap the higher valuation given to high-growth stocks.

The fourth measure, the 1-year EPS growth, shows that there is a significant difference between the two samples for all three periods. For year −3, the tracking stock sample grows an average rate of 11.2%. This is on a significantly slower pace than the matched sample's growth of 14.8%. For year −2, the difference in growth is equally remarkable. The tracking stock firms grow an average of just 0.7%, significantly slower than 13.3% of the matched firms. For the 1-year period immediately preceding tracking stock announcement, the issuers have a mean EPS growth of −42.0%. This is significantly lower than the 19.3% of the matched firms.

The 1-year EPS growth numbers for the postissuance year +1 show that the TS firms have an average decrease of EPS of 87.76%, whereas over the same period, the matched firms' average EPS grows 156.57%. For year +2, the TS firms have a 37.01% EPS growth, versus 24.41% for the matched firms. And for year +3, both samples have EPS decline, with the matched firms declining by significantly greater degree.

The last measure in Table 4 examines the market price-to-book ratio of the two samples. For all three preissuance periods, the ratio for both samples has ranged from 4.25 to 8.52. There is no statistically significant difference between the two samples. And for the postissuance periods, the matched firms continue to have higher price-to-book ratio, although only the ratio for year +2 is significantly different.

5. Summary and Conclusions

Using a sample of 42 tracking stock issuance announcements during 1984–1999, this paper examines the long-run, as well as announcement period, stock returns for firms that issue tracking stock. It finds that firms that issue tracking stock suffer significant common stock underperformance in the preceding 3-year period, as measured against an appropriate benchmark. This underperformance reaches its highest level, with an average long-term excess return of −36.9%, in the immediate 12 months prior to the issuance announcement. We also find that the announcement of the issuance of tracking stock is associated with a significant excess return of 1.18%. A negative price drift characterizes the postissuance period, with the third year showing a significant underperformance.

Using net profit margin and return on equity as proxies, further analysis reveals that these tracking stock issuers are less profitable than their industry peers. Repeating the pattern seen in the stock return results, this profitability comparison worsens in the immediate 12 months prior to the announcement. Net profit margin for this period is 7.3% for the issuers and 12.5% for the control group. The issuers' average return of equity of 20.3% also is significantly lower than the 32.1% produced by the match sample.

Closely related to declining profitability is declining growth in earnings per share. The pattern of inferior performance is evident in the EPS numbers. The tracking stock issuers' shrank by 42.0%. Over the same period, the control group has a mean EPS growth of 19.3%.

An interesting finding: even though company executives often cite uncovering hidden value as a principal reason for tracking stock issuance, we could find no evidence to support this argument. When measured over the 3-year preissuance period, the tracking stock issuers have price-to-book ratios that are comparable to the matched firms.

This paper contributes to the existing literature in the following ways. First, it finds support for the overreaction hypothesis, which suggests that, because of overconfidence, the investors overreact to the news of issuance, when they mistakenly extrapolate the good news far into the future, without realizing that corporate performance tend to mean-revert. And as predicted by the hypothesis, this overreaction is followed by a long period of underperformance.

Second, this paper finds evidence to support the proposition that deteriorating operating results are, at least in part, responsible for a company's decision to issue tracking stock. In this respect, tracking stock issuance is similar to other voluntary corporate restructurings, such as equity carve-outs, spin-offs, and interfirm asset sales, which are designed to reverse declining corporate performance.

Acknowledgment

The Rider University Summer Fellowship provided financial support for this project.

References

Barber, B. M. and J. D. Lyon, "Detecting Long-run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics." *Journal of Financial Economics* 43, 341–372 (1997).

Billett, M. T. and D. C. Mauer, "Diversification and the Value of Internal Capital Markets: The Case of Tracking Stock." *Journal of Banking & Finance* 24, 1457–1490 (2000).

Billett, M. T. and A. M. Vijh, "*The Wealth Effects Of Tracking Stock Restructurings*." Working Paper, University of Iowa (2002).

Billett, M. T. and A. M. Vijh, "The Wealth Effects of Tracking Stock Restructurings." *Journal of Financial Research* 27(4), 559–583 (2004).

Clayton, M. J. and Y. Qian, "Wealth Gains from Tracking Stocks: Long-run Performance and Ex-date Returns." *Financial Management* 33, 83–106 (2004).

D'Souza, J. and J. Jacob, "Why Firms Issue Targeted Stock." *Journal of Financial Economics* 56, 459–483 (2000).

Daley, L., V. Mehrotra and R. Sivakumar, "Corporate Focus and Value Creation: Evidence from Spinoffs." *Journal of Financial Economics* 45, 257–281 (1997).

Desai, H. and P. C. Jain, "Firm Performance and Focus: Long-run Stock Market Performance following Spinoffs." *Journal of Financial Economics* 54, 75–101 (1999).

Elder, J. and P. Westra, "The Reaction of Securities Prices to Tracking Stock Announcements." *Journal of Economics and Finance* 20, 36–55 (2000).

Hass, J. J., "Directorial Fiduciary Duties in a Tracking Stock Equity Structure: The Need for a Duty of Fairness." *Michigan Law Review* 94, 2089–2177 (1996).

Kothari, S. P. and J. B. Warner, "Measuring Long-horizon Security Price Performance." *Journal of Financial Economics* 43, 301–339 (1997).

Lang, L., A. Poulsen and R. Stulz, "Asset Sales, Firm Performance, and the Agency Costs of Managerial Discretion." *Journal of Financial Economics* 37, 3–38 (1995).

Logue, D. E., J. K. Seward and J. P. Walsh, "Rearranging Residual Claims: A Case for Targeted Stock." *Financial Management* 25, 43–61 (1996).

Ofek, F., "Capital Structure and Firm Response to Poor Performance: An Empirical Analysis." *Journal of Financial Economics* 34, 3–30 (1993).

Rosenfeld, J. D., "Additional Evidence on the Relation between Divestiture Announcements and Shareholder Wealth." *Journal of Finance* 39, 1437–1448 (1984).

Shleifer, A. and R. Vishny, "Asset Liquidity and Debt Capacity." *Journal of Finance* 47, 1343–1366 (1992).

Spiess, K. and J. Affleck-Graves, "Underperformance in Long-run Stock Returns following Seasoned Equity Offerings." *Journal of Financial Economics* 38, 243–268 (1995).

Vijh, A. M., "Long-term Returns from Equity Carveouts." *Journal of Financial Economics* 51, 273–308 (1999).

Zuta, S., "*Diversification Discount and Targeted Stock: Theory and Empirical Evidence*." Working paper, New York University, New York (1997).

Zuta, S., "*Diversification Discount and Targeted Stock: Theory and Empirical Evidence*." Working paper, University of Maryland, Maryland (2002).

This page intentionally left blank

Chapter 11

Differences in Underpricing Returns Between REIT IPOs and Industrial Company IPOs

William Dimovski*
Deakin University, Australia

Robert Brooks
Monash University, Australia

This paper follows Chan, Stohs, and Wang (2001), which argues that the underlying value of the real estate is not by itself the reason for the very substantial differences in underpricing returns between real estate investment trust (REIT) IPOs and industrial company IPOs. We use variables identified in previous studies that have helped explain the underpricing of industrial company IPOs to help explain the underpricing of property trust IPOs. We find that the prospectus forecast dividend yield is a critical variable in the valuation and hence underpricing of REIT IPOs compared to industrial company IPOs. The sentiment towards the market and whether or not the issue is underwritten also impact the underpricing of REITs but the impact is much less than on industrial company IPOs.

Keywords: IPOs; REITs; underpricing.

1. Introduction

From the earliest reported studies, the magnitude of the underpricing returns for property trust or real estate investment trust (REIT) IPOs has been substantially different to the magnitude of the underpricing returns for industrial company IPOs. The low single-digit REIT underpricing returns are reported in the US in Ling and Ryngaert (1997), in European IPOs in Brounen and Eichholtz (2001), and in Australian IPOs in Dimovski and Brooks (2006). Handsome double-digit industrial company underpricing returns are reported in Loughran and Ritter (2004) for US IPOs, in Levis (1993) for UK IPOs, and in Lee *et al*. (1996) for Australian IPOs.

Wang *et al*. (1992) have offered three explanations for these differences in returns. First, more uninformed investors subscribe to REIT IPOs compared to industrial company IPOs; secondly, because REIT IPOs (before 1989) had

*Corresponding author.

to liquidate their holdings at some future point in time, this restricted their ability to grow, and thirdly REITs hold underlying real assets which support the IPO valuation. The first two of these explanations have been found not to be factors in post-1990 US REITs (Wang *et al.*, 1992; Ling and Ryngaert, 1997), leaving only the third explanation. However, Chan *et al*. (2001) investigated 56 Hong Kong real estate IPOs and 343 nonreal estate IPOs and find that the mean underpricing return of the real estate IPOs is comparable to the nonreal estate IPOs. They suggest that the underlying real estate explanation is not the whole solution either.

Although Jenkinson and Ljundqvist (1996) summarize many theoretical explanations for underpricing, Beatty and Ritter (1986) argue that the "need" for underpricing is the result of uncertainty about the value of the IPO before it lists. This has permitted researchers to examine underpricing by examining the financial and nonfinancial characteristics of IPOs. This study compares some financial and nonfinancial attributes of REITs and industrial company IPOs in Australia for the period 1994–1999. It specifically tests these attributes (or variables) identified in previous studies that have helped explain the underpricing of industrial company IPOs to help explain the underpricing of REIT IPOs. What this paper contributes is the finding that the prospectus forecast dividend yield makes a great deal of difference with the magnitude of the underpricing of REIT IPOs compared with industrial company IPOs.

The structure of this paper is as follows. Section 2 briefly summarizes some previous IPO research. Section 3 identifies the data and their sources. In Section 4, we report the results of our analysis. Section 5 contains some concluding remarks.

2. Some Previous IPO Research

One of the first major papers in the area of why the underpricing of REITs is different to the underpricing of industrial stocks was by Wang *et al.* (1992). In their study of 87 US REIT IPOs for the 1971–1988 period, they find a negative 2.82% underpricing return to the initial subscribers. It is hard to understand why subscribers invested in this primary equity capital. Even Wang *et al.* (1992) suggest that it may have been ignorance on the part of the subscribers. Industrial company IPO studies in the US have reported substantially higher underpricing returns. Some US studies of industrial company IPOs (Ibbotson, 1975; Ritter, 1987; Ibbotson *et al.* 1994; Loughran and Ritter, 2004) have reported underpricing returns of between 3.6% (1984) and 71.7% (1999).

Ling and Ryngaert (1997) extend Wang *et al.*'s (1992) work to investigate US REIT IPOs during 1991–1994. They report a 3.60% underpricing return on REIT IPOs. Ling and Ryngaert (1997) attribute this return to the greater involvement of the better informed institutional investors. They suggest that Rock's (1986) "winner's curse" may have operated in the property IPO market. The winner's curse theory suggests that better informed investors buy underpriced issues and do not offer to buy overpriced ones. Because of the rationing of the new issue capital, the better informed (and more influential) investors are able to buy a larger proportion of the "good" IPOs, whereas the less well informed (less influential) are able to buy a smaller proportion of the "good" issues and a larger proportion of the "not so good" issues—hence the winner's curse.

In the European context, Brounen and Eichholtz (2001) investigate the performance of 83 European property share IPOs over the period 1990–2000. They find an average market-adjusted underpricing return of 3.43%. Brounen and Eichholtz (2002) investigate a sample of 54 UK, French, and Swedish property share IPOs listed over the period 1984–1999. They find an average market-adjusted underpricing return of 2.55%. A range of initial day returns for industrial company IPOs in Europe varies from Leleux and Muzyka's (1993) study of French IPOs returning 4.2%, to Levis' (1993) 12% for UK IPOs to Alpalhao's (1992) 54.4% for Portuguese IPOs. Loughran *et al.* (1994) have compiled a comprehensive list of average initial day returns from industrial company IPOs around the world. These returns are updated on Jay Ritter's Web site, http://bear.cba.ufl.edu/ritter. Even in the European environment, the magnitude of the underpricing differences between industrials and REITs is substantial.

In Australia, Dimovski and Brooks (2006) report on 37 Australian property trust (REIT) IPOs from 1994 to 1999 and report an average 1.2% underpricing return. They find that some of the property trust IPOs have extremely low volumes of shares traded on the first day and so the simple use of a closing price at the end of the first day to determine underpricing returns (without reference to the volume of trading) may not always be the optimum method of calculating underpricing returns.

They suggest that a market-adjusted first-day closing price can be calculated using the second-day closing price discounted or inflated by the change in the property trust market index on the second day. Dimovski and Brooks (2006) find that the underpricing of property trusts can in part be explained by prospectus forecast profit distributions (or dividends) and the market sentiment towards property from the date of the prospectus to the date of listing. They argue that higher dividend forecasting trusts are riskier and hence higher underpricing returns are found in such trusts.

Industrial company IPO underpricing studies in Australia, however, report significantly higher average underpricing returns. Dimovski and Brooks (2004a) for the 1994–1999 industrial IPOs find an average 27.0% underpricing return. Earlier Australian industrial company IPO studies find average underpricing returns of around 20% (Finn and Higham, 1988; How *et al.* 1995; Lee *et al.*, 1996).

In another major study of property IPOs, Chan *et al.* (2001) investigate the Hong Kong IPO market during 1986–1997. Their sample includes 56 real estate IPOs and 343 nonreal estate IPOs. They find that the mean underpricing return of the real estate IPOs (16.21%) is comparable with the nonreal estate IPOs (18.96%) and argue that neither the underlying real estate holding nor uninformed investors explain the returns of REIT IPOs compared to industrial company IPOs. It is worth noting that the real estate IPOs may not have been specifically REIT IPOs, hence the substantially higher than previously reported underpricing returns from REITs.

3. Data and Methods

The *Connect 4 Company Prospectuses* database was used to identify the Australian IPOs from 1994 to 1999 and also to ascertain some financial and nonfinancial characteristics that might help explain underpricing. These characteristics are explained below. A common measure of underpricing return is the closing price of the shares or units (in REITs) on the first day of listing divided by the issue price to the public, minus 1. Because of the possible low-volume issue with the REITs, for robustness, we also calculate an underpricing return using the market-adjusted first-day closing price as discussed above for the 37 REITs. The closing prices were obtained from *IRESS* and the *Netquote Information Services* databases and some were checked with the *Financial Review* newspaper.

The regression model with the level of underpricing return (using day 1 and market-adjusted first-day closing prices for the REITs) as the dependent variable is

$$\begin{aligned}\text{RETURN} = {} & \beta_0 + \beta_1\,\text{REIT} + \beta_2\,\text{DIVYLD} + \beta_3\,\text{REITDIVYLD} \\ & + \beta_4\,\text{MKTSENTI} + \beta_5\,\text{REITMKTSENTI} + \beta_6\,\text{UWRITTEN} \\ & + \beta_7\,\text{REITUWRITTEN} + \beta_8\,\text{INDEPACC} \\ & + \beta_9\,\text{REITINDEPACC} + \beta_{10}\,\text{LNTOTMIL} \\ & + \beta_{11}\,\text{REITLNTOTMIL} + \beta_{12}\,\text{ISSUEPRI} \\ & + \beta_{13}\,\text{REITISSUEPRI} + \varepsilon, \end{aligned} \tag{1}$$

where RETURN is the underpricing return; REIT is a dummy variable to identify an REIT IPO, DIVYLD is the prospectus forecasted dividend yield for all IPOs, REITDIVYLD is the prospectus forecasted dividend yield for REIT IPOs only, MKTSENTI measures the change in the all-ordinaries index from the date of the prospectus to the date of listing for all IPOs, REITMKTSENTI measures the change in the all-ordinaries index from the date of the prospectus to the date of listing for REIT IPOs only, UWRITTEN is a dummy identifying if the IPO is underwritten, REITUWRITTEN is a dummy identifying the REIT IPO is underwritten, INDEPACC is a dummy identifying if the IPO had a top 5 accounting firm act as the independent accountant, REITINDEPACC is a dummy identifying the REIT IPO had a top 5 independent accountant, LNTOTMIL is the natural log of the capital raised by the IPO, REITLNTOTMIL is the natural log of the capital raised by the REIT, ISSUEPRI is the issue price of all IPOs; the REITISSUEPRI is the issue price of the REIT, the β's are unknown parameters to be estimated, and ε is assumed $\sim N\ (0, \sigma^2)$.

Essentially the interactive variables allow Equation (1) to be interpreted in the following manner. Each of the variables has impacts on the underpricing returns of the industrial IPOs with the partial coefficients β_2, β_4, β_6, β_8, β_{10}, and β_{12}. Each of the variables has impacts on the underpricing returns of the REIT IPOs with the sum of the two partial coefficients linked to each variable, that is, $\beta_2 + \beta_3$ to DIVYLD, $\beta_4 + \beta_5$ to MKTSENTI, $\beta_6 + \beta_7$ to UWRITTEN, $\beta_8 + \beta_9$ to INDEPACC, $\beta_{10} + \beta_{11}$ to LNTOTMIL, and $\beta_{12}+\beta_{13}$ to ISSUEPRI. The *difference* in underpricing returns between REITs and industrials is therefore explained by the coefficients β_3, β_5, β_7, β_9, β_{11}, and β_{13}.

The DIVYLD variable is included because it was found significant in the Dimovski and Brooks (2004a) study of industrial IPOs and in the Dimovski and Brooks (2006) study of REIT IPOs. The MKTSENTI variable is included because it was found significant in Dimovski and Brooks (2004a). It was found that as the MKTSENTI (all-ordinaries index) increased from the date of the prospectus to the date of listing, so too did the underpricing return increase. The UWRITTEN variable is included because it was found significant and positive in the Dimovski and Brooks (2004b) study investigating the influence of underwriters on the amount of money left by industrial company IPOs. Money left is defined as the underpricing return multiplied by the number of shares issued in the IPO. Although Beatty (1989) identified that underpricing was significantly reduced where issuers used high-reputation auditors, auditors were not always nominated in our data set. This study instead includes an INDEPACC variable identifying a top 5 independent or investigating accountant's report in the prospectus. LNTOTMIL is included because both Ibbotson *et al.* (1994) and Michaely and Shaw (1994) report larger IPOs experience lower underpricing returns. ISSUEPRI is included because Chalk and Peavy (1987) and Ibbotson *et al.* (1994) find lower issue price IPOs are more underpriced.

4. Results

Table 1 reports some summary descriptive results. For the 1994–1999 period, 37 property trust IPOs and 262 industrial company IPOs were listed on the Australian Stock Exchange. The total amount of public equity capital raised by the property trusts was around A$5.7 billion, whereas around A$22.1 billion

Table 1. Characteristics of Australian industrial and property trust IPOs, 1994–1999.

Category	IPOs	Public equity raised A$ billions		Day 1 % return available to investors	Mkt-Adjusted day 1 % return available to investors	Forecast distrib/ divident yield (%)
Property trusts	37	5.7	Mean	1.2	−0.6	8.9
			Median	0.0	0.0	9.2
			Minimum	−20.0	−20.0	0.0
			Maximum	15.8	15.8	12.5
Industrials	262	22.1	Mean	27.0		3.3
			Median	10.0		2.4
			Minimum	−67.3		0.0
			Maximum	370.0		14.0

was raised by the industrials. The mean first-day return available to the subscribers was 1.2% for the property trusts and 27.0% for the industrials. The median returns were 0.0% and 10%, respectively. The maximum first-day returns in this sample period were 15.8% and 370.0% and the minimum returns were −20% and −67.3%, respectively. The market-adjusted first-day return for the REIT IPOs shows a −0.6% return, whereas the median, minimum, and maximum figures are as on day 1. Forecast distribution (or dividend) yields for the full year following the IPO are also included. The average dividend yield for the REITs is 8.9%, whereas it is only 3.3% for the industrials.

Table 2 reports the multiple ordinary least-squares regression results between day 1 and market-adjusted day 1 underpricing returns and the selected explanatory variables for the 299 IPOs (262 industrials and 37 REITs) in the 6-year period. To reduce the influence of outlier return results, Table 2 also reports winsorized results (those IPOs whose returns exceeded the 99th percentile have been scaled back to a return at the 99th percentile) and results where outliers have been removed totally from the OLS regression. There were three industrial company outliers whose underpricing returns exceeded 3.5 standard deviations from the average (consistent with How, 2000) that were removed from the data set. Standard regression diagnostics were calculated. In testing for nonnormal errors, a Jarque–Bera statistics is applied to the data. In testing for heteroscedasticity, a White test is applied and reported where necessary. In testing for omitted variables or model misspecification, a Ramsey Reset test is applied and reported.

Our results suggest that without the winsorizing of the outliers, or their removal, there are no significant differences between REITs and industrials in terms of the characteristics that may influence their underpricing. Once the outliers are winsorized or removed, however, there are differences with respect to the impact of DIVYLD, MKTSENTI, and UWRITTEN. For DIVYLD, the effect is a sign change and lower absolute magnitude. For MKTSENTI and UWRITTEN, the effect is a reduction in magnitude.

It appears that although underpricing is expected from industrial company IPOs (because of uncertainty about their valuation which in turn is due to uncertainty about their future cash flows) and that the disclosure of certain variables helps decrease this uncertainty and underpricing, REIT IPOs are viewed differently. The assets that the REITs own are likely to be income-producing already and also likely to have longer-term lease agreements with tenants in place. Such characteristics suggest some certainty about an REIT's

Table 2. Underpricing of Australian industrial and property trust IPOs, 1994–1999 OLS results.

Variable	All 299 IPOs All 299 IPOs				Outliers winsorized				Outliers removed			
	Day 1		Mkt Adj Day 1		All 299 IPOs winsorized day 1		All 299 IPOs wins'd mkt-adj day 1		296 IPOs day 1		296 IPOs mkt-adj day 1	
	Coef.	Pr.	Coef.	Pr.	Coef. *	Pr. *	Coef. *	Pr. *	Coef. *	Pr. *	Coef. *	Pr. *
C	0.277	0.003	0.277	0.003	0.229	0.005	0.229	0.005	0.212	0.003	0.212	0.003
REIT	−0.411	0.556	−0.548	0.432	−0.362	0.001	−0.499	0.000	−0.345	0.001	−0.482	0.000
DIVYLD	−3.061	0.000	−3.061	0.000	−2.643	0.000	−2.643	0.000	−2.388	0.000	−2.388	0.000
REITDIVYLD	3.747	0.340	4.533	0.249	3.330	0.000	4.116	0.000	3.074	0.000	3.861	0.000
MKTSENTI	2.058	0.000	2.058	0.000	1.904	0.000	1.904	0.000	1.769	0.000	1.769	0.000
REITMKTSENTI	−1.752	0.369	−1.682	0.389	−1.598	0.002	−1.527	0.004	−1.463	0.003	−1.392	0.007
UWRITTEN	0.165	0.028	0.165	0.028	0.178	0.003	0.178	0.003	0.189	0.000	0.189	0.000
REITUWRITTEN	−0.149	0.646	−0.156	0.631	−0.162	0.012	−0.169	0.011	−0.173	0.004	−0.180	0.003
INDEPACC	−0.008	0.907	−0.008	0.907	−0.047	0.441	−0.047	0.441	−0.080	0.185	−0.080	0.185
REITINDEPACC	0.059	0.868	0.054	0.881	0.099	0.215	0.093	0.269	0.131	0.095	0.126	0.132
LNTOTMIL	−0.035	0.216	−0.035	0.217	−0.012	0.599	−0.012	0.599	0.004	0.811	0.004	0.811
REITLNTOTMIL	0.047	0.589	0.064	0.466	0.025	0.296	0.041	0.105	0.009	0.651	0.025	0.238
ISSUEPRI	0.025	0.671	0.025	0.671	−0.002	0.970	−0.002	0.970	−0.033	0.368	−0.033	0.368
REITISSUEPRI	−0.060	0.851	−0.074	0.817	−0.033	0.521	−0.047	0.431	−0.002	0.963	−0.016	0.770
R^2/Adj R^2	0.140	0.101	0.145	0.106	0.170	0.133	0.176	0.139	0.169	0.131	0.175	0.137
J_B test	4238.835	0.000	4229.516	0.000	328.272	0.000	326.971	0.000	453.364	0.000	451.400	0.000
White test	27.689	0.149	27.675	0.150	37.723	0.014	37.656	0.014	33.772	0.038	33.710	0.039
Reset test	0.877	0.235	0.873	0.237	0.641	0.241	0.636	0.244	0.739	0.207	0.732	0.241

* White heteroscedasticity consistent coefficient and p-values reported.

future cash flows and therefore its valuation. These “more certain” cash flow characteristics would help explain the reduction in magnitude of the MKT-SENTI and UWRITTEN coefficients for REITs.

In addition, Australian REITs need to pay out their profits or suffer punitive tax rates of 50% on any undistributed profits, whereas there is no obligation on industrials to pay any dividends at all. Compared with industrial company IPOs, the combination of more certain cash flows and likely consistently larger dividend payments by REITs would help explain the lower magnitude of the DIVYLD coefficient for REITs. The sign change for REITs to positive for the DIVYLD coefficient is consistent with Beatty and Ritter (1986). With REITs, the dividend yield is essentially the earnings yield (because of the punitive tax rates on nondistributed earnings). The earning yield represents the return from the investment, at least in the short term. Given the relationship between return and risk, it would be expected that the higher the DIVYLD, the greater the risk about an IPO, the higher the underpricing. Again this is different to industrial company IPOs because industrials do not have to pay dividends (let alone their full earnings), this risk return argument cannot be used on them. Indeed, because forecasted dividend payments by industrials are not common and not large, they are viewed as a proxy for reducing uncertainty and hence underpricing.

5. Conclusion

This study investigated 37 property trust and 262 industrial company IPOs in Australia for the period 1994–1999. The magnitude of the difference in underpricing returns for both classifications of IPO is investigated and like previous international studies, we report that the underpricing returns from each type of IPO are quite different and that there are differences in the impact of the DIVYLD, MKTSENTI, and UWRITTEN variables on the level of underpricing.

What this paper examines is the perilously low underpricing returns (or even negative returns using market-adjusted first-day closing prices) made by the subscribers to REIT IPOs. What this paper contributes is an explanation of the magnitude of importance the prospectus forecast dividend makes to the underpricing of REIT IPOs compared with industrial company IPOs. We find that the forecast dividend yield is an important element of the underpricing return in REIT IPOs, whereas it is clearly not as vital for industrial company issues. If, as Chan *et al.* (2001) argue, the underlying value of the real estate may not provide a clear base of support for an REIT IPOs valuation, we suggest that the forecast dividend yield will help guide the valuation.

Acknowledgment

The authors thank an anonymous referee for his helpful comments on previous versions of this paper.

References

Alpalhao, R. M., "*Opertas Publicas Iniciais: O Caso Portugues*." Working Paper, Univesidade Nova De Lisboa (1992).

Beatty, R. P., "Auditor Reputation and the Pricing of Initial Public Offerings." *The Accounting Review* 64, 693–709 (1989).

Beatty, R. P. and J. R. Ritter, "Investment Banking, Reputation, and the Underpricing of Initial Public Offerings." *Journal of Financial Economics* 15, 213–232 (1986).

Brounen, D. and P. Eichholtz, "Going Public in the European Property Share Market." *Real Estate Finance* 18, 55–60 (2001).

Brounen, D. and P. M. A. Eichholtz, "Initial Public Offerings: Evidence From British, French and Swedish Share Markets." *Journal of Real Estate Finance and Economics* 23, 103–117 (2002).

Chalk, A. J. and J. W. Peavy III, "Initial Public Offerings: Daily Returns, Offering Types and the Price Effect." *Financial Analysts Journal* 43, 65–69 (1987).

Chan, S. H., M. H. Stohs and K. Wang, "Are Real Estate IPOs a Different Species? Evidence from Hong Kong IPOs." *The Journal of Real Estate Research* 21, 201–220 (2001).

Dimovski, W. and R. Brooks, "The Pricing of Property Trust IPOs in Australia." *Journal of Real Estate Finance and Economics* 32, 185–199 (2006).

Dimovski, W. and R. Brooks, "Initial Public Offerings in Australia 1994 to 1999, Recent Evidence of Underpricing and Underperformance." *Review of Quantitative Finance and Accounting* 22, 179–198 (2004a).

Dimovski, W. and R. Brooks, (2004b), "Do You Really Want to Ask an Underwriter How Much Money You Should Leave on the Table?" *Journal of International Financial Markets, Institutions and Money* 14, 267–280 (2004b).

Finn, F. J. and R. Higham, "The Performance of Unseasoned New Equity Issues-Cum Stock Exchange Listings in Australia." *Journal of Banking and Finance* 12, 333–351 (1988).

How, J. C. Y., H. Y. Izan and G. S. Monroe, "Differential Information and the Underpricing of Initial Public Offerings: Australian Evidence." *Accounting and Finance* 35, 87–105 (1995).

How, J. C. Y., "The Initial and Long Run Performances of Mining IPOs in Australia." *Australian Journal of Management* 25(1), 95–118 (2000).

Ibbotson, R. G., "Price Performance of Common Stock New Issues." *Journal of Financial Economics* 2, 235–272 (1975).

Ibbotson, R. G., J. L. Sindelar and J. R. Ritter, "The Market's Problems with the Pricing of Initial Public Offerings." *Journal of Applied Corporate Finance* 7, 66–74 (1994).

Jenkinson, T. and A. Ljundqvist, *Going public: The Theory and Evidence on How Companies Raise Equity Finance*, Clarendon Express, Oxford (1996).

Lee, P. J., S. L. Taylor and T. S. Walter, "Australian IPO Pricing in the Short and Long Run." *Journal of Banking and Finance* 20, 1189–1210 (1996).

Leleux, B. F. and D. F. Muzyka, "The Demise of European IPO Market: a Post-issue Performance Study." Working Paper, INSEAD, USA (1993).

Levis, M., "The Long-Run Performance of Initial Public Offerings: the UK Experience 1980–1988." *Financial Management* 22, 28-41 (1993).

Ling, D. C. and M. Ryngaert, "Valuation Uncertainty, Institutional Involvement, and the Underpricing of IPOs: The Case of REITs." *Journal of Financial Economics* 43, 433–456 (1997).

Loughran, T. and J. Ritter, "Why Has IPO Underpricing Changed Over Time." *Financial Management* 33, 5–37 (2004).

Loughran, T., J. R. Ritter and K. Rydqvist, "Initial Public Offerings: International Insights." *Pacific-Basin Finance Journal* 2, 165–199 (1994).

Michaely, R. and W. H. Shaw, "The Pricing of Initial Public Offerings: Tests of Adverse-Selection and Signaling Theories." *Review of Financial Studies* 7, 279–319 (1994).

Ritter, J. R., "The Costs of Going Public." *Journal of Financial Economics* 19, 269–281 (1987).

Rock, K., "Why New Issues are Underpriced." *Journal of Financial Economics* 15, 187–212 (1986).

Wang, K., S. H. Chan and G. Gau, "Initial Public Offerings of Equity Securities: Anomalous Evidence Using REITs." *Journal of Financial Economics* 31, 381–410 (1992).

White, H., "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." *Econometrica* 48, 817–838 (1980).

This page intentionally left blank

Chapter 12

Performance of Canadian Mutual Funds and Investors

Rajeeva Sinha*
University of Windsor, Canada

Vijay Jog
Carleton University, Canada

The study examines the performance of a comprehensive sample of Canadian open-end equity mutual funds and investors. Our results show that while the majority of funds outperform their self-selected benchmarks, the performance is lackluster in comparison with some well-recognized benchmarks like the TSE 300 and the 90-day T-Bill rates. These returns are even lower when one accounts for the timing of entry and exit by mutual fund investors. We also find that returns of mutual funds are adversely affected by active trading and advisory and non-advisory expenses are negatively related to performance. Accordingly, we conclude that investors are likely to be better off by following a passive and index-based investment approach in the long term.

Keywords: Open-end mutual funds; mutual fund performance; investor returns.

JEL Classification: D14, G23.

1. Introduction

With nearly $440 billion in assets and 51 million account holders by the end of year 2003 in Canada, mutual funds now occupy a prominent position among financial intermediaries. The 1990s witnessed an explosive growth in mutual funds in Canada; the number of accounts grew nearly 10-fold during this period. Similar growth in mutual fund assets has been reported in many countries around the world.

This phenomenal growth notwithstanding, there continue to be serious concerns about the value added by actively managed mutual funds and the ability of investors to earn superior risk-adjusted returns from their mutual fund investments. This has been documented by the pioneering work of Jensen

*Corresponding author.

(1968) in the US followed by Malkiel (1995), Elton *et al*. (1996), and Gruber (1996). Studies by Odean (1998) and Barber *et al*. (2003) of investors' trading activity, and Sirri and Tufano (1998) of fund flows also suggest that investors perform even worse because of their proclivity to chase winners and their reluctance to let go of losers in their choice of mutual funds. This study examines the performance of Canadian mutual funds and mutual fund investors. The objective is to assess the value added by money managers and the rates of returns realized by mutual fund investors. We also analyze a number of fund related characteristics, like expenses, trading behavior, loads and portfolio composition of mutual funds. This study also provides evidence on the magnitude and implications of the difference between returns to mutual fund investors (IRR) and the returns reported by mutual funds (RR). The study also augments the rather limited systematic evidence on Canadian mutual fund industry. Earlier studies have shown that mutual fund lacks performance persistence (Malkiel, 1995); investors chase past winners (Sirri and Tufano, 1998); and are reluctant to exit loosing funds (Odean, 1998). The cumulative effect of these attributes of mutual funds and investor behavior will be RR is likely to be greater than IRR.

Our study also contributes by providing non-US evidence on mutual fund performance. The need for a widening of the evidence on managerial value added in the mutual fund industry beyond the numerous US studies using overlapping data is noted in the academic literature. As Khorana *et al*. (2005) point out, academic studies of mutual funds have remained geographically narrow.

The study is organized as follows: Section 2 provides a literature review of the discussion on performance and the trading behavior of mutual fund investors. We also review the existing literature on Canadian mutual funds. Section 3 examines some methodological and measurement issues that underpin the validity of the findings. Section 4 discusses the sample. Section 5 reports the findings. Section 6 concludes the study.

2. Literature on Performance and Trading Behavior of Mutual Funds

Despite the growth and popularity of the mutual fund industry, studies focused on US mutual fund industry are unable to conclude whether the active money management adds value to individual investors net of risk and expenses. In a pioneering study, Jensen (1968) concluded that mutual funds significantly underperformed the market after expenses and those investors would be better off pursuing a passive investment strategy by following a comparable market proxy. However, later studies by Grinblatt and Titman (1989a, b, 1992) and Ippolito (1989) on the net performance of mutual funds concluded that mutual fund managers did add value net of expenses because of the private information that money managers possessed. These studies were however, criticized for their choice of benchmarks in assessing performance and for survivorship bias in that they included in their sample only current and existing funds. Later studies including Malkiel (1995), Elton *et al.* (1996), Gruber (1996), and Elton *et al.* (1993) concluded that the findings of Grinblatt and Titman (1989a,b, 1992) and Ippolito (1989) on positive value added by money managers did not hold when more representative benchmarks are used and adjustments are made for the potential survivorship bias. In a recent paper, Bhargava *et al.* (2001) evaluated the performance of 114 US international equity managers and found that international equity managers, on average, were also unable to outperform the MSCI World market proxy during the sample period 1988–1997.

While these studies have typically concentrated on the reported returns by mutual funds, three strands of literature claim that returns to mutual fund investors (IRR, hereafter) may be even lower than the returns reported by mutual funds (RR, hereafter). The first group of studies by Chevalier and Ellison (1997), Sirri and Tufano (1998) analyze the sensitivity of capital flows into funds as a function of past performance and provide extensive evidence in support of an inverse relationship between past performance and current fund flows. Barber *et al.* (2003), in a study of trading behavior of more than 30,000 households, find that investors use past returns as a positive signal of fund quality and future performance but do not necessarily receive higher returns.[1]

The strategy of investing in outperforming funds has also been described as the "hot hands" phenomenon. Hendricks *et al.* (1993), Goetzmann and Ibboston (1994), and Brown and Goetzmann (1995) suggest that mutual funds that show above average performance in one period will also follow it up with above average performance in the following period. Thus, according to these

[1]This is also termed representative heuristic in behavioral finance. An above average performance by a mutual fund in the previous year is likely to induce greater inflow of funds in the current year. As good performance is often followed by sub par returns, chasing past performance has led to subsequent sub par performance in the 1990s.

studies mutual fund investors will get higher returns if they were to choose mutual fund investors that are past winners. However, Malkiel (1995) in a study of US mutual funds found that while there appeared to be persistence of returns in the 1970s, there was no significant persistence of returns during the 1980s. In the 1980s, the performance decay was the typical characteristic and past performance was no predictor of future performance. The evidence on persistence can be tested directly as performed in this paper. IRR will be greater than RR if there is performance persistence and less than RR in the absence of performance persistence.

Finally, a study by Odean (1998) documents the reluctance by investors to realize losses. This loss aversion also has the implication of widening the gap between RR and IRR. Using a unique data set on the trading behavior of 30,000 households, Odean (1998) found that investors are reluctant to realize losses by selling underperforming funds. This is an example of the disposition effect (Shefirin and Statman, 1985). The combined implication of the evidence on investors chasing past winners, lack of performance persistence, and reluctance to realize losses implies that the IRR would be lower than RR. Investors are likely to buy into funds that have performed well in the past, fail to find persistence in its performance, and are unwilling to book losses by exiting the funds. In addition, the evidence that most investors who sell shares are likely to sell them for reasons unrelated to portfolio asset reallocation; it also seems highly likely that IRR will be less than RR for most investors.

Nesbitt (1995) examined this impact of market timing by mutual fund investors by compiling the dollar-weighted returns of 17 categories of mutual funds and found that the dollar-weighted returns (equivalent to IRR) were less than the time-weighted returns (equivalent to RRs) for every category of mutual funds. Nesbitt (1995) concluded that investors suffer a shortfall in return because of ill-timed movement of funds.

The studies on mutual funds in countries other than the US, report mixed evidence on their performance relative to the market. Otten and Barms (2002) find that in four out of five European countries mutual funds displayed positive after cost alphas and significantly outperformed the market after adding costs.[2] Blake and Timmermann (1998) in their study of UK mutual funds found some evidence of underperformance (by 1.8%) on a risk-adjusted basis for the average mutual fund manager in line with the findings of studies on

[2]Alphas represent excess returns based on the Jensen (1968) article.

US mutual funds. They report an interesting relationship between termination dates, establishment dates of funds and their performance. The underperformance of mutual funds intensified as the funds termination dates approached and they weakly outperformed the market in the first year of their formation. Dahlquist *et al*. (2000) in their study of Swedish mutual funds report that small equity funds, low fee funds and funds having a high trading activity performed better than the market. Cai *et al*. (1997), in their study of Japanese Mutual funds, found significant underperformance and attributed this to institutional factors and tax provisions in the mutual funds industry in Japan. They also report that the turnover rate in mutual funds was twice that of the Tokyo Stock Exchange.

These studies using data on non-US mutual funds cannot be used to make comparisons with the findings on US mutual funds as they represent a different institutional reality. The studies on Japanese mutual funds (Cai *et al.* 1997; Brown *et al*. 2001) and on Swedish funds (Dahlquist *et al.* 2000) are examples of the effects the institutional environment can have on mutual fund behavior and performance in that country. Blake and Timmermann (1998) and Otten and Barms (2002) suggest that because of their smaller market importance the European mutual funds excluding the UK may be in a better position as a group to beat the market than US funds. This makes the findings of these studies specific to the mutual fund industry in that country.

Our study presents out of sample evidence from a market similar to the US on mutual funds and can be used to generalize the findings on US funds. The mutual fund industry in Canada faces a market and regulatory framework that is similar to US mutual funds in that the two markets function under very similar legal and regulatory framework; the characteristics of the retail investing public and the trading costs and turnover of the equity market have a close resemblance.[3] Some of the supply side characteristics like the structure

[3]Typically, Canada follows a very similar regulatory regime as that of the US although by a lag of one or two years. Many large Canadian companies are interlisted on the US stock exchanges and thus are subject to the SEC regulations. In Ontario (which is Canada's largest province), the Ontario Securities Commission (OSC) closely follows the SEC changes so that Canadian regulatory framework remains competitive to that in the US Similar observations can be made about the closeness of the stock markets. Typical correlations between the TSE 300 and the S&P 500 range from 0.78 for monthly returns from January 1995 to December 2004 and 0.80 monthly returns from January 1998 to December 2002.

of the banking industry are different but the breadth of the distribution channels and other characteristics as outlined by Khorana *et al.* (2005) are similar. Thus, this study will facilitate further generalizations on the role of mutual funds as financial intermediaries.

There have been some earlier studies on Canadian mutual finds. For example, Brown and Young (2002), in their evaluation of a small sample of Canadian mutual funds, report that on average only one in 17 investors actually make money based on the timing of their entry and exit from mutual funds.[4] Deaves (2004) provides an evaluation of 300 Canadian equity funds for the period 1988–1998. He finds that on average, Canadian mutual fund managers underperform benchmarks net of expenses but finds evidence of performance persistence. More specifically, Deaves (2004) finds that funds with alphas above their mean values are likely to follow with alphas above mean values in subsequent years.

Our study extends and enriches this evidence by making three important contributions. First, we believe that the sample period of 1988–1998 used in Deaves (2004) study may simply be a reflection of the 1990s decade which was a period of above average returns in the stock markets where choosing a single strategy (e.g., tilting the portfolio in growth stocks) could have delivered superior and persistent performance. This could explain the high levels of performance persistence in Canadian equity mutual funds in Deaves (2004) study. To be able to draw observations on the performance of mutual funds that are robust of business cycles and stock market fluctuations, we need studies based on longer sample period and a comprehensive sample. Our analysis of 914 Canadian equity funds represents a near complete sample of all Canadian equity funds established until 2001. The extended and near complete sample period makes it more representative and allows us to draw inferences that are robust to different stages of the business cycles and market conditions. Second, we analyze the investor-realized returns. Third, we examine turnover, expenses age of funds, portfolio allocation and a variety of other fund related characteristics, and their relationship with mutual fund performance.

[4]The older Canadian studies include those by Calvett and Lefolle (1980) who studied the quarterly returns of 17 mutual funds for the period 1966–1975 and found that 16 of these funds performed worse than the market index as proxied by the TSE 300. In a later study, Bishara (1988) examined the performance of Canadian mutual funds for the period 1967–1984 and reported that only the growth funds outperformed the market although the balanced and income funds did beat the market during one period of rising market.

3. Assessing Performance and Persistence

A key objective of our study is to investigate the performance and the persistence in performance of mutual funds in Canada. Accordingly, we compute a number of performance measures to evaluate performance of Canadian equity funds. Performance is measured until the year ending 2002 for alive funds and for the last year of operation for dead funds. Averages of 2, 3, 5, and 15 years for all the performance measures have been computed, backtracking from the year 2002. As our sample comprises of all funds established until 2001 and we calculate both long-term and short-term performance based on averages over 2, 3, 5, 10, and 15 years, the inferences drawn from the tables are robust to changes in market sentiments and different stages of the business cycles. This allows us to test the extent of performance persistence in Canadian equity mutual funds. In addition, we present data on alive or current funds and all funds (inclusive of dead funds) separately throughout the discussion to test the importance of the potential survivorship bias.

3.1. *Performance of mutual funds*

We evaluate performance of mutual funds by computing a number of performance measures. First, we report raw returns or RR defined as the percentage change in the fund's value for the period, including dividends and net of expenses. The use of raw returns or RR is in line with Brown *et al*. (1996) and Chevalier and Ellison (1997) who have shown that peer group or within sector comparisons of raw returns provide a valid basis for the assessment of managerial effort in the mutual fund industry. In addition to raw returns, we also use Jensen's alpha and Sharpe's risk adjusted measure of performance.[5]

3.2. *Performance of mutual fund investors*

As pointed out in Section 2, we plan to test for the effects of the timing of investors purchase and sales of mutual fund units through the use of IRR.

[5]Alpha is a measure of the difference between a fund's actual monthly excess return and its expected monthly excess return. Alpha is calculated by using a single factor model using data for returns net of expenses. Sharpe's risk-adjusted measure of performance is the ratio of a fund's excess return to its standard deviation.

The formula for calculating *IRR* is

$$\sum_{0\to ncf}^{n} \frac{CF_n}{(1+IRR)^n} = 0$$

where, CF_n is the Cash Flow in period n, *IRR* the internal rate of return, and n is the number of Periods.

The above formula provides the monthly *IRR*. To annualize *IRR*, the following calculation is used[6]: Annualized $IRR = (1 + IRR)^{12} - 1$. As in the case of *RR*, *IRR* is calculated for the years 1 and the average of years 2, 3, 5, 10, 15.

3.3. *Performance persistence*

We use the approach of Goetzmann and Ibboston (1994) and Malkiel (1995) in assessing performance persistence. A winner (loser) is defined as a fund that has achieved a rate of return over the calendar year that exceeds (is less than) the median fund return. Performance persistence or "h*ot hands*" occurs when winning is followed by winning in the subsequent year(s). Thus if a winner continues to post returns greater than the median returns in the years 2, 3, and 5, we include it among repeat winners. We follow each fund across up to 5 years to investigate the persistence in performance. We measure performance persistence for each of the years 1970–2001.

4. Data

Our data come from a comprehensive sample of 914 Canadian open-end mutual equity funds with a market value of $103.95 billion at the end of year 2002. The data set provided by *Fundata* and *Fundmonitor.com* includes both alive and dead funds.[7] The oldest fund for which we have record was

---

[6]As an example, suppose an investor made just two transactions in his portfolio over a 12-year period. The initial investments of $10,000 were made on January 1, 1986 and let us assume that the portfolio grew by 15% per year for the next eight years. Subsequently, another $500,000 was added on January 1, 1994. Let us assume that in the two years following the second investment, the portfolio fell in value by a total of 20%. On January 1, 1996, the overall value of the portfolio would stand at $424,472. The cumulative (simple) return would read −17% while the Internal rate of return (IRR) would be a much lower −58%. The IRR figure reflects the fact that most of the money was invested at a high and a large portion of it was lost over a relatively short period of time.

[7]We gratefully acknowledge the support of *Fundmonitor.com* for the data on IRR and the example cited in footnote 6.

established in 1950. There is no establishment date available for 111 of the 914 funds in the sample. However, a closer examination of the data set leads us to conclude that most of these 111 funds were established prior to 1988, and 69% or 62% of these funds are dead. We have the establishment dates of 114 dead and alive funds between 1950 and 1987. We believe that our sample covers nearly all equity funds established in Canada, dead or alive, until the end of the year 2001.

In addition to returns, we use a number of variables that denote fund characteristics such as front and back loads, fee options, management fees, and management expenses. The list of fund-related characteristics along with their definitions is provided in the Appendix. Fund-related characteristics are measured for the year-end 2002 for "alive" funds and for their last year of operation for dead funds.

An examination of the growth in assets and number of funds shows that there is a clear divide between funds established in the 1970s and 1980s when compared with funds established in the 1990s. Out of the 800 funds with establishment dates available in the sample, 559 funds were established between the years 1989 and 2001. The average asset size of mutual funds had grown from $19.5 million funds in the 1970s to $175.67 million in the 1990s. The 1990s was also a period of rapid expansion in both the number of funds and the assets invested into mutual funds. Mutual fund assets grew at an impressive annual rate of 27% and the number of accounts grews annually by 39% during the 1990s. Largely because of the stock market crash in the year 2000 and the onset of recession, in 2002 there was for the first time a decline in the market value of assets and in 2003, there was a decline in the number of mutual fund accounts (IFIC, 2004). In our sample of Canadian equity funds (dead and alive), the annual rate of growth in the number of funds was slightly higher at 29%. The assets of alive Canadian equity funds grew at a rate of 31% and the assets of dead fund grew annually at 12%.

5. Empirical Findings

5.1. *Fund performance and returns to investors*

The principal focus of the empirical analysis is on the performance of mutual funds. In the discussion that follows, we examine performance of mutual funds from two different perspectives: fund returns and investor returns. For the fund perspective, we compare their reported results against their self-chosen benchmarks. For the investor perspective, we use IRRs and benchmarks that are easily available to individual investors. As can be seen, we find in Table 1 that the majority of funds outperform their self-selected associated indices in the long run, on a risk-adjusted basis. However, the performance of the mutual funds outperforming their chosen index may be satisfactory from the perspective of the managers but not from the perspective of the investors as the self-selected benchmarks may be unavailable for investment for the individual investor.

Accordingly, we examine returns to investors in two stages. First, we report on long-term comparisons of returns of RR with TSE 300 and T-Bill returns. Second, we examine the relationship between RR and IRR. Table 2 profiles the performance of Canadian mutual funds and compares it to two benchmarks, the TSE 300 index and the 3-month T-Bill rates. The table shows that the majority of mutual funds did not outperform the 3-month T-Bill rates for the 1- to 5-year period and the TSE 300 index over the 5- to 15-year horizon. The performance of the mutual funds is superior to TSE 300 in the 1- to 3-year horizon ending in year 2002. This was also a period that was more turbulent than any time in the history of the TSE 300 and where movement in one stock (Nortel) accounted for 35% of the movement in the TSE 300 index at its peak. It is possible that by simple underweighting in Nortel stocks due to internal policy constraints on maximum allowable weighting for one specific stock, many funds outperformed the TSE 300. However, these percentages fall sharply when we look at 5-, 10-, and 15-year returns. In the very

Table 1. Percentages of funds and investors outperforming their benchmark.

	Year 1	Year 2	Year 3	Year 5	Year 10
Alpha[a]					
Alive funds	23.9	45.2	58.1	59.5	72.0
All funds	31.0	48.8	57.1	55.9	71.0
Alpha net of (RR–IRR)					
Alive funds	12.40	25.90	24.20	20.50	21.70
All funds	17.50	28.50	26.00	20.10	22.70

[a] Alpha is a measure of the difference between a fund's actual monthly excess return and its expected monthly excess return, which in turn is based on that fund's sensitivity (beta) to the excess return for the benchmark index. The methodology of calculating alpha is given in the Appendix.

Table 2. Net returns of Canadian equity mutual funds (%).

	Year ending 2002											
	1-year Returns		2-year Returns		3-year Returns		5-year Returns		10-year Returns		15-year Returns	
	RR	IRR	RR	IRR	RR	IRR	RR	IRR	RR	IRR	RR	IRR
Alive funds												
Returns	−8.29	−10.67	−7.28	−7.93	−0.77	−2.58	1.90	0.03	8.11	5.14	7.05	4.82
Difference (RR−IRR)	2.38		0.65		1.81		1.87		2.97		2.23	
No. of funds	634	585	565	475	475	390	280	215	146	110	105	73
% of funds above TSE 300	50.79	35.90	96.28	95.58	82.11	73.33	11.43	6.51	13.01	5.46	25.71	13.70
% of funds above T-Bill rates	6.78	6.84	10.27	7.79	28.21	19.74	30.71	19.07	89.04	63.64	69.52	35.62
All funds												
Returns	−7.03	−9.70	−6.02	−6.96	−0.40	−2.02	2.26	0.42	8.08	5.41	6.94	4.80
Difference (RR−IRR)	2.67		0.94		1.62		1.84		2.67		2.14	
No. of Funds	734	672	640	540	534	442	312	242	166	128	115	82
% of funds above TSE 300	54.09	39.29	96.25	94.81	82.96	74.21	13.78	8.26	12.05	5.47	23.48	12.20
% of funds above T-Bill rates	12.40	11.01	14.84	12.41	32.02	23.08	33.65	21.49	88.55	64.06	66.09	35.37
TSE 300	−9.13		−2.12		7.00		6.82		12.08		8.97	
T-bill rate	2.50		3.66		4.18		4.31		4.76		6.52	
% change in CPI 1992 = 100	2.2		2.4		2.5		2.02		1.75		2.56	

long run (10- to 15-year horizon) we find that most funds outperform the 3-month T-Bill rates but not the TSE 300. Clearly, for the funds alive as of year 2002, their performance, both long term and short term has been less than stellar.

These results have to be interpreted in the context of the studies on holding period behavior of mutual fund investors. The Investment Company Institute (ICI) study of US mutual fund investors (ICI, 2001) found that the median holding period of a typical mutual fund investor is 7 years. Barber *et al.* (2003) study of 30,000 households found that 25% of the investors never sold shares during the five and half years of their study. Choi *et al.* (2000) study found that more than half of the participants in 401 (K) plans never made a trade during the three years covered by the study. Ameriks and Zeldes (2000) found that 73% of participants in large employee-sponsored retirement plan made no changes in their asset allocations over the 10-year period covered by the study. Only 3% of the participants made six or more transactions during the sample period. The ICI survey further found that 45% of the small percentage of shareholders who sold shares did not do so for the purposes of adjusting their portfolio but because they needed the money to buy their house or to pay for educational or other expenses. There is no study documenting the portfolio behavior of the Canadian mutual fund investors. Table 2 shows that only 12% of the mutual funds over a 5-year period, 13% over a 10-year period, and 26% funds over a 15-year period outperformed the TSE 300.

IRR assesses the returns accruing to investors. Figures 1 and 2 show the frequency distribution of returns for mutual fund investors and mutual funds. These figures show a consistent pattern of return difference. As can be seen, reported returns to mutual funds (RR) are consistently higher than returns accruing to individual investors (IRR) for all years. The mean level of differences between (RR – IRR) is nearly 2% on average and tends to increase for long-term average performance. The impact of this consistent pattern of RR being greater than IRR can also be seen in Tables 1 and 2. While it is true that the majority of fund managers outperform their chosen indexes on an RR basis, when we take alpha net of the difference between RR and IRR, we find less than a quarter of the investors outperform their associated indexes. Thus performance may be superior on a risk-adjusted basis from the perspective of mutual find managers but not from the perspective of investors as only a quarter of funds outperform the adjusted alpha.

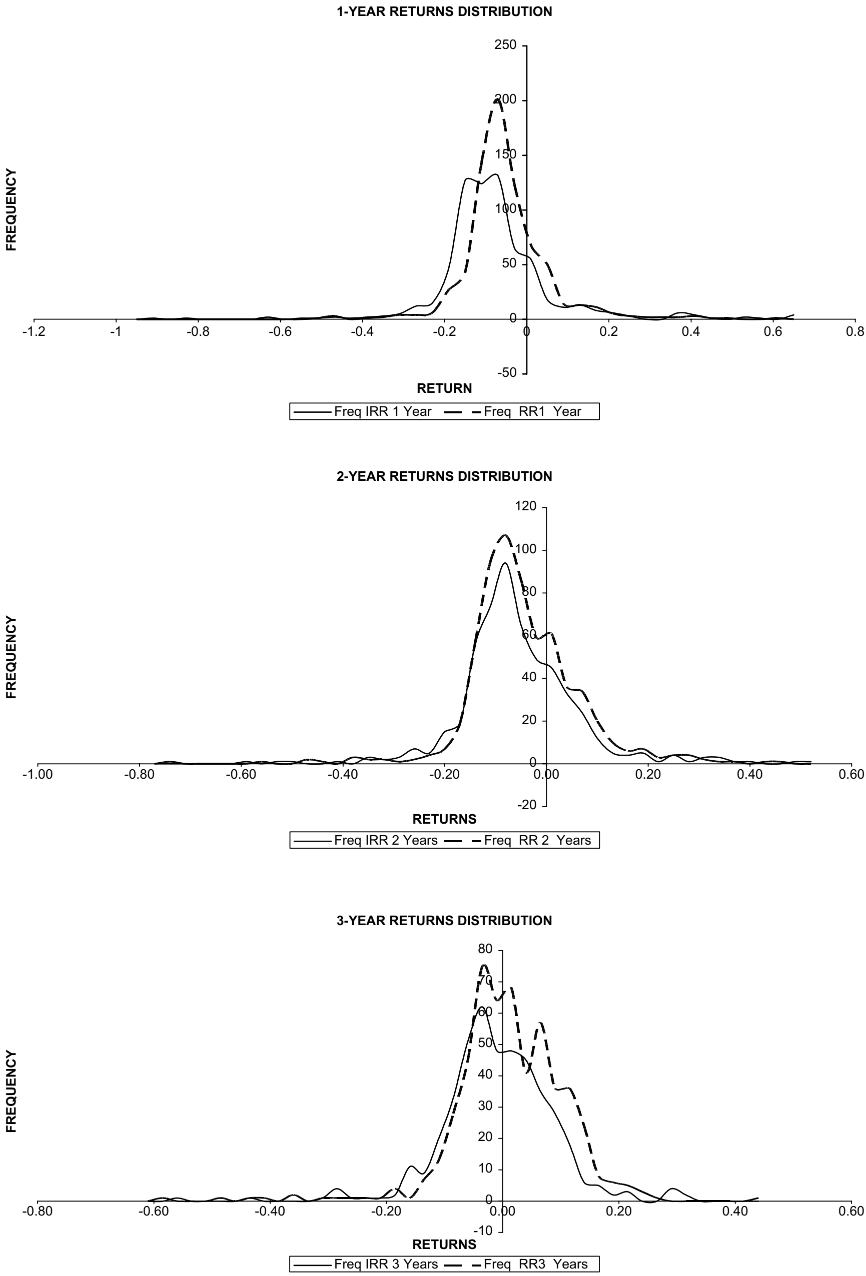


Figure 1. Distribution of dollar and time weighted returns.

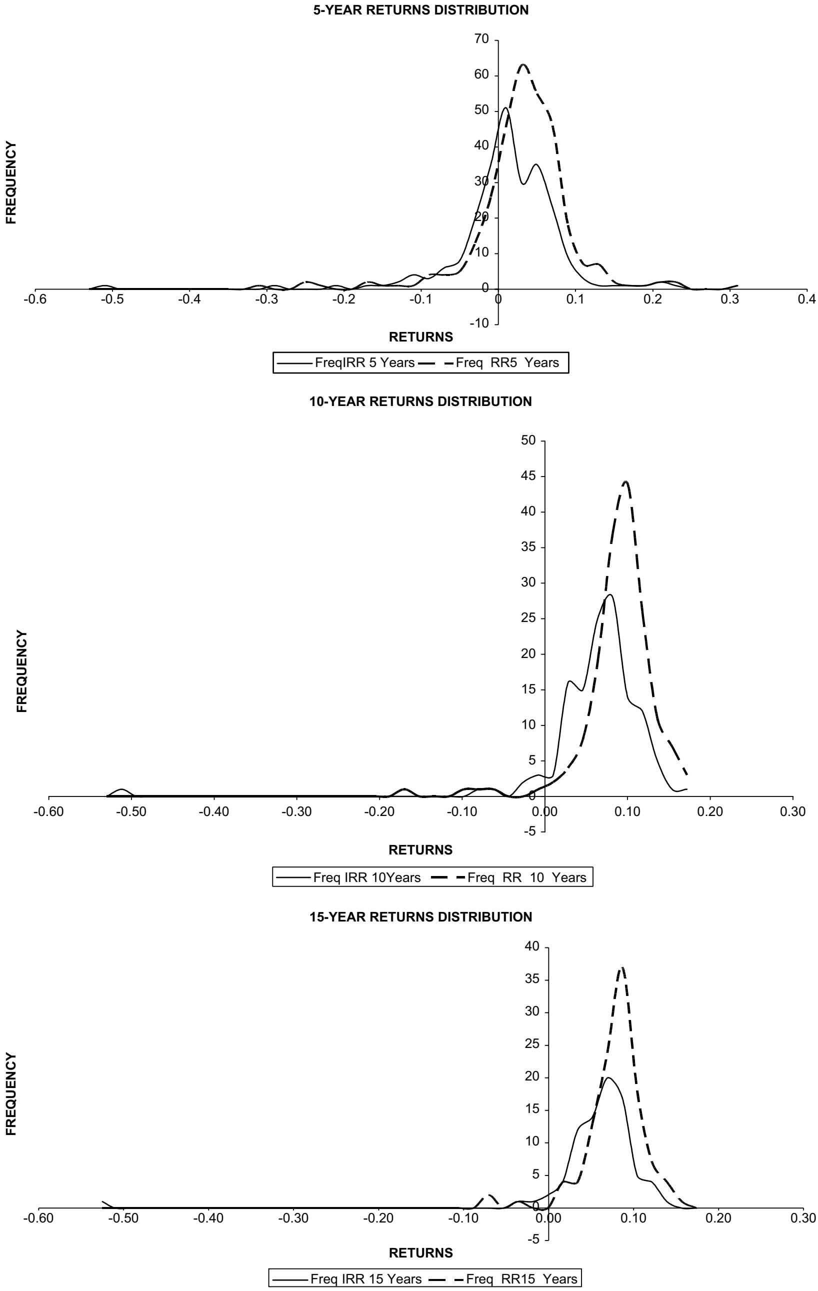


Figure 2. Distribution of dollar and time weighted returns.

Table 3 documents the implications of chasing winners and reluctance on the part of the investors to book losses when they arise. We find that IRR is greater than RR only in the case of funds that are underperforming their associated indexes. We also find evidence that funds with IRR greater than RR have significantly lower alphas, Sharpe values and conversely alphas and Sharpe values are higher for funds where RR is greater than IRR. The funds with IRR greater than RR are also significantly more likely to be in the bottom half of the performance ranking. In fact, the performance of funds with IRR greater than RR is more likely to be in the bottom quarter. Funds with IRR greater than RR are significantly more likely to be reporting negative returns and significantly less likely to be reporting returns in excess of 10%. The general conclusion is that underperforming funds are more likely to have IRR greater than RR and funds that outperform the market are more likely to have RR greater than IRR. Investors returns (IRR) are lower than reported returns of mutual funds (RR) for funds that are outperforming the market and have significantly superior performance characteristics because investors are buying these funds after their performance has peaked and a track record of superior performance established. Their decision to invest in a fund is taken typically after the mutual fund performance (RR) has peaked. Investors also display a distinct loss aversion. IRR greater than RR is more for funds that are underperforming. They tend to stay with funds even after a consistent track record of underperformance by these funds. These findings support the conclusion of the Odean (1998) study of shareholding accounts that investors are reluctant to book losses.

Table 4 provides potential explanation for the difference between IRR and RR. We find that typically for funds that are alive in the year 2002 investors have a 1 in 2 chance of choosing a repeat winner in the second year; a 1 in 4 chance of chance of choosing a repeat winner in the third year; and a 1 in 20 chance of picking a repeat winner in the fifth year. The strategy of moving funds to past winners is further compromised by the likelihood that the average investor will reallocate funds in her portfolio once over several years. This evidence clearly shows that the performance of mutual fund investors is not persistent and that the “winners” do not repeat. Moreover, frequency distribution of return data in Figures 1 and 2 shows that the investors choose the wrong time to invest in a fund (after the funds peak performance) and are reluctant to move out of underperforming funds. The performance decay of dead funds over the years, as expected is much higher than that of “alive”

Table 3. Entry and exit decisions of mutual fund investors and firm characteristics.

	Year ending 2002											
	1-year return		2-year return		3-year return		5-year return		10-year return		15-year return	
	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR
Alpha	−0.0010	−0.0023	0.0000	0.0005	−0.0010	0.0009	−0.0017	0.0004	−0.0003	0.0014	0.0012	0.0014
Beta	0.80	0.84	0.85	0.85	0.73	0.69	0.81	0.74	0.83	0.77	0.79	0.79
Sharpe	−0.40	−0.42	−0.21	−0.19	−0.18	−0.12	−0.08	−0.05	0.03	0.09	0.04	0.04
R^2	0.78	0.88	0.87	0.86	0.76	0.71	0.81	0.74	0.79	0.78	0.78	0.79
Tax Eff	−101.98	−100.45	−81.85	−83.53	−72.01	−44.99	−79.36	−79.04	63.88	92.54	92.91	94.89
MER	2.22	2.17	2.01	2.34	2.63	2.31	2.62	2.33	2.89	2.21	3.11	2.26
% In bottom half	53.67	50.78	53.04	47.66	49.55	46.24	47.66	44.02	48.62	43.18	48.76	42.68
% In bottom quarter	35.31	29.16	33.00	25.47	25.33	23.77	21.19	19.61	20.87	17.34	25.42	15.84
Rolling 12-month periods with a return over 10%	42.17	38.03	36.76	39.77	38.17	40.49	43.31	45.07	47.57	47.41	46.14	48.50
Rolling 12-month periods with a return below 0%	46.89	49.95	55.14	42.64	46.90	39.57	38.48	33.54	29.85	28.58	27.22	29.06

(*Continued*)

Table 3. (*Continued*).

	Year ending 2002											
	1-year return		2-year return		3-year return		5-year return		10-year return		15-year return	
	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR	IRR > RR	RR > IRR
% In Cdn. equity	72.04	74.46	72.96	74.71	71.40	75.12	77.22	76.21	67.49	77.52	74.32	74.66
% In US equity	10.32	11.20	10.41	11.17	11.78	10.75	11.52	10.55	13.56	10.78	12.67	11.32
% In Int. equity	4.46	5.00	4.72	4.82	6.29	4.36	2.38	4.24	8.52	3.86	3.69	6.39
% In bonds	4.48	1.31	3.03	1.51	1.18	2.13	1.42	1.45	1.29	1.81	2.08	1.02
% In cash	8.70	8.04	8.88	7.79	9.35	7.63	7.47	7.54	9.12	6.04	7.25	6.61
No load	25.34	30.16	29.02	30.56	32.85	32.69	38.71	41.70	25.71	45.95	18.18	41.89
Fee option	56.85	41.87	50.26	40.34	45.26	38.78	41.94	34.08	40.00	31.53	54.55	29.73
Max. front fee	4.21	3.83	3.90	3.95	4.33	3.86	4.30	4.22	5.26	3.95	5.56	4.51
Max. backend fee	5.10	4.92	4.99	5.04	5.01	4.93	5.02	4.69	4.28	4.46	4.96	4.54
Mgt. fee	2.08	2.06	2.15	2.06	2.11	2.02	2.02	1.94	1.82	1.84	1.86	1.82
Switching fee	39.53	26.84	26.53	28.40	30.93	25.51	31.91	27.47	33.33	30.23	50.00	29.31
Trailer fee	53.49	56.53	41.50	61.73	54.64	56.46	59.57	54.40	66.67	58.14	62.50	62.07

Note: Shaded area indicates that the respective means between the two groups are significantly different at the 5% level.

Table 4. Performance persistence of mutual funds.

Persistence in performance (alive funds only)				Persistence in performance (dead funds only)			
Year no. of Funds	Repeat wins if performance greater than median[a] (%)	Repeat wins for 3 years (%)	Repeat wins for 5 years (%)	Year no. of Funds	Repeat wins for 2 years (%)	Repeat wins for 3 years (%)	Repeat wins for 5 years (%)
1970/5	100	50	50	1970/1	—	—	—
1971/5	33	33	33	1971/1	0	0	0
1972/8	75	75	0	1972/2	100	0	0
1973/8	75	50	0	1973/8	50	0	0
1974/31	40	33	7	1974/3	33	33	0
1975/34	35	29	6	1975/7	50	0	0
1976/39	63	32	16	1976/7	25	0	0
1977/39	47	21	16	1977/7	50	25	0
1978/40	55	50	0	1978/7	50	50	0
1979/41	60	35	0	1979/7	75	25	0
Decade (1970s) average	58	41	13	Decade (1970s) average	48	15	0
1980/47	48	13	0	1980/7	25	0	0
1981/55	52	18	4	1981/7	67	33	0
1982/59	34	17	3	1982/7	40	20	0
1983/64	56	22	3	1983/7	50	50	25
1984/68	68	53	15	1984/7	75	25	0
1985/72	75	22	6	1985/7	25	25	0
1986/75	37	16	5	1986/10	60	20	0
1987/81	47	30	15	1987/12	33	0	0
1988/102	46	29	8	1988/13	33	0	0
1989/113	45	25	9	1989/16	38	25	0

(*Continued*)

Table 4. (*Continued*).

Persistence in performance (alive funds only)				Persistence in performance (dead funds only)			
Year no. of Funds	Repeat wins if performance greater than median[a] (%)	Repeat wins for 3 years (%)	Repeat wins for 5 years (%)	Year no. of Funds	Repeat wins for 2 years (%)	Repeat wins for 3 years (%)	Repeat wins for 5 years (%)
Decade (1980s) average	51	25	7	Decade (1980s) average	45	20	3
1990/123	51	33	5	1990/18	33	0	0
1991/127	63	38	6	1991/19	20	20	0
1992/132	59	15	3	1992/20	60	50	0
1993/139	43	19	6	1993/21	55	0	0
1994/150	40	16	4	1994/25	42	8	0
1995/169	56	32	3	1995/29	36	7	0
1996/182	57	37	5	1996/30	40	7	0
1997/209	60	24	5	1997/35	39	17	6
1998/251	42	27	5	1998/47	33	8	0
1999/318	62	27	—	1999/54	52	15	—
Decade (1990s) average	53	27	5	Decade (1990s) average	41	13	1
2000/429	29	18	—	2000/64	31	3	—
2001/495	73	—	—	2001/52	4	—	—

[a]Winner if greater than median return, loser if less than median.

funds. These findings are in conformity with Malkiel (1995) study of US funds that find evidence of declining performance persistence.

5.2. *Turnover and Performance*

In this section, we conduct cross-sectional analysis between performance and fund characteristics. We first start with evaluating turnover as a measure of trading activity by fund managers. Higher levels of trading activity can be attributed to the investment style of the money managers. It has also been noted that greater trading levels may be triggered by the flow of funds to winning mutual funds forcing money managers to shift funds to rebalance their portfolios. Table 5 profiles the consequences of turnover for mutual funds. We find that low turnover funds are in fact superior to the rest of the funds in the sample. We split the sample on the basis of a turnover dummy that takes the value 1 if the turnover is low. We find that firms with low turnover or trading activity have significantly higher average returns than the rest of the funds in the sample. It can also be seen that the management expense ratio of the low turnover funds is, on average, significantly lower than the rest of the funds. Low turnover funds are also significantly less likely to be in the bottom half and bottom quartile of funds ranked by performance. It is also the case that low turnover funds are significantly less likely to have rolling 12-month periods of negative returns.

Further analysis of low turnover of funds shows that low turnover funds have a significantly higher percentage of funds that have charges for switching funds between funds within the family and that they display significantly higher levels of maximum backend and frontend fees that are also statistically higher when compared with the rest of the funds in the sample. We also note that trailer fees are more likely in low turnover funds.[8] It appears that low turnover funds perform two roles. First, with statistically significant higher levels of backend and frontend fees and switching charges they increase the costs of moving funds between mutual funds and within a family of mutual funds. Second, they reduce the incentives for churning funds by brokers by paying them on a continuous basis with statistically significant higher occurrences of trailer fee charges.

[8]Trailer fees are a continuing commission paid to brokers to compensate for their ongoing services to their investment clients.

Table 5. Turnover and open-end equity mutual funds in Canada over the years.

Total No. of Funds	Alive funds only (year ending 2002)	
	Low turnover funds 177	Rest 457
1-year RR	−8.28%	−8.30%
2-year RR*	−5.77%	−7.87%
3-year RR*	0.49%	−1.26%
5-year RR*	3.36%	1.19%
10-year RR*	9.16%	7.48%
15-year RR*	8.07%	8.07%
1-year IRR	−9.52%	−11.15%
2-year IRR*	−5.85%	−8.96%
3-year IRR*	−0.56%	−3.48%
5-year IRR*	1.09%	−0.68%
10-year IRR*	6.47%	4.15%
15-year IRR*	6.63%	3.48%
RSQ1	86.11%	87.36%
RSQ2	84.74%	86.81%
RSQ3	71.10%	72.09%
RSQ5*	71.54%	76.51%
RSQ 10	75.60%	79.88%
RSQ 15*	73.87%	83.45%
MER*	1.87	2.18
Years of experience*	21.3077	25.6000
Perc in bottom half*	53.36%	60.00%
Perc in bottom quartile*	34.16%	44.18%
Rolling 12-month periods with a return over 10%	39.65%	37.90%
Rolling 12-month periods with a return below 0%*	51.81%	56.30%
Perc In Cdn Eq	73.1473	74.0971
Perc In US Eq	11.4397	10.7787
Perc In Int Eq	4.6122	5.0287
Perc in bonds	2.8824	1.8946
Perc in cash	7.9206	8.2032
RRSP eligibility	92.63%	95.67%
NOLOAD	29.47%	29.72%
Fee option	45.79%	45.67%
Max. front fee*	4.8144	4.1261
Max. backend fee*	5.6316	5.3231
Mgt. fee	1.9856	2.0689
Switching fee*	40.00%	25.57%
Trailer fee*	67.89%	43.44%

*The respective means between the two groups are significantly different at the 5% level.

5.3. *Expenses and performance*

Next we focus on the role of fund-related expenses on the performance of funds. Figure 3 shows that the relative share of management expense ratio (MER) is rising for funds over the years.[9] When we classify the funds by decade of establishment, we find that while the MER is declining for newer funds their share in management fees is increasing. Table 6 examines whether the high and growing levels of MERs is justified by the value added by money managers. We examine the correlation between fund expenses and fund performance/characteristics and find that both non-advisory and total expenses are significantly and negatively related to fund performance and to alphas. The correlation between percentage of the time the fund is in the bottom half/bottom quartile of performance and expenses is significant and negative. Non-advisory and total expenses are also significantly and negatively correlated with the percentage of rolling 12-month periods of negative returns. These statistically significant results suggest a perverse relationship between fund expenses, (growing) management fees, and a broad range of performance measures. Among other significant results, advisory expenses are negatively

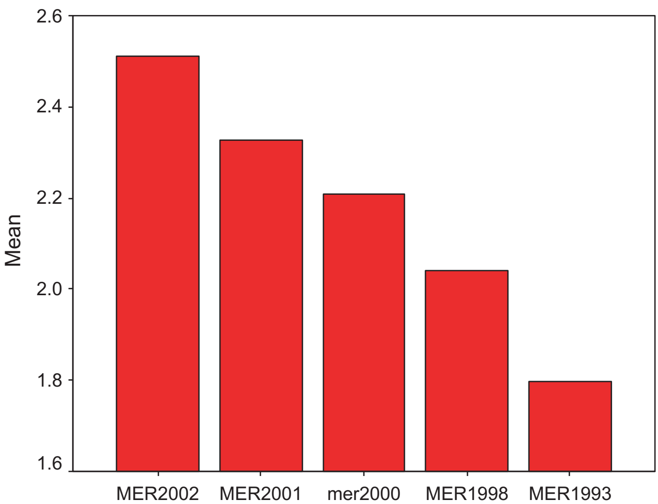


Figure 3. Management expense ratio over the years.

[9]Ruckman (2003) notes that fund-related expenses are nearly 50% higher for Canadian mutual funds when compared with US funds.

Table 6. Correlation between advisory/non-advisory expenses and performance/characteristics of open-end equity mutual funds in Canada over the years.

	Year ending 2002					
	Advisory Expenses[a]		Non Advisory Expenses[b]		Total Expenses[c]	
	Alive Funds	All Funds	Alive Funds	All Funds	Alive Funds	All Funds
1-year RR	−0.01	−0.08*	−0.05	−0.03	−0.04	−0.05
2-year RR	0.05	−0.01	−0.10*	−0.07	−0.08	−0.07
3-year RR	−0.02	−0.05	−0.18*	−0.15*	−0.17*	−0.15*
5-year RR	−0.10	−0.12*	−0.40*	−0.37*	−0.39*	−0.37*
10-year RR	−0.28*	−0.28*	−0.49*	−0.46*	−0.52*	−0.50*
15-year RR	−0.31*	−0.31*	−0.62*	−0.60*	−0.64*	−0.63*
1-year IRR	−0.03	−0.04	0.03	0.00	0.02	−0.01
2-year IRR	0.03	−0.03	−0.13*	−0.10	−0.10*	−0.10
3-year IRR	0.01	−0.03	−0.10*	−0.09	−0.10*	−0.11*
5-year IRR	−0.10	−0.12	−0.34*	−0.32*	−0.35*	−0.34*
10-year IRR	−0.20*	−0.20*	−0.10	−0.09	−0.16	−0.15
15-year IRR	−0.18	−0.17	−0.16	−0.15	−0.20	−0.20
APLHA1	−0.20*	−0.19	0.05	0.07	−0.05	−0.02*
APLHA2	−0.11*	−0.13*	0.03	0.03	−0.02	−0.03
APLHA3	−0.10	−0.10*	−0.03	−0.02	−0.10	−0.08
APLHA5	−0.19*	−0.17*	−0.12	−0.13*	−0.21*	−0.22*
APLHA10	−0.39*	−0.40*	−0.01	−0.03	−0.28*	−0.28*
APLHA15	−0.37*	−0.40*	−0.02	−0.07	−0.29*	−0.35*
RSQ1	−0.05	−0.05	0.00	0.00	−0.03	−0.03
RSQ2	−0.01	−0.01	−0.01	−0.01	−0.02	−0.02
RSQ3	−0.02	−0.02	−0.01	−0.01	−0.04	−0.04
RSQ5	−0.06	−0.06	−0.01	−0.01	−0.07	−0.07
RSQ 10	0.16	0.16	−0.21	−0.21	−0.09	−0.09

(*Continued*)

Table 6. (*Continued*).

	Year ending 2002					
	Advisory Expenses [a]		Non Advisory Expenses[b]		Total Expenses[c]	
	Alive Funds	All Funds	Alive Funds	All Funds	Alive Funds	All Funds
RSQ 15	−0.01	−0.01	0.09	0.09	−0.01	−0.01
MER	0.11*	0.14*	0.93*	0.92*	1.00	1.00
Years of experience	0.15	0.11	0.02	0.01	0.04	0.02
Perc in bottom half	−0.08*	−0.08*	−0.10*	−0.10*	−0.13*	−0.14*
Perc in bottom quartile	−0.07	−0.09*	−0.19*	−0.19*	−0.20*	−0.21*
Rolling 12-month periods with a return over 10%	−0.20*	−0.24*	0.00	0.01	−0.06	−0.07
Rolling 12-month periods with a return below 0%*	0.03	0.05	−0.18*	−0.20*	−0.13*	−0.15*
Perc in Cdn Eq	−0.05	−0.03	−0.08	−0.08	−0.06	−0.07
Perc in US Eq	0.07	0.07	0.09	0.08	0.11*	0.10*
Perc in Int Eq	0.00	−0.02	0.12	0.13*	0.10*	0.11*
Perc In bonds	−0.12*	−0.10*	−0.01*	−0.01	−0.10*	−0.09*
Perc In cash	0.13*	0.09*	−0.06	−0.03	0.00	0.00
RRSP eligibility	0.06	−0.03*	0.05	−0.08	0.08*	0.08*
No load	−0.41*	−0.42*	0.03	0.03	−0.15*	−0.16*
Fee option	0.18*	0.19*	0.04	0.04	0.13*	0.13*
Max. front fee	−0.10	0.15*	0.34*	0.22	0.32*	0.28*
Max. backend fee	0.13*	0.37*	−0.05	−0.12*	0.01	0.05
Mgt. fee	1.00	1.00	−0.26*	−0.26*	0.11*	0.14*
Switching charges	−0.07	−0.07	0.14*	0.14*	0.14*	0.14*
Trailer fee	−0.03	−0.03	0.12*	0.12*	0.10*	0.10*
Non-advisory expenses	−0.26*	−0.26*	1.00	1.00	0.93*	0.92*

*The respective means between the two groups are significantly different at the 5% level.

[a]Advisory expenses = Management fees.

[b]Non-Advisory expenses = MER − Management fees.

[c]Total Expenses = MER.

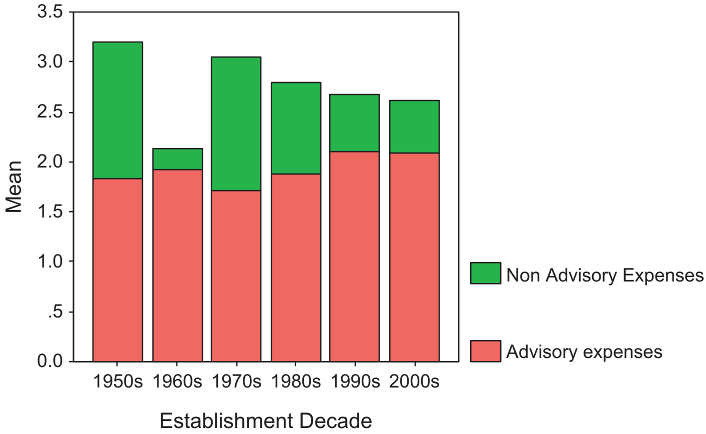


Figure 4. Advisory and non-advisory expenses and the decade fund established.

correlated with the no-load option and switching charges are positively correlated with non-advisory and total expenses.

5.4. *Age of funds and performance*

In this section, we attempt to evaluate the difference in performance between old and new funds. Table 7 looks at the funds established in the 1990s and compares these funds with the funds that were established prior to this decade.

We find that funds established after 1989 are significantly superior in their performance when compared with funds established prior to 1989, a result that may simply reflect the overall superior performance of the equity markets in the 1990s. Though the average returns of the pre-1989 funds are lower than those reported for funds established after 1989, the pre-1989 funds are significantly less likely to be in the bottom half or quartile of funds ranked by returns. The funds established before 1989 are significantly more likely to post returns above 10% and significantly less likely to post negative returns.

Funds established prior to 1989 also have a significantly higher MER. However, management fees of funds established after 1989 are significantly higher. While overall expenses are lower for funds established after 1989 the level of advisory expenses have gone up. Funds established after 1989 invest significantly less in Canadian equity and hold a significantly higher proportion of their investment portfolio in cash when compared with funds established prior to 1989. Thus "alive" funds established prior to 1989 are likely to give

Table 7. Open-end equity mutual funds in Canada over the years.

	Year ending 2002			
	Funds established till 1989		Funds established after 1989	
	Alive funds Only[a]	All funds[a]	Alive funds only[a]	All funds[a]
No. of funds	121	150	497	568
1-year RR	−8.59%	−7.18%	−8.21%	−6.99%
2-year RR	−8.78%*	−6.64%	−6.77%*	−5.81%
3-year RR	−2.10%*	−0.82%	−0.21%*	0.30%
5-year RR	0.96%*	1.71%	2.76%*	2.80%
1-year IRR	−9.98%	−8.69%	−10.84%	−9.99%
2-year IRR	−8.40%	−6.45%	−7.78%	−7.13%
3-year IRR	−2.08%	−0.84%	−2.08%	−2.52%
5-year IRR	−0.49%	0.28%	0.57%	0.59%
APLHA1	−0.0018*	−0.0002	−0.0036*	−0.0022
APLHA2	0.0003	0.0012*	-0.0003	0.0005*
APLHA3	0.0011	0.0011	0.0003	0.0004
APLHA5	0.0002	0.0000	0.0002	-0.0001
RSQ1	89.94%*	89.94%*	86.16%*	86.16%*
RSQ2	88.66%*	88.66%*	85.44%*	85.44%*
RSQ3	74.64%	74.64%	70.82%	70.82%
RSQ5	78.42%*	78.42%*	72.21%*	72.21%*
MER	2.52%*	2.48%*	1.99%*	1.99%*
Years of experience	20.53*	19.62*	24.97*	24.76*
In bottom half	52%%*	65%	60%%*	61.19%
In bottom quarter	31%*	48%	45%*	45.36%
Rolling 12-month periods with a return over 10%	47%*	46*	35%*	37.08%*
Rolling 12-month periods with a return below 0%	33%*	34%*	61%*	59.17%*
% In Cdn equity	78.98*	78.76*	72.36*	72.82*
% In US equity	10.01	9.77	11.23	11.09
% In Int equity	4.62	4.63	5.00	5.06
% In bonds	1.57	1.46	2.34	2.32
% In cash	4.83*	5.38*	9.07*	8.72*
RRSP eligible	97%	96.43%	94%	94.55%
NOLOAD	37%*	38.57%*	28%*	29.39%*
Fee option	36%	35.00%*	48%	47.12%*
Max. front fee	5.04%*	4.40*	4.18%*	3.74*
Max. backend fee	5.27%	4.45*	5.42%	4.96*

(*Continued*)

Table 7. (*Continued*).

	Year ending 2002			
	Funds established till 1989		Funds established after 1989	
	Alive funds only[a]	All funds[a]	Alive funds only[a]	All funds[a]
No. of funds	121	150	497	568
Mgt. fee*	1.85%*	1.85*	2.09%*	2.06*
Switching fee	25%	25.00%	31%	31.15%
Trailer fee	53%	43.75%	44%	52.58%

*The means for the funds established up to 1989 are significantly different at the 5% level from the means of the funds established after 1989.

lower but largely positive returns while funds established after 1989 are likely to give higher returns with greater fluctuation in returns. The management fees of the post-1989 funds are also higher and are able to outperform pre-1989 funds with almost double the asset levels, which suggest that they are able to better time the market.

6. Conclusions

In this paper, we document the performance of a comprehensive sample of Canadian mutual funds. We can draw a number of conclusions from the findings of this study. First, the value added by money managers on a long-term basis is meager and inconsistent. For a holding period of 5 years or more, only a quarter of mutual funds outperform the market and a mere of 5% of the funds perform higher than the median levels of returns of the sample for more than 5 years.[10] Second, the performance of funds that record low levels of trading is superior to the rest of the funds in the sample. Although we cannot conclude, whether the low level of trading was the choice of managers or it was because of the costs of moving in and out of these funds due to high fees. However, the evidence favors a buy-and-hold strategy compared with an active-trading strategy. Third, MER has risen and the share of management fees in MERs has increased. However, higher MERs and management fees

[10]We believe that the persistence of performance found in the late 1990s when compared with the TSE 300 may have simply been a result of underweighting in some technology stocks (e.g., Nortel) due to diversification restrictions imposed by the investment policy.

do not seem to have generated better performance. Fourth, performance of newer funds (establishment date — post-1989) is superior to older funds and the newer funds have a lower MER. Fifth, we find that the returns earned by mutual fund investors is less than that reported by mutual funds. Further research using detailed data on the timing of purchase and sale decisions of mutual fund investors from other countries including the US would further enhance our understanding of the timing capability of mutual fund investors. This is of particular interest given the current thrust in social security reform toward greater individual control over the investment decisions if retirement funds.

These results have some important implications for the mutual fund managers, regulators, and investors. We believe that continued disclosure by individual mutual funds, that include trading volumes (turnover), returns (IRR) earned by investors, in addition to reporting the returns earned by mutual funds (RR), management expense and other expenses as a percentage of gross returns, and comparison of their performance to passive indices (adjusting for transaction costs) would benefit the entire industry and investors. We also hope that these results prompt mutual funds' self-regulatory body and the Ontario Securities Commission to ensure that such disclosures become a norm in annual reports of the mutual fund industry. We also believe that our results would dissuade individual investors from investing in funds simply based on their past returns and that they would seek to invest in lower turnover funds with low MERs and management fees or alternately invest in passive index funds with low MERs.

References

Ameriks, J. and S. Zeldes, *How do Household Portfolios Shares Vary with Age*, 2000, unpublished working paper.

Barber, B. M., T. Odean and L. Zheng, "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows" (December 2003). http://ssrn.com/abstract=496315, accessed, February, 2005.

Bhargava, R., J. G. Gallo and P. E. Swanson, "The Performance, Asset Allocation and Investment Style of International Equity Managers." *Review of Quantitative Finance and Accounting* 17, (2001) 377–395.

Bishara, H., "Evaluation of the Performance of Canadian Mutual Funds, 1967–84," *Proceedings of the Administrative Sciences Association of Canada*, pp. 1718–1726, 1988.

Blake, D. and A. Timmermann, "Mutual Fund Performance: Evidence for the UK." *European Finance Review* 2, 57–77 (1998).

Brown, S. J. and W. N. Goetzmann, "Performance Persistence." *Journal of Finance* 50, 679–698 (1995).

Brown, K. C., W. V. Harlow and L. T. Starks, "Of Tournaments and Temptations: an Analysis of Managerial Incentives in the Mutual Fund Industry." *Journal of Finance* 51, 85–109 (1996).

Brown, S. J., W. N. Goetzmann, T. Hiraki, T. Otsuki and N. Noriyoshi, "The Japanese Open-End Fund Puzzle." *Journal of Business*, Vol. 71, 59–77 (2001).

Brown, A. and D. Young, *Only One in 17 Investors Make Money on their Investments*, Fundmonitor.com, accessed March, 2004, 2002.

Cai, J., K. C. Chan and T. Yamada, "The Performance of Japanese Mutual Funds." The *Review of Financial Studies* 10, 237–273 (1997).

Calvett, A. L. and J. Lefolle, "The CAPM under inflation and the performance of Canadian mutual funds." *Journal of Business Administration* 12, Fall (1980).

Carhart, M., "On persistence in Mutual Fund Performance." *Journal of Finance* 52, 57–82 (1997).

Chen, Z. and P. J. Knez, "Portfolio Performance Measurement: Theory and Applications." *Review of Financial Studies* 9, 511–556 (1996).

Chevalier, J. and G. Ellison, "Risk Taking by Mutual Funds as a Response Incentives." *Journal of Political Economy* 105, 1167–1200 (1997).

Choi, J., D. Laibson and A. Metrick, "How Does the Internet Affect Trading? Evidence from Investor Behavior in 401 (K) Plans." *Journal of Financial Economics*, Vol. 64, No. 3, 397–421 (2002).

Dahlquist, M., S. Engstorm and P. Soderlind, "Performance and Characteristics of Swedish Mutual Funds." *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 3, 409–423 (2000).

Deaves, R., "Data Conditioning Biases, Performance, Persistence and Flows: The Case of Canadian Equity Funds." *Journal of Banking and Finance* 28, 673–694 (2004).

Elton, E., M. Gruber and C. Blake, "Survivorship Bias and Mutual Fund Performance." *Review of Financial Studies* 9, 1097–1120 (1996).

Elton, E., M. Gruber, S. Das and M. Hlavka, "Efficiency with Costly Information: a Reinterpretation of Evidence from Managed Portfolios." *Review of Financial Studies* 6, 1–22 (1993).

Ferson, W. and R. Schadt, "Measuring Fund Strategy and Performance in Changing Economic Conditions." *Journal of Finance* 51, 425–462 (1996).

Goetzmann, W. and R. Ibboston, "Do Winners Repeat? Patterns in Mutual Fund Behaviour." *Journal of Portfolio Management*, Winter, 9–18 (1994).

Goetzmann, W. and M. Massa, *Daily Momentum and Contrarian Behaviour of Index Fund Investors*, Working Paper No., 99–13, Yale International Center for Finance, 1999.

Grinblatt, M. and S. Titman, "Mutual Fund Performance: an Analysis of Quarterly Portfolio Holdings." *Journal of Business* 62, 393–416 (1989a).

Grinblatt, M. and S. Titman, "Portfolio Performance Evaluation: Old Issues and New Insights." *Review of Financial Studies* 2, 393–421 (1989b).

Grinblatt, M. and S. Titman, "The Persistence of Mutual Fund Performance." *Journal of Finance* 47, 1977–1984 (1992).

Gruber, M., "Another Puzzle: the Growth in Actively Managed Mutual Funds." *Journal of Finance* 51, 783–807 (1996).

Hendricks, D., J. Patel and R. Zeckhauser, "Hot Hands in Mutual Funds: Short-run Persistence of Relative Performance, 1974–1988." *Journal of Finance* 48, 93–130 (1993).

IFIC, *Historical Overview*, The Investment Funds Institute of Canada, available at, http://www.ific.ca/eng/frames.asp?l1=Statistics, accessed, November, 2004.

ICI, *Redemption Activity of Mutual Fund Owners*, Investment Company Institute, *Fundamentals*, Vol. 10, No. 1, 2001.

Ippolito, R., "Efficiency with Costly Information: a Study of Mutual Fund Performance." *Quarterly Journal of Economics* 104, 1–23 (1989).

Jensen, M., "The Performance of Mutual Funds in the Period 1945–1964." *Journal of Finance* 23, 389–416 (1968).

Khorana, A., H. Servaes and P. Tufano, "Explaining the Size of the Mutual Fund Industry Around the World." *Journal of Financial Econom*ics, Vol. 78, No.1, 145–185 (2005).

Malkiel, B., "Returns from Investing in Equity Mutual Funds 1971–1991." *Journal of Finance* 50, 549–573 (1995).

Nesbitt, S. L., "Buy High, Sell Low: Timing Errors in Mutual Fund Allocations." *Journal of Portfolio Management* 22(1), 57–60 (1995).

Odean, T., "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance* 53, 1775–1798 (1998).

Otten, R. and D. Bams, "European Mutual Fund Performance." *European Financial Management*, Vol. 8, No.1, 75–101 (2002).

Pesaran, M. H. and A. Timmermann, "Predictability of Stock Returns: Robustness and Economic Significance." *Journal of Finance* 50, 1201–1228 (1995).

Ruckman, K., "Expense Ratios of North American Mutual Funds." *Canadian Journal of Economics* 36, 192–223 (2003).

Shefrin, H. and M. Statman, "The Disposition Effect." *Journal of Finance* 40, 777–790 (1985).

Sirri, E. R. and P. Tufano, "Costly Search and Mutual Fund Flows." *Journal of Finance* 53, 1589–1622 (1998).

Appendix. List of Variables Used in the Study.

Variable	Variable description
Load and expense-related variables	
BLOAD	A dummy variable takes the value 1 if a sales fee is charged when mutual fund units are sold takes the value 0 otherwise
FLOAD	A dummy variable takes the value 1 if a sales fee is charged when mutual fund units are purchased takes the value 0 otherwise
BORFLOAD	A dummy variable takes the value 1 if a sales fee is charged when mutual fund units are purchased or sold takes the value 0 otherwise
NO LOAD	A dummy variable takes the value 1 if no sales fee is charged when mutual fund units are purchased or sold
MANFEE	The management fee is a certain percentage deducted form a fund's net assets to pay the fund manager. Often the percentage declines as the fund's net assets grow. The management fee might also be amended by or be primarily composed of a performance fee, which raises or lowers based on the fund's returns when compared with an established benchmark
MER	The aggregate of all expenses related to the fund operation, including management fee, custodian fee, transaction fee, etc., on a percentage term over the net asset value of the fund
SWITCHFEE	A dummy variable takes the value 1 if within the same fund family, there is a fee for switching investment among different funds takes the value 0 otherwise
TRAILERFEE	A dummy variable takes the value 1 for the no-load mutual fund, a continuing commission is paid to brokers to compensate for their on going services to their investment clients takes the value 0 otherwise

Appendix (*Continued*).

Other acronyms and variables	
RRSP	Registered Retirement Savings Plan (RRSP). A dummy variable takes the value 1 if the fund RRSP Eligible (yes, no, or foreign) takes the value 0 otherwise
TURNOVER	A measure of the fund trading activity level determined by the lesser of purchases or sales, and then divided by average monthly assets. A dummy variable takes the value1 if turnover is in the first quartile 0 otherwise
CPI	Consumer Price Index
TAXEFF	The percentage of the after-tax return over pre-tax return
TSE	Toronto Stock Exchange
TBILL	Treasury Bill

Chapter 13

Identifying Major Shocks in Market Volatility and Their Impact on Trading Strategies

Pauline Shum
York University, Canada

Kevin X. Zhu*
Ibbotson Associates/Morningstar, USA

Major political and financial events, such as the invasion of Kuwait by Iraq and the 1997 Asian Financial Crisis, create great uncertainty in the economy. Accordingly, they have a strong influence on the stock market. In addition, the unexpected shifts in market volatility can last for days or months. The main goal of this paper is to identify the periods in which investors experience major shocks in market volatility in the US. As an interesting extension, we also study the impact of these shocks on the performance of two popular trading strategies.

Keywords: Stock market volatility; trading strategies.

1. Introduction

Major political and financial events, such as the invasion of Kuwait by Iraq and the 1997 Asian Financial Crisis, create great uncertainty in the economy. Accordingly, they have a strong impact on stock prices. In addition, the unexpected shifts in market volatility can last for days or months. The main goal of this paper is to identify the periods in which investors experience unexpectedly large shocks in market volatility. As an interesting extension, we also study the impact of these shocks on the performance of two popular trading strategies.

The market index that we use in this study is the Centre for Research in Security Prices (CRSP) value-weighted total returns index. In an efficient market, security prices change upon the arrival of new information. Hence, periods of unexpectedly high volatility are associated with the arrival of new information (see, for example, Daniel *et al.*, 1998). To identify these periods, we employ a three-step procedure. Together, the three steps help determine the timing and the duration of an unexpectedly large shift in stock market volatility.

*Corresponding author.

The first step involves a statistical algorithm, the Iterated Cumulative Sums of Squares (ICSS), developed by Inclan and Tiao (1994). It is designed to detect variance change points in a time series. If volatility jumps with the arrival of new information, then this method offers an objective determination of the dates when the market receives new information. We define a period as the length of time between two variance change points. At the end of this step, we are able to identify periods with above average volatility.

In the second step, we construct a measure of the "intensity" of news arrival during each period identified by the ICSS algorithm. Fair (2002) documents all of the 5-minute changes in the S&P 500 futures that were greater than 0.75% (in absolute value), for the period April 21, 1982 to October 29, 1999. Fair considers changes of this magnitude to be significant, and indicative of news arrival. If this is the case, market volatility will increase accordingly. Using Fair's data, we calculate a "news intensity" variable for each period. This step serves as a check for the results in the first step: there should be a positive correlation between volatility and the intensity of news arrival in each period.

In order to gauge whether the shift in volatility at the beginning of a high volatility and high news intensity period was unexpected, the third step involves a comparison between realized market volatility and implied volatility obtained from a corresponding options market. This comparison provides a sufficient condition that the market has received unexpected and dramatic news that changes the investors' information set, and alters their expectations.

The results of the three steps above are combined to generate the set of periods in which stock market volatility shifts up substantially and unexpectedly. During our sample period from January 1, 1986 to December 31, 1999,[1] we identify six periods of unexpectedly high market volatility. These dates coincide with periods of great uncertainty in the US stock market: the stock market crash of 1987, the failure of the UAL buyout in 1989, the invasion of Kuwait by Iraq in 1990, the market correction in 1996, the Asian Financial Crisis in

[1]This sample period is chosen based on data availability. The year 1986 was the first year in which the CBOE's Volatility Index (VIX) is available, and 1999 is the last year in Fair's (2002) study. VIX data are employed in the third step of our identification procedure, and Fair's results are employed in the second step of the procedure.

1997, and the default of Russian debt and the failure of Long Term Capital in 1998.[2] In all six periods, the stock market performed poorly. As volatility was high during these periods, our findings confirm the well-known leverage effect in the time-series finance literature: negative stock price changes have a larger impact on volatility than positive changes.

A natural question to ask is how investors are affected by such shifts in volatility. As an interesting extension, we explore how popular trading strategies such as contrarian and momentum portfolios do during and after periods of unexpectedly high volatility. To do this, we construct a set of loser and winner portfolios by capitalization, and assess their performance using two popular measures: conditional Jensen's alpha and cumulative risk-adjusted returns. Because the periods vary in length, we cannot group them together in the tests. Instead, we have to assess the evidence for each period, and then consider them collectively. Based on a sample of 180 portfolios, our findings suggest that a contrarian strategy slightly dominates a momentum strategy in loser portfolios during these periods. In the case of small-cap losers, the returns to trading them in a contrarian portfolio could be substantial. In other portfolios, there is no clear pattern. We further demonstrate that an investor would have to recognize the shift in volatility right away in order to profit from buying small-cap loser portfolios. This is very hard to do without the benefit of hindsight. Hence, in the context of an efficient market, we conclude that contrarian and momentum strategies do not necessarily outperform a diversified passive portfolio during times of great uncertainty.

This paper is organized as follows. In Section 2, we introduce the three complementary procedures we employ to identify periods of unexpected high market volatility. In Section 3, we discuss the tests of the contrarian and momentum strategies. In Section 4, we provide highlights of our findings and their implications. Section 5 concludes the paper.

---

[2]On October 13, 1989, the NYSE dropped to 6.9%. On July 17, 1996, the NYSE experienced the second largest decline since the crash of 1987, and broke the record in trading volume. Much like the crash of 1987, this movement cannot be traced to any one particular event. There was a general feeling at the time that the market was due for a correction. There was a series of disappointing profit reports from high-tech companies, and the fear that interest rates were on the rise.

2. Identifying Unexpectedly Large Shifts in Market Volatility

We employ three complementary procedures to identify periods in which the US stock market experiences major, unanticipated episodes of market volatility. A *major* episode is defined as one that causes a large upward shift in volatility and generates high news intensity (to be defined). An *unanticipated* episode is determined by a comparison between realized volatility and implied volatility from options prices. In the rest of this section, we describe the three procedures, and discuss the results.

2.1. *Variance change points*

Inclan and Tiao (1994) develop a statistical methodology, called the ICSS algorithm, to detect variance change points in a time series. (For an application of this methodology to emerging markets, see Aggarwal *et al.*, 1999). The length of time between two variance change points is defined as a period.

The basic ideas of the methodology are follows. Suppose we have a time series $\{\alpha_t\}$with independent observations from a normal distribution with zero mean and unconditional variance σ^2. Suppose further that the corresponding variance process of $\{\alpha_t\}$ has two regimes, i.e., one change point. Let: $C_k = \sum_{t=1}^{k} \alpha_t^2$, $k = 1, \ldots, T$, be the cumulative sum of square (mean-centered) observations from the start of the series to the kth observation, where T is the total number of observations in the sample. Define a statistic, $D_k = C_k/C_T - k/T$, $k = 1, \ldots, T$ with $D_0 = D_T = 0$. Under the null hypothesis of homogeneous variable, i.e., there is only one regime in the variance process, the D_k statistic oscillates around zero because both $C_{k/k}$ and $C_{T/T}$ are sample estimates of σ^2 within the corresponding sample period. In contrast, if there is a regime change in the variance process — the alternative hypothesis — D_k will depart from zero. The asymptotic distribution of D_k under the null hypothesis is:

$$D_k = \frac{k(T-k)(1-F_{T-k,k})[(k/T)+(T-k)F_{T-k,k}/T]^{-1}}{T^2},$$

where $F_{T-k,k}$ is the F distribution with $T-k$ and k degrees of freedom. The variable k is determined by the optimization, $\max_k |D_k|$. Based on the above distribution, the critical values for the upper and the lower boundaries of D_k can be chosen by giving a known level of probability to detect a significant change in the variance. When $\max_k |D_k|$ goes out of the predetermined boundaries, the null hypothesis of no regime change in the variance process is rejected, and the k at which $\max_k |D_k|$ is attained is taken as an estimate of the change point. To understand how this method can be extended to accommodate *multiple* change points, see Inclan and Tiao (1994).

We apply the ICSS algorithm to the CRSP value-weighted daily total returns index from January 2, 1986 to December 31, 1999. The results are presented in Table 1.

Table 1. CRSP daily total returns results from the ICSS algorithm[a] and news intensity[b].

Variance change dates		Standard deviation between start and end dates[c]	Occurrence	Occurrence/days %
Start	End			
2-Jan-86	13-Oct-87	0.0082	17	2.6
13-Oct-87	30-Oct-87	0.0648*	11	61.1
30-Oct-87	25-Jan-88	0.017*	13	14.8
25-Jan-88	2-Sep-88	0.0083	1	0.5
2-Sep-88	12-Oct-89	0.0057	1	0.2
12-Oct-89	19-Oct-89	0.0305*	3	37.5
19-Oct-89	1-Aug-90	0.0068	3	1
1-Aug-90	12-Nov-90	0.0127*	12	11.5
12-Nov-90	30-Dec-91	0.0078	11	2.7
...	...	...	...	...
18-Mar-96	3-Jul-96	0.0055	2	1.9
3-Jul-96	2-Aug-96	0.012*	2	6.5
2-Aug-96	5-Dec-96	0.0047	2	1.6
5-Dec-96	26-Mar-97	0.0073	4	3.6
26-Mar-97	7-May-97	0.012*	4	9.3
7-May-97	15-Oct-97	0.0077	9	5.6
15-Oct-97	3-Nov-97	0.0253*	3	15
3-Nov-97	2-Feb-98	0.0098	3	3.3
2-Feb-98	29-Jul-98	0.0074	1	0.6
29-Jul-98	15-Oct-98	0.02*	15	19
15-Oct-98	31-Dec-99	0.0106	14	3.7

[a]The ICSS algorithm is applied to the CRSP value-weighted total return index (daily returns) from January 2, 1986 to December 31, 1999, with a 95% confidence level. The standard deviation is calculated based on daily returns between the start and end dates.
[b]Occurrence is the number of days in which five-minute price changes of at least 0.75% in the S&P futures (as identified by Fair, 2002) were recorded. News intensity is defined as occurrence divided by the number of days within a period. Since Fair's sample period ends on October 29, 1999, we calculate the news intensity in the last period from October 15, 1998 to October 29, 1999.
[c]An asterisk denotes standard deviations that are greater than 0.12.

In total, the ICSS algorithm detected 26 variance change points, corresponding to 27 periods over the 14 years in our sample.[3] The sample daily standard deviation is 0.0091 or 0.1448 annualized. (The last two columns in Table 1 are discussed in the next section.)

2.2. *Intensity of news arrival*

In an efficient market, security prices change upon the arrival of new information. Hence, periods of unexpectedly high volatility are associated with significant news. As a check, we examine the "intensity" of news arrival during each period identified by the ICSS algorithm. We expect a positive correlation between volatility and the intensity of news arrival.

Fair (2002) documents all of the 5-minute changes in the S&P 500 futures that were greater than 0.75% (in absolute value), for the period April 21, 1982 to October 29, 1999. Fair considers changes of this magnitude to be significant and indicative of news arrival. Using Appendix A in Fair (2002), we report, for each period, the number of days in which at least one such changes are recorded. In Table 1, we refer to this number as "occurrence." As each period varies in length, we also normalize occurrence by the length of the period. This is our "news intensity" measure.

From Table 1, we can see that there is a good match between the results in the first two steps. There are eight periods that stand out with a standard deviation equal to or greater than 0.012 (marked with an asterisk). These periods also have the highest news intensity, as measured by occurrence/days, in the sample. However, it is not clear whether these eight periods of high volatility were all unanticipated. We look for confirmation in the next step.

2.3. *Comparison between realized volatility and implied volatility*

We employ the CBOE Volatility Index (VIX) as our measure of implied volatility. The VIX is calculated using real-time S&P 100 index option (OEX) bid/ask quotes, and is a weighted average of the implied volatility of eight OEX calls

[3]For robustness, we apply the algorithm to sub-samples with different end dates. Similar results are obtained. Specifically, the change points for the periods with high volatility do not change when the sample is cut, only the change points close to the middle and the two tails of the sample may change.

and puts.[4] The average time to maturity of the options is 30 days. Hence, the VIX can be regarded as an *ex-ante* forecast of the average volatility in the US market over the next 30 days. Realized volatility for the same time period is defined as follows:

$$\mathrm{RV}_t = \sqrt{\frac{\sum_{i=0}^{n-1} R_{t+i}^2 - \left(\frac{1}{n}\sum_{i=0}^{n-1} R_{t+i}\right)^2}{n-1}},$$

where RV_t = realized volatility at time t, R_t = stock return at time t, and $n = 30$. The computation of realized volatility is similar to that in Christensen and Prabhala (1998). As we only observe stock returns on trading days, we set n to 22.

The relationship between implied and realized volatility is an important issue to practitioners and to academics. Option pricing theory implies that option prices are positively correlated with the volatility of the underlying asset. It is widely regarded that implied volatility is an efficient — albeit biased[5] — forecast of future returns volatility of the underlying asset over the remaining life of the option (see Poterba and Summers, 1986; Day and Lewis, 1988; Sheikh, 1989; Harvey and Whaley, 1992; Christensen and Prabhala, 1998). If implied volatility is an efficient forecast of realized volatility, the relationship between implied and realized volatility should be consistent over time. During our sample period, we find that implied volatility was generally greater than realized volatility.[6]

The relationship between implied and realized volatility will change, if the market is hit with major news that has not been incorporated into investors' expectations. Specifically, the impact of the news will be reflected in realized volatility immediately. The change in implied volatility will follow with a

[4]There are no options on the CRSP value-weighted total return index. The correlation coefficient between the daily CRSP value-weighted returns and the daily S&P 100 returns was 0.9624 during our sample period.

[5]This bias, as we explain later, does not affect our results, as we are not interested in estimating volatility. Our goal is simply to study the relationship between implied and realized volatility over time.

[6]This observation is consistent with previous findings. It suggests that investors tend to overestimate future market volatility. One explanation that has been offered is that in an incomplete market, risks that option underwriters bear cannot be perfectly hedged. As a result, the underwriters require a risk premium.

delay, because the news was not part of investors' original information set. As long as the surprise element is large enough, realized volatility will temporarily exceed implied volatility. This sudden reversal in the relationship between implied and realized volatility can be used to detect if a change in volatility is unanticipated.

The difference between implied and realized volatility is shown graphically in Figures 1 and 2. In Figure 1, we calculate realized volatility using the daily CRSP index. In Figure 2, we calculate realized volatility using the daily S&P 100 index, the underlying index of the option.

Because of the high correlation (0.9624) between CRSP and the S&P 100, Figures 1 and 2 look virtually the same. Implied volatility is generally higher than realized volatility, hence the difference is mostly positive. However, around the following months: October 1987, October 1989, August 1990, July 1996, October 1997, and September 1998, realized volatility clearly exceeded implied volatility. (see Figures 3 and 4.) During these months, investors were likely to have been "surprised."

More importantly, these six months coincide with periods of high volatility and high news intensity identified in Table 1 (the eight are marked with an asterisk). The only exception is the period between March 26, 1997 and May 7, 1997. While the standard deviation and the news intensity are both high during

January 2, 1986 – December 31, 1999

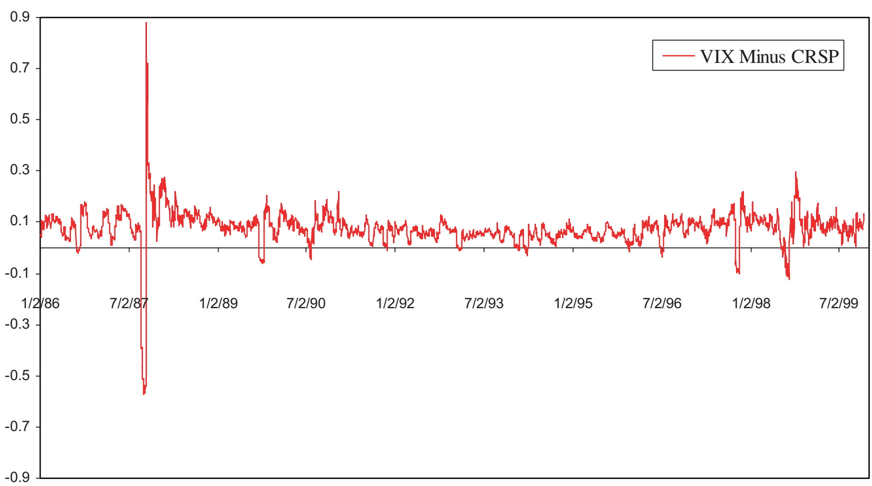


Figure 1. The implied volatility (VIX) versus the realized volatility (CRSP).

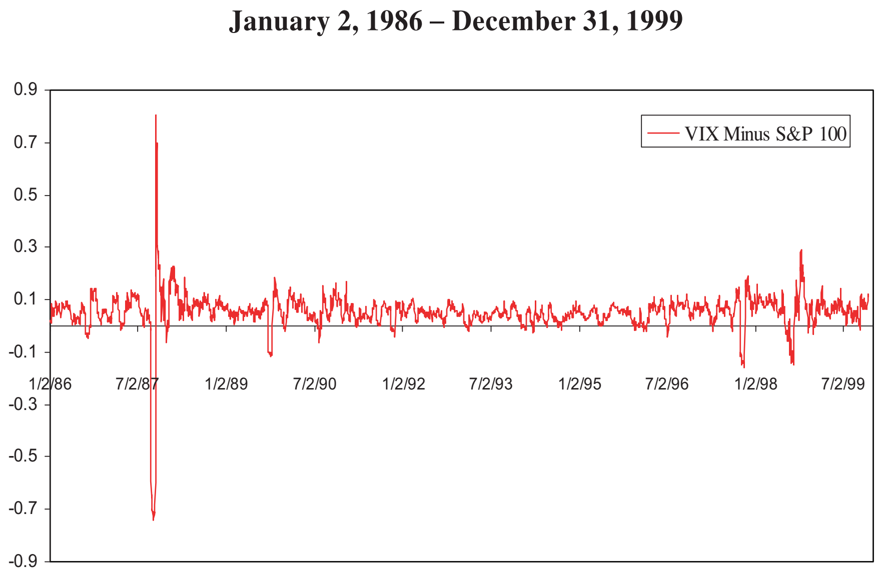


Figure 2. The implied volatility (VIX) versus the realized volatility (S&P 100).

this period, the relationship between implied and realized volatility does not indicate that there was any news that surprised the market. Hence, this period is excluded. Also, as the first two periods of high volatility[7] are consecutive periods, we combine them into one.

Therefore, using the corresponding variance change points as start and end dates, we have identified the following six major, unanticipated episodes of market volatility: (1) October 13, 1987 to January 25, 1988; (2) October 12, 1989 to October 19, 1989; (3) August 1, 1990 to November 12, 1990; (4) July 3, 1996 to August 2, 1996; (5) October 15, 1997 to November 3, 1997; and (6) July 29, 1998 to October 15, 1998. For ease of exposition, we denote these episodes as vol87, vol89, vol90, vol96, vol97, and vol98; that is, by the year in which the unexpectedly high volatility occurred.

These dates coincide with periods of great uncertainty in the US market, such as the stock market crash of 1987, the failure of the UAL buyout in 1989, the invasion of Kuwait by Iraq in 1990, the market correction in 1996, the Asian Financial Crisis in 1997, and the default of Russian debt and the

[7]These two periods are October 13, 1987 to October 30, 1987, and October 30, 1987 to January 25, 1988, respectively.

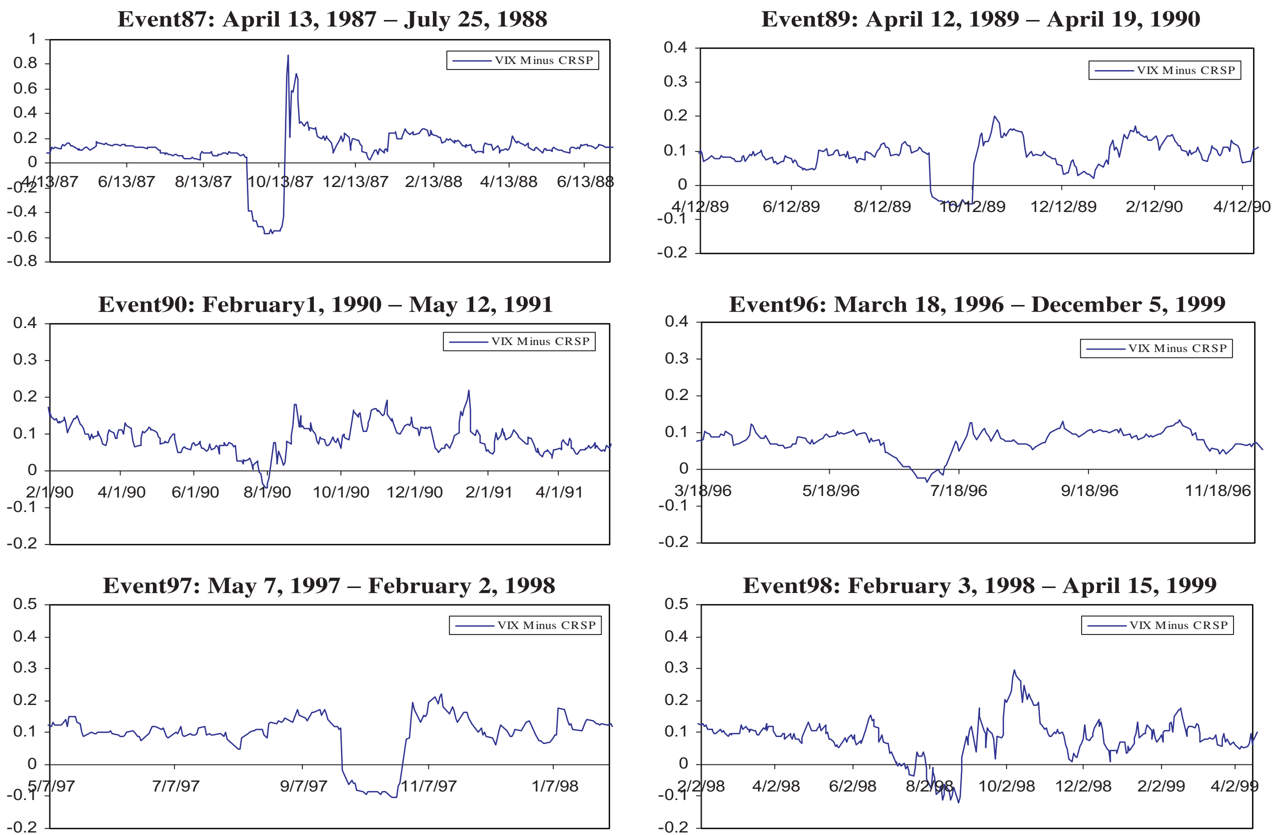


Figure 3. VIX–CRSP, examination periods.

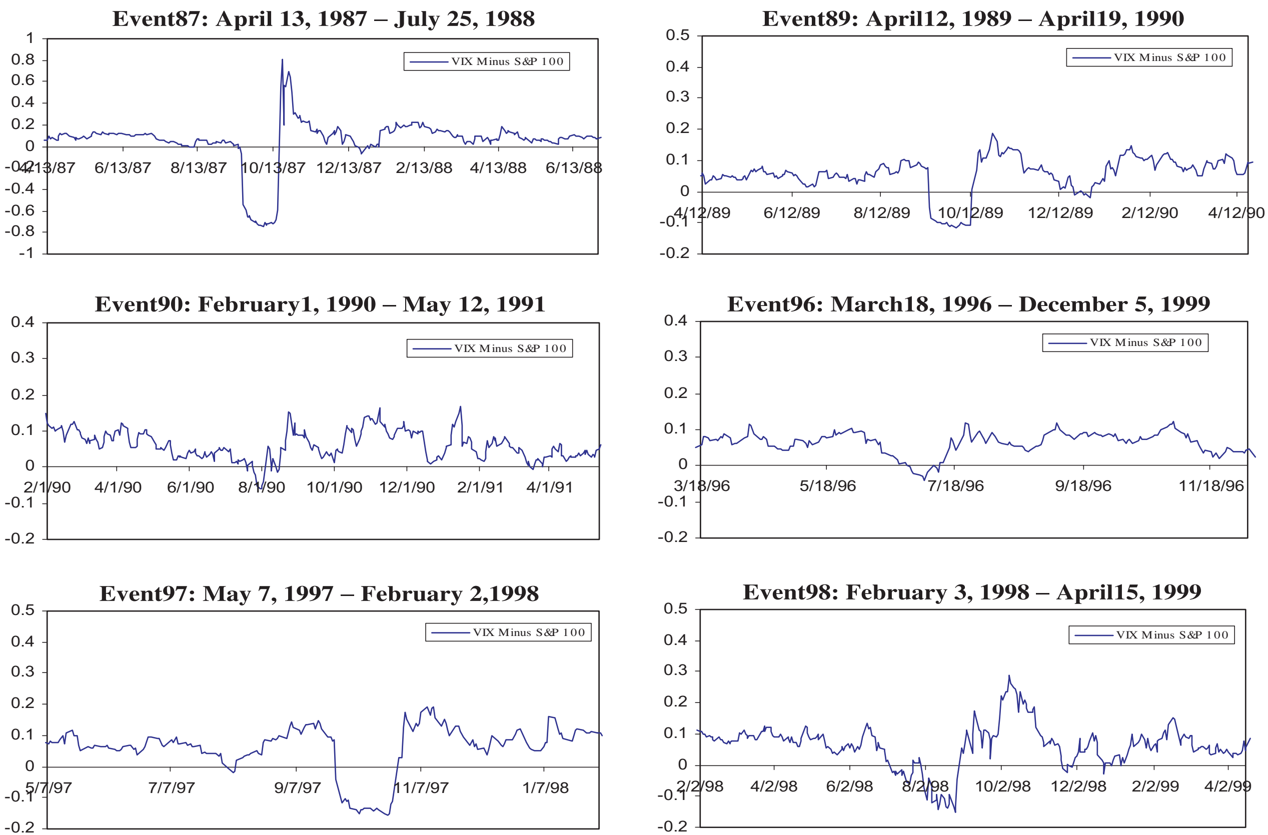


Figure 4. VIX–S&P 100, examination periods.

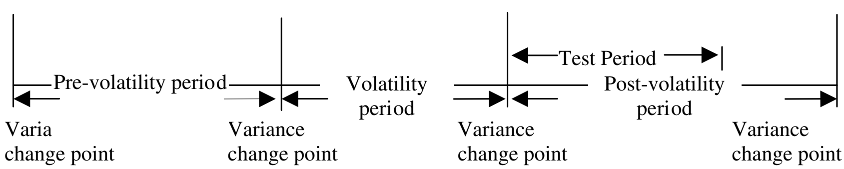
failure of Long Term Capital in 1998. In all six periods, the stock market performed poorly. As volatility was high during these periods, our findings confirm the well-known leverage effect in the time-series finance literature: negative stock price changes have a larger impact on volatility than positive changes.

3. Contrarian and Momentum Strategies

In this section, we test how well popular trading strategies do during periods in which stock market volatility shifts up substantially and unexpectedly. We focus on the performance of contrarian and momentum strategies. The former sells winners and buys losers. The latter sells losers and buys winners. To assess performance, we employ two measures: conditional Jensen's alpha and market-adjusted cumulative average returns (MACARs).

We should note that we do not use volatility-based trading strategies, such as taking positions on options, because the shifts in volatility that we focus on in this study are *unanticipated*. Rather, we examine how investors who routinely employ contrarian and/or momentum strategies are affected by the sudden shifts in volatility.[8]

In the tests, the timeline of *each* of the six volatility periods (vol87, vol89, vol90, vol96, vol97, and vol98) is divided as follows:



In the market efficiency literature, the portfolio formation period and the test period are often designed to be symmetric. We also make the same assumption that the test period has the same length as the volatility period. Further,

[8]It is conceivable that some elaborate volatility-based trading strategies may be able to benefit from volatility shifts. However, as the shifts in our paper are unknown in advance, it is difficult to apply these strategies. Furthermore, the momentum and contrarian strategies that we *do* consider work in completely opposite directions, i.e., one buys winners and the other buys losers. If there is a consistent pattern of the market reaction to major volatility shocks, then we can expect one of the strategies to show some degree of gains consistently.

if a pre-volatility period or a post-volatility period is longer than six months, we cut off the period at six months. The reason is that for volatility periods that are close together, for example, vol89 and vol90, we have to separate the post-volatility period of the former from the pre-volatility period of the latter. The dates of the timeline for each volatility period are presented in Table 2.

To estimate normal returns, we adopt a conditional CAPM that allows the risk characteristics of the portfolios to vary across the pre-volatility, volatility, and post-volatility periods.

3.1. *Sample selection and portfolio formation*

Because illiquid stocks with low trading volume may not be very sensitive to changes in market conditions, we focus on actively traded stocks in the CRSP index. To exclude illiquid stocks, we implement the following sample selection process to the pre-volatility and the volatility periods. To be included in the sample, a stock must (1) trade on at least 90% of the trading days; (2) have an average daily trading volume larger than 6,000 shares in vol87, vol89, and vol90, and 10,000 shares in vol96, vol97, and vol98. These cut-off levels represent approximately the bottom 10th percentile for the corresponding periods (i.e., 1987–1990, and 1996–1998); and, (3) have an average price higher than one dollar. The last criterion excludes penny stocks from the sample, because these stocks have large bid-ask spreads relative to their prices, and are usually illiquid. This restriction has been used in other studies, such as Jegadeesh and Titman (1995). Note that since we only impose these restrictions on the pre-volatility and the volatility periods, and not on the test periods, our results should be relatively free from the sample selection bias or the data-snooping bias.

Table 2. Start and end dates of the pre-event, event, and post-event periods.

Event	Pre-event period		Event period		Post-event period	
	Start	End	Start	End	Start	End
Event87	87-Apr-13	87-Oct-12	87-Oct-13	88-Jan-25	88-Jan-26	88-Jul-25
Event89	89-Apr-12	89-Oct-11	89-Oct-12	89-Oct-19	89-Oct-20	90-Apr-19
Event90	90-Feb-01	90-Jul-31	90-Aug-01	90-Nov-12	90-Nov-13	91-May-12
Event96	96-Mar-18	96-Jul-02	96-Jul-03	96-Aug-02	96-Aug-03	96-Dec-05
Event97	97-May-07	97-Oct-14	97-Oct-15	97-Nov-03	97-Nov-04	98-Feb-02
Event98	98-Feb-02	98-Jul-28	98-Jul-29	98-Oct-15	98-Oct-16	99-Apr-15

To allow for differences across firm size, we run our tests separately for portfolios in the small-cap, middle-cap, and large-cap categories. At the end of our sample period (i.e., 1999), a market capitalization of less than \$1 billion is typically defined to be small-cap. A middle-cap company is one with a market capitalization between \$1 billion and \$3 billion. A large-cap company is one with a capitalization that is larger than \$3 billion. Since our data span 14 years, the above criteria cannot be employed for the entire sample period due to stock price inflation. In 1999, \$1 billion and \$3 billion correspond to roughly the bottom 10th and 20th percentiles, respectively, in market capitalization in the CRSP database. We employ these percentiles to classify stocks: stocks are first sorted by their average market capitalization in the pre-volatility period. The small-cap group consists of stocks in the bottom 10th percentile. The middle-cap group consists of stocks above the bottom 10th percentile up to the bottom 20th percentile. The rest of the stocks belong to the large-cap group.

To form our set of loser and winner portfolios, we rank the stocks in each capitalization group by their MACARs (following De Bondt and Thaler, 1985) during each volatility period.[9] The CRSP value-weighted index is again used as the market, and the three-month US Treasury yield is used as the risk-free rate. In the tests described below, we focus on portfolios of stocks in the bottom 5th, 10th, 15th, 20th, and 25th percentiles (the losers), as well as in the top 5th, 10th, 15th, 20th, and 25th percentiles (the winners).

3.2. *Conditional Jensen's alpha*

The conditional CAPM can be expressed as:

$$E(r_{p,T+1}|\phi_T) = \beta_{p,T+1}(\phi_T)E(r_{m,T+1}|\phi_T),$$

where T denotes the volatility period; $T+1$ denotes the test period; ϕ_T is the information set prior to the test period $T+1$; $r_{p,T+1}$ is the excess return of portfolio p in the test period; $r_{m,T+1}$ is the excess return of the market; and $\beta_{p,T+1}(\phi_T)$ is the portfolio beta in the test period and it is a function of ϕ_T. Removing the expectation operators, the previous equation in its ex-post

[9]Conrad and Kaul (1993) promote the use of holding-period returns instead of MACARs. One problem with holding-period returns is that they are extremely sensitive to the start and end dates of the holding period. Therefore, they are better suited to studies where the investment horizon is fixed, e.g., three months, six months, etc.

format is:

$$r_{p,T+1} = \alpha_{p,T+1}(\phi_T) + \beta_{p,T+1}(\phi_T) r_{m,T+1} + u_{T+1},$$

where $\alpha_{p,T+1}(\phi_T)$ is the conditional alpha in the test period.

We condition the pricing model on time, recognizing that when volatility shifts substantially, the risk characteristics of the portfolios will change. Specifically, we allow alpha and beta to vary across sub-periods.[10]

$$r_{pt} = \alpha + \alpha^E D_t^E + \alpha^T D_t^T + (\beta + \beta^E D_t^E + \beta^P D_t^P) r_{mt} + \varepsilon_t,$$

where D^E, D^T, and D^P are dummy variables corresponding to the volatility, test, and post-volatility periods. In other words, there are three possible regimes. Each sub-period can have a different alpha and a different beta. Using this specification, our tests can be stated as follows:

The contrarian strategy works if alpha is negative (positive) in the volatility period, and positive (negative) in the test period.
The momentum strategy works if alpha is negative (positive) in the volatility period, and negative (positive) in the test period.

For vol87, we use a slightly different specification, because it consists of two consecutive volatility periods: October 13, 1987 to October 30, 1987, and November 1, 1987 to January 25, 1988. We modify the model for vol87 as follows:

$$\begin{aligned} r_{pt} &= \alpha + \alpha^{E_1} D_t^{E_1} + \alpha^{E_2} D_t^{E_2} + \alpha^T D_t^T + \Big(\beta + \beta^{E_1} D_t^{E_1} \\ &\quad + \beta^{E_2} D_t^{E_2} + \beta^P D_t^P\Big) r_{mt} + \varepsilon_t, \end{aligned}$$

where E_1 corresponds to the volatility period October 13, 1987 to October 30, 1987, and E_2 corresponds to the volatility period November 1, 1987 to January 25, 1988. D^{E_1} and D^{E_2} are the corresponding dummy variables.

[10]Jagannathan and Wang (1996), Ferson and Schadt (1996), and Christopherson *et al*. (1998), among others, employ economic/financial factors in their conditional CAPM studies. However, these factors cannot be used here because we are working with daily data over a relatively short horizon.

3.3. *Market-adjusted cumulative average returns*

The second test focuses on the MACAR of each portfolio. If a loser (winner) portfolio experiences a subsequent price reversal in the test period, then the contrarian strategy works. If a loser (winner) portfolio encounters further downward (upward) pressure during the test period, then the momentum strategy works. In the results, we look for MACARs that are significantly different from zero.

3.4. *Results*

3.4.1. *Conditional Jensen's alphas*

To assess the results, we compare α^E with α^T for each portfolio (small-/mid-/large-cap, winner/loser, percentile), focusing on the α values that are statistically significant at the 5% level. If the signs of α^E and α^T are opposite, then the results are in favor of a contrarian strategy. If the signs are the same, then the results are in favor of a momentum strategy. The results are summarized in Table 3.

Within each market capitalization, there are five portfolios, corresponding to the bottom (top) 5th, 10th, 15th, 20th, and 25th percentile stocks for the loser (winner) category. The bottom panel in Table 3 shows that in aggregate, 41% of the portfolios exhibit a reward for adopting either a contrarian or a momentum trading strategy. Within this set of portfolios, a contrarian strategy is more than twice as likely (29% vs. 12%) to generate abnormal returns. The top panel shows the results for the individual categories. For example, a contrarian strategy produces abnormal returns in 60% of the small-cap loser portfolios, while a momentum strategy produces abnormal returns in only 17%. Put differently, a contrarian strategy produces normal returns or better in 83% of the small-cap loser portfolios. Results for the other types of portfolios are more scattered.

3.4.2. *MACARs*

Here, we focus on MACARs that are significantly different from zero statistically. Results based on the average daily MACARs[11] for each portfolio during

[11] To be consistent across periods, we report the average daily MACARs because the test periods differ in length.

Table 3. Summary of results — Jensen's alphas.

Volatility period	Investment strategy	Loser			Winner			Total	Percentage[a]	Percentage[b]
		Small	Middle	Large	Small	Middle	Large			
vol87	Contrarian	5	2	0	0	0	0	7	23	23
	Momentum	0	0	0	0	0	0	0	0	
vol89	Contrarian	0	0	0	4	0	0	4	13	43
	Momentum	0	5	3	0	0	1	9	30	
vol90	Contrarian	5	4	5	0	0	0	14	47	50
	Momentum	0	0	0	1	0	0	1	3	
vol96	Contrarian	3	4	0	0	0	0	7	23	23
	Momentum	0	0	0	0	0	0	0	0	
vol97	Contrarian	0	0	0	5	5	0	10	33	50
	Momentum	5	0	0	0	0	0	5	17	
vol98	Contrarian	5	5	0	0	0	0	10	33	53
	Momentum	0	0	0	4	2	0	6	20	
Total	Contrarian	18	15	5	9	5	0			
	Momentum	5	5	3	5	2	1			
Percentage[c]	Contrarian	60	50	17	30	17	0			
	Momentum	17	17	10	17	7	3			

Aggregate

Investment strategy	Total	Percentage[d]	Percentage[e]
Contrarian	52	29	41
Momentum	21	12	

Notes: In each volatility period, we examine a total of 30 portfolios, 5 in each category (5th, 10th, 15th, 20th, 25th percentile).
[a]Percentage for a given type of investment strategy out of 30 portfolios.
[b]Percentage for both types of investment strategies out of 30 portfolios.
[c]Percentage for a given type of portfolios out of 30 portfolios.
[d]Percentage for a given type of investment strategies out of 180 portfolios (30 × 6 volatility periods).
[e]Percentage for both types of investment strategies out of 180 portfolios.

the test period confirm, by and large, our findings based on the conditional Jensen's alphas. For example, most of the action can again be found in the small-cap loser portfolios: a contrarian strategy produces abnormal returns in 67% of these portfolios (see Table 4).

Studying cumulative returns provides additional insight from an investment perspective. As an illustration, we consider the MACARs for the loser portfolios in vol87, vol90, vol96, and vol98.[12] Loser portfolios have daily MACARs that range from 0.23% to 0.58% in the small-cap group (not shown). Hence, even after adjusting for transaction costs, the profits from buying small-cap loser portfolios are still economically significant. Assuming a proportional transaction cost of 0.1%[13] of the trade, the annualized return for a MACAR of 0.23% is equal to $(0.23\% - 0.1\%) \times 252 = 32.76\%$, at the low end.

3.4.3. *Discussion*

Despite the findings above, we should make clear that buying small-cap losers after a major, unanticipated jump in market volatility do not always lead to sizable positive abnormal returns. Such was the case in vol89 and vol97, when abnormal returns were not statistically different from zero. In addition, an investor would have to recognize the shift in volatility right away in order to profit from this strategy of buying small-cap loser portfolios. For example, when we delay the start date of the portfolio formation to half way through a volatility period (with the test period adjusted accordingly), the profit opportunity disappears. The results are shown in Tables 5 and 6. Immediate recognition of a significant shift in volatility is very difficult without the benefit of hindsight — recall that the volatility periods we identified are determined retroactively. Hence, in the context of an efficient market, contrarian and momentum strategies do not necessarily outperform a diversified passive portfolio during the types of periods in question.

---

[12]To conserve space, we do not report the value of the MACARs for each portfolio in each period here. The tables are available from the authors upon request.

[13]Bhardwaj and Brooks (1992) report that when the dollar volume (V) is between \$50,001 and \$500,000, the typical commission is \$134 + 0.001V. In other words, the variable cost is 0.1%. For on-line investors, the cost is even lower.

Table 4. Summary of results — Market-adjusted cumulative average returns (MACAR).

Volatility period	Investment strategy	Loser			Winner			Total	Percentage[a]	Percentage[b]
		Small	Middle	Large	Small	Middle	Large			
vol87	Contrarian	5	3	2	0	0	4	14	47	47
	Momentum	0	0	0	0	0	0	0	0	
vol89	Contrarian	0	0	0	0	0	0	0	0	23
	Momentum	0	5	2	0	0	0	7	23	
vol90	Contrarian	5	4	5	0	0	3	17	57	60
	Momentum	0	0	0	1	0	0	1	3	
vol96	Contrarian	5	2	0	0	0	0	7	23	43
	Momentum	0	0	0	5	1	0	6	20	
vol97	Contrarian	0	0	0	4	0	0	4	13	27
	Momentum	4	0	0	0	0	0	4	13	
vol98	Contrarian	5	1	0	0	0	1	7	23	23
	Momentum	0	0	0	0	0	0	0	0	
Total	Contrarian	20	10	7	4	0	8			
	Momentum	4	5	2	6	1	0			
Pecentage[c]	Contrarian	67	33	23	13	0	27			
	Momentum	13	17	7	20	3	0			

Aggregate

Investment strategy	Total	Percentage[d]	Percentage[e]
Contrarian	49	27	37
Momentum	18	10	

Notes: In each volatility period, we examine a total of 30 portfolios, 5 in each category (5th, 10th, 15th, 20th, 25th percentile).
[a]Percentage for a given type of investment strategy out of 30 portfolios.
[b]Percentage for both types of investment strategies out of 30 portfolios.
[c]Percentage for a given type of portfolios out of 30 portfolios.
[d]Percentage for a given type of investment strategies out of 180 portfolios (30 × 6 volatility periods).
[e]Percentage for both types of investment strategies out of 180 portfolios.

Table 5. Summary of results — Jensen's alphas (portfolios formed half way through a volatility period).

Volatility period	Investment strategy	Loser			Winner			Total	Percentage[a]	Percentage[b]
		Small	Middle	Large	Small	Middle	Large			
vol87	Contrarian	0	0	0	0	0	0	0	0	0
	Momentum	0	0	0	0	0	0	0	0	
vol89	Contrarian	0	0	0	5	0	0	5	17	20
	Momentum	1	0	0	0	0	0	1	3	
vol90	Contrarian	5	5	0	0	0	0	10	33	50
	Momentum	0	0	0	3	2	0	5	17	
vol96	Contrarian	1	5	0	0	0	0	6	20	20
	Momentum	0	0	0	0	0	0	0	0	
vol97	Contrarian	0	0	0	5	1	0	6	20	40
	Momentum	5	1	0	0	0	0	6	20	
vol98	Contrarian	0	0	0	0	0	0	0	0	20
	Momentum	0	0	0	4	0	2	6	20	
Total	Contrarian	6	10	0	10	1	0			
	Momentum	6	1	0	7	2	2			
Percentage[c]	Contrarian	20	33	0	33	3	0			
	Momentum	20	3	0	23	7	7			

Aggregate

Investment strategy	Total	Percentage[d]	Percentage[e]
Contrarian	27	15	25
Momentum	18	10	

Notes: In each volatility period, we examine a total of 30 portfolios, 5 in each category (5th, 10th, 15th, 20th, 25th percentile).
[a]Percentage for a given type of investment strategy out of 30 portfolios.
[b]Percentage for both types of investment strategies out of 30 portfolios.
[c]Percentage for a given type of portfolios out of 30 portfolios.
[d]Percentage for a given type of investment strategies out of 180 portfolios (30 × 6 volatility periods).
[e]Percentage for both types of investment strategies out of 180 portfolios.

Table 6. Summary of results — market-adjusted cumulative average returns (macar) (portfolios formed half way through a volatility period).

Volatility period	Investment strategy	Loser			Winner			Total	Percentage[a]	Percentage[b]
		Small	Middle	Large	Small	Middle	Large			
vol87	Contrarian	0	0	0	0	0	0	0	0	0
	Momentum	0	0	0	0	0	0	0	0	
vol89	Contrarian	0	0	0	0	0	0	0	0	0
	Momentum	0	0	0	0	0	0	0	0	
vol90	Contrarian	5	0	0	0	0	0	5	17	47
	Momentum	0	0	0	5	4	0	9	30	
vol96	Contrarian	0	5	0	0	0	0	5	17	17
	Momentum	0	0	0	0	0	0	0	0	
vol97	Contrarian	0	0	0	5	0	0	5	17	27
	Momentum	3	0	0	0	0	0	3	10	
vol98	Contrarian	0	0	0	0	0	0	0	0	3
	Momentum	0	0	0	0	0	1	1	3	
Total	Contrarian	5	5	0	5	0	0			
	Momentum	3	0	0	5	4	1			
Percentage[c]	Contrarian	17	17	0	17	0	0			
	Momentum	10	0	0	17	13	3			

Aggregate

Investment strategy	Total	Percentage[d]	Percentage[e]
Contrarian	15	8	16
Momentum	13	7	

Notes: In each volatility period, we examine a total of 30 portfolios, 5 in each category (5th, 10th, 15th, 20th, 25th percentile).
[a]Percentage for a given type of investment strategy out of 30 portfolios.
[b]Percentage for both types of investment strategies out of 30 portfolios.
[c]Percentage for a given type of portfolios out of 30 portfolios.
[d]Percentage for a given type of investment strategies out of 180 portfolios (30 × 6 volatility periods).
[e]Percentage for both types of investment strategies out of 180 portfolios.

4. Conclusion

In this paper, our primary goal is to implement a three-stage procedure to identify large, unanticipated shifts in stock market volatility, which we show, that can last days or months. Between January 2, 1986, and December 31, 1999, we have identified six major shocks in the US market volatility. The contributions of our procedure are: (1) in identifying shifts in volatility that were *unanticipated*, (2) in showing that high volatility periods coincide with periods of high news intensity, and (3) in linking the major shocks to macroeconomic events. In addition, in all six periods, the stock market performed poorly. As volatility was high during these periods, our findings confirm the well-known leverage effect in the time-series finance literature: negative stock price changes have a larger impact on volatility than positive changes.

A natural question to ask is how investors are affected by such shifts in volatility. As an interesting extension, we explore how popular trading strategies such as contrarian and momentum portfolios do during and after periods in which stock market volatility shifts up substantially and unexpectedly. Based on a sample of 180 portfolios, we find that a contrarian strategy slightly dominates a momentum strategy in loser portfolios during these periods. In the case of small-cap losers, the returns could be substantial. In other portfolios, there is no clear pattern. We further demonstrate that an investor would have to recognize the shift in volatility right away in order to profit from buying small-cap loser portfolios. This is very hard to achieve without the benefit of hindsight. Hence, in the context of an efficient market, we conclude that contrarian and momentum strategies do not necessarily outperform a diversified passive portfolio during the times of great uncertainty.

References

Aggarwal, R., C. Inclan and R. Leal, "Volatility in Emerging Stock Markets." *Journal of Financial and Quantitative Analysis* 34, 33–55 (1999).

Christensen, B. J. and N. R. Prabhala, "The Relation Between Implied and Realized Volatility." *Journal of Financial Economics* 50, 125–150 (1998).

Conrad, J. and G. Kaul, "Long-Term Market Overreaction or Biases in Computed Returns?" *Journal of Finance* 48, 39–63 (1993).

Daniel, K. D., D. Hirshleifer and A. Subrahmanyam, "Investor Psychology and Security Market Under- and Overreactions." *Journal of Finance* 53, 1839–1886 (1998).

Day, T. and C. Lewis, "The Behaviour of the Volatility Implicit in Option Prices." *Journal of Financial Economics* 22, 103–122 (1988).

De Bondt, W. F. M. and R. H. Thaler, "Does the Stock Market Overreact?" *Journal of Finance* 40, 793–808 (1985).

Fair, R, "Events that Shook the Market." *Journal of Business* 75, 713–731 (2002).

Harvey, C. R. and R. Whaley, "Market Volatility Prediction and the Efficiency of the S&P 100 Index Option Market." *Journal of Financial Economics* 31, 43–73 (1992).

Inclan, C. and G. C. Tiao, "Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance." *Journal of the American Statistical Association* 89, 913–923 (1994).

Jegadeesh, N. and S. Titman, "Overreaction, Delayed Reaction, and Contrarian Profits." *Review of Financial Studies* 8, 973–993 (1995).

Poterba, J. and L. Summers, "The Persistence of Volatility and Stock Market Fluctuations." *American Economic Review* 76, 1142–1151 (1986).

Sheikh, A, "Stock Splits, Volatility Increases and Implied Volatility." *Journal of Finance* 44, 1361–1372 (1989).

This page intentionally left blank

Chapter 14

The September Phenomenon of US Equity Market

Anthony Yanxiang Gu
SUNY College at Geneseo, USA

John T. Simon*
Governors State University, USA

Mean September return of the US stock market is significantly negative and is the lowest among the calendar months. The phenomenon is more apparent for large stocks and looks strengthened recently, particularly for large stocks. September performance of the stock market is directly connected to GDP growth, inflation rate, and stock market performance of the year, and inversely related to interest rate. Tax-loss selling, "window dressing", and macroeconomic seasonality could also contribute to the poor September performance.

Keywords: September performance; equity market.

Stock market anomalies have been one of the most intriguing issues in finance. The most puzzling empirical findings include the January effect, the abnormally large returns on common stocks in most months of January (Wachtel, 1942; Rozeff and Kinney, 1976), and the weekend effect, the abnormally high average Friday returns and significantly negative average Monday returns (Cross, 1973; French, 1980).

Some practitioners in the US stock market believe that stock market generally performs the worst in the month of September (Browning, 2003; Baldwin, 2003). In this study, we examine each month's return of major US stock indices and test whether September is generally the worst month for the US equity markets.

1. The Data

We have used three indices for the analyses: (1) the Dow Jones Industrial Average (DJIA) price-weighted index from 1896 (the market was closed from August to December in 1914 due to the First World War) obtained from Dow

*Corresponding author.

Jones, Inc.; (2) the S&P 500 value-weighted index from 1950 obtained from Standard and Poor; and (3) the NASDAQ Composite value-weighted index from 1971 (the year when NASDAQ started) obtained from NASDAQ. All data are up to the year 2005. The returns are calculated as the natural logarithm differentials between two successive closes. We have also used the data of GDP growth, inflation, and Treasury bill rate from 1930 to 2002, obtained from the International Monetary Fund.

The test periods for the DJIA include 1896–2005, 1896–1949, 1950–1989, and 1990–2005. The test periods for the S&P 500 Index include 1950–2005, 1950–1989, and 1990–2005. The test periods for the NASDAQ Composite Index include 1971–2005, 1971–1989, and 1990–2005. The year 1990 was chosen as a cut-off point because it approximately demarcates the high-tech era, in which most of the individual investors obtained access to personal computers, Internet, and e-mail, and the investment environment was characterized by more information, faster communication, and real-time order execution. The division of the periods is somewhat arbitrary, because it is impossible to determine exactly in which year an era's major impacts on stock market start and end, that is, the impact of highly advanced technological development on stock market behavior may not start exactly in 1990. However, as we examine the averages, the results can provide meaningful indications for the purpose of this study.

2. Empirical Findings

The mean monthly returns of the indices during the specified periods are depicted in Figures 1–3 and the statistics of the empirical findings are presented in Table 1. In Table 1, the mean and standard deviation of each month's return for the three indices for their entire observation periods and subperiods are reported. The t-statistics in the column of September indicate the comparison of September return with 0, and the number below it is the statistical significance of the t-statistics for the observation periods. The test is one-tailed, the null hypothesis is that September return is greater than or equal to zero, and the alternative hypothesis is that September return is less than zero. The t-statistics in the columns of the months other than September indicate the pairwise comparison of that month's return with September return, and the number below it is the statistical significance of the t-statistics for the observation periods. The test is one-tailed, the null hypothesis is that September

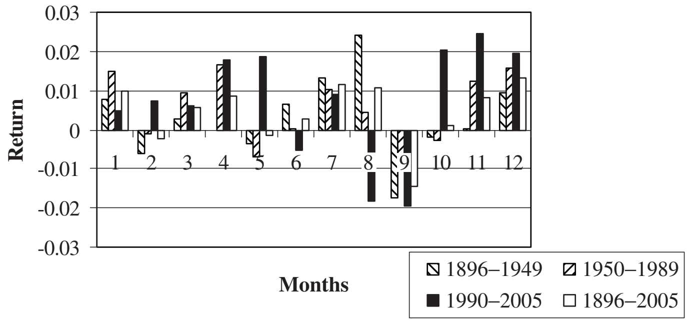


Figure 1. DJIA monthly returns.

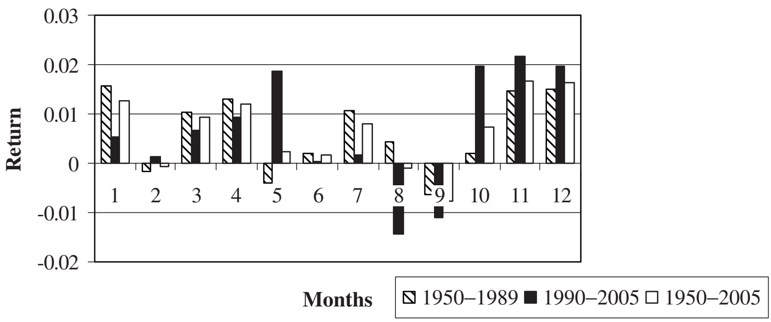


Figure 2. S&P 500 monthly returns.

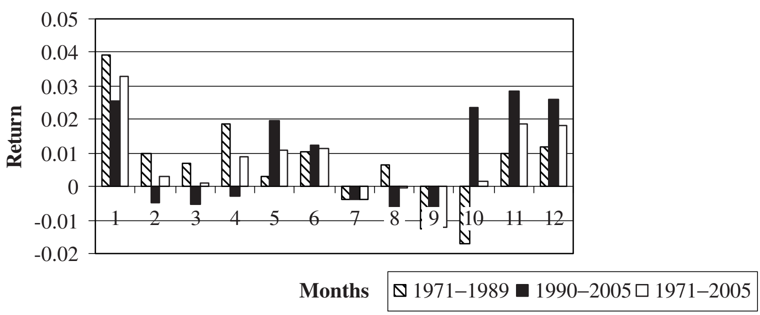


Figure 3. NASDAQ monthly returns.

return is at least as much as that month's return, and the alternative hypothesis is that September return is less than that month's return.

Mean September return is negative for all the three indices for all their observation periods. As shown in Table 1, for the entire observation period

Table 1. Comparison of September return with return in other months.

		January	February	March	April	May	June	July	August	September	October	November	December
DJIA													
1896–2005	Mean	0.0099	−0.0022	0.0056	0.0085	−0.0016	0.0025	0.0114	0.0108	−0.0143	0.0012	0.0083	0.0133
	St. dev.	0.0450	0.0413	0.0532	0.0620	0.0569	0.0532	0.0544	0.0573	0.0641	0.0612	0.0571	0.0419
	t-stat[1]	3.2864	1.8023	2.7469	3.1753	1.8212	2.0475	3.8840	2.9266	−2.3320[2]	1.9388	2.9831	4.3080
	Significance	0.0007***	0.0372**	0.0035***	0.0010***	0.0357**	0.0215**	0.0001***	0.0021***	0.0108**	0.0276**	0.0018***	0.0000***
1896–1949	Mean	0.0076	−0.0061	0.0027	−0.0002	−0.0037	0.0067	0.0130	0.0243	−0.0175	−0.0017	0.0004	0.0095
	St. dev.	0.0411	0.0480	0.0691	0.0793	0.0739	0.0679	0.0674	0.0620	0.0805	0.0682	0.0660	0.0521
	t-stat[1]	1.9844	1.0570	1.6417	1.4562	1.1698	1.5839	2.6260	2.8338	−1.5864[2]	1.2613	1.4062	2.3773
	Significance	0.0263**	0.1478	0.0534*	0.0757*	0.1238	0.0596*	0.0057***	0.0033***	0.0594*	0.1064	0.0828*	0.0106**
1950–1989	Mean	0.0151	−0.0010	0.0094	0.0165	−0.0069	0.0001	0.0101	0.0046	−0.0080	−0.0025	0.0123	0.0159
	St. dev.	0.0530	0.0326	0.0296	0.0371	0.0344	0.0347	0.0413	0.0449	0.0398	0.0579	0.0472	0.0269
	t-stat[1]	2.3986	0.7851	2.0875	3.2905	0.1363	0.9592	2.3962	1.3841	−1.2693[2]	0.4652	2.0564	3.5177
	Significance	0.0107**	0.2186	0.0217**	0.0011***	0.4462	0.1717	0.0107**	0.0871*	0.1059	0.3222	0.0232**	0.0006***
1990–2005	Mean	0.0047	0.0075	0.0061	0.0176	0.0187	−0.0054	0.0090	−0.0183	−0.0195	0.0203	0.0247	0.0195
	St. dev.	0.0359	0.0374	0.0383	0.0421	0.0278	0.0314	0.0323	0.0586	0.0532	0.0408	0.0451	0.0352
	t-stat[1]	1.8731	2.1997	1.5310	2.3024	2.6512	1.2807	2.4085	0.0570	−1.4644[2]	2.0048	2.5817	2.6062
	Significance	0.0403**	0.0220**	0.0733*	0.0180**	0.0091***	0.1099	0.0147**	0.4776	0.0819*	0.0317**	0.0104**	0.0099***
S&P 500													
1950–2005	Mean	0.0128	−0.0007	0.0093	0.0120	0.0025	0.0016	0.0082	−0.0012	−0.0076	0.0072	0.0167	0.0164
	St. dev.	0.0473	0.0328	0.0336	0.0386	0.0361	0.0345	0.0408	0.0484	0.0436	0.0538	0.0435	0.0319
	t-stat[1]	2.7569	1.0135	2.1866	3.0267	1.3559	1.3939	2.6070	0.7563	−1.3101[2]	1.4455	3.1205	3.9108
	Significance	0.0040***	0.1576	0.0165**	0.0019***	0.0903*	0.0845*	0.0059***	0.2264	0.0978*	0.0770*	0.0014***	0.0001***
1950–1989	Mean	0.0158	−0.0015	0.0103	0.0131	−0.0040	0.0020	0.0108	0.0042	−0.0063	0.0022	0.0148	0.0151
	St. dev.	0.0509	0.0303	0.0309	0.0386	0.0363	0.0351	0.0412	0.0447	0.0413	0.0596	0.0431	0.0301
	t-stat[1]	2.4124	0.5387	1.9137	2.5969	0.2611	0.9833	2.4208	1.1841	−0.9582[2]	0.6644	2.3082	3.1242
	Significance	0.0103**	0.2966	0.0315**	0.0066***	0.3977	0.1658	0.0101**	0.1218	0.1719	0.2552	0.0132**	0.0017***

(*Continued*)

Table 1. (*Continued*)

		January	February	March	April	May	June	July	August	September	October	November	December
1990–2005	Mean	0.0052	0.0014	0.0068	0.0094	0.0186	0.0005	0.0016	−0.0145	−0.0111	0.0197	0.0216	0.0197
	St. dev.	0.0370	0.0396	0.0407	0.0398	0.0310	0.0341	0.0403	0.0560	0.0502	0.0336	0.0453	0.0370
	t-stat[1]	1.2943	1.2529	1.0626	1.5214	2.1907	1.1604	1.0412	−0.1640	−0.8828[2]	1.8381	2.1284	2.3109
	Significance	0.1076	0.1147	0.1524	0.0745*	0.0223**	0.1320	0.1571	0.4359	0.1956	0.0430**	0.0251**	0.0177**
NASDAQ													
1971–2005	Mean	0.0330	0.0032	0.0011	0.0088	0.0107	0.0111	−0.0040	−0.0006	−0.0120	0.0016	0.0186	0.0184
	St. dev.	0.0604	0.0721	0.0568	0.0595	0.0524	0.0489	0.0548	0.0703	0.0651	0.0887	0.0762	0.0556
	t-stat[1]	2.9781	1.0319	0.9280	1.5247	1.7580	1.6942	0.7525	0.7413	−1.0916[2]	0.6777	1.8774	2.6135
	Significance	0.0027***	0.1547	0.1800	0.0683*	0.0439**	0.0497**	0.2285	0.2318	0.1414	0.2513	0.0345**	0.0066***
1971–1989	Mean	0.0394	0.0100	0.0067	0.0186	0.0030	0.0101	−0.0040	0.0064	−0.0127	−0.0171	0.0101	0.0120
	St. dev.	0.0614	0.0434	0.0553	0.0442	0.0442	0.0344	0.0475	0.0558	0.0484	0.1054	0.0583	0.0402
	t-stat[1]	2.6468	1.3342	1.0168	2.2656	0.9968	1.8903	0.8684	1.4612	−1.1488[2]	−0.1527	1.2828	2.0347
	Significance	0.0085***	0.0994*	0.1614	0.0180**	0.1660	0.0375**	0.1983	0.0806*	0.1328	0.4402	0.1079	0.0284**
1990–2005	Mean	0.0257	−0.0049	−0.0055	−0.0028	0.0198	0.0123	−0.0040	−0.0089	−0.0111	0.0238	0.0286	0.0259
	St. dev.	0.0603	0.0969	0.0598	0.0736	0.0609	0.0631	0.0641	0.0856	0.0824	0.0593	0.0942	0.0705
	t-stat[1]	1.5473	0.2426	0.2606	0.3303	1.4393	0.8753	0.3479	0.0723	−0.5402[2]	1.2741	1.3614	1.7389
	Significance	0.0713*	0.4058	0.3990	0.3729	0.0853*	0.1976	0.3664	0.4716	0.2985	0.1110	0.0967*	0.0513*

Note: The data for this table were obtained from Dow Jones, Inc., Standard and Poor, and NASDAQ. Return is calculated as natural logarithm differential between the close of a month's last trading day and the close of the previous month's last trading day. In addition to the mean return and standard deviation of the months, the column for each month other than September shows the t-statistics of pairwise comparison of that month's return with September return, and the significance level. The column for September shows the t-statistics of comparing September return with zero, and the significance level.

[1] t-statistics for the hypothesis that September mean is less than the given month's mean.

[2] t-statistics for the hypothesis that September mean is less than zero.

*Significance at the 10% level.

**Significance at the 5% level.

***Significance at the 1% level.

of each index, September return is significantly negative for the DJIA Index from 1896 to 2005 (mean = −1.43%, t-statistics = −2.33), for the S&P 500 Index from 1950 to 2005 (mean = −0.76%, t-statistics = −1.31), and for the NASDAQ Composite Index from 1971 to 2005 (mean −1.20%, t-statistics = −1.09). In general, the level of significance declines as the number of observations decreases.

September return is generally the lowest compared with that of all the other months for all three indices and for all their observation periods. For this comparison, the null hypothesis is that September return is larger than the return in the other months, and the alternative hypothesis is that September return is smaller than the return in the other months. As shown in Table 1, for the DJIA Index during the period 1896–2005, each of the other 11 months' performance exceeds that of September's at a 1% or 5% level of statistical significance. For example, the mean June return is 0.25% and the standard deviation is 0.0532; and in September, the mean return is −1.43% and the standard deviation is 0.0641. A pairwise comparison of September versus June return for all the years in that period shows a t-statistics of 2.0475 whose p-value is 0.0215, which is statistically significant at the 5% level. There are two exceptions, however: for the period 1990–2005, the S&P 500 Index's mean August return is −1.45%, which is below the mean September return of −1.11%; and for the period 1971–1989, the NASDAQ Composite Index's mean October return is −1.71%, which is below the mean September return of −1.27%. But neither of these differences is statistically significant.

September return in the year 2001 is an outlier and is excluded in Table 2. Because of the September 11 attacks, the US stock market closed until September 17 and the market experienced significant losses during the week. Without the outlier, September mean return is still the lowest for the three indices for their entire observation periods; however, in the subperiod 1990–2005, mean August return is lower than mean September return for the DJIA and the S&P 500 indices, and April, July, and August mean returns are lower than that of September for the NASDAQ Composite Index. Nonetheless, the differences are not statistically significant.

The comparison results are summarized in Table 3. In this table, column 3 shows the number of the 11 months in the year in which the month's mean return exceeded September mean return, and column 4 shows the number of months whose return exceeded September return at the 1% level of significance. For example, for the S&P 500 Index during the period 1950–2005,

Table 2. Comparison of September return with return in other months (data for year 2001 is excluded).

		January	February	March	April	May	June	July	August	September	October	November	December
DJIA													
1896–2005	Mean	0.0099	−0.0019	0.0062	0.0079	−0.0018	0.0029	0.0115	0.0114	−0.0134	0.001	0.0076	0.0133
	St. dev.	0.0452	0.0414	0.0530	0.0619	0.0572	0.0533	0.0547	0.0572	0.0636	0.0614	0.0569	0.0421
	t-stat[1]	3.1573	1.7077	2.6811	3.0038	1.6748	1.9646	3.7546	2.8630	−2.1830[2]	1.7948	2.8078	4.1652
	Significance	0.0010***	0.0453**	0.0043***	0.0017***	0.0485**	0.0260**	0.0001***	0.0025***	0.0156**	0.0378**	0.0030***	0.0000***
1990–2005	Mean	0.0044	0.0105	0.0106	0.0133	0.0188	−0.0032	0.0095	−0.0157	−0.0129	0.0200	0.0209	0.0196
	St. dev.	0.0372	0.0367	0.0351	0.0397	0.0288	0.0312	0.0334	0.0597	0.0480	0.0422	0.0439	0.0365
	t-stat[1]	1.4815	1.8660	1.3258	2.0644	2.3041	0.9022	2.0654	−0.1256	−1.0452[2]	1.6538	2.3225	2.2542
	Significance	0.0803*	0.0416**	0.1031	0.0290**	0.0185**	0.1911	0.0290**	0.4509	0.1568	0.0602*	0.0179**	0.0204**
SP & 500													
1950–2005	Mean	0.0124	0.0010	0.0106	0.0109	0.0024	0.0021	0.0085	0.0000	−0.0062	0.0070	0.0157	0.0166
	St. dev.	0.0476	0.0304	0.0323	0.0381	0.0364	0.0346	0.0411	0.0480	0.0427	0.0543	0.0432	0.0322
	t-stat[1]	2.5449	1.0449	2.1426	2.8120	1.1617	1.2442	2.4251	0.7169	−1.0810[2]	1.2814	2.9029	3.7184
	Significance	0.0069***	0.1504	0.0183**	0.0034***	0.1252	0.1094	0.0093***	0.2383	0.1423	0.1028	0.0027***	0.0002***
1990–2005	Mean	0.0033	0.0079	0.0116	0.0051	0.0195	0.0022	0.0024	−0.0110	−0.0061	0.0199	0.0182	0.0205
	St. dev.	0.0374	0.0307	0.0369	0.0371	0.0319	0.0345	0.0415	0.0561	0.0477	0.0347	0.0448	0.0382
	t-stat[1]	0.8362	1.3402	0.9900	1.0742	1.8544	0.8272	0.6982	−0.2214	−0.4973[2]	1.5143	1.7663	1.9684
	Significance	0.2085	0.1008	0.1695	0.1505	0.0424**	0.2110	0.2482	0.4140	0.3133	0.0761*	0.0496**	0.0346**
NASDAQ													
1971–2005	Mean	0.0305	0.0107	0.0058	0.0049	0.0111	0.0108	−0.0022	0.0028	−0.0069	−0.0019	0.0152	0.0186
	St. dev.	0.0595	0.0574	0.0505	0.0558	0.0531	0.0495	0.0547	0.0684	0.0585	0.0875	0.0747	0.0565
	t-stat[1]	2.7821	1.1782	0.8682	1.1176	1.4525	1.3704	0.4465	0.6141	−0.6868[2]	0.2662	1.5430	2.3469
	Significance	0.0045***	0.1236	0.1958	0.1359	0.0779*	0.0899*	0.3291	0.2717	0.2485	0.3959	0.0662*	0.0125**
1990–2005	Mean	0.0197	0.0116	0.0045	−0.0123	0.0213	0.0115	−0.0001	−0.0018	0.0005	0.0173	0.0217	0.0269
	St. dev.	0.0573	0.0732	0.0458	0.0652	0.0627	0.0653	0.0643	0.0836	0.0703	0.0554	0.0932	0.0728
	t-stat[1]	1.1235	0.4146	0.1745	−0.8920	1.0261	0.4342	−0.0293	−0.0737	0.0297[2]	0.7643	0.8790	1.3384
	Significance	0.1401	0.3424	0.4320	0.1937	0.1611	0.3354	0.4885	0.4711	0.4884	0.2287	0.1971	0.1011

Note: The data in this table are the same as that in Table 1, except that year 2001 is excluded as an outlier because of the September 11 attacks.

[1] t-statistics for the hypothesis that September mean is less than the given month's mean.

[2] t-statistics for the hypothesis that September mean is less than zero.

*Significance at the 10% level.

**Significance at the 5% level.

***Significance at the 1% level.

there were five months (January, April, July, November, and December) whose return exceeded September return at the 1% level of significance. Column 5 shows the number of months whose return exceeded September return at the 5% level of significance. This includes those months where their returns exceeded September return at the 1% level of significance. Taking the S&P 500 Index as an example again, during the period 1950–2005, March return exceeded September return at the 5% level of significance, making a total of six months where their returns exceeded September return at the 5% level of significance. Column 6 shows the number of months whose return exceeded September return at the 10 % level of significance. This includes those months where their returns exceeded September return at the 1% and 5% levels of significance. For the S&P 500 Index from 1950 to 2005, there are three more months (May, June, and October) that have their returns exceeding September return at the 10% level of significance, making a total of nine months whose

Table 3. Comparative performance of September return with return in other months.

Index	Time period	Number of months whose mean return exceeded September mean return	Number of months whose return exceeded September return at 1% level of significance	Number of months whose return exceeded September return at 5% level of significance	Number of months whose return exceeded September return at 10% level of significance
DJIA	1896–2005	11	7	11	11
	1896–1949	11	2	4	8
	1950–1989	11	2	6	7
	1990–2005	11	2	8	9
S&P 500	1950–2005	11	5	6	9
	1950–1989	11	2	6	6
	1990–2005	10	0	4	5
Nasdaq	1971–2005	11	2	5	6
	1971–1989	10	1	4	6
	1990–2005	11	0	0	4

Note: The data for this table are taken from Table 1. Column 3 is the number of the 11 months whose mean return exceeded September mean return in the given time period, and Columns 4, 5, and 6 are the number of months whose return exceeded September return at 1%, 5%, and 10% levels, respectively.

return exceeded September return at the 10% level of significance. As mentioned above, the level of significance increases as the number of sample points increases.

The September phenomenon can be further evidenced by the frequency of negative September returns, which is compared with the frequency of negative returns in other months during the respective time periods. The statistics are presented in Table 4.

In Table 4, column 3 shows the number of years when September has a negative return, out of the total number of years in the sample, and column 4 shows the percentage of years when September has negative return. For example, for the S&P 500 Index, during 1950–2005, 32 out of the 56 Septembers in that time period have negative returns, which is 57.1%. Column 5 shows the number of months in that time period which have negative returns, as well as the number of months included in that time period. For the S&P 500 Index, there were 276 months out of the 672 months when the return was negative. Column 6 shows this as a percentage, which in this case is 41.1%. It is clear that for every index, on a percentage basis, September return has more often been negative compared with all of the other months' returns.

The September phenomenon can be further revealed by the number of years in which September return is the worst, and the evidence is presented in Table 5. In this table, column 3 shows the number of years in which September was the worst performing month during that year. For example, for the

Table 4. Comparison of the frequency of negative return in September with negative return in other months.

Index	Time period	Number of years September had negative return	% of times September had negative return	Number of times monthly return were negative	% of times monthly return were negative
DJIA	1896–2005	65 out of 110	59.1	558 out of 1310	42.6
S&P 500	1950–2005	32 out of 56	57.1	276 out of 672	41.1
NASDAQ	1971–2005	17 out of 35	48.6	175 out of 419	41.8

Note: The data for this table are from Dow Jones, Inc., Standard and Poor, and NASDAQ. Column 3 shows the number of years September had a negative return for the given index and time period, column 4 expresses this as a percentage, column 5 gives the number of months where return was negative for the given index and time period, and column 6 expresses that as a percentage.

NASDAQ Composite Index during the 1971–2005, in six years out of the 35 years, September was the month with the worst return. Column 4 gives this as a percentage, which is 17.1%.

If all months had an equal likelihood of being the worst performing month of the year over the sample period, September would be the worst about one-twelfth of the time, or 8.3% of the time. As can be seen from Table 5, in the case of the DJIA Index and the NASDAQ Composite Index, September held that distinction much more often, with almost twice the frequency as we would normally expect. In the case of the S&P 500 Index, September is still the worst performing month more often than chance would allow, although not as strikingly often as in the case of the other two indices.

Mean September return of large stocks is more significantly negative than that of small stocks. The DJIA Index is composed of the 30 largest stocks; the S&P 500 Index contains the 30 largest stocks in the DJIA Index and other 470 smaller stocks, whereas the NASDAQ Composite Index contains a few thousand much smaller stocks. For each index's entire observation period, as presented in Table 1, the mean September return for the DJIA Index, the S&P 500 Index, and the NASDAQ Composite Index are −1.43%, −0.76%, and −1.2%, respectively. For the period 1950–1989, the mean September return for the DJIA and the S&P 500 indices are −0.8%, and −0.63%, respectively. For the period 1990–2005, the mean September returns for the three indices are −1.95%, −1.11%, and −1.11%, respectively. Further, as exhibited in Table 4, the percentage of months with negative return seems to be consistent among

Table 5. Frequency of September being the worst month of the year.

Index	Time period	Number of times in which September was the worst month	% of times September was the worst month
DJIA	1896–2005	17 out of 110	15.5
S&P 500	1950–2005	6 out of 56	10.7
NASDAQ	1971–2005	6 out of 35	17.1

Note: The data for this table are from Dow Jones, Inc., Standard and Poor, and NASDAQ. Column 3 shows the number of times in which September performance was the worst in the year during the time period for the three indices, and column 4 expresses this as a percentage. September would have one-twelfth or 8.3% of the time being the worst month if all months had an equal likelihood of being the worst month.

all indices, which is close to 42%; however, the percentage of Septembers with negative return is much higher for the DJIA and the S&P 500 indices, which is close to 58%, whereas for the NASDAQ Composite Index it is near 49%. As for the January effect, small firm stocks were noticed to perform even better than large firm stocks in the best month (Rozeff and Kinney, 1976; Reinganum, 1981; Keim, 1983; Roll, 1983); for the worst month, that is, September, small firm stocks also perform better than large ones, or in other words, the negative performance is less pronounced for small firm stocks.

We also examined which week(s) in September contribute(s) more to the month's poor performance, similar to the examination by Wang, Li, and Erickson (1997) who found that Monday effect occurs primarily in the fourth and fifth weeks of the month. The first week of the month is defined as the week that contains the first trading Friday of the month. The second week of September 2001 from Monday to Friday has been excluded because of the September 11 attacks. In that year, the market reopened on Monday, that is, on September 17. Return of the third week of September 2001 is calculated as the natural logarithm differential between the closes of September 21 and September 10. We found that there is not any one week that is significantly worse than others, although for the DJIA and NASDAQ Composite indices, the second and fourth weeks look to have worse return, but not significantly worse. The first two weeks look worse than the last three weeks for the NASDAQ Composite index, but again the difference is not significant.

Finally, the poor September market performance looks more apparent in recent years, particularly for the large stocks. For the three subperiods 1950–1970, 1971–1989, and 1990–2005, the mean September returns are −0.269%, −1.383%, and −1.95%, respectively, for the DJIA Index; 0.02%, −1.343, and −1.11%, respectively, for the S&P 500 Index; and for the two subperiods 1971–1989 and 1990–2005 the mean September returns are −1.27% and −1.11%, respectively, for the NASDAQ Composite Index. The differences between the mean returns of the latest period 1990–2005 and the previous periods are the greatest for the DJIA Index, followed by the S&P 500 Index, and the difference is the smallest for the NASDAQ Composite Index. In contrast, the January effect has been diminishing over the last 25 years as it became well known (Gu, 2003). As for the poor September performance of the US stock market, further observation is required to find whether the anomaly will also diminish as it becomes well known in the future.

3. Factors Related to the September Effect

No evidence was found that the US stock market's poor September performance is related to any microeconomic factors. It may be helpful to note the suggested possible reasons for the January effect, which include tax-loss selling effects (Ritter, 1988), "window dressing" (Haugen and Lakonishok, 1988), and seasonality in risk premium or expected returns (Chang and Pinegar, 1989, 1990; Kramer, 1994). Tax-loss selling and window dressing could also contribute to the generally poor September performance if investors, particularly institutional investors, sell large quantities of stocks in September that they do not expect to perform well. Unfortunately, data of institutional trading activities of each month are not available. Macroeconomic seasonality could also be a reason.

A regression analysis was carried out to find possible relations between September return and some of the macroeconomic factors, such as annual GDP growth, inflation, interest rate, and the index's annual return. The first three explanatory variables, real GDP growth, inflation, and Treasury bill rate, capture the equity market's exposure to macroeconomic forces. The last explanatory variable, the index's annual return of the year, relates the size of the September anomaly to the annual performance of the index. September return is the dependent variable. Two sets of regression analyses for each of the three indices were undertaken to reveal the relationship between September return and the four explanatory variables. The first set uses the contemporaneous values of the three macroeconomic explanatory variables; the second set uses the next year's values of the three explanatory variables; both sets use the index's contemporaneous annual return. The expected values of the three macroeconomic factors may have bigger impacts on the returns because stock price is based on expectations. Assuming that investors' expectations are rational and accurate (as in the conventional theory of financial economics), we use the next year's values of the three factors as proxies to the expected values. Results of the regression analysis are presented in Table 6.

From the data in Table 6, we see that the September return is significantly positively connected to GDP growth for all three indices for actual values of GDP growth; for the DJIA Index, the estimated connection is more significant for next year's values. The month's return is positively related to inflation rate, but significant only for the DJIA Index for the expected values. September performance is negatively related to T-bill rate, but the relation is weak and insignificant. September return of each index is positively connected to the index's annual performance; the connection is significant for the NASDAQ Composite and the S&P 500 indices, but not significant for the DJIA Index. In comparison, empirical findings about the relation between January effect

Table 6. Estimates from regressing September return on selected explanatory variables.

Variable	GDP			Annual return t	T-bill rate	Adj. R^2
	Intercept	Growth$_t$	Inflation$_t$			
The DJIA						
Coefficient	−0.0359	0.0466	0.0015	0.0063	−0.0003	
t-value	(−2.852)***	(3.222)***	(0.809)	(1.273)	(−0.12)	0.1453
		Growth$_{t+1}$	Inflation$_{t+1}$	Return$_t$	Rate$_{t+1}$	
Coefficient	−0.054	0.0583	0.0053	0.0027	−0.0001	
t-value	(−4.793)***	(4.560)***	(3.251)***	(0.629)	(−0.046)	0.3632
The NASDAQ						
Coefficient	−0.0502	0.0863	0.0031	0.143	−0.0027	
t-value	(−1.438)	(1.778)*	(0.799)	(3.819)***	(−0.561)	0.3158
		Growth$_{t+1}$	Inflation$_{t+1}$	Return$_t$	Rate$_{t+1}$	
Coefficient	−0.0624	0.0715	0.0038	0.1101	−0.0458	
t-value	(−1.705)*	(1.263)	(0.967)	(2.415)**	(−0.010)	0.2411
The S&P 500						
Coefficient	−0.0390	0.0404	0.0028	0.1885	−0.0016	
t-value	(−2.722)***	(1.992)*	(1.334)	(6.083)***	(−0.715)	0.4166
		Growth$_{t+1}$	Inflation$_{t+1}$	Return$_t$	Rate$_{t+1}$	
Coefficient	−0.0372	0.0368	0.0032	0.1373	−0.001	
t-value	(−2.678)***	(1.364)	(1.48)	(3.366)***	(−0.399)	0.3540

Note: The data for this table are from Dow Jones, Inc., Standard and Poor, NASDAQ, and International Monetary Fund. Return is calculated as natural logarithm differential between the close of a month's last trading day and the close of the previous month's last trading day. The four explanatory variables are annual GDP growth, inflation, interest rate, and the index's annual return. Two sets of regression analyses are undertaken. The first set uses the contemporaneous values of the three macroeconomic explanatory variables; the second set uses the next year's values of the three explanatory variables, both sets use the index's contemporaneous annual return. The next year's values of the three macroeconomic factors are used as proxies to the expected values. September return is the dependent variable.

*Significant at the 10% level.

**Significant at the 5% level.

***Significant at the 1% level.

and the factors are mixed. Kohers and Kohli (1992) reported that the January effect existed in the expansionary phases but not during the recession phases for the S&P Composite Index from 1948 to 1988. Kramer (1994) points out that higher January returns correlate to inflation, and Ligon (1997) finds that January return is negatively related to real interest rates. However, Gu (2003) has revealed that January effect is negatively connected to both actual and expected real GDP growth, inflation, and the index's return of the year.

The regression power is larger when using the expected values of the explanatory variables for the DJIA Index, as reflected by the significantly larger adjusted R^2 value, but that is not the case for the NASDAQ Composite and the S&P 500 indices. A possible explanation could be that the larger companies are closely analyzed by more professionals and hence more closely reflect investors' expectations. These companies have the capacity to better predict the future and adjust their operation than small companies, and hence their performances can be more closely related to expectations.

4. Conclusion

A disproportionately large number of Septembers have experienced negative returns for the DJIA, NASDAQ Composite, and the S&P 500 indices, and the mean September return is the lowest among all the months. Also, the phenomenon of generally poor September performance has been more pronounced in recent decades. September return of the indices is positively connected to GDP growth, inflation rate, and stock market performance of the year, and negatively related to interest rate. Tax-loss selling, "window dressing", and macroeconomic seasonality could also contribute to the poor September performance. Further research is required to find the reasons why September is the worst month for the US equity market.

References

Baldwin, A., *Historically, September is Market's Worst Month*, New York, Associated Press, 2003.

Browning, E. S., "After Hot Summer for Stocks, Time to Study." The *Wall Street Journal*, September 2, C1 (2003).

Chang, E. C. and L. M. Pinegar, "Seasonal Fluctuations in Industrial Production and Stock Market Seasonals." *Journal of Financial and Quantitative Analysis* 24, 59–74 (1989).

Chang, E. C. and L. M. Pinegar, "Stock Market Seasonal and Prespecified Multifactor Pricing Relations." *Journal of Financial and Quantitative Analysis* 25, 517–533 (1990).

Cross F., "The Behavior of Stock Prices on Fridays and Mondays." *Financial Analysts Journal* 29, 67–69 (1973).

French, K., "Stock Returns and the Weekend Effect." *Journal of Financial Economics* 8, 55–69 (1980).

Gu, A. Y., "The Declining January Effect: Evidence from US Equity Markets." *Quarterly Review of Economics and Finance* 43(2), 395–404 (2003).

Haugen, R. A. and J. Lakonishok, *The Incredible January Effect: The Stock Market's Unsolved Mystery*, Dow Jones-Irwin, Homewood, IL, 1988.

Keim, D. B., "Size-related Anomalies and Stock Return Seasonality: Further Empirical Evidence." *Journal of Financial Economics* 12, 13–32 (1983).

Kohers, T. and R. K. Kohli, "The Year-end Effect in Stock Returns Over Business Cycles: A Technical Note." *Journal of Economics and Finance* 16, 61–68 (1992).

Kramer, C., "Macroeconomic Seasonality and the January Effect." *Journal of Finance* 49, 1883–1891 (1994).

Ligon, J., "A Simultaneous Test of Competing Theories Regarding the January Effect." *Journal of Financial Research* 20, 13–32 (1997).

Reinganum, M. R., "Misspecification of Capital Asset Pricing: Empirical Anomalies based on Earnings' Yields and Market Values." *Journal of Financial Economics* 9, 19–46 (1981).

Ritter, J. R., "The Buying and Selling Behavior of Individual Investors at the Turn of the Year." *Journal of Finance* 43, 701–717 (1988).

Ritter, J. R. and N. Chopra, "Portfolio Rebalancing and the Turn-of the-Year Effect." *Journal of Finance* 44, 149–166 (1989).

Roll, R., "On Computing Mean Returns and the Small Firm Premium." *Journal of Financial Economics* 12, 371–386 (1983).

Rozeff, M. S. and W. R. Kinney Jr., "Capital Market Seasonality: The Case of Stock Returns." *Journal of Financial Economics* 3, 379–402 (1976).

Wachtel, S. B., "Certain Observations on Seasonal Movements in Stock Prices." *Journal of Business* 15, 184–193 (1942).

Wang, K., Y. Li and J. Erickson, "A New Look at the Monday Effect." *Journal of Finance* 52(5), 2127–2186 (1997).

This page intentionally left blank

Chapter 15

Fundamental Drivers of Electricity Prices in the Pacific Northwest

Chi-Keung Woo*
Energy and Environmental Economics Inc., USA
and
Hong Kong Baptist University, Hong Kong

Ira Horowitz
University of Florida, USA
and
Hong Kong Polytechnic University, Hong Kong

Nate Toyama
Sacramento Municipal Utilities District, USA

Arne Olson
Energy and Environmental Economics Inc., USA

Aaron Lai
Bank of America, USA

Ray Wan
InfoAtlas, Inc., USA

We estimate an AR(1)/GARCH(1, 1) model that shows the impact of natural-gas prices, hydro conditions, and temperatures on wholesale on-peak electricity prices at the Mid-Columbia (Mid-C) trading hub in the Pacific Northwest of the United States. After controlling for the effects of these three factors, prices are seen to exhibit a weak seasonal pattern, but a strong day-of-week pattern. It is also shown that price spikes can persist for several days. Finally, in support of the GARCH hypothesis, Mid-C prices are seen to have a time-dependent variance that primarily moves with natural-gas prices, and that large price variances tend to persist. Thus, even though buyers might cross hedge using natural-gas futures and temperature-based weather

*Corresponding author.

futures, the effectiveness of any hedge is compromised by randomness in hydro conditions. To be sure, a buyer can eliminate the electricity price risk by entering into a forward contract, but only at the expense of what is likely to be a large risk premium embodied in the forward price.

Keywords: GARCH; electricity; price volatility.

1. Introduction

In the 1990s, electricity markets became a focal point of state regulatory efforts with the end in mind of separating the heretofore integrated generation, transmission, and distribution functions, albeit subject to continuing but different regulatory oversight (Borenstein, 2002; Stoft, 2002; Helman, 2006). On the West Coast, the emerging industry structure created wholesale markets in which generators compete for the right to supply electricity. The electrical energy goes through a transmission network, with power being injected by a selling party, e.g., a generator, and power being withdrawn by a buying party, e.g., a local distribution company (LDC). There are also marketers and brokers that may act as intermediaries, trading electricity with generators and LDCs. Under the obligation to serve, an LDC resells the power to end-users, often at fixed rates, so as to fulfill their electricity needs in real time upon demand.

Commodity spot-market prices are inherently unstable, and electricity spot markets are no exception. In fact, wholesale electricity prices are the most volatile among all energy prices. The purpose of this paper is to isolate and quantify the impact of the supply-and-demand sources of that instability for one particularly important spot market: notably, the Mid-Columbia (Mid-C) market, a major trading hub located in the Pacific Northwest of the United States.

There are two supply-side considerations: technical and pecuniary. From a technical standpoint, equipment can fail randomly: generators must occasionally be taken off-line for maintenance and repair; transmission networks can get congested; and electricity can incur line losses while discharging through the power grid. Moreover, except for hydroelectric generation with reservoir storage, electricity cannot be economically stored. From a pecuniary standpoint, inadequate precipitation results in low hydro generation, and production costs in natural-gas-fueled plants are subject to the vagaries of natural-gas markets and changing prices for emission permits required in a state such as California with tight emission regulation. We do not, however, include emission cost in our estimation, because suitable data are unavailable for the Pacific Northwest. Finally, restrictions on water discharged by thermal plants can severely limit power generation during a heat wave.

The demand for electricity depends upon the time of day, the day of the week, and the season of the year; and even within each of those well-defined settings, it is stochastic and dependent upon the weather. Demand may also be impacted by the vagaries of the natural-gas markets, as natural gas is an alternative source of energy for some electricity end-uses such as space and water heating.

A reliably functioning electricity grid requires that supply and demand balance in real time. Thus, the technical and pecuniary factors that can shift the electricity supply curve upward along a price–insensitive demand curve, and the inability of suppliers to increase short-term supply in response to rising prices evolving from unanticipated increases in demand in a tight market, may in tandem result in non-transitory price spikes. Layered into this is the possibility of the abuse of any market power held by generators, and the potential for marketers to bend the rules and game the market, thereby exacerbating and extending what might otherwise have been a short-term price problem (Borenstein, 2002; Borenstein *et al*., 2002; Joskow and Kahn, 2002; Hurlbut *et al*., 2004; Helman, 2006). The worst-case scenario became an all-too-clear reality during the Western energy crisis that was characterized by soaring wholesale electricity prices over a 13-month period beginning in May 2000 (Jurewitz, 2002).

The paper contributes to the deregulation policy debate because high and volatile prices cast doubt on the ability of electricity-market reforms to deliver reliable service at stable and reasonable prices (Woo *et al*., 2003; Trebilcock and Hrab, 2005; Chao, 2006). In addition, determining the fundamental drivers of those prices has important implications for: risk management (Kleindorfer and Li, 2005; Deng and Oren, 2006; Woo *et al*., 2006b); the pricing of electricity options, futures, forwards, and generation assets (Deng *et al*., 1999, 2001; Woo *et al*., 2001; Kamat and Oren, 2002; Lucia and Schwartz, 2002; Keppo and Lu, 2003; Fleten and Lemming, 2003; Eydeland and Wolyniec, 2003); the evaluation of power contracts (Woo *et al*., 2006b); the detection of market-power abuse and price manipulation (Borenstein *et al*., 2002; Joskow and Khan, 2002; Helman, 2006); the investigation of generation investment behavior (Neuhoff and De Vries, 2004); the assessment of wholesale market integration (De Vany and Walls, 1996; Woo *et al*., 1997; Park *et al*., 2006); and the assessment of the effect of retail competition on forward-contract pricing (Green, 2003).

We estimate an AR(1)/GARCH(1, 1) model to show that deregulated electricity prices can spike easily due to daily fluctuations in three demand-supply fundamentals: natural-gas price, hydro conditions, and temperature.[1] After controlling for the effects of those three factors, and this is to the best of our knowledge the very first econometric model to do so,[2] Mid-C prices are seen to exhibit a weak seasonal pattern, but a strong day-of-week pattern. It is also shown that price spikes can persist for several days. Finally, in support of the GARCH hypothesis, Mid-C prices are seen to have a time-dependent variance that moves with natural-gas prices, and that large price variances tend to persist. Thus, even though an LDC might cross hedge using natural-gas futures and temperature-based weather futures, the effectiveness of any hedge is compromised by random hydro conditions. Assuredly, the LDC can eliminate the price risk by entering into a forward contract, but only at the expense of what is likely to be a large risk premium embodied in the forward price.

[1] As such, our study is not a time-series analysis whose primary goal is to identify and estimate the stochastic process of spot-price dynamics (e.g., De Vany and Walls, 1996; Lucia and Schwartz, 2002; Eydeland and Wolyniec, 2003; Goto and Karolyi, 2003; Haldrup and Nielsen, 2004; Mount *et al*., 2006). In one of many constructive comments provided to us an earlier draft, however, one referee called our attention to the extremely seasonal nature of the Mid-C price series, which implies that a seasonal ARIMA model would provide a useful approach for any future analysis of the Mid-C price-generating process.

[2] There are extant econometric analyses of electricity price movements in response to demand and supply fundamentals. Vucetic *et al*. (2001), for example, find that California electricity-market prices vary with the state's total demands. Johnsen (2001) shows Nordic prices to be responsive to the level of demand, reservoir inflow from rainfall and snow melting, and temperature. Bessembinder and Lemmon (2002) explain how deregulated electricity prices move with total demands in the California and Pennsylvania-Maryland-Jersey (PJM) markets. GAO (2002) indicates that electricity prices in California rise with in-state generation, decline with net imports, and vary systematically with seasonality, day-of-the-week and time-of-day. Mount *et al*. (2006) find that prices in the PJM market move with total demand and reserve margin and that they adjust relatively quickly to their equilibrium levels after random shocks.

2. The Mid-C Hub

The Mid-C hub is physically located at several substations along the Columbia River in central Washington. Like some other wholesale markets, Mid-C is captive to the substantial influence of hydroelectric production. The area itself is an intersection point for several regional transmission systems, the most prominent of which belongs to the federal Bonneville Power Administration (BPA). The area also houses several large hydroelectric dams, including Grand Coulee.

Mid-C is a bilateral market in which trading for physical delivery takes place primarily through telephone calls among a vast array of diverse market participants. The latter include independent power producers, power marketers, investor-owned utilities, municipal utilities and cooperatives, Canadian utilities, and the BPA. Mid-C is one of several important pricing points in the Western Electricity Coordinating Council (WECC). With over 180,000 MW of generation capacity in 2004, the WECC is the largest of the 10 reliability councils in North America, encompassing a geographic area containing the western part of the continental United States, the two Canadian provinces of British Columbia and Alberta, and portions of one Mexican state (WECC, 2004).

Mid-C prices reflect the underlying physical realities of the Pacific Northwest electricity system, most particularly the strong role played by hydro generation in the Northwest Power Pool (NWPP) of the WECC. In 2004, the NWPP had 81,018 MW of generating capacity, 60% of which was hydro (WECC, 2004, Table 11). These generation units in the NWPP disperse across the Canadian provinces of British Columbia and Alberta and the upper western half of the United States (WECC, 2004, p. 24). Below-normal precipitation can severely impede hydro generation, thus increasing dispatch of the more costly thermal generation to meet a given level of demand, which subsequently raises the electricity market price.

The two largest hydro systems are the Federal Columbia River Power System (FCRPS) which produces the power marketed by BPA, and that operated by the British Columbia Hydro and Power Authority. Both systems feature vast storage reservoirs that store the spring runoff for future use. There are also a number of smaller systems that provide significant amounts of energy but do not have substantial reservoir storage.

The role of hydro power affects market prices in a number of ways. First, a hydroelectric system is energy constrained, not capacity constrained. That is, a hydroelectric system with storage may have very large peaking capability but can only run at a low average rate due to a limited amount of water stored behind the dams. For example, the federal dams alone comprise some 20,000 MW of nameplate capacity, enough to serve 40% of California's 50,000 MW peak demand; yet they produce only about 9,300 MW of power on average. A hydroelectric system's ability to produce a large amount of power with stored water, though only for few hours a day, dampens daily and hourly price fluctuations. Hence, we expect the Mid-C price level and variance to be inversely related to river flows, a testable hypothesis in our econometric analysis.

Second, the Pacific Northwest system is interconnected with California, Arizona and New Mexico. These inter-ties were constructed during the 1950s through the 1980s to take advantage of seasonal diversity of loads and resources: surplus hydropower in the Pacific Northwest flows southbound during the summer months when air-conditioning loads are high in California, Arizona and New Mexico; and surplus thermal power in the Pacific Southwest flows northbound during the winter months when heating loads are high in Oregon, Washington and British Columbia. Wholesale electricity prices tend to be lower at Mid-C than at other western hubs during the spring and summer when river flows are high, and tend to be higher at Mid-C during the winter when river flows are low.

Third, in addition to a seasonal pattern, the Pacific Northwest has a diurnal pattern. Hydro production is concentrated during on-peak hours, defined in the Western Systems Power Pool Agreement as 06:00–22:00, Monday through Saturday (excluding WECC holidays), when hourly electricity demands are high and hydro facilities' fast-ramping capability is most valuable. The daily off-peak period comprises the remaining hours when hourly demands are low and the hydro system is ramped down, and electric energy is imported to allow the reservoirs to replenish for next day's on-peak production. Consequently, Mid-C prices are lower than prices at other western hubs during on-peak hours but they are higher during off-peak hours.

While daily price fluctuations are less at Mid-C during any given year, hydro generation can vary substantially from one year to the next, largely depending on winter snow pack. Market prices are very low during years of abundant hydroelectric supply such as the late 1990s; even low-cost coal and nuclear plants in the NWPP are sometimes dispatched off for long periods. Drought conditions, however, can cause price spikes to sustain over long periods due to hydro-energy shortage. This became evident during the now well-documented 2000–2001 energy crisis when prices were actually higher in the Pacific Northwest than in California.

Electricity for Friday and Saturday delivery is traded together on Thursday; energy for Sunday and Monday delivery is traded on Friday. While on-peak trading is active, off-peak trading can be thin. Thus our analysis focuses on daily on-peak prices.

3. The Data Set

The linchpin for our data is the time series of daily Mid-C on-peak prices from Platts for the 39-month period from January 2002 through March 2005 when we initiated this study. The data comprise a sample of 972 daily observations that characterize the relatively calm market environment that followed the Western energy crisis.

Let y_t denote the Mid-C on-peak price on delivery day t. This price is the regressand in a partial-adjustment equilibrium-price regression model in which the lagged price, y_{t-1}, is one of seven time-series regressors that are hypothesized to "explain" the day-t price. The lagged price helps capture the price effect of a random event, such as an unexpected generation plant outage or transmission line failure, that may last many days. The other six regressors reflect demand-and-supply conditions and along with the constant term, denoted $x_{0t} \equiv 1$, form the row vector $\boldsymbol{x}_t$.

There are three supply-side factors, one of which is the natural-gas price, the data for which come from Platts; the other two reflect hydro conditions. Specifically, the second regressor x_{1t} is the day-t natural-gas price at the Henry Hub in Louisiana, measured in US$ per million British thermal units (MMBtu). Henry Hub is the largest cash market in North America and is also the delivery point for the New York Mercantile Exchange's (NYMEX) natural-gas futures contracts. We specify the Henry Hub price rather than the local natural-gas price at Sumas (Washington) because the Sumas price may be endogenous and driven by local thermal generation, while that at Henry Hub is exogenous to the Pacific Northwest and is useful for cross hedging and risk management (Woo *et al.*, 2006a). For the exogenous Henry Hub price to be a reasonable proxy for the potentially endogenous Sumas price, the two prices should be highly correlated. Happily, this is indeed the case ($R = 0.94$) during our sample period.

The possibility of seller market power leads us to conjecture that the natural-gas price may have a varying marginal effect on the electricity price. In a competitive market without market power, the market equilibrium price tracks short-run marginal costs. Thus, a dollar increase in marginal fuel cost due to a natural-gas price increase translates into a dollar increase in the market price, implying a linear relationship between the market price and the natural-gas price, with the marginal effect being the marginal generation unit's rate of converting fuel to electricity (i.e., heat rate). But the marginal effect is increasing in the natural-gas price when a generation owner possesses sufficient market power that it can pass on to a buyer at an increasing rate any marginal-cost increase it sustains due to a natural-gas price increase. Empirical evidence of an increasing marginal effect helps explain why wholesale electricity prices can become increasingly spiky and volatile in a rising natural-gas price environment. To subject our conjecture of a varying marginal effect of the natural-gas price to the empirical test, the Henry Hub price squared, denoted $x_{1t}^2 = x'_{1t}$, is introduced as a third regressor.

The first of the two daily hydro conditions, denoted x_{2t}, is measured by a precipitation index formed as a weighted average of values ranging from 1 for the driest conditions, through 4 for normal precipitation, and 7 for the wettest conditions in the state of Washington. The values are determined through the US Geological Survey of a daily-varying sample of the state's hydro stations. If, for example, a sample of 10 stations yielded four "driest" and six "normal" conditions on day t, the resulting index would be $x_{2t} = 0.4 \times 1 + 0.6 \times 4 = 2.8$. The second hydro condition, denoted $x_{3t,}$ is the Columbia River flow (00000 ft.^{3} per second) at The Dalles Dam on the Washington-Oregon border. Columbia River flow at The Dalles is the most closely-watched indicator of FCRPS energy potential, and is the subject of forecasts published by the Northwest River Forecast Center during the winter months. While the flow may depend on hydro-generation dispatch in response to the Mid-C price (Johnsen, 2001; Bushnell, 2003), we contend that the flow is exogenous as water release serves several competing purposes, including flood control, recreation, irrigation, and navigation (WECC, 2005, p. 5).

The two hydro conditions define the fourth and fifth regressors. Favorable hydro conditions are hypothesized to increase supply and hence reduce the Mid-C price.

Finally, there are two demand-side factors, both of which reflect the daily West Coast weather pattern that drives the region's aggregate electricity demand. These demand factors are measured by the daily maximum and minimum temperatures for Portland and Sacramento in Fahrenheit degrees from the National Oceanic and Atmospheric Administration. These two cities are chosen because they represent Oregon/Washington and Northern California weather and have futures traded on the Chicago Mercantile Exchange. To differentiate weather-sensitive demands, we convert the temperature data into two scalar variables: (1) x_{4t} is the cooling-degree-days (CDD) or x_{4t}= max(daily maximum temperature $-$ 65, 0); and (2) x_{5t} is the heating-degree-days (HDD), or $x_{5t} = \max(65 -$ daily minimum temperature, 0). The Chicago Mercantile Exchange also uses the 65-degree threshold to compute the CDD and HDD variables. Extreme weather conditions are hypothesized to increase the demand for energy and hence lead to higher spot prices. Only the Portland CDD and HDD, however, survived the statistical-significance test for inclusion in our final regression model.

Eleven monthly dummy variables representing January through November are introduced to capture any seasonal variations in demand, and five daily dummy variables representing Monday through Friday are introduced to capture any daily variations in an on-peak delivery week of Monday through Saturday. These dummy variables form the row vectors $\boldsymbol{M}$ and $\boldsymbol{D}$, respectively. Our use of dummy variables facilitates easy interpretation of the month-of-year and day-of-week effects, which is the tack taken by Goto and Karolyi (2003) and GAO (2002).

Figures 1A–C provide a visual display of the sample-period data. Figure 1A shows that daily Mid-C electricity prices move with Henry Hub natural-gas prices. Moreover, the Mid-C price series exhibit a seasonal pattern, with relatively high prices in the winter months of December through February and in the summer months of July through September. Mid-C prices, however, are relatively low in the spring runoff months of March through June and in the fall mild-weather months of October and November. Figure 1B plots the time series of the Washington hydro index and Columbia River flow, suggesting high generation from Columbia River dams during the spring and summer months, while generation at other hydro sites with little storage is highest during the winter rainy months. Figure 1C plots the Portland CDD and HDD series, implying the seasonal pattern of weather-sensitive electricity demands. Taken together, Figures 1A–C suggest that the Mid-C price increases with the

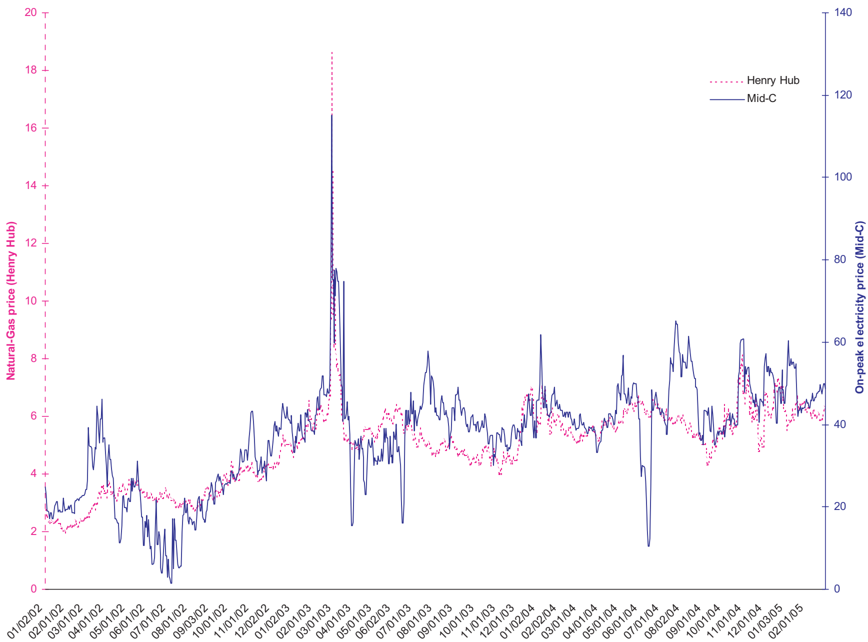


Figure 1A. Daily natural-gas price (US\$/MMBtu) at Henry Hub and daily on-peak electricity price (US\$/MWh) at Mid-Columbia (Mid-C) for the period of January 2002–March 2005.

Henry Hub natural-gas price, declines with hydro flow, and rises with extreme temperatures.

The summary statistics of Table 1 show that both sets of energy prices are highly skewed to the right and give impetus to our *ex ante* concern with the peaking-energy-prices problem. The maximum Mid-C on-peak price, for example, is some 6 standard deviations above the mean, while that for the Henry Hub natural-gas price is more than 7 standard deviations above the mean. The simple correlation coefficients in the last column of the table lend initial support to our hypothesis that Mid-C prices move in concert with natural-gas prices and counter to favorable hydro conditions. The low correlations with regard to temperature extremes are less encouraging.

4. The Empirical Results

4.1. *The GARCH model*

Woo *et al*. (2006b) develop a partial-adjustment equilibrium-price model that yields a parsimonious specification for an electricity spot-price regression.

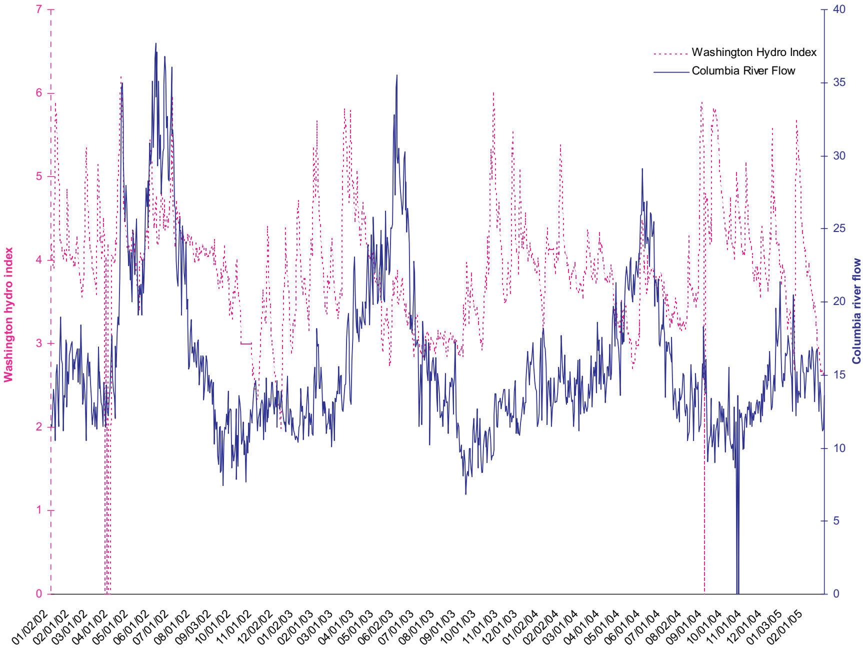


Figure 1B. Daily Washington hydro index (1 = driest, . . . , 4 = normal, . . . , 7 = wettest) and daily Columbia River flow (00000 ft.^{3} per second) for the period of January 2002–March 2005.

Let $\boldsymbol{\beta}$, $\boldsymbol{\eta}$, and $\boldsymbol{\delta}$ denote column vectors of parameters, let λ denote a parameter, and let μ_t denote a random-error term. In the present context the spot-price regression translates into the following model:

$$y_t = \boldsymbol{x}_t\boldsymbol{\beta} + \boldsymbol{M\eta} + \boldsymbol{D\delta} + \lambda y_{t-1} + \mu_t. \tag{1}$$

The vectors of parameter estimates are denoted $\boldsymbol{b}$, $\boldsymbol{m}$, and $\boldsymbol{d}$, respectively; l is the estimate of λ. Thus, for example, b_0 is the estimate of the intercept, b_1' is the estimated coefficient attached to $x_{1t}' = x_{1t}^2$, m_3 is the estimated coefficient attached to $M_3 = 1$ in March, and d_4 is the estimated coefficient attached to $D_4 = 1$ on Thursday.

Before proceeding with the estimation, it is necessary to assure that all time series are stationary. The time series are of sufficient length to allow us to do so using the very general Phillips–Perron test without having to be concerned about its less-than-optimal small-sample properties (Greene, 2004, p. 645). The results of the test, both for a single mean and a trend, and with lags of six periods, are given in Table 2. The statistics show the unit-root hypothesis to

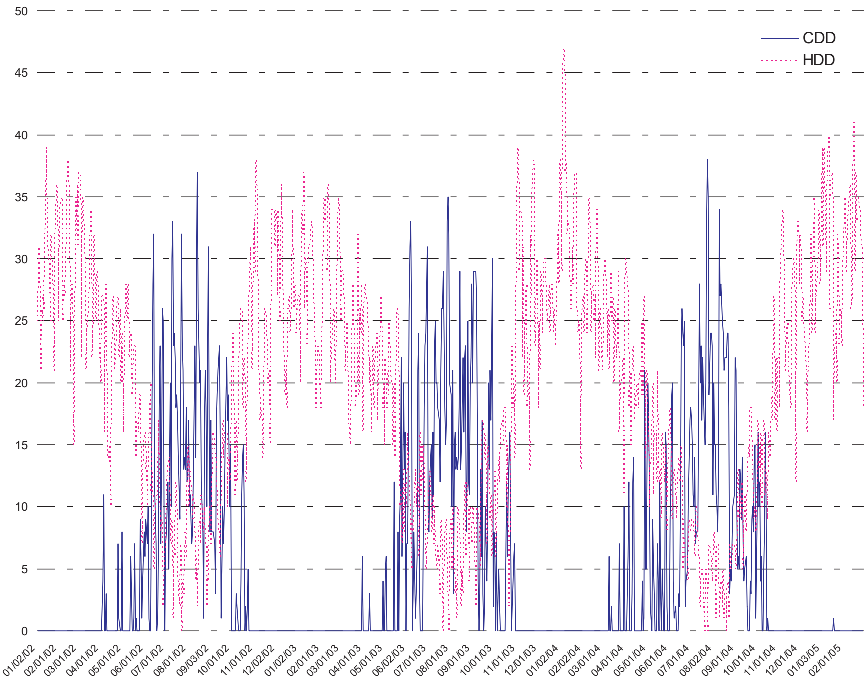


Figure 1C. Daily cooling-degree-day (CDD) and heating degree day (HDD) at Portland for the period of January 2002–March 2005.

be firmly rejected at the 1% significance level (the standard that we adopt in the paper) for each of the seven time series.

We acknowledge, however, the dual possibility of serially-correlated disturbances with time-dependent variances. We are also aware of the fact that that this variance may depend upon one or more of the regressors (Alexander, 2001; Engle, 2001).

First, as regards serially-correlated disturbances, a preliminary analysis revealed the AR(2) parameter estimate to be highly insignificant under an AR(2)/GARCH(1, 1) specification.[3] We therefore hypothesize the disturbance

[3]A referee recommended using diagnostic tests for model specification and comparisons. Following the referee's suggestion, we have tested a number of alternative stochastic specifications, including: (a) AR(2)/GARCH(1, 1); (b) AR(1)/GARCH(1, 1) and its variants; (c) AR(1); and (d) OLS. The results for (b), (c) and (d) are reported in Table 3. In addition, with regard to the parametric specification of the Mid-C price-regression model, we have estimated both linear and log-linear functional forms, both with and without a lagged dependent variable. The linear and log-linear specifications are found to yield qualitatively identical results. Hence, in the

Table 1. Summary statistics for daily observations in January 2002–March 2005.

Variable	Mean	Minimum	Maximum	Standard deviation	Skewness	Kurtosis	Correlation with Mid-C on-peak price
Mid-C on-peak price (US$/MWh)	37.06	1.43	115.18	12.98	−0.0552	1.4811	1.0
Henry Hub price (US$/MMBtu)	4.96	1.98	18.60	1.42	0.9074	9.1332	0.7789
Henry Hub price squared (US$/MMBtu)	26.60	3.92	345.96	17.10	7.51	129.33	0.6931
Washington hydro index (1 = driest, . . . , 7 = wettest)	3.89	1.99	6.19	0.69	0.4494	0.3873	−0.2969
Columbia River flow (00000 ft.^{3}/s)	15.99	6.83	37.70	5.53	1.4662	2.0236	−0.4423
Portland cooling-degree-days (CDD)	5.41	0	38	8.53	1.5237	1.3268	−0.0367
Portland heating-degree-days (HDD)	18.96	0	47	9.66	0.0556	−0.9054	0.1223

Table 2. Phillips–Perron statistic (τ) for testing if a time series has a unit root.

Variable	Lags	Phillips–Perron statistic*	
		Single mean	Trend
Mid-C price (US$/MWh)	6	−4.2736	−6.0500
Henry Hub price (US$/MMBtu)	6	−4.1451	−6.4596
Henry Hub price squared (US$/MMBtu)	6	−10.12	−12.89
Washington hydro index	6	−6.1479	−6.1887
Columbia River flow (00000 ft.^{3}/s.)	6	−4.5359	−4.7239
Portland CDD	6	−8.7018	−8.6954
Portland HDD	6	−6.3125	−6.3096

*The unit root hypothesis can be rejected at the 1% significance level.

term to follow an AR(1) process:

$$\mu_t = \rho\mu_{t-1} + \varepsilon_t. \tag{2a}$$

The estimate of ρ is denoted r.

Second, the disturbance term in the AR(1) process, ε_t, is hypothesized to be normally distributed about a zero mean and to have a GARCH(1, 1) time-dependent variance, σ_t^2, that is also a linear function of the vector z_t, a subset of $\boldsymbol{x}_t$. Any such time dependency for the variance of ε_t will carry over to that of μ_t, and subsequently plague the spot-market price y_t, too.

Specifically, $z_t = (x_{1t}, x_{3t})$ includes the natural-gas price and Columbia River flow. When the Washington hydro index and Portland CDD and HDD were included in a preliminary analysis, their estimated coefficients were highly insignificant; hence these variables were eliminated from the final estimation.

Let γ denote a column vector of parameters whose estimate will be denoted $\boldsymbol{g}$, and let α_j (j= 0, 1, 2) denote a parameter estimated by a_j. The expanded GARCH(1, 1) specification is:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\sigma_{t-1}^2 + \mathbf{z}_t\boldsymbol{\gamma}. \tag{2b}$$

interests of a parsimonious presentation, Table 3 focuses exclusively on the linear specification. Finally, the t-statistics in Table 3 decisively reject the null hypothesis that the Mid-C price does not depend on its one-day-lagged value.

The final regression model to be fitted is written as:

$$y_t = x_t\boldsymbol{\beta} + M\boldsymbol{\eta} + D\boldsymbol{\delta} + \lambda y_{t-1} + \rho\mu_{t-1} + \varepsilon_t \tag{3}$$

Here, ε_t has a zero mean and a conditional variance given by Equation (2b).

4.2. *The GARCH estimates and their implications*

The first column of figures in Table 3 are the maximum likelihood estimates of the parameters for Equation (3), which is the AR(1)/GARCH(1, 1) specification. The second column is the estimated regression excluding z_t, while the third and fourth columns duplicate the first two, but now without the AR(1) term. The fifth column shows the estimates with the AR(1) term, but without the GARCH terms, and the sixth column shows the ordinary-least-squares (OLS) estimates. The estimated coefficients for the regressors are generally quite robust over all specifications, although not necessarily with the same level of significance; no specification ostensibly outperforms the others in terms of goodness of fit. Based on the Akaike Information Criterion, we focus our discussion on the estimates for Equation (3) and relegate the others to the reader's perusal.

As regards the natural-gas price, both $1 > b_1 = 0.8514 > 0$ and $b_1' = 0.0643 > 0$ are statistically significant. Thus an increase in the natural-gas price has an increasing marginal effect on the electricity spot-market price.[4] The implication is that high natural-gas prices presage exploding electricity prices in the Pacific Northwest, because generation owners possess and may exercise what the parameter estimates suggest to be some degree of market power.[5]

The estimated parameters for both hydro conditions are statistically significant and have the expected negative sign, which implies that greater hydro

[4]Specifically, $\mathrm{d}y_t/\mathrm{d}x_{1t} = 0.8514 + 0.1286x_{1t} > 1$ when $x_{1t} > \$1.16$ per MMBTU, below the minimum Henry Hub price in the sample period; see Table 1.

[5]We do not believe, however, that the extent of market power abuse is so significant in our sample period that it distorts our parameter estimates of electricity price response to its underlying fundamentals. Our judgment is based on the fact that the period was characterized by a relatively calm market environment. Moreover, in the wake of the Enron scandal and with the regulators on the alert for post-Enron market-power abuses and market-gaming strategies (see, for example, California Attorney General's Office, 2004), it is unlikely that significant collusive activity took place during this period, or that such market power as individual generators might have possessed was sufficiently exploited as to invalidate our spot-price regression results.

Table 3. Maximum likelihood estimates (t ratios) of the Mid-C electricity-market-price regressions.

Variable	AR(1)/GARCH(1, 1)		GARCH(1, 1)		AR(1)	OLS
	Assumption (a)	Assumption (b)	Assumption (a)	Assumption (b)		
MSE	12.53	12.48	11.59	11.86	11.39	11.40
Akaike Information Criterion	4468.92	4483.01	4537.44	4586.62	5043.09	5042.57
Total R^2	0.9275	0.9278	0.9319	0.9303	0.9348	0.9347
Log likelihood	−2205.46	−2213.51	−2240.72	−2266.31	−2496.54	−2497.28
x_{0t} = intercept	6.0711	6.3706	4.7179	5.9689	6.9221	7.4580
	(5.75)*	(6.09)*	(4.70)*	(5.72)*	(4.65)*	(4.88)*
x_{1t}= Henry Hub Price	0.8514	0.8576	0.7528	0.3301	0.6338	0.7120
	(3.49)*	(3.49)*	(3.90)*	(1.65)	(2.25)	(2.56)
$x_{1t}^2 = x_{1t}$'= (Henry Hub Price)^2	0.0643	0.0673	0.0660	0.0902	0.0981	0.1003
	(4.95)*	(5.25)*	(6.60)*	(8.82)*	(5.40)*	(5.47)*
x_{2t}= Washington Hydro Index	−0.5833	−0.6165	−0.5763	−0.5417	−0.5931	−0.6353
	(−4.40)*	(−4.68)*	(−4.97)*	(−4.31)*	(−3.24)*	(−3.35)*
x_{3t} = Columbia River Flow	−0.1644	−0.1663	−0.1101	−0.1393	−0.2305	−0.2407
	(−4.65)*	(−4.85)*	(−3.49)*	(−4.22)*	(−5.60)*	(−5.71)*
x_{4t}= Portland CDD	0.0608	0.0605	0.0656	0.0657	0.1095	0.1074
	(3.51)*	(3.53)*	(4.14)*	(4.03)*	(5.19)*	(4.96)*
x_{5t}= Portland HDD	0.0152	0.0150	0.0017	0.0180	0.0434	0.0378
	(0.87)	(0.87)	(0.11)	(1.05)	(1.84)	(1.58)
M_1 =1 if Jan; else 0	−0.0870	−0.0911	0.1293	−0.1580	0.1974	0.2417
	(−0.23)	(−0.25)	(0.39)	(−0.47)	(0.40)	(0.46)

(*Continued*)

Table 3. (*Continued*).

Variable	AR(1)/GARCH(1, 1)		GARCH(1, 1)		AR(1)	OLS
	Assumption (a)	Assumption (b)	Assumption (a)	Assumption (b)		
M_2 =1 if Feb; else 0	0.1862	0.1591	0.2006	0.0256	0.1781	0.2143
	(0.43)	(0.39)	(0.59)	(0.07)	(0.35)	(0.41)
M_3 =1 if Mar; else 0	3.5113	3.2562	−0.5272	2.2656	0.5138	0.5597
	(8.36)*	(7.89)*	(−1.38)	(6.61)*	(0.94)	(0.98)
M_4 =1 if Apr; else 0	−0.1193	−0.1065	−0.1138	−0.0201	0.7397	0.7313
	(−0.21)	(−0.19)	(−0.26)	(−0.04)	(1.27)	(1.20)
M_5 =1 if May; else 0	0.4823	0.6068	0.0202	0.4824	1.0075	0.9887
	(0.81)	(1.06)	(0.04)	(0.89)	(1.42)	(1.34)
M_6 =1 if Jun; else 0	−0.8744	−0.9870	−1.3321	−0.8235	−0.4347	−0.5442
	(−1.21)	(−1.39)	(−2.01)	(−1.24)	(−0.51)	(−0.62)
M_7 =1 if Jul; else 0	0.2906	0.2638	−0.1924	0.0504	−0.0164	−0.0527
	(0.45)	(0.41)	(−0.34)	(0.09)	(−0.02)	(−0.06)
M_8 =1 if Aug; else 0	−0.6443	−0.6243	−0.9539	−0.8744	−0.5547	−0.5948
	(−1.10)	(−1.07)	(−1.95)	(−1.66)	(−0.74)	(−0.77)
M_9 =1 if Sep; else 0	−0.6955	−0.6568	−0.6612	−0.6615	−0.4868	−0.5259
	(−1.24)	(−1.14)	(−1.41)	(−1.31)	(−0.73)	(−0.76)
M_{10} =1 if Oct; else 0	−0.6640	−0.6718	−0.5132	−0.5521	−0.2311	−0.2905
	(−1.70)	(−1.72)	(−1.60)	(−1.61)	(−0.40)	(−0.48)
M_{11} =1 if Nov; else 0	−0.3090	−0.2950	−0.0044	−0.2507	0.1290	0.1434
	(−0.72)	(−0.67)	(−0.01)	(−0.65)	(0.24)	(0.26)

(*Continued*)

Table 3. (*Continued*).

Variable	AR(1)/GARCH(1, 1)		GARCH(1, 1)		AR(1)	OLS
	Assumption (a)	Assumption (b)	Assumption (a)	Assumption (b)		
D_1 =1 if Mon; else 0	1.4801 (5.91)*	1.4288 (6.35)*	1.4687 (4.84)*	1.4840 (4.80)*	1.5496 (3.93)*	1.5599 (4.05)*
D_2 =1 if Tue; else 0	0.1595 (0.63)	0.1253 (0.51)	0.4081 (1.42)*	0.2156 (0.72)	0.3931 (1.03)	0.4178 (1.10)
D_3 =1 if Wed; else 0	0.7879 (2.66)*	0.7661 (2.66)*	0.9140 (2.87)*	0.8527 (2.64)*	1.0265 (2.67)*	1.0446 (2.72)*
D_4 =1 if Thu; else 0	0.5497 (1.87)	0.5373 (1.92)	0.5122 (1.67)	0.4885 (1.50)	0.3245 (0.84)	0.3583 (0.93)
D_5 =1 if Fri; else 0	−1.0178 (−3.70)*	−1.0001 (3.82)*	−0.9977 (−3.00)*	−1.0767 (−3.09)*	−1.2941 (−3.29)*	−1.2703 (−3.31)*
y_{t-1} = Lag(Mid-C Price)	0.7864 (50.28)*	0.7801 (49.40)*	0.8187 (60.54)*	0.8246 (61.80)*	0.7707 (40.72)*	0.7559 (42.94)*
μ_{t-1} = lagged error in the AR(1) process	0.1713 (3.86)*	0.1919 (4.42)*			0.0510 (1.37)	
α_0 = intercept of the GARCH (1,1) process	0 (5294)*	0.5479 (5.51)*	0 (10477)*	0.6130 (5.79)*		
ε^2_{t-1} = lagged error squared in the GARCH (1,1) process	0.3411 (6.89)*	0.3392 (7.46)*	0.2325 (7.87)*	0.2974 (6.81)*		
σ^2_{t-1} = lagged conditional variance in the GARCH (1,1) process	0.6063 (16.16)*	0.6555 (20.87)*	0.7598 (43.70)*	0.6810 (21.04)*		

(*Continued*)

Table 3. (*Continued*).

Variable	AR(1)/GARCH(1, 1)		GARCH(1, 1)		AR(1)	OLS
	Assumption (a)	Assumption (b)	Assumption (a)	Assumption (b)		
x_{1t} = Henry Hub Price in the GARCH(1, 1) process	0.0973 (3.02)*		0.0380 (2.18)			
x_{3t} = Columbia River Flow in the GARCH(1, 1) process	0.0222 (1.76)		0.0115 (1.78)			

Note: The large $t-$statistics for the intercept estimates for the GARCH(1, 1) process under assumption (a) in the first and third columns of figures are the result of (1) the coefficient estimates are less than 10^{-7} at convergence, and (2) they have standard errors below 10^{-11}. Assumption (a) refers to the expanded GARCH(1, 1) specification: $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 + \mathbf{z_t} \boldsymbol{\gamma}$. Assumption (b) refers to the standard GARCH(1, 1) specification: $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2$. "*" The hypothesis of zero coefficient is rejected at the 1% significance level.

flow increases electricity supply, thereby effecting a reduction in the spot price. Specifically, $b_2 = -0.5833 < 0$ and $b_3 = -0.1644 < 0$.

A 1-unit increase in CDD ($b_4 = 0.0608 > 0$) has a larger price effect than the same increase in HDD ($b_5 = 0.0152 \approx 0$), implying that summer cooling demand for electricity has a greater impact on Mid-C prices than winter heating demand. This result is understandable because air-conditioning on a hot summer day almost exclusively relies on electricity, while a household can meet its heating needs on a cold winter day using alternative fuels (e.g., natural gas and firewood).

The estimate of $l = 0.7864 > 0$, which is statistically significant at virtually any level of confidence, verifies the applicability of the partial-adjustment equilibrium-price model that is the basis for the present analysis, to the Mid-C market. A US\$1/MWh increase in yesterday's Mid-C price would raise today's price by US\$0.79/MWh. This l estimate also implies that after a random shock, the Mid-C market resumes its price equilibrium in $1/(1-0.7864) = 4.6$ days. Hence, the l estimate helps explain the empirical fact that high electricity prices cluster and persist.

After controlling for weather effects and the seasonal pattern of natural-gas prices and hydro conditions, we did not anticipate the dearth of statistically significant estimated coefficients for the monthly dummy variables, with March being the sole and notable exception ($m_3 = 3.5113 > 0$). There are, however, striking day-of-the-week effects, with Tuesday being the prominent exception and Thursday being border-line. Specifically, $d_1 = 1.4801$ and $d_3 = 0.7879$ signal greater demand that effects higher spot prices on Monday and Wednesday, and $d_5 = -1.0178$ signals lower demand that effects lower prices on Friday. This day-of-week pattern of electricity market prices mirrors the one for daily on-peak electricity demands in the Pacific Northwest.

Finally, $r = 0.1713$, $a_1 = 0.3411$, and $a_2 = 0.6063$, each of which is statistically significant, verify the AR(1)/GARCH(1, 1) hypothesis of serially-correlated disturbances and a time-dependent variance. The relatively small size of r implies that a change in yesterday's random disturbance has a small impact on today's random disturbance, which directly affects today's price level. The proximity of $a_1 + a_2 = 0.3411 + 0.6063 = 0.9474$ to unity suggests, however, that a jump in yesterday's random disturbance and therefore yesterday's variance can have a long-lasting effect on both today's and future variances. Moreover, even if the daily natural-gas price and Columbia River flow were to remain constant, it would take 1/(1 – 0.9474) = 19 days for the daily conditional variance to converge to its unconditional value.

The estimate of $g_1 = 0.0973$ implies that this variance increases with increases in the natural-gas price. But the small and statistically insignificant estimate $g_3 = 0.0222 (p < 0.10)$ suggests that while an increase in Columbia River flow reduces the Mid-C price level due to $b_3 = -0.1644 < 0$, it only has a negligible impact on the price variance. Taken as a whole, the set of GARCH estimates supports the notion that once a large variance is realized, there is a strong impetus for it to persist.[6]

5. Conclusions

Deregulation at both the national and state levels has had mixed results, some of which were unintended and some of which were unwelcome. Deregulation, however, is here to stay and management within those industries must not simply live with and adjust to the dynamics of unregulated markets, but must also seek advantage from the opportunities that they present. In the electricity industry, wholesale spot markets exhibit high and volatile prices that tend to cluster and persist, including regions with substantial hydro production. The only real issue for the market agents and LDC managements in particular is how best to manage the underlying spot-market price risk that they confront as buyers, when the retail rates they can charge as sellers are fixed over the short and intermediate terms, and they are committed to satisfy any and all customer demands at those fixed rates. In response to that issue, LDC managements have considered whether the financial instruments that have been successfully employed in other spot markets could be applied with comparable success in electricity markets. The GARCH analysis that we have undertaken in the present paper, with specific application to one important spot market in the

---

[6] We also estimated Equation (3) using a log-linear specification with the variables appropriately redefined. Except for the squared value of ln(Henry Hub price) with a statistically insignificant coefficient estimate, the log-linear regression results are qualitatively identical to that of the linear specification. Finally, with a GARCH-in-mean specification the coefficient estimate for the added regressor σ_t is not statistically significant at the 20% level. Hence, we rely on the AR(1)/GARCH(1, 1) model for our analysis and inferences. Because our focus is on the fundamental drivers of spot electricity prices, we have not explored other time-dependent variance specifications such as CGARCH, EGARCH, QGARCH, STGARCH, or TGARCH listed in Poon and Granger (2003, pp. 508–509) and Alexander (2001, pp. 63–116).

Pacific Northwest, highlights the difficulties that those efforts are certain to encounter.

In particular, with Mid-C as an exemplar, we have confirmed that electricity spot-market prices can be high and volatile, and subject to persisting price spikes. Moreover, due in large measure to randomness in daily natural-gas prices, hydro conditions, and temperatures, the prices cannot be forecast accurately. Assuredly, cross hedging via natural-gas and weather futures can yield a market-based forecast and mitigate electricity price risks (Woo *et al.*, 2006b), but hydro-related price risk remains.

While buyers can eliminate the local (e.g., Mid-C) price risk by entering into a forward contract with local delivery, as forward contracts are traded bilaterally for delivery in every major hub in the WECC, forward-contract purchases can be expensive because of the large risk premium often embodied in the forward price (Woo *et al.*, 2001; Bessembinder and Lemmon, 2002; Longstaff and Wang, 2004). Our results therefore provide practitioners and researchers with a sobering cold-water reality check into the problems they confront in their on-going efforts to deal with pervasive electricity spot-market price uncertainty.

Acknowledgment

We thank two referees and Professor Cheng Few Lee for their helpful comments. Without implications, all errors are ours.

References

Alexander, C., *Market Models: A Guide to Financial Data Analysis*, John Wiley, New York, 2001.

Bessembinder, H. and M. Lemmon, "Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets." *Journal of Finance* 57, 1347–1382 (2002).

Borenstein, S., "The Trouble with Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspective* 16, 191–211 (2002).

Borenstein, S., J. B. Bushnell and F. A. Wolak, "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review* 92, 1376–1405 (2002).

Bushnell, J. B., "A Mixed Complementarity Model of Hydrothermal Electricity Competition in the Western United States." *Operations Research* 51, 80–93 (2003).

California Attorney General's Office, *Attorney General's Energy White Paper*, Sacramento, CA, 2004.

Chao, H. P., "Global Electricity Liberalization: the Critical Need for Market Design Research." *Energy* 31, 923–939 (2006).

Deng, S. J., B. Johnson and A. Sogomonian, "Exotic Electricity Options and Valuation of Electricity Generation and Transmission." *Decision Support Systems* 30, 383–392 (2001).

Deng, S. J., B. Johnson and A. Sogomonian, "Spark Spread Options and the Valuation of Electricity Generation Assets." *IEEE Proceedings of the 32$^{\text{nd}}$ Hawaii International Conference on System Sciences*, 1999.

Deng, S. J. and S. S. Oren, "Electricity Derivatives and Risk Management." *Energy* 31, 940–953 (2006).

De Vany, A. and W. D. Walls, "Price Dynamics in a Network of Decentralized Power Markets." *Journal of Regulatory Economics* 15, 123–140 (1996).

Engle, R., "GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives* 15, 157–168 (2001).

Eydeland, A. and K. Wolyniec, *Energy and Power Risk Management: New Development in Modeling, Pricing and Hedging*, John Wiley, New York, NY, 2003.

Fleten, S.-E. and J. Lemming, "Constructing Forward Price Curves in Electricity Markets." *Energy Economics* 25, 409–424 (2003).

GAO, *Restructured Electricity Markets: California Market Design Enabled Exercise of Market Power*, Report GAO-02-828, United States General Accounting Office, Washington D.C., 2002.

Goto, M. and G. A. Karolyi, *Understanding Electricity Price Volatility Within and Across Markets*, Working Paper, Fisher College of Business, Ohio State University, 2003.

Green, R., *Retail Competition and Electricity Contracts*, Department of Applied Economics, CMI Working Paper 33, University of Cambridge, UK, 2003.

Greene, W. H., *Econometric Analysis*, 4th edn, Prentice Hall, Upper Saddle River, NJ, 2004.

Haldrup, N. and M. O. Nielsen, *A Regime Switching Long Memory Model for Electricity Prices*, Working Paper, Department of Economics, University of Aarhus, Denmark, 2004.

Helman, U., "Market Power Monitoring and Mitigation in the US Wholesale Power Markets." *Energy* 31, 877–904 (2006).

Hurlbut, D., K. Rogas and S. Oren, "Protecting the Market from 'Hockey Stick' Pricing: How the Public Utility Commission of Texas is Dealing with Potential Price Gouging." *Electricity Journal* 17, 26–33 (2004).

Johnsen, T. A., "Demand, Generation and Price in the Norwegian Market for Electric Power." *Energy Economics* 23, 227–251 (2001).

Joskow, P. L. and E. Khan, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000." *Energy Journal* 23, 1–35 (2002).

Jurewitz, J. L., "California Electricity Debacle: A Guided Tour." *Electricity Journal* 15, 10–29 (2002).

Kamat, R. and S. Oren, "Exotic Options for Interruptible Electricity Supply Contracts." *Operations Research* 50, 835–850 (2002).

Keppo, J. and H. Lu, "Real Options and a Large Producer: The Case of Electricity Markets." *Energy Economics* 25, 459–472 (2003).

Kleindorfer, P. R. and L. Li, "Multi-period VAR-constrained Portfolio Optimization with Applications to the Electric Power Sector." *Energy Journal* 26, 1–26 (2005).

Longstaff, F. A. and A. W. Wang, "Electricity Forward Prices: A High-Frequency Empirical Analysis" *Journal of Finance* 59, 1877–1900 (2004).

Lucia, J. J. and E. S. Schwartz, "Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange." *Review of Derivatives Research* 5, 5–50 (2002).

Mount, T. D., Y. Ning and X. Cai, "Predicting Price Spikes in Electricity Markets Using a Regime-Switching Model with Time-Varying Parameters." *Energy Economics* 28, 62–80 (2006).

Neuhoff, K. and L. De Vries, "Insufficient Incentives for Investing in Electricity Generation." *Utilities Policy* 12, 253–267 (2004).

Park, H., J. W. Mjelde and D. A. Bessler, "Price Dynamics among US Electricity Spot Markets." *Energy Economics* 28, 81–101 (2006).

Platts, *Methodology and Specification Guide: Emissions*, Platts, Boulder, CO, 2005.

Poon, S. H. and C. W. J. Granger, "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* XLI, 478–539 (2003).

Stoft, S., *Power System Economics*, IEEE, Piscataway, NY, 2002.

Trebilcock, M. J. and R. Hrab, "Electricity Restructuring in Ontario." *Energy Journal* 26, 123–146 (2005).

Vucetic, S., K. Tomsovic and Z. Obradovic, "Discovering Price-Load Relationships in California's Electricity Market." *IEEE Transactions on Power Systems* 16, 280–286 (2001).

WECC, *10-year Coordinated Plan Summary*, Western Electricity Coordinating Council, Salt Lake City, UT, 2004.

WECC, *2005 Summer Assessment*, Western Electricity Coordinating Council, Salt Lake City, UT, 2005.

Woo, C. K., A. Olson, I. Horowitz and S. Luk, "Bi-directional Causality in California's Electricity and Natural-Gas Markets." *Energy Policy* 34, 2060–2070 (2006a).

Woo, C. K., I. Horowitz, A. Olson, B. Horii and C. Baskette, "Efficient Frontiers for Electricity Procurement by an LDC with Multiple Purchase Options." *OMEGA* 34, 70–80 (2006b).

Woo, C. K., D. Lloyd and A. Tishler, "Electricity Market Reform Failures: UK, Norway, Alberta and California." *Energy Policy* 31, 1103–1115 (2003).
Woo, C. K., I. Horowitz and K. Hoang, "Cross Hedging and Forward-Contract Pricing of Electricity." *Energy Economics* 23, 1–15 (2001).
Woo, C. K., D. Lloyd-Zannetti and I. Horowitz, "Electricity Market Integration in the Pacific Northwest." *Energy Journal* 18, 75–101 (1997).

This page intentionally left blank

Index

Chapter 12

Chapter 13

Chapter 14

Chapter 15