Public Administration and Public Policy/97

Handbook of Monetary Policy

edited by

Jack Rabin

Glenn L. Stevens

Handbook of Monetary Policy

PUBLIC ADMINISTRATION AND PUBLIC POLICY

A Comprehensive Publication Program

Executive Editor

JACK RABIN
Professor of Public Administration and Public Policy
School of Public Affairs
The Capital College
The Pennsylvania State University—Harrisburg
Middletown, Pennsylvania

1. *Public Administration as a Developing Discipline (in two parts),* Robert T. Golembiewski
2. *Comparative National Policies on Health Care,* Milton I. Roemer, M.D.
3. *Exclusionary Injustice: The Problem of Illegally Obtained Evidence,* Steven R. Schlesinger
4. *Personnel Management in Government: Politics and Process,* Jay M. Shafritz, Walter L. Balk, Albert C. Hyde, and David H. Rosenbloom
5. *Organization Development in Public Administration (in two parts),* edited by Robert T. Golembiewski and William B. Eddy
6. *Public Administration: A Comparative Perspective, Second Edition, Revised and Expanded,* Ferrel Heady
7. *Approaches to Planned Change (in two parts),* Robert T. Golembiewski
8. *Program Evaluation at HEW (in three parts),* edited by James G. Abert
9. *The States and the Metropolis,* Patricia S. Florestano and Vincent L. Marando
10. *Personnel Management in Government: Politics and Process, Second Edition, Revised and Expanded,* Jay M. Shafritz, Albert C. Hyde, and David H. Rosenbloom
11. *Changing Bureaucracies: Understanding the Organization Before Selecting the Approach,* William A. Medina
12. *Handbook on Public Budgeting and Financial Management,* edited by Jack Rabin and Thomas D. Lynch
13. *Encyclopedia of Policy Studies,* edited by Stuart S. Nagel
14. *Public Administration and Law: Bench v. Bureau in the United States,* David H. Rosenbloom
15. *Handbook on Public Personnel Administration and Labor Relations,* edited by Jack Rabin, Thomas Vocino, W. Bartley Hildreth, and Gerald J. Miller
16. *Public Budgeting and Finance: Behavioral, Theoretical, and Technical Perspectives, Third Edition,* edited by Robert T. Golembiewski and Jack Rabin
17. *Organizational Behavior and Public Management,* Debra W. Stewart and G. David Garson
18. *The Politics of Terrorism: Second Edition, Revised and Expanded,* edited by Michael Stohl
19. *Handbook of Organization Management,* edited by William B. Eddy
20. *Organization Theory and Management,* edited by Thomas D. Lynch
21. *Labor Relations in the Public Sector,* Richard C. Kearney
22. *Politics and Administration: Woodrow Wilson and American Public Administration,* edited by Jack Rabin and James S. Bowman
23. *Making and Managing Policy: Formulation, Analysis, Evaluation,* edited by G. Ronald Gilbert

24. *Public Administration: A Comparative Perspective, Third Edition, Revised,* Ferrel Heady
25. *Decision Making in the Public Sector,* edited by Lloyd G. Nigro
26. *Managing Administration,* edited by Jack Rabin, Samuel Humes, and Brian S. Morgan
27. *Public Personnel Update,* edited by Michael Cohen and Robert T. Golembiewski
28. *State and Local Government Administration,* edited by Jack Rabin and Don Dodd
29. *Public Administration: A Bibliographic Guide to the Literature,* Howard E. McCurdy
30. *Personnel Management in Government: Politics and Process, Third Edition, Revised and Expanded,* Jay M. Shafritz, Albert C. Hyde, and David H. Rosenbloom
31. *Handbook of Information Resource Management,* edited by Jack Rabin and Edward M. Jackowski
32. *Public Administration in Developed Democracies: A Comparative Study,* edited by Donald C. Rowat
33. *The Politics of Terrorism: Third Edition, Revised and Expanded,* edited by Michael Stohl
34. *Handbook on Human Services Administration,* edited by Jack Rabin and Marcia B. Steinhauer
35. *Handbook of Public Administration,* edited by Jack Rabin, W. Bartley Hildreth, and Gerald J. Miller
36. *Ethics for Bureaucrats: An Essay on Law and Values, Second Edition, Revised and Expanded,* John A. Rohr
37. *The Guide to the Foundations of Public Administration,* Daniel W. Martin
38. *Handbook of Strategic Management,* edited by Jack Rabin, Gerald J. Miller, and W. Bartley Hildreth
39. *Terrorism and Emergency Management: Policy and Administration,* William L. Waugh, Jr.
40. *Organizational Behavior and Public Management: Second Edition, Revised and Expanded,* Michael L. Vasu, Debra W. Stewart, and G. David Garson
41. *Handbook of Comparative and Development Public Administration,* edited by Ali Farazmand
42. *Public Administration: A Comparative Perspective, Fourth Edition,* Ferrel Heady
43. *Government Financial Management Theory,* Gerald J. Miller
44. *Personnel Management in Government: Politics and Process, Fourth Edition, Revised and Expanded*, Jay M. Shafritz, Norma M. Riccucci, David H. Rosenbloom, and Albert C. Hyde
45. *Public Productivity Handbook*, edited by Marc Holzer
46. *Handbook of Public Budgeting*, edited by Jack Rabin
47. *Labor Relations in the Public Sector: Second Edition, Revised and Expanded*, Richard C. Kearney
48. *Handbook of Organizational Consultation*, edited by Robert T. Golembiewski
49. *Handbook of Court Administration and Management*, edited by Steven W. Hays and Cole Blease Graham, Jr.
50. *Handbook of Comparative Public Budgeting and Financial Management*, edited by Thomas D. Lynch and Lawrence L. Martin
51. *Handbook of Organizational Behavior*, edited by Robert T. Golembiewski
52. *Handbook of Administrative Ethics,* edited by Terry L. Cooper
53. *Encyclopedia of Policy Studies: Second Edition, Revised and Expanded,* edited by Stuart S. Nagel
54. *Handbook of Regulation and Administrative Law,* edited by David H. Rosenbloom and Richard D. Schwartz
55. *Handbook of Bureaucracy,* edited by Ali Farazmand
56. *Handbook of Public Sector Labor Relations*, edited by Jack Rabin, Thomas Vocino, W. Bartley Hildreth, and Gerald J. Miller
57. *Practical Public Management*, Robert T. Golembiewski
58. *Handbook of Public Personnel Administration*, edited by Jack Rabin, Thomas Vocino, W. Bartley Hildreth, and Gerald J. Miller
59. *Public Administration: A Comparative Perspective, Fifth Edition*, Ferrel Heady
60. *Handbook of Debt Management*, edited by Gerald J. Miller

61. *Public Administration and Law: Second Edition*, David H. Rosenbloom and Rosemary O'Leary
62. *Handbook of Local Govemment Administration*, edited by John J. Gargan
63. *Handbook of Administrative Communication*, edited by James L. Garnett and Alexander Kouzmin
64. *Public Budgeting and Finance: Fourth Edition, Revised and Expanded*, edited by Robert T. Golembiewski and Jack Rabin
65. *Handbook of Public Administration: Second Edition*, edited by Jack Rabin, W. Bartley Hildreth, and Gerald J. Miller
66. *Handbook of Organization Theory and Management: The Philosophical Approach*, edited by Thomas D. Lynch and Todd J. Dicker
67. *Handbook of Public Finance*, edited by Fred Thompson and Mark T. Green
68. *Organizational Behavior and Public Management: Third Edition, Revised and Expanded*, Michael L. Vasu, Debra W. Stewart, and G. David Garson
69. *Handbook of Economic Development*, edited by Kuotsai Tom Liou
70. *Handbook of Health Administration and Policy*, edited by Anne Osborne Kilpatrick and James A. Johnson
71. *Handbook of Research Methods in Public Administration*, edited by Gerald J. Miller and Marcia L. Whicker
72. *Handbook on Taxation*, edited by W. Bartley Hildreth and James A. Richardson
73. *Handbook of Comparative Public Administration in the Asia-Pacific Basin*, edited by Hoi-kwok Wong and Hon S. Chan
74. *Handbook of Global Environmental Policy and Administration,* edited by Dennis L. Soden and Brent S. Steel
75. *Handbook of State Govemment Administration,* edited by John J. Gargan
76. *Handbook of Global Legal Policy,* edited by Stuart S. Nagel
77. *Handbook of Public Infomation Systems,* edited by G. David Garson
78. *Handbook of Global Economic Policy,* edited by Stuart S. Nagel
79. *Handbook of Strategic Management: Second Edition, Revised and Expanded,* edited by Jack Rabin, Gerald J. Miller, and W. Bartley Hildreth
80. *Handbook of Global Intemational Policy,* edited by Stuart S. Nagel
81. *Handbook of Organizational Consultation: Second Edition, Revised and Expanded,* edited by Robert T. Golembiewski
82. *Handbook of Global Political Policy,* edited by Stuart S. Nagel
83. *Handbook of Global Technology Policy,* edited by Stuart S. Nagel
84. *Handbook of Criminal Justice Administration*, edited by M. A. DuPont-Morales, Michael K. Hooper, and Judy H. Schmidt
85. *Labor Relations in the Public Sector: Third Edition*, edited by Richard C. Kearney
86. *Handbook of Administrative Ethics: Second Edition, Revised and Expanded*, edited by Terry L. Cooper
87. Handbook of Organizational Behavior: Second Edition, Revised and Expanded, edited by Robert T. Golembiewski
88. *Handbook of Global Social Policy,* edited by Stuart S. Nagel and Amy Robb
89. *Public Administration: A Comparative Perspective, Sixth Edition,* Ferrel Heady
90. *Handbook of Public Quality Management*, edited by Ronald J. Stupak and Peter M. Leitner
91. *Handbook of Public Management Practice and Refom*, edited by Kuotsai Tom Liou
92. *Personnel Management in Govemment: Politics and Process, Fifth Edition,* Jay M. Shafritz, Norma M. Riccucci, David H. Rosenbloom, Katherine C. Naff, and Albert C. Hyde
93. *Handbook of Crisis and Emergency Management*, edited by Ali Farazmand
94. *Handbook of Comparative and Development Public Administration: Second Edition, Revised and Expanded*, edited by Ali Farazmand
95. *Financial Planning and Management in Public Organizations*, Alan Walter Steiss and 'Emeka O. Cyprian Nwagwu
96. *Handbook of Intemational Health Care Systems,* edited by Khi V. Thai, Edward T. Wimberley, and Sharon M. McManus
97. *Handbook of Monetary Policy,* edited by Jack Rabin and Glenn L. Stevens

98. *Handbook of Fiscal Policy,* edited by Jack Rabin and Glenn L. Stevens

Additional Volumes in Preparation

Principles and Practices of Public Administration, edited by Jack Rabin, Robert F. Munzenrider, and Sherrie M. Bartell

Public Administration: An Interdisciplinary Critical Analysis, edited by Eran Vigoda

Handbook of Developmental Policy Studies, edited by Stuart S. Nagel

ANNALS OF PUBLIC ADMINISTRATION

1. *Public Administration: History and Theory in Contemporary Perspective*, edited by Joseph A. Uveges, Jr.
2. *Public Administration Education in Transition*, edited by Thomas Vocino and Richard Heimovics
3. *Centenary Issues of the Pendleton Act of 1883*, edited by David H. Rosenbloom with the assistance of Mark A. Emmert
4. *Intergovernmental Relations in the 1980s*, edited by Richard H. Leach
5. *Criminal Justice Administration: Linking Practice and Research*, edited by William A. Jones, Jr.

Handbook of Monetary Policy

edited by

Jack Rabin
The Pennsylvania State University–Harrisburg
Middletown, Pennsylvania

Glenn L. Stevens
Franklin & Marshall College
Lancaster, Pennsylvania

MARCEL DEKKER

MARCEL DEKKER, INC. NEW YORK • BASEL

ISBN: 0-8247-0578-5

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-261-8482; fax: 41-61-261-8896

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

Preface

A market economy depends on government for its very existence. In the United States, the federal and state governments establish the institutions, laws, and regulations that are necessary for the market economy to function. Capitalist systems are inherently unstable. They are often characterized by periods of economic prosperity followed by periods of contraction and recession. This has been the experience in the United States.

Following enactment of the Employment Act of 1946, the government of the United States has been committed to maintaining full employment and price stability. Since 1979, economic policymakers in the United States, and in many other countries, have been committed to a policy of relative price stability. Today, central banks not only agree, more or less unanimously, that price stability should be the main goal of monetary policy, but most of them have in fact achieved it. The average inflation rate in the rich economies is currently just above 1%, its lowest in almost half a century.

The purpose of this handbook is to explain the development and implementation of monetary policy. We first examine theories and issues related to the preservation of economic activity, and include chapters that explore the business cycle, how it has changed over the years, and why the preservation of economic stability is a principal goal of public policy. In addition, several contributions provide a historical perspective on the development of economic theories and government economic policies. Moreover, we do not neglect the political dimensions of economic policy and how government and private organizations use the tools of economics to forecast and to measure economic activity.

Arguably, monetary policy is the most powerful weapon available to government for the management of economic activity. Certainly, that has been the experience in the United States in recent decades. Thus, the second part of the handbook reviews the development of monetary policy and its institutions. It also explores the challenge of inflation and how it has been the principal target of monetary policy. Other articles in this part examine the development and role of financial markets and institutions, issues associated with the implementation of monetary policy, and the management of interest rates.

The companion volume, the *Handbook of Fiscal Policy*, contains several articles that explain the development of government fiscal policymaking and the legacy of John Maynard Keynes. Other selections examine taxes and tax policies, government budgeting and accounting, and issues associated with government debt management. Several articles discuss the role of government in the formulation of economic policies for growth and for full employment. It concludes by reviewing issues associated with the implementation of fiscal policies.

Jack Rabin
Glenn L. Stevens

Contents

Part C: Financial Markets and Institutions

Part D: Implementation of Monetary Policy

Part E Monetary Policy and Interest Rates

Contents of Companion Volume
Handbook of Fiscal Policy

This Page Intentionally Left Blank

1

Beyond Shocks What Causes Business Cycles? An Overview

Jeffrey C. Fuhrer
Vice President and Economist, Federal Reserve Bank of Boston, Boston, Massachusetts

Scott Schuh
Economist, Federal Reserve Bank of Boston, Boston, Massachusetts

In the summer of 1997, when the Federal Reserve Bank of Boston selected the topic for its forty-second annual economic conference, many pundits were asking: "Is the business cycle dead, or at least permanently dampened?" By the time the Bank's conference convened in June 1998, the same pundits queried: "What caused the massive recessions in Asia?" and "Can the United States remain 'an oasis of prosperity,' as Fed Chairman Alan Greenspan termed it, while economies worldwide are under siege from financial crises?" How quickly things change!

Beyond Shocks: What Causes Business Cycles? turned out to be a particularly timely conference. Of course, the answers to the pundits' questions are inextricably tied to an underlying fundamental question: What makes economies rise and fall? To determine whether the business cycle is dead, one must first determine whether economic fluctuations arise from the decisions of governments, financial market participants, and businesses, or simply from unexpected events (that is, "shocks"). To determine why Asian economies plunged into severe recession, it is necessary to understand how external pressures on vulnerable financial markets can lead to a sudden collapse, with severe consequences for nonfinancial sectors. And to determine whether the robust economic expansion in the United States will continue, it is necessary to evaluate how a slew of adverse economic factors, financial and real, could interact to end it.

So, what caused the Asian crisis, the recessions of the 1970s and 1980s, and even the Great Depression? According to many modern macroeconomists, shocks did. This unsatisfying answer lies at the heart of a currently popular framework for analyzing business cycle fluctuations. This framework assumes that the macroeconomy usually obeys simple behavioral relationships but is occasionally disrupted by large "shocks," which force it temporarily away from these relationships and into recession. The behavioral relationships then guide the orderly recovery of the economy back to full employment, where the economy remains until another significant shock upsets it.

Reprinted from: *New England Economic Review*, Federal Reserve Bank of Boston (Nov./Dec. 1998) 3–23.

Attributing fluctuations to shocks—movements in important economic variables that occur for reasons we do not understand—means we can never predict recessions. Thus, a key goal of the conference was to try to identify *economic* causes of business cycles, rather than attributing cycles to "shocks." The greater the proportion of fluctuations we can classify as the observable and explainable product of purposeful economic decisions, the better chance we have of understanding, predicting, and avoiding recessions.

Several themes emerged during the conference. One was the concept of "vulnerability." It was especially prominent in discussions of the recent Asian crises and bears on the distinction between shocks and systematic economic behavior. Rudiger Dornbusch perhaps put it best in the following analogy. Consider the collapse of a building during an earthquake. While the proximate cause of the collapse was the earthquake, the underlying cause may better be attributed to poor construction techniques. Because of its structural defects, the building was going to collapse when the right "shock" came along. So it goes with financial and real economic collapses, Dornbusch and many others would argue.

While it will always be difficult to anticipate the particular event that precipitates a collapse, it is important to constantly assess the vulnerability of financial, product, and labor markets to potential shocks. Macroeconomists and forecasters tend to focus primarily on the overall health of the economy as measured by aggregate demand or by the unemployment rate; they may be able to improve their economic models by incorporating vulnerability. Likewise, policymakers should be vigilant against vulnerability. To do so, they will need to develop new tools. In Asia, for example, policymakers should have had a better assessment of the ability of the financial system to absorb shocks to currency valuations.

Developing such an assessment would likely have been hampered, many conference participants pointed out, by the inability to obtain key data on the debt portfolios of financial institutions, the performance of bank loans, and the exposure of the country as a whole to exchange rate risk. Proposals abounded for more accessible banking data and new indexes of risk exposure. Although little agreement was reached on exactly what information would be most useful, most agreed that policymakers and investors need new and more timely measures to adequately assess the vulnerability of economies to severe disruptions.

A second theme of the conference discussion was the role of systematic monetary policy in causing and preventing business cycles. Many have blamed the bulk of recessions on monetary policy. But as pointed out by Peter Temin, Christina Romer, and Christopher Sims, in assigning blame, it is important first to distinguish the systematic response of monetary policy to existing conditions from policy regime shifts and exogenous policy shocks. To take a leading example, did the Fed cause the Great Depression by raising domestic interest rates to maintain the gold standard, or was the outflow of gold from the United States following Great Britain's abandonment of the gold standard the cause, and the response of the Fed a "business as usual" response to that triggering event? Such questions are very difficult to answer, but a careful attempt to do so must be made if we are to understand the role of monetary policy in cycles.

Most participants agreed that the Fed played a significant role in causing many of the recessions of the past century, largely in the pursuit of its goal of long-run price stability. The degree to which monetary policy did or could moderate the effects of cyclical downturns was less clear. Many pointed to the apparent diminution of the amplitude of business cycles in the postwar period as evidence of the Fed's ability to lessen the severity of contractions.

Interestingly, Sims's more formal analysis of this question raised doubts that the systematic component of monetary policy either causes fluctuations or can offset them, at least through interest rate movements. Using econometric substitution of modern interest rate policy back into the Great Depression era, Sims found that modern policy would have had little effect on employment or prices. While this finding met with a good deal of skepticism from participants, one skeptic who tried to prove Sims wrong—discussant Lawrence Christiano—reported that he could not. In any case, the suggestion that conventional interest rate policy is limited in its ability to offset major recessions is thought-provoking. Of course, the limitations of interest rate policy do not preclude alternative policies, such as deposit insurance and acting as lender of last resort in financial crises. These policies may be at least as important as interest rate policy.

A third conference theme was the importance of a deeper understanding of the contribution of changes in the efficiency and structure of production to business cycle fluctuations. Recently, some macroeconomists have advanced the idea that shocks to these supply-side or "real" factors cause many, if not most, of the ups and downs in the economy. This idea contrasts sharply with the traditional macroeconomic notion that changes in aggregate demand cause most fluctuations, and the two views generate quite different policy implications.

Two real shocks were evaluated. One is a shock to the technological efficiency of firms' production of goods and services. Technological changes are very positively correlated with output and business cycles, a relationship that has led many observers to conclude that technology shocks cause fluctuations. Susanto Basu, however, demonstrates that more detailed and sophisticated estimates of technological change substantially reduce, if not completely eliminate, the correlation between technology shocks and the business cycle. He also shows how modern macroeconomic models, especially those that rely primarily on technology shocks, have difficulty fitting the data. Proponents of technology-oriented models were predictably skeptical of his results.

The second real shock is a change in the desired distribution or allocation of economic resources across firms, industries, and regions. Restructuring involves the costly and time-consuming reallocation of factors of production, especially workers, between firms, industries, and regions through the processes of job creation and destruction. It also typically involves lower output, higher unemployment, and often even recessions. In fact, job reallocation and job destruction rise sharply during recessions, leading some to surmise that shocks to the process of reallocation itself may be responsible for recessions and should therefore be taken into consideration by macroeconomic models. Scott Schuh and Robert Triest discover strong correlations between job reallocation and the primary determinants of how jobs are allocated across firms and industries: prices, productivity, and investment. Correlations between these determinants and job reallocation suggest that it is not mysterious allocative shocks that cause business cycles, but significant changes in observable economic variables.

Together, the two studies of real shocks reaffirm the fact that the production and employment behavior of firms is subject to substantial variation over the business cycle, but they deepen doubts that the variation is due to real shocks. Instead, the correlations between output and simple measures of real shocks reflect the failure of conventional analyses to incorporate a sufficiently detailed specification of production and market structure. As more and more of firms' behavior is accounted for in macroeconomic models, less and less scope remains for real shocks to generate business cycles. However, much is still to be learned about business cycles from the behavior of factor utilization, investment, prices, productivity, and the like.

OPENING ADDRESS: HISTORY OF THOUGHT ON THE ORIGINS OF BUSINESS CYCLES

Paul Samuelson's opening address begins with the question "Is the business cycle dead?" While the macroeconomy appears to have stabilized over the past 50 years, perhaps owing to successful countercyclical macropolicy, Samuelson sees no evidence of a trend toward the elimination of business cycle fluctuations. He notes that after most periods of extended expansion, especially those accompanied by outstanding performance in asset markets, suggestions of a "new era" of recession-proof prosperity have arisen, and they have been received "with increasing credulity" as the expansion rolls on. Acknowledging this historical association between healthy economies and booming asset markets, Samuelson takes a more realistic view, stressing also the intertwined histories of business cycle downturns and bubbles and crashes in asset markets.

Samuelson cites Victor Zarnowitz's recent observation that in the seven decades between 1870 and World War II, the United States suffered six major depressions. In the past 50 years, we have had no declines of comparable severity. Samuelson attributes this improved performance to changes in "policy ideology, away from laissez-faire and toward attempted countercyclical macropolicy." But despite the gains in policy's management of the economy, Samuelson sees no "convergence towards the disappearance of non-Pareto-optimal fluctuations. We are not on a path to Nirvana." The scope for improved performance arising from better government policies appears marginal today.

So pronounced fluctuations in production, prices, and employment are here to stay, despite the best efforts of policymakers. But why? In the end, Samuelson argues, fluctuations are usually the product of two factors. First, on the upside, asset price bubbles will always be with us, because individuals have no incentives to eliminate "macromarket inefficiency." While we have made tremendous progress toward "micro-efficiency"—making individual financial markets more efficient through the widespread use of options and other derivatives, for example—little evidence can be found, either in economic history or in economic theory, that "macromarket inefficiency is trending toward extinction." One can make money by correcting any apparent mispricing of a particular security, but one cannot make money attempting to correct apparent macro inefficiencies in the general level of stock market prices.

Economists and financial market participants simply have no theory that can predict when a bubble will end. As a result, an individual investor will be perfectly rational in participating in a bubble, as he will make money from the bubble so long as it continues, which could be indefinitely. As Samuelson puts it, "You don't die of old age. You die of hardening of the arteries, of all the things which are actuarially . . . associated with the process. But that's not the way it is with macro inefficiency." Bubbles go on until they stop, and no one has ever been able to predict when that will be.

Downturns can develop from the asset markets themselves, and they can develop quite quickly. Because asset prices are based on the "prudent ex ante expectations" of market participants, swings in market expectations can produce large and rapid swings in asset prices, causing massive revaluation of asset-holder's wealth. This was in part the cause of the ongoing Asian crisis, according to Samuelson. Market participants reasonably reassessed the valuation of investments (and therefore currencies) in Asia and quickly altered the direction of capital flow, precipitating a currency and banking crisis there.

Given the lack of private incentive to restrain the stimulative effects of this "oldest business cycle mechanism," we come to the second factor that contributes to business cycle fluctuations: government policy. Samuelson noted that he has often said, "When the next recession arrives, you will find written on its bottom, 'Made in Washington.' " This is not, as he points out, because the Fed is a sadistic organization. Rather, "if the central bank and fiscal authorities did not step on the brakes of an overexuberant economy *now,* they might well have to overdo that later." When persistent macromarket inefficiencies threaten both employment and price stability and private incentives fail to encourage financial markets back into line, only policymakers can take the systemic view necessary to guide the economy back into balance.

HISTORICAL EVIDENCE ON BUSINESS CYCLES: THE U.S. EXPERIENCE

Peter Temin examines the causes of U.S. business cycles over the past century. In developing his taxonomy of causes, Temin points out three inherent problems with the effort. First, the idea of a "cause" is fraught with ambiguity. In part, this ambiguity arises from the difficulty in distinguishing the endogenous, or "normal response" component of government policies and private actions, from the exogenous, or out-of-the-ordinary actions of private and public agents. In Temin's view, only exogenous events should be seen as causal. He uses oil prices and the 1973–75 recession to illustrate the dilemma: Was the recession following the oil shock "caused" by the oil shock, or by the monetary policy response to the oil shock? The imputation of causes depends on one's model of economic history, and particularly on the degree to which one makes behavior endogenous or exogenous.

Second, the Great Depression should be treated as a unique event. As Temin notes, output lost during this enormous downturn was almost one-half of the sum of output lost in all other downturns in the past century. The body of writing on the Great Depression is larger than that on all other business cycles combined. Consideration of the causes of the Great Depression provides useful lessons about the causes of the less prominent cycles of the past century. For example, it seems implausible that a single "shock" in 1929 pushed the U.S. economy into massive depression. Instead, Temin argues, the Great Depression was likely the result of a sequence of contractionary influences. Prominent among these were the fear that the hyperinflationary pressures in Eastern Europe following the First World War would spread to the United States, the adoption by industrialized countries of the relatively inflexible gold standard in response to these pressures, and the breakdown of banking and legal systems. The Great Depression was really a sequence of smaller recessions large and persistent enough, given policy responses, to throw the world into depression.

Third, Temin cautions that his assignment of causes relies on the existing literature on the subject. The literature on recessions other than the Great Depression is quite sparse, with earlier recessions receiving considerably less attention than more recent ones. And within this limited set of sources, most authors focus on the *transmission* of cycles, rather than on the causes. Finally, most of the available sources do not highlight expectations and do not clearly distinguish anticipated from unanticipated changes.

Temin classifies the reported causes of recessions as either domestic or foreign, and either real or monetary. Changes in the relative prices of assets, both real and financial, are classified as real phenomena. Temin finds that the preponderance of cycles in the past century may be attributed to domestic causes, with the split between real and monetary causes roughly equal for the entire period. Monetary causes of recessions were more prevalent in the pre–World War I period than during the post–World War II period, however.

Temin focuses on the larger downturns. The cause of the Great Depression of 1931 is classified in Temin's taxonomy as a foreign monetary phenomenon. The action of the Fed to maintain the gold value of the dollar by raising interest rates was to behave as a "traditional and responsible central banker" or, in other words, to follow a normal and expected endogenous policy course. Thus, the Fed's behavior cannot be viewed as an exogenous cause of the Great Depression, in Temin's view. The search for causes then reverts to the question of what produced this monetary policy response. Temin suggests that U.S. monetary policy was responding to the external gold drain that arose from Britain's departure from the gold standard, which threatened to weaken the dollar. The Fed's reaction in increasing interest rates, and the bank panics and failures that followed, were endogenous responses to the gold drain.

In assessing the causes of the four largest downturns of the century—the Great Depression, and the recessions of 1920, 1929, and 1937—Temin concludes first that no single cause explains all four downturns. Three of the four possible causes in Temin's taxonomy appear as causes of the downturns. Second, three of the four recessions appear to be responses to domestic shocks. Most often, we cannot blame our downturns on foreign causes.

Taking all of the cycles studied into consideration, Temin offers the following conclusions: (1) "It is not possible to identify a single type of instability as the source of American business cycles." Thus, Dornbusch's statement, "None of the U.S. expansions of the past 40 years died in bed of old age; every one was murdered by the Federal Reserve," is not supported by Temin's analysis. (2) Domestic real shocks—ranging from inventory adjustments to changes in expectations—were the most frequent source of fluctuations. (3) Other than the two oil shocks of 1973 and 1979, foreign real shocks were not an important source of U.S. cycles. (4) Monetary shocks have decreased in importance over time. (5) When measured by the loss of output, domestic sources have loomed larger than foreign sources; real sources have caused about the same losses as monetary sources.

Christina Romer takes issue both with Temin's classification scheme and with his interpretation of the literature on the causes of recessions. She suggests that an improved classification scheme and a different reading of the literature would yield a more critical role for domestic monetary shocks, particularly in the inter- and postwar periods.

Romer suggests that Temin's methodology is biased toward finding very few monetary causes of recessions. Whereas Temin classifies most Fed behavior as a fairly typical response to prevailing conditions and therefore not the ultimate cause of the recession, Romer would prefer a more practical classification of monetary policy actions. If the monetary policy action was the inevitable or highly likely result of a trigger, then we should consider the policy action endogenous and therefore not a cause. If, however, "a conscious choice was made" or if "alternative policies were . . . discussed at the time," then the policy should be considered at least partly exogenous, and monetary policy should get some blame for the recession.

Romer shows that, using this criterion, many more of the twentieth-century recessions have an important monetary policy aspect. Monetary factors would likely be given an important causal role in the 1931 recession, for example, as "reasonable men *at the time* were urging the Fed to intervene" in the face of financial panics. Thus, the choice not to intervene but to raise the discount rate was not inevitable or even most likely. Romer also questions the extent of the constraint imposed by the gold standard, as U.S. gold reserves in 1931 were probably adequate to have allowed the Fed to pursue expansionary open market operations while maintaining the gold value of the dollar, as in fact it did in 1932.

Turning to the 1973 recession, for which Temin ascribes no monetary role, Romer argues that the central bank was not simply acting as "a respectable central bank [that] resists inflation," and therefore responding only as expected. Romer points out that the decision to tighten in 1974 was not a foregone conclusion but rather a conscious choice, as "the economy was already in a downturn and many were calling for loosening." Thus, "monetary policy and the oil shock share responsibility for the 1973 recession."

Romer also challenges Temin's attribution of the 1957 and 1969 recessions to declines in government spending. She points out that the high-employment budget surplus actually *falls* throughout the late 1950s, suggesting a net stimulative impulse from the federal government for the 1957 recession. For both recessions, Romer asserts that the Federal Reserve made a conscious decision to tighten in order to reduce inflation.

As Romer sees it, "the key change has not been from monetary to real shocks or vice versa, but from random shocks from various sources to governmental shocks." Since the Second World War, the government has been more effective at counteracting most shocks, accounting for the diminished frequency of cycles. However, the combination of a tendency toward over expansion and a few large supply shocks caused inflation to get out of hand. In sum, Romer would agree with the thrust of Dornbusch's statement, which is that monetary policy has played a vital role in postwar recessions. She might re-cast the role of the Fed, however, as "more like a doctor imposing a painful cure on a patient with an illness than a murderer."

HISTORICAL EVIDENCE ON BUSINESS CYCLES: BUSINESS CYCLES ABROAD

Michael Bergman, Michael Bordo, and Lars Jonung examine the broad cyclical properties of GDP, using a newly compiled data set of annual observations for a sample of "advanced" countries. Their data set spans the years 1873 to 1995. The authors show that the duration of business cycles (the calendar time from peak to peak or trough to trough) has been fairly similar across countries and fairly stable over time. The average duration rose from about four years in the pre–World War I period to about five and one-half years during the interwar period, falling back to just under five years in the period following World War II. The most severe recessions appear to have occurred prior to 1946, and the magnitude of all fluctuations in GDP seems to have decreased in the postwar period.

Formal statistical tests of diminished cyclical fluctuations in the postwar period generally confirm the visual evidence. This observation has often been interpreted as evidence that countercyclical policy has been more effective in the postwar period. However, an alternative explanation is that the increased integration of the world economy serves to mitigate the negative influence of any one country's disruptions on other countries.

Conventional wisdom holds that downswings are sharper and "steeper," whereas upswings are more gradual. Bergman, Bordo, and Jonung test this proposition and find that, for the United States, upswings are indeed more gradual than downswings. The evidence for other countries is more mixed, however, with most exhibiting this asymmetry prior to World War II but only a minority displaying asymmetry in the postwar period.

The authors then attempt to determine the extent to which different components of GDP—including consumption, investment, government expenditures and revenues, exports, and imports—account for its cyclical volatility. For virtually all countries and time periods, all components of GDP except consumption generally are more volatile than GDP. This finding is consistent with the presence of a consumption-smoothing motive, that is, the desire of consumers to maintain a relatively smooth stream of consumption over time in the face of volatility in their income and wealth.

The authors find that larger countries experience deeper recessions; the average decline in GDP below trend is larger for large countries than for small, open European countries. For most countries, the downturn in GDP during a recession is accounted for by declines in consumption, investment, and net exports.

Finally, Bergman, Bordo, and Jonung consider the patterns of international co-movement of output and prices in their data. They find that the correlations among real output in the 13 countries have increased over time, suggesting a more integrated world economy and possibly a stronger coherence of the business cycle across countries. During the gold standard, real GDP for most countries exhibited little or no correlation with real GDP in other countries. During the interwar period, U.S. GDP was significantly correlated with seven other countries, but corresponding correlations between other countries were not evident. The authors suggest that this correlation arises from the role of the United States as the "epicenter" of the Great Depression. Output linkages among European countries strengthened considerably in the postwar period, perhaps the result in part of the establishment of the European common market and in part of the common influence of the oil shocks in the 1970s.

Price levels appear to be much more consistently correlated across countries. Like output, price levels have become increasingly correlated over time, perhaps consistent with "increased global integration of goods markets," the authors suggest.

Richard Cooper offers a different perspective on Bergman, Bordo, and Jonung's conclusion that "the cyclical pattern . . . appears to remain surprisingly stable across time, regimes, and countries" and on the broad question of the international origin and transmission of the business cycle. He examines years in which the raw data for real GDP declined, for a set of nine countries during the periods 1873 to 1913 and 1957 to 1994. Cooper prefers this approach, as the author's results may depend on the filtering and detrending methods that they used in constructing their data.

The conclusions that he draws for the earlier period are as follows: First, "most downturns are domestic in origin, and are not powerfully transmitted to the other important trading nations." Second, if one were interested in international transmission, one would focus on 1876, a year in which the Continent and Canada experienced declines in GDP, and on 1879 and 1908, years in which several countries experienced output declines. Third, Belgium exhibits only one downturn during these periods, a suspicious finding given the 12 downturns in neighboring Netherlands and 14 in France. As a result, Cooper calls into question the reliability of the annual data for any of these countries prior to 1914.

For the period 1960 to 1995, Cooper notes that the few recessions have been concentrated in five years: 1958, 1975, 1981–82, and 1993. This suggests strong international transmission, in contrast to the earlier period. All of the recessions in the United States were accompanied by recessions elsewhere. The greater coherence may be attributed to the importance of the oil price shocks in these recessions, Cooper notes.

Cooper goes on to question the detrending method used by Bergman and his coauthors. Only 60 percent of their recessions match NBER reference dates. The issue of appropriate filtering is important when considering the welfare implications of business cycles, Cooper suggests. A departure of output below its (rising) trend may imply relatively little lost income or underutilized resources, whereas an absolute decline in output would almost surely entail significant welfare losses.

Cooper outlines a number of broad changes in industrial economies that would lead one to question Bergman, Bordo, and Jonung's conclusion about the stability of the business cycle over long spans of time. He suggests that "the most dramatic by far . . . is the reduction in the fraction of the labor force required for food production." The decline in this number from about one-half in 1880 to below 5 percent by 1995 for all of these countries is likely to have altered the dynamics of the business cycle significantly, according to Cooper. Other important secular changes include the increased participation of women in the paid work force, the growth in the importance of government expenditures, and major technological innovations, including electricity, automobiles, and aircraft. "A relatively unchanged economic cycle that survived these dramatic secular changes in modern economies would be robust indeed," Cooper suggests.

GOVERNMENT POLICY AND BUSINESS CYCLES

Christopher Sims examines one of the most contentious questions in macroeconomics: the role of monetary policy in twentieth-century business cycles. Sims points out that one cannot determine the influence of monetary policy simply from observed changes in interest rates and output. The observation that a rise in interest rates precedes each postwar recession does not show that policy-induced interest rate movements *caused* the recession. If, for example, rapid expansion of private demand for credit systematically causes all interest rates to rise near the end of an expansion, this rise in interest rates should not be interpreted as the cause of a subsequent slowdown; it is a consequence of previous strong demand. Because such "eyeball" interpretations of the data can lead to confusion about the role of monetary policy, Sims advocates examining the interactions among many economic variables in order to obtain a clear picture of the role of any one of them in economic fluctuations.

Sims employs a methodology that allows each of six variables (industrial production, consumer prices, currency, a monetary aggregate, the discount rate, and commodity prices) to respond to lags of the other variables, and to the contemporaneous values of *some* of the other variables. The restrictions on the contemporaneous interactions among variables reflect common-sense notions about policy, goods market, and financial market behavior. Monetary policy-induced interest rate changes affect prices, output, and monetary aggregates only with a one-month lag; monetary policy responds to output and prices only with a lag, reflecting data availability; and commodity prices respond to everything contemporaneously, reflecting their auction-market, flexible nature.

This simple model is estimated on monthly data for the postwar years 1948 to 1997. Sims uses the model to show that most of the variation in the Fed's discount rate represents systematic policy responses rather than unanticipated shifts in policy. The discount rate responds primarily to movements in production, commodity prices, and M1. These three determinants of interest rate movements in turn cause the largest increases in CPI inflation, suggesting that the Fed responds to these as signals of future inflationary pressures.

When Sims estimates this same model on the interwar period from 1919 to 1939, he finds similarities but also some important differences in monetary policy responses and influences. One key difference is that the effect of interest rate changes in the early period is roughly double the effect in the later period. On the other hand, monetary policy in the early period appears to be more accommodative toward unanticipated increases in output, raising the discount rate less in response to output and thereby allowing greater inflation in commodity and in final goods prices. Interestingly, the model shows that when depositors' worries caused a rush into currency in the interwar period, the Fed typically *raised* the discount rate, accelerating the shrinkage of money.

This first set of exercises establishes that the systematic responses of policy to output and prices represent the dominant source of interest rate fluctuations in Sims's model, and that these interest rate movements are likely the most important source of policy's effects on the rest of the economy. Noting that economic fluctuations have been smaller in the postwar period, Sims proposes using his model to answer a key question: whether better systematic monetary policy is responsible for the improved economic performance of the postwar period.

To answer this question, Sims transplants the estimated monetary policy equation for one period into the other period, then observes the estimated behavior of output, prices, and monetary aggregates under this counterfactual monetary regime. The results from these exercises are remarkable. In the first variant, the (estimated average) policy judgment of Burns, Volcker, and Greenspan is imposed on the 1920s and 1930s. Overall, Sims finds the outcomes—particularly the Great Depression—would have been little changed by this more responsive postwar policy. The drop in production from 1929 to 1933 is "completely unaffected by the altered monetary policy." Postwar policy would have made the 1920–21 and 1929–33 deflations less severe, but not by much. The upheaval of the 1920s and 1930s would have been the same, even if modern monetary policymakers had been at the reins. Sims notes that his methodology leaves the banking runs, panics, and currency speculations that plagued the Depression era as unexplained non-monetary shocks. To the extent that a persistent commitment to monetary ease would have alleviated such disruptions, the drop in output might have been less severe, he suggests.

The effects of substituting interwar monetary policy into the postwar economy are qualitatively the same. Even though the discount rate responds much more slowly to the postwar economic fluctuations, resulting in a markedly different interest rate pattern, the influence of this altered policy on industrial production and consumer prices is quite small at business cycle frequencies. The implications for output and inflation at longer horizons are what one would expect with a more accommodative policy: Output and inflation both rise higher in the 1970s, resulting in a larger recession in the 1980s, although Sims is careful to point out that these findings may well be statistically unreliable. Overall, he reaches the startling conclusion that "the size and timing of postwar U.S. recessions had little to do with either shocks to monetary policy or its systematic component."

Lawrence Christiano focuses on Sims's surprising conclusion that monetary policy played little or no role in the Great Depression. He disagrees with the methodology that Sims uses to reach this conclusion, but upon employing what he considers a superior method, he confirms Sims's results.

One criticism of Sims's methodology revolves around the assumption that private agents behaved the same in the postwar period after the creation of the Federal Deposit Insurance Corporation (FDIC) as they did during the interwar period prior to the FDIC. Christiano suggests that the frequency with which interwar depositors converted deposits to currency at the slightest sign of bad news, in contrast to the virtual absence of such bank runs in the postwar period, suggests that the presence of the FDIC fundamentally changed private agents' behavior. In particular, they may have viewed the commitment of Federal Reserve policy to maintain banking system liquidity quite differently in the postwar period, and in a way that cannot be captured by the simple "reaction functions" or interest rate equations in Sims's analysis.

The more important flaw in Sims's analysis, according to Christiano, is the characterization of the postwar monetary policy rule. Under this rule, after all, the Fed would have *contracted* the money supply by 30 percent in the 1930s. Christiano cannot conceive of a sensible policymaker who would pursue a contractionary monetary policy during a widely recognized, worldwide depression. So Christiano proposes instead to use a monetary policy equation that keeps money (M1) from falling during the episode.

Using this more plausible counterfactual policy in Sims's model for the interwar period, Christiano finds that a stable M1 path for the early 1930s would have prevented the dramatic price declines that actually occurred. Surprisingly, however, even under the more realistic policy response, which implies a more realistic path of money growth, "the basic course of the Great Depression would not have been much different," as shown by the similarity between the path of output in Christiano's simulation and the actual path of output.

Benjamin Friedman is also skeptical of the empirical results developed in Sims's paper, stating: "If the model he presents has succeeded in identifying Federal Reserve actions and measuring their economic effects, these findings should force us to reconsider many aspects of economics and economic policy." Friedman finds troubling Sims's result that postwar monetary policy would not have significantly altered the course of the Great Depression, and he views as even more problematic the finding that Depression-era monetary policy would have worked just the same in the postwar period as did actual policy. Friedman notes that the general price level was approximately the same at the onset of World War II as at the onset of the Civil War, while prices since that time have risen approximately tenfold. That the monetary policy that delivered the interwar *de*flation is the same one that delivered the "historically unprecedented phenomenon of a half century of sustained *in*flation" would make inflation, even over periods of several decades, never and nowhere a monetary phenomenon.

Friedman suggests that Sims's model delivers its surprising results because it fails to adequately identify the Fed's monetary policy actions or the effects of those actions on the macroeconomy. If so, then the model's "implied irrelevance of monetary policy" for the postwar inflation translates further into irrelevance for assessing monetary policy's role in causing or cushioning business cycles. One indication that Sims's postwar policy rule does not accurately represent Fed actions, Friedman argues, is the difference between the Sims model's policy prescriptions for the Depression era and John Taylor's policy rule prescriptions for the same period. Friedman finds that Taylor's rule would imply nominal interest rates "an order of magnitude more negative than what Sims reports," casting some doubt on how well Sims's policy rule reflects all of postwar Fed behavior.

Finally, Friedman notes that the assumption that Fed policy can be characterized by one unchanging rule over the entire postwar period is implausible. He asks, "Are we really to equate Paul Volcker's tough stance against inflation with the see-no-evil regime of Arthur Burns?" While Friedman recognizes that Sims tests for a shift in monetary policy in 1979, Sims does so by testing for a shift in all 279 of his model's parameters. Friedman notes that Sims could have more narrowly focused this test to detect only shifts in the parameters that summarize monetary policy.

FINANCIAL MARKETS AND BUSINESS CYCLES: LESSONS FROM AROUND THE WORLD

A panel composed of Rudiger Dornbusch, Maurice Obstfeld, and Avinash Persaud analyzed recent financial market crises, most notably the turmoil in Asia, and drew lessons on how to reduce the likelihood and severity of future crises. Generally speaking, the panelists agreed more on why the crises occurred than on what should be done to prevent future crises.

Dornbusch believes that recent financial crises in Asia, Russia, and Mexico differed from most preceding crises because they centered on capital markets rather than on the balance of payments. Both types of crises often are associated with currency crises as well, but the vulnerability or risk imposed on an economy by a capital market crisis is fundamentally different. He explains that financial systems experiencing a capital market crisis exhibit five characteristics: (1) borrowing short and lending long generates a *mismatching of maturities* between liabilities and assets; (2) borrowing in foreign currency units and lending in domestic currency units generates a *mismatching of denominations;* (3) borrowing to carry assets exposed to large fluctuations in price generates *market risk;* (4) high risk exposure throughout a country generates a *national credit risk;* and (5) the central bank is weakened by *gambling away foreign exchange reserves.*

According to Dornbusch, the capital market crisis in Asia made the regional economy vulnerable, or at risk, to adverse external factors. And two such factors happened. First, "Japan went into the tank." Just as the Japanese economy was starting to show signs of emerging from several years of sluggish growth, the Japanese government tightened fiscal policy and the economy slumped again. This time the weakened economy exposed underlying banking problems that exacerbated the situation so much that the Japanese economy eventually began to contract. Because Japan is the largest economy in the region and the leader in regional export and import markets, the Japanese slump put stress on the foreign trade structure of the entire region, which is characterized by extensive export and import linkages.

A second adverse factor was the sharp depreciation of the yen vis-à-vis the U.S. dollar, "leaving the dollar peggers high and dry." Asian economies that were dependent on robust exports to Japan but had pegged their currencies to the dollar suddenly found their exports priced too high, in yen terms. Export demand fell sharply among Asian trading partners, and almost overnight domestic economies throughout the region began experiencing severe contractions. Together these adverse external factors turned vulnerable economies into collapsing economies. Thus, Dornbusch attributes the Asian economic downturn to a confluence of capital market vulnerability and adverse external factors.

Obstfeld also believes that the primary source of economic vulnerability in recent financial crises was capital markets, but he emphasizes shifts in expectations as the central factor driving the economic fluctuations. He notes that "exogenous fluctuations in capital flows have become a dominant business cycle shock" for developing countries in the modern era, and that similar financial crises were quite common prior to World War II.

Obstfeld describes two main types of crises—exchange rate (currency) crises, and national solvency crises—and explains that although they can occur separately, they often "interact in explosive ways." The main linkage between them is self-fulfilling expectations. An economy with a weak and vulnerable capital market can avoid crisis so long as there is no expectation of one. But when expectations change, the desirable but tenuous equilibrium will give way abruptly to a crisis. A sudden new expectation of currency depreciation can start the process rolling, once speculators perceive the threat that public debt will be paid through inflation. He cites Indonesia as an example of this phenomenon.

In Persaud's view, moral hazard and inadequate oversight were key factors in generating the underlying capital market vulnerability. "Moral hazard [induced by International Monetary Fund bailouts] . . . probably played a role in the exponential rise in foreign bank lending to Emerging Asia," and "crony capitalism" may have further "impaired the proper allocation of resources." Furthermore, Asia's economic success was "unbalanced" in the sense that lending went toward overinvestment that was concentrated in a limited number of sectors. Inadequate supervision and unreliable information about this worsening capital situation allowed the rise in risky lending and overinvestment to go unchecked until it was too late.

Persaud also cites the weakened Japanese economy and depreciating yen as important factors, but he identifies the collapse of the Thai baht on July 2, 1997 as the "trigger" that set off the Asian crisis. The effect of this trigger was amplified as investors suddenly realized new or mispriced risks in the region and greatly reduced their "appetites for risk"; this led to widespread and simultaneous capital outflows from the region.

A key factor contributing to this capital flight, says Persaud, was the sudden discovery that domestic corporate investment positions were highly concentrated. When the crisis emerged, heavyweight investors in the region discovered that their peers were also deeply vested in the same small number of collapsing Asian economies. Thus, these influential investors not only wanted to get out of Asia because of the inherent financial problems, they also wanted to get out first, because they knew that a massive capital outflow would dramatically reduce asset prices in the region.

The panelists generally agreed that unwise economic decisions had promoted an environment of vulnerability, and that Japan's economic weakness and other events turned a precarious situation into turmoil. However, their recommendations about how to respond to the current crisis, and how to prevent future crises, were notably different.

Dornbusch believes that the key to preventing future capital market crises is to control financial risk. He proposes using model-based value-at-risk ratings and disseminating "right thinking" within the international financial community regarding controlling and pricing such risk. Controlling capital flows themselves, however, is not appropriate. He advocates International Monetary Fund (IMF) inspections of financial market conditions during country consultations, but he is doubtful the IMF will become sufficiently forward-looking and preemptive, because IMF member countries will resist such changes. For this reason, he particularly opposes an Asian IMF. Dornbusch advocates moving toward regional currencies like the euro. Regarding the appropriate response to current developments, Dornbusch is adamant that tight money policies are required to restore financial stability; debt restructuring can be negotiated later. Fiscal policy is not a viable tool because of the fiscal deterioration associated with the recent crises.

Obstfeld asserts that "policy must counteract the severe capital-account shocks by creating a new expectational climate" that will restore confidence in these economies. He sees no economic prescription for this change "short of infeasibly extensive official financial support from abroad." In contrast to Dornbusch, Obstfeld concludes that fiscal expansion is the least risky policy prescription, particularly in Japan. Monetary expansion in Japan might also help, but it carries the risk of further yen devaluation and is insufficient until Japan resolves its banking problems. He ends by warning that monetary tightening now by the Federal Reserve and the new European Central Bank to fight domestic inflation "would be an error of perhaps historic proportions."

Persaud highlights the need to develop policies that "work with financial markets and not against them." He views many actual and proposed policies as counterproductive. Capital controls intended to curb outflows would implicitly curb much-needed inflows. Looking to the IMF for faster and more lucrative assistance is also unwise. He doubts that the IMF loans can keep pace with the magnitude of required private capital flows, and in any case further IMF assistance worsens the moral hazard problem.

Instead, Persaud wants an international financial system that permits countries access to an international pool of foreign exchange reserves if—and only if—they meet certain "selectivity criteria" intended to reflect sound and prudent financial operations. The criteria, which must be "public, clear, and transparent," would consider the extent of external debt, the productivity of capital inflows, the competitiveness of exchange rates, the soundness of government finances, and the openness of governance. Countries or financial institutions that do not meet these criteria should be allowed to fail. Indeed, Persaud believes that selective assistance is a critical requirement for eliminating moral hazard.

PRODUCTION, TECHNOLOGY, AND BUSINESS CYCLES

Susanto Basu tackles another of the most contentious questions among modern macroeconomists: Do fluctuations in technological change or productivity growth actually cause business cycle fluctuations? Some prominent neoclassical macroeconomists assert not only that the answer is yes, but that technology change is the *primary* determinant of such fluctuations. This assertion is contested by macroeconomists like Basu who adhere to the Keynesian tradition of emphasizing fluctuations in aggregate demand as the primary contributor to business cycles. Because these two views of the sources of business cycles lead to radically different macroeconomic models and prescriptions for government policy, resolution of this debate is critical.

Basu argues that neoclassical economists have misinterpreted the link between technological change and business cycles by misusing the standard measure of technological change: the Solow residual, named after M.I.T. economist Robert Solow. Solow's methodology is simple: measure the growth of output; subtract the appropriately weighted growth of all observable inputs such as labor, capital, and materials; and the difference, or residual, is an estimate of unobserved technological change. Economists use this sensible but indirect measure because they do not have direct data measures of technological change.

Thus far, most attempts to construct Solow residuals with conventional data on inputs yield a measure that is positively correlated with output, giving rise to the claim that technological changes cause business cycles. But Basu argues the Solow residual was only intended to estimate the long-run impact of technology on the economy, not the cyclical impact. He notes that Solow warned long ago that his measure would be spuriously correlated with output and the business cycle because firms adjust to fluctuations in demand by varying the rates at which they utilize capital and labor.

Basu has developed a new measure of technological change that adjusts for features that could lead to an excessively positive correlation between technological change and output. Basu's methodology, developed in earlier research with John Fernald and Miles Kimball (henceforth the BFK technology measure), adjusts for four factors: (1) variable utilization of capital and labor; (2) variable worker effort; (3) imperfect competition and other special advantages firms may have in production; and (4) different characteristics of firms across industries. In other words, it adjusts for many of the demand-side features

Solow was concerned about. The BFK methodology requires relatively few controversial restrictions or assumptions; indeed, previous measures of technological change are special cases of it.

The salient and distinguishing feature of the new BFK technology measure is that it is essentially uncorrelated with output and the business cycle. Unlike the Solow residual, which is positively correlated with output and the business cycle, it exhibits no simple statistical evidence of causing business cycle fluctuations. Moreover, the BFK measure is much less variable than the Solow residual. Together, these features reduce, if not eliminate, the likelihood that unexpected technological changes cause business cycles. Basu shows that this conclusion holds up in simple statistical models of the production process.

Another potentially important characteristic exhibited by the BFK technology measure is that it suggests what all workers fear: that technological improvements reduce employment. At least initially, the BFK measure is very negatively correlated with factor inputs, such as labor and factor utilization. In other words, when firms improve their technical efficiency by installing the latest and greatest machines, they are able to produce the same output with fewer inputs, so they reduce costs by cutting their work force rather than reducing their prices and producing more. Only much later, as profits rise, do they expand their output and hire workers. This interpretation of the data stands in stark contrast to interpretations based on the conventional Solow residual, in which employment and other factor inputs rise with technological improvements.

In the second part of his investigation, Basu uses his technology measure to evaluate whether the dynamic properties of two state-of-the-art macroeconomic models match the postwar data. One is the real business cycle (RBC) model, which features technological change as the main source of business cycle fluctuations. It also assumes complete, competitive markets with fully adjustable prices. The other model is basically similar but introduces slowly adjusting or "sticky" prices. Sticky prices are a common feature of macroeconomic models that emphasize fluctuations in aggregate demand as the main source of business cycles.

The result of Basu's evaluation is quite discouraging for state-of-the-art macroeconomic models. He finds that neither the RBC nor the sticky price model generally fits the data very well. The RBC model, in particular, does not match the dynamic properties of the data, and it cannot reproduce the essentially zero correlation that exists between the BFK technological change and output or the negative correlation between factor inputs and output. These models also fail to reflect the generally sluggish response of output changes in the economy. Basu reports that the sticky price model is qualitatively better because it approximately reproduces these two correlations, although it does not do so well. The prognosis for these models becomes even bleaker when he evaluates the models with both technological change and various specifications of monetary policy.

Basu concludes that the defining cyclical feature of technological change is a short-run reduction in inputs and factor utilization, and that business cycle models face the challenge of reproducing that feature. At present, standard RBC and sticky price models cannot do the job, and variable factor utilization does not impart enough rigidity to generate sufficient sluggishness. He projects that the sticky-price models modified to include other sources of rigidities, "show some promise of being able to match the data, but clearly have a long way to go."

Mark Bils questions whether Basu's technology measure adjusts *too much* for the positive correlation between factor utilization and output. He hypothesizes that the proportions of capital and labor used in production are likely to be fixed in the very short run.

Thus, when capital utilization rises slightly, labor hours will rise in equal proportion. If so, total factor productivity should be positively correlated with output but labor productivity should be approximately uncorrelated with output. Bils finds exactly these correlations in data on detailed manufacturing industries. Because the BFK methodology infers movements in capital utilization from movements in materials prices, and because materials prices are more positively correlated with output than labor costs, Bils believes the BFK measure makes capital utilization more positively correlated with output than labor utilization is.

Other aspects of Basu's methodology make Bils skeptical of the results. He doubts that labor quality (effort) is positively correlated with output, as in the BFK measure, because there is evidence that workers hired during expansions are paid less and therefore of lower quality. Moreover, he thinks the relationship between effort and hours will vary depending on the stickiness of wages and the type of shock. Bils also argues that factor utilization will vary more if shocks are transitory rather than permanent. Basu's methodology relies more on variables associated with transitory shocks, so it may yield estimates of utilization that are too positively correlated with output.

Finally, Bils assesses the plausibility of price stickiness in two empirical exercises. One exercise is based on the theory that if prices are sticky, then firms with significant inventory holdings should be less likely to reduce inputs and output when technology increases, because they can inventory unsold output. He reports evidence that "labor hours are much less likely to decline for industries that hold significant inventories," but points out that this evidence does not conclusively determine the actual flexibility of prices. So in a second exercise he provides more direct evidence from models of relative prices. Prices are significantly negatively correlated with current total factor and labor productivity but not with past productivity, a relation Bils interprets as evidence that prices are not sticky.

Thomas Cooley is also cautious about interpreting Basu's results as evidence against the idea that technological change is an important source of business cycle fluctuations. Like Bils, Cooley has reservations about the methodology underlying the BFK technology measure, although he embraces Basu's finding that firms do not enjoy market power from technological advantages in production. In particular, he notes that the correlation of the BFK technology measure with output is sensitive to the exact form of the econometric methodology used to construct the measure and to the identifying assumptions of the modeling framework.

However, granting the validity of Basu's results, Cooley directs his critique at the logic of Basu's inferences about the implications for macroeconomic models. First, he questions Basu's conclusion that the results necessarily rule out RBC-type models. He argues that RBC models no longer rely on artificially sluggish technology shocks to obtain sluggish output responses. Sluggishness can arise from factor utilization as well as financial market imperfections, differences among firms, and other features. As for the RBC model's inability to generate a negative correlation between technology and factor inputs, he suspects that this result is not robust.

Cooley also questions whether the evidence should lead one to conclude that prices are sticky. Basu provides no direct evidence of sticky prices, and economic theory does not make clear predictions about the direction in which capital and labor should respond to technology changes. The response will depend, among other things, on the nature of the technology change, market structure, and the sensitivity of demand to prices. This point calls into question Basu's assertion that he does not need to consider the behavior of profits and product markets.

Cooley thinks Basu's results suggest that technological change is embodied in new capital investment—a characteristic absent from the BFK methodology. With technology embodied in capital, the short-run responses of output and factor inputs to technological change are different from those of a standard RBC model and are capable of yielding the patterns Basu finds in the data. Moreover, in this case the nature of depreciation matters for interpreting the effects of cyclical factor utilization.

REALLOCATION, RESTRUCTURING, AND BUSINESS CYCLES

Scott Schuh and Robert Triest investigate the idea that business cycles might be caused by the shuffling of jobs as firms restructure the way they do business. New data produced during the past decade show that firms are continuously changing. Some expand and create jobs while others contract and destroy jobs. The pace of change is rapid; one in 10 jobs is newly created and one in 10 jobs newly destroyed in manufacturing each year. The sources of these ups and downs of particular firms include product demand and innovation, prices and wages, regional economic conditions, technological change, and other factors idiosyncratic to each firm, rather than factors common across all firms. Job creation and destruction together represent job reallocation, a measure of job turnover or churning in the economy.

Traditionally, macroeconomists looking at the labor market have ignored job reallocation and have focused solely on total employment growth (or the total unemployment rate). However, Schuh and Triest point out that a given rate of employment growth can occur with either low or high rates of job reallocation. More important, the intensity of job reallocation has significant consequences for unemployment, wage growth, and productivity growth.

For example, if changes alter the desired distribution of jobs across firms, industries, and regions, job reallocation must intensify to keep productive efficiency high. More intense reallocation usually means higher job destruction that forces many workers into unemployment. These unemployed workers lose any skills they had that were unique to their previous job (such as knowledge of firm operating procedures), have a hard time finding a comparable new job, and stay unemployed longer. Eventually they may have to accept a job entailing sizable reductions in their wages. Such issues are linked inherently to the determination of aggregate unemployment, wage growth, and productivity.

Schuh and Triest point out that job reallocation and the pace of restructuring rise markedly during recessions. Traditional macroeconomic models cannot explain why because they do not incorporate the phenomenon of job reallocation. But in light of the potentially negative economic consequences of job reallocation, it is important to know whether an identifiable connection exists between reallocation and business cycles, and whether the correlation between them is of no consequence and can continue to be ignored.

Schuh and Triest ask the following fundamental question: Does job reallocation cause business cycles, or do business cycles cause job reallocation? Evidence on job reallocation has sparked an interest in building theoretical models capable of explaining the observed patterns in the data, and they classify these theories into two types. One type stresses the role of factors that primarily determine the desired allocation of economic resources, such as workers, across firms. The other type stresses the role of aggregate factors, such as monetary policy, that primarily determine the overall level of economic activity. Both types of theories aim to explain why job reallocation rises during recessions. Yet both types of theories tend to rely on vaguely defined aggregate and allocative "shocks" rather than observable variables.

Schuh and Triest argue that these theories do not and cannot answer their fundamental question, for two reasons. First, although the two-way classification of factors may be conceptually sensible, in practice the definitions of allocative and aggregate factors become hopelessly muddled. Second, these theories have little to say about what *causes* business cycles—that is, *why* they occur—because they focus more on *how* they occur.

Schuh and Triest present results from three empirical exercises that extend research by Schuh with Steven Davis and John Haltiwanger on job creation, destruction, and reallocation (henceforth referred to as DHS). One exercise analyzes the behavior of job reallocation during the 1990s using newly available data. A second exercise attempts to learn what kinds of plants destroy and reallocate jobs and how, in hope of discovering clues about the causes of recessions. The third exercise looks for evidence of causal relationships between job reallocation, the fundamental determinants of reallocation, and the business cycle. Each of these exercises uses data from the U.S. Bureau of the Census on individual manufacturing plants (the Longitudinal Research Database (LRD)).

The new data show that the 1990–91 recession was much less severe in manufacturing than preceding recessions, as evidenced by a relatively modest decline in employment. Nevertheless, job destruction and job reallocation both increased in a manner similar to that in previous recessions. The ensuing expansion was unusual in that job destruction and reallocation remained above average, rather than declining quickly after the recession. In addition, job creation experienced two large surges that were *not* preceded by surges in job destruction, as creation surges typically are. The authors interpret these surges as evidence of favorable allocative shocks, in contrast to the unfavorable allocative shocks of the 1970s and 1980s.

Regarding the nature of job creation and destruction, Schuh and Triest take a deeper look at two areas: (1) the magnitude, permanence, concentration, and cyclicality of job flows; and (2) the differences in job flows between larger, older, and higher-wage plants (henceforth, simply "large") and smaller, younger, lower-wage plants (henceforth, simply "small"). Previous DHS research concluded that job flows are large, permanent, and concentrated in a minority of plants with large employment changes. Also, large plants account for most of the increases in job destruction and reallocation during recessions. Together these DHS findings suggest that during recessions only a small fraction of really large plants experience really large and permanent rates of job destruction, and thus they imply that the cause of job destruction and recessions is related to large plants.

The Schuh and Triest findings significantly refine this DHS view. They find that small plants tend to have much higher rates of job creation and destruction than large plants, and that high rates of job creation and destruction—especially plant start-ups and shutdowns—are much more likely to be permanent. Thus, even though large plants account for most of the increase in job destruction during recessions, these large-plant job destruction rates are likely to be much smaller in percentage terms and less permanent. In fact, Schuh and Triest find that almost one-half of all jobs destroyed by plants experiencing relatively mild contractions are ultimately restored within five years. In other words, all plants are adversely affected by recessions but large plants appear to be more resilient than small plants, which expand and contract more dramatically and permanently.

Finally, Schuh and Triest uncover some evidence that suggests allocative factors cause business cycles. Their evidence is based on the premise that there are observable determinants of the allocation of jobs across firms, industries, and regions—prices, productivity, and investment—and that changes in those determinants cause job reallocation to increase, which in turn causes recessions. One key finding is that when relative prices and productivity growth across detailed industries change dramatically, job destruction and job reallocation also increase dramatically shortly afterward. Another key finding is that increases in job reallocation generally are *not* associated with increases in trend productivity and investment growth, as some recent theoretical models seem to imply.

Ricardo Caballero regards some of the Schuh-Triest results as "potentially promising," but he challenges two fundamental tenets. He questions the central premise that job reallocation is countercyclical, and he doubts that reallocation shocks actually cause fluctuations. In addition, he objects to the author's characterization and testing of theories of job reallocation.

Caballero contends that the term "job reallocation" is a misnomer. He does not dispute the fact that Schuh and Triest's measure of job reallocation is countercyclical. However, he argues that the main feature of job reallocation over time is a significant fluctuation in total destruction that is unconnected with the process of total job creation. Thus while individual jobs are destroyed and created at the plant level, thereby generating worker reallocation, it is what he calls a "dynamic fallacy of composition" to infer that a link exists between total job destruction and creation that could be characterized as total job "reallocation." Put another way, job "reallocation" would be higher if job destruction rose now and fell later while job creation stayed constant, but it would not be true in this case that job losers were reallocated to new jobs.

Caballero cites evidence from his own research that the surge in total job destruction during recessions is more than offset by a decline in destruction during the subsequent expansion. He calls this latter effect "chill," where job destruction falls below the rate associated with the "normal" underlying level of job turnover in the economy. He argues that it is important to understand that this chill can arise from market imperfections and produce technological sclerosis as a result of insufficient turnover. This argument contrasts with theories earlier this century that suggested that all job turnover is healthy for the economy.

Caballero believes "it is a large leap to claim that reallocation shocks are a substantial *source* of business cycles, at least in the United States," although he thinks they might be important elsewhere such as Eastern Europe, for example. He argues that plausible statistical models show that reallocation shocks are "substantially" less important than aggregate shocks, at least for net employment growth. He also demonstrates that such models can produce confusion about the relative importance of job reallocation, and asks whether the "fragile decomposition" of shocks as aggregate versus allocative is worthwhile, compared to focusing on observable shocks such as prices or interest rates.

In general, Caballero thinks it is a mistake at this point to focus on trying to discover whether or not reallocation shocks cause business cycles. Instead, effort should be directed toward the less debatable issue of whether "the churn [ongoing processes of creation and destruction] has a significant effect on the economy at *business cycle* frequencies."

Steven Davis shares the ambition of Schuh and Triest to develop new evidence on the connection between job reallocation and the business cycle. Indeed, he devotes a significant portion of his comments to explaining why this endeavor is important. But Davis, too, challenges the claim that reallocation activity is countercyclical, and he argues further that total job reallocation is inappropriate for this analysis. He also suggests a more effective methodology for summarizing the relationship between job flows and plant characteristics.

Davis provides a detailed description of the dynamic nature of job and worker flows and then advances several reasons why it is important to study these flows. First, "the extent to which the reallocation and matching process operates smoothly determines . . . the difference between successful and unsuccessful economic performance," with European unemployment serving as a prime example. Second, successful conduct of policy requires accounting for the reallocation and matching process. Third, recent modeling of reallocation frictions and heterogeneity makes it evident that aggregate shocks have allocative consequences, and shocks to factor demand can drive fluctuations in economic aggregates. Fourth, "models with reallocation frictions also help to address some well-recognized shortcomings in prevailing theories of the business cycle."

Davis believes that Schuh and Triest err in treating gross job reallocation "as equivalent to the intensity of reallocation activity." His criticism is that gross job reallocation does not account for the fact that movements in job creation and destruction merely may be achieving changes in total employment instead of reflecting a fundamental reallocation of labor across plants. Davis argues that the amount of job reallocation in excess of the change in total employment is a more suitable measure of reallocation intensity. He reports evidence that, unlike total job reallocation, excess job reallocation is uncorrelated with the business cycle.

POLICY IMPLICATIONS

In the closing session, leading economists from the public and private sectors discussed the implications for government policies of the conference's analysis of the causes of recessions. Panelists focused especially on the important role of vulnerability in setting the stage for unanticipated or adverse events. Each argued that governments should implement policies to reduce the economy's vulnerability and exposure to risk, provide more and accurate information to private agents about the extent of risk, and—if necessary—aid the recovery of economies that plunge into crises.

Henry Kaufman believes that sweeping structural changes to financial markets in recent years have significantly altered the linkages between financial markets and the real economy. Among the developments he identifies are securitization, derivatives, globalization, and leveraged investing. Several themes pervade his analysis. First, global financial markets are becoming increasingly sophisticated and complete. Second, this maturation process increasingly makes financing available to borrowers who would not have been able to obtain it previously. Third, and a consequence of the first two points, financial markets are becoming increasingly volatile, as risk-taking becomes easier while accurate risk assessment becomes more difficult. Altogether, these changes increase the likelihood that financial market turbulence will make economies more vulnerable to shocks and recessions.

Kaufman believes the changes increase the difficulty and reduce the efficacy of monetary policy. Monetary policy is more difficult because traditional monetary factors—monetary aggregates, debt aggregates, and the like—have become less reliable indicators of the stance of monetary policy and the state of money markets. Monetary policy is less effective because increased availability and easier acquisition of credit mean that short-term interest rates must increase more to achieve the same real response. Furthermore, increased volatility in asset prices (wealth) leads to greater volatility in aggregate economic behavior. Thus, he argues, the Federal Reserve should take asset price developments explicitly into account in formulating monetary policy.

Internationally, Kaufman sees a need for increased supervision of financial markets. Paradoxically, he notes, when financial markets become deregulated and "freewheeling," the need for more accurate, timely, and complete information increases, particularly about the risks in which financial entities are engaging. He decries the poor job of oversight and information gathering done by official institutions thus far and proposes several reforms. In particular, he recommends a new body he calls a Board of Overseers of Major Institutions and Markets, which would set a code of conduct, supervise risk-taking, and harmonize capital requirements.

Kaufman also favors reforms to two international economic organizations. First, the IMF should be reorganized to specialize in a narrower set of core functions. The new IMF would continue to facilitate lending to countries in financial distress and to press for reform in government policies in these countries. But it would also be charged with rating the credit-worthiness of countries, by assessing economic and financial conditions, reviewing extant government policies, and demanding remedial action where needed. Kaufman also argues that the G-7 must be restructured to account for the European Monetary Union and its euro currency.

Martin Zimmerman provides perspective from one of the largest and most cyclical components of the U.S. economy: the automobile industry. He explains how the auto industry, specifically Ford Motor Company, views the unfolding of a recession—how consumers postpone their car purchases, how auto makers respond to weakening sales, and how interest rate policy is an important determinant of the economic fortunes of the auto industry. But ultimately he argues against the central theme of the conference. That is, Zimmerman believes it is impossible to go "Beyond Shocks."

The economy is always subject to shocks, according to Zimmerman. For the auto industry, a shock is anything that causes consumers to suddenly alter their normal plans to purchase new cars. Zimmerman tells the story of how the 1990–91 recession unfolded. As late as June 1990, economic forecasters were predicting confidently that there would be no recession, only a slowdown. But Iraq's invasion of Kuwait and the U.S. military response caused a precipitous drop in consumer confidence and sales of cars to consumers. The shock of the Kuwait invasion, like all shocks, by definition was not forecastable, says Zimmerman (an assessment that was not well-received by his employers, he adds wryly).

Although shocks are pervasive, the central question is whether the shocks will tip the economy over into recession. Here, he asserts that not all shocks do, in fact, trigger recessions. The economy must already be vulnerable when the shocks hit. Absent this vulnerability, the economy may be able to withstand shocks. Likewise, absent shocks, vulnerability may never result in a recession.

What is the role of policy in a world of vulnerability and inevitable shocks? Zimmerman notes that every precipitous drop in auto sales has been associated with an increase in interest rates, so he tends to associate monetary tightening with the emergence of economic vulnerability (weak growth). But because not every increase in interest rates was followed by a recession, he surmises that a shock is required to turn vulnerability into recession. He asserts that monetary policy cannot prevent shocks because they are inherently unpredictable. Instead, policy should minimize vulnerability of the economy.

Agustin Carstens contributes a view of recessions and policy from the perspective of emerging economies such as Mexico. He identifies five characteristics of business cycles in emerging economies that distinguish them from business cycles in industrialized economies. First, business cycles in emerging countries are closely synchronized with the fortunes of industrialized countries: "When the United States gets a cold, Mexico gets pneumonia." Second, business cycles are more volatile in emerging economies. Third, emerging economies are susceptible to additional sources of volatility, such as terms of trade fluctuations. Fourth, and more recently, increasing globalization of markets has encouraged massive capital flows into emerging countries like Mexico. But these capital flows are very unstable, so emerging countries can experience sudden and massive capital outflows that devastate their economies. Finally, emerging economies have to deal with exchange-rate regimes and their failures.

These characteristics force emerging economies to adopt very different policies to deal with business cycles. Industrialized countries, as leaders of the world economic engine, follow policies designed to manage aggregate demand so as to achieve low inflation and full employment. Such policies are countercyclical. In contrast, emerging countries follow policies designed to avoid or mitigate economic crises that break out there, often because industrialized countries are slumping and reducing their demand for emerging country exports. One essential goal of these policies is to reestablish the credibility of emerging economies, especially the credibility of their currencies and financial markets. Often this means reestablishing the credibility of governments that have made bad policy decisions. These types of policies, then, are usually procyclical.

Carstens offers four specific policy recommendations for emerging economies to help them to reduce vulnerability and follow a more stable path. First, they must reduce their vulnerability to changes in the international prices of exports, by adopting more open trade and investment regimes. Second, they should allow market determination of interest and exchange rates so these rates can accomplish their purpose of absorbing shocks. Third, they must ensure the robustness of their financial institutions to macroeconomic fluctuations. Fourth, they should push forward with structural changes in order to achieve central bank autonomy, privatization of production, labor market flexibility, and reduced dependence on foreign saving. In each case, more complex policies are required beyond the traditional demand management schemes followed by industrialized countries, Carstens notes.

Michael Mussa, as a leading official at the International Monetary Fund, offered an informed, practical—and oftentimes contrarian—view of the conference papers, the conventional wisdom about the ongoing global economic crises, and recent criticisms of international policy responses to the crises.

Mussa infers from Sims's paper that systematic monetary policy *does* have a significant, positive effect on the real economy, despite Sims's claim to the contrary. He says Sims understates the effect of monetary policy, citing Sims's own results showing that industrial output would have been nearly one-fifth higher if the Fed had followed modern monetary policies during the Great Depression. He also points out that Sims omits the positive role monetary policy can play in avoiding banking and financial panics by subsidizing and reforming weak banks, and by reassuring depositors that their accounts were safe. Had Sims accounted for this, and for the fact that fiscal policy should have been more aggressive, one-half to three-quarters of the impact of the Great Depression could have been avoided.

Mussa finds the two long historical analyses of business cycles to be inherently valuable. He particularly agrees with Temin's premise that recessions "have a multiplicity of causes," although he doubts that it is possible—or useful—to try to quantitatively separate causes into different categories of influence. Like Romer, Mussa believes that Temin underestimates the contribution of monetary policy to recessions. However, Mussa is cautious about the quality of older economic data and what we can reliably infer from them, particularly data for countries other than the United States.

Regarding the paper by Schuh and Triest on labor reallocation and business cycles, Mussa is "skeptical that labor reallocation is itself an independent cause of most U.S. business cycles." He suggests that the authors focus more on the relationship between labor reallocation and the NAIRU (non-accelerating-inflation rate of unemployment). Regarding the central issue addressed in Basu's paper, Mussa believes that "the notion that adverse downward movements in total technology cause recessions [because workers don't work as hard] is just plain silly. This is the theory according to which the 1930s should be known not as the Great Depression but as the Great Vacation."

Mussa then turned to a discussion of current economic developments and the appropriateness of policy. On the domestic economy, Mussa likens recent monetary policy performance to the movie, "As Good As It Gets." Aside from some minor quibbles, Mussa judges U.S. monetary policy management during the last decade to be "remarkable" by any standard. But he notes that it has been "very good management with very good luck." Moreover, he warns, to say that monetary policy has been as good as it gets implies that monetary policy is better than it is normally expected to be—in other words, it is likely to get worse, not better. Ultimately, the monetary authority cannot avoid all recessions; it can only be expected to avoid "big" ones.

On the international situation, Mussa likens catastrophic economic events such as the Great Depression and the current worldwide financial crisis to the movie "Titanic." What caused the Titanic to sink, he asks? Perhaps an exogenous shock (the iceberg), he quips. But it was more than that. Errors in the design and operation of the ship, inadequate preparation for the sinking, and other factors all contributed. In the same way, the current financial crisis has many complex causes and contributing factors.

However, reasons Mussa, the *real* tragedy of the Titanic was not that it sank and 1,500 lives were lost, but that *800 of the Titanic passengers were saved that day!* Clearly this policy mistake discouraged shipbuilders from spending money on improving designs and shipping lines from bearing the cost of conducting safe navigation of future cruises across the Atlantic. The Titanic rescue demonstrated that entrepreneurs in the shipping industry didn't need to worry about safety—they knew that the government would be there to save them from their imprudence!

Mussa employs this tongue-in-cheek argumentation to rebut those who argue that moral hazard problems should prevent the international community from responding to the current financial crisis. Despite moral hazard problems, saving 800 Titantic passengers *was* the right thing to do. And despite clear moral hazard problems, Mussa says the IMF attempts to rescue Korea and other besieged economies *is* the right thing to do. He argues that IMF support is not a gift but a loan, and that the IMF's earlier financial support of Mexico has been validated by Mexico's successful servicing of IMF debt.

CONCLUSION

In the end, most participants agreed that the business cycle is *not* dead but is likely here to stay. No one championed the ideas that a "new," recession-proof economy has emerged, that unanticipated adverse economic events have stopped buffeting the economy, or that government policy has become so adroit that it can offset every dip in the aggregate economy. If anything, the mere mention of these ideas drew disdainful remarks, and even served as "proof" that the ideas were without merit. Indeed, the general premise among participants was that the right question was *when,* not *if,* the next recession occurs, what will have caused it? The consensus answer is it is likely to be not one but many things, with government policy and vulnerability playing important—but still not fully understood—roles.

Most participants also agreed that policymakers in a world continually subject to business cycles should adopt certain goals to improve their ability to deal with fluctuations. First, policymakers must learn how to recognize and address the economy's vulnerability to disruptions and unanticipated events. Second, policy institutions should conduct and support research that shows the contribution of deliberate actions of economic agents to economic fluctuations. Finally, and most important, policymakers should understand that they cannot prevent every recession, but they should concentrate their efforts on averting The Big Ones, such as the Great Depression.

2

The Business Cycle
It's Still a Puzzle

Lawrence J. Christiano
Professor of Economics, Northwestern University, Evanston, Illinois

Terry J. Fitzgerald
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

INTRODUCTION AND SUMMARY

Good fiscal and monetary policy requires a clear understanding of the workings of the economy, especially what drives the business cycle—the periodic ups and downs in economic activity. Since at least the late 1800s, a full swing from the start of an economic expansion to a recession and back to the start of another expansion has generally taken between two and eight years. Every citizen is keenly aware of the state of the economy, whether it is in prosperity or recession.

Everyone is so conscious of the business cycle because most sectors of the economy move up and down together.[1] This phenomenon, referred to as *comovement,* is a central part of the official definition of the business cycle. The definition is set by the National Bureau of Economic Research (NBER), which decides when recessions begin and end. Under the NBER's definition, ". . . a recession is a [persistent] period of decline in total output, income, employment, and trade, usually lasting from six months to a year, and *marked by widespread contractions in many sectors of the economy.*"[2]

Even though comovement is a defining characteristic of the business cycle, in recent decades macroeconomists have tended to focus on understanding the persistence in the ups and downs of aggregate economic activity. They have generally been less concerned with understanding the synchronized nature of this pattern across sectors. In part, the omission reflects the conceptual difficulties inherent in thinking about an economy with many sectors.[3] Standard models of business cycles assume there is only one good being produced and so they consider only one economic sector. These models do not encourage thinking about the comovement of economic activity across many sectors. Since these models were first introduced, in the late 1970s and early 1980s, the state of macroeconomics has advanced rapidly. The conceptual and computational barriers to thinking about multiple sec-

Reprinted from: *Federal Reserve Bank of Chicago Economic Perspectives* (Fourth Quarter 1998) 56–83.

tors are quickly falling away. As a result, recent years have witnessed a renewed interest in understanding comovement.

We have two objectives in this article. The first is to document business cycle comovement. We examine data on hours worked in a cross section of economic sectors. We examine the business cycle components of these data and show that the degree of comovement is substantial. Our second objective is to analyze explanations for this comovement. We find that none are completely satisfactory. Still, this is a growing area of research, and we are seeing some progress.

IDENTIFYING COMOVEMENT

To study comovement across sectors over the business cycle, we need the following two things: a measure of the level of economic activity in the various sectors of the economy; and a precise definition of what we mean by the business cycle component of the data. Below, we address these two issues. After that, we present our results, characterizing the degree of comovement in the data.

The Data

We measure economic activity in a given sector by the number of hours worked in that sector. Table 1 lists the sectors we consider. The hours worked measure that covers the most sectors is *total private hours worked.*[4] This measure covers all sectors of the economy, except government and agriculture. It is broken into hours worked in goods-producing industries and in service-producing industries. Goods-producing industries are further broken into mining, manufacturing, and construction. Similarly, service-producing industries are broken into five subsectors. The subsectors of manufacturing, durable goods and nondurable goods, are broken into yet smaller sectors. The data in the third column give an indication of the relative magnitude of each subsector. In particular, any given row reports the average number of people employed in that sector, divided by the average number of people employed in the sectoral aggregate to which that sector belongs. Thus, for example, 58 percent of manufacturing employment is in the durable goods sector and 42 percent is in the nondurable goods sector. Also, the largest goods-producing industry, by far, is manufacturing, which has 80 percent of all goods-producing employees.

Next, we try to characterize how much business cycle comovement there is across the economic sectors we consider. That is, if we limit ourselves to the business cycle range of fluctuations in the data—fluctuations that last between two and eight years—to what extent do the data move up and down together?[5]

Business Cycle Component of the Data

A detailed discussion of our notion of the business cycle component of the data is in technical appendix 1. Figure 1 illustrates the basic idea behind what we do. The choppy line in panel A of Figure 1 displays total private hours worked. The reported data are the logarithm of the raw data. The advantage of using the logarithm of the data in this way is that the resulting movements correspond to percent changes in the underlying raw data. The deviations between the actual data and the trend line in panel A of Figure 1 are graphed in panel B. Those deviations contain the rapidly varying, erratic component, inherited from the choppy portion of the data that is evident in panel A. The smooth curve in panel B is our

Table 1 Properties of the Business Cycle Components of Hours Worked

Variable number	Hours worked variable	Relative magnitude	Relative volatility	Business cycle comovement
1	Total private	1.00	1.00	.00
2	Goods-producing industries	.33	3.91	.99
3	Mining	.03	5.46	.38
4	Construction	.17	6.75	.88
5	Manufacturing	.80	3.92	.97
6	Durable goods	.58	6.90	.97
7	Lumber and wood products	.06	10.18	.89
8	Furniture and fixtures	.04	8.14	.94
9	Stone, clay, and glass products	.05	4.98	.95
10	Primary metal industries	.09	9.89	.86
11	Fabricated metal products	.13	7.21	.96
12	Machinery, except electrical	.19	11.10	.93
13	Electrical and electronic equipment	.15	8.75	.88
14	Transportation equipment	.17	7.83	.89
15	Instruments and related products	.08	5.03	.76
16	Miscellaneous manufacturing	.04	3.23	.90
17	Nondurable goods	.42	1.39	.91
18	Food and kindred products	.21	.16	.50
19	Tobacco manufactures	.01	1.83	.08
20	Textile mill products	.11	3.92	.76
21	Apparel and other textile products	.15	2.64	.85
22	Paper and allied products	.09	1.97	.85
23	Printing and publishing	.16	.91	.90
24	Chemicals and allied products	.13	1.01	.80
25	Petroleum and coal products	.02	2.02	.16
26	Rubber and misc. plastics products	.09	7.82	.89
27	Leather and leather products	.03	2.71	.64
28	Service-producing Industries	.67	.25	.93
29	Transportation and public utilities	.10	.87	.95
30	Wholesale trade	.10	.65	.87
31	Retail trade	.31	.36	.87
32	Finance, insurance, and real estate	.10	.35	.48
33	Services	.38	.19	.49

Notes: The column labeled "Relative magnitude" reports an indication of the relative magnitude of each sector. Any given row reports the average number of people employed in that sector divided by the average number of people employed in the sectoral aggregate to which that sector belongs, for example, 58 percent of manufacturing employment is in the durable goods sector and 42 percent is in the nondurable goods sector. The columnlabeled "Relative volatility" reports the variance of the business cycle component of the logarithm of hours worked in the indicated row variable divided by the variance of the business cycle component of the logarithm of total private hours worked. The column labeled "Business cycle comovement" is calculated using the process described in note 6 of the article.
Source: Authors' calculations from data of DRI Basic Economics database. 1964–86.

measure of the business cycle component of the total private hours worked data. Specifically, that measure excludes both the trend part of the data and the rapidly varying, erratic component. It includes only the component of the data that contains fluctuations in the range of two to eight years. According to our approach, the economy is in recession when our business cycle measure is negative and in prosperity when it is positive.

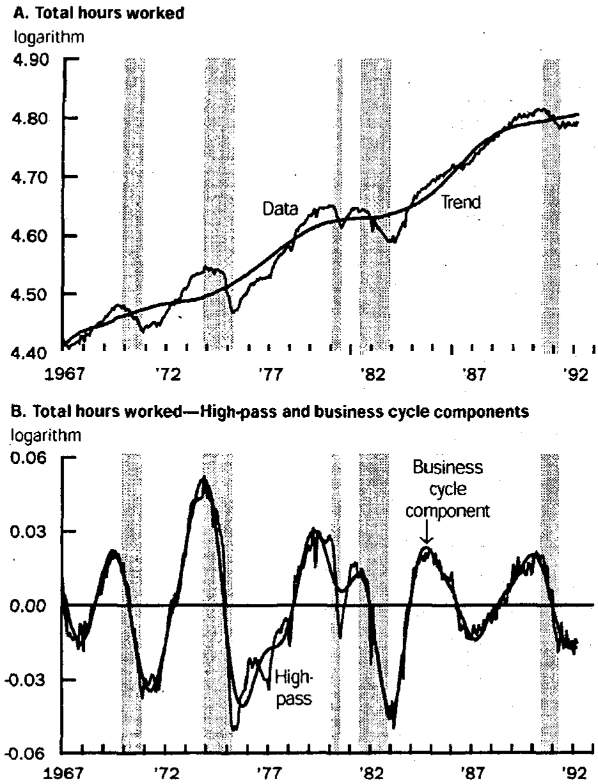


Figure 1 Total hours worked and its trend. Note: Shaded areas indicate recessions as determined by the National Bureau of Economic Research. Source: Authors' calculations from data of DRI Basic Economics database, 1964–96.

Figure 1 also compares our measure of the business cycle with the one produced by the NBER. The start of each shaded area indicates the date when, according to the NBER, the economy reached a business cycle peak. The end of each shaded area indicates a business cycle trough. Note how total private hours worked fall from peak to trough and then generally grow from trough to peak. An obvious difference in the two business cycle measures is that ours is a continuous variable, while the NBER's takes the form of peak and trough dates. As a result, our measure not only indicates when a recession occurs, but also the intensity of the recession. Apart from these differences, however, the two measures appear reasonably consistent. For example, note that near the trough of every NBER recession, our measure of the business cycle is always negative. But the two measures do not always agree. According to our measure, the economy was in recession in 1967 and in 1987, while the NBER did not declare a recession during those periods. In part, this is because there must be several months' negative employment growth before the NBER declares a recession. However, our procedure only requires a temporary slowdown.

Figure 1 provides informal evidence in support of the facts we wish to document. As noted in the introduction, the NBER must see a broad-based decline before declaring a recession. Thus, the NBER dates in figure 1 indicate periods when many economic sectors showed weakness. Since these dates roughly coincide with periods of weakness in total private hours worked, this is consistent with the view that most sectors move up and down together, at least in the two- to eight-year frequency range. We stress, however, that the NBER's dating procedures are informal. Our objective in this section is to provide a formal, quantitative assessment of the degree of comovement among economic sectors.

We computed a business cycle component for each of the 33 series listed in Table 1. As we anticipated, we find that the business cycle components in most of the series move together closely. This is true, despite a striking lack of uniformity in other respects. For example, note how different the trends in Figure 2 are. The first two columns report data for the goods-producing industries and its major components. The second two columns report the analogous data for the service-producing industries. Generally, trend employment is down in the goods-producing industries and up in the service-producing industries. The levels of volatility in the business cycle components of the various series are also very different. The fourth column of Table 1 reports the variance of the business cycle component of a variable, divided by the variance of aggregate hours worked. The relative variance of hours worked in goods-producing industries is typically quite high, substantially above 2, and it is quite low for service-producing industries. That goods-producing industries are volatile relative to the service-producing industries is well known.

MEASURING BUSINESS CYCLE COMOVEMENT

Despite the very substantial differences in the trends of the data series shown in Figure 2, their movements over the business cycle are quite similar. Figure 3 illustrates the business cycle components of the same variables used in Figure 2. In each case, we computed the business cycle component using exactly the same method underlying the calculations in panel B of Figure 1. Each graph contains the business cycle component of the variable indicated and the business cycle component for total private hours. This was taken directly from panel B of Figure 1.

In most of the series in Figure 3, the data move up and down closely with the business cycle component of total hours worked. There are some exceptions. For example, the business cycle movements in mining bear little resemblance to the business cycle movements in total hours worked. At the same time, mining represents a very small part of the private economy and employs only 3 percent of workers in the goods-producing industry. Another exception is the finance, insurance, and real estate (FIRE) industry, whose business cycle component exhibits reasonably high comovement with aggregate employment until the 1980s, after which this relationship breaks down.

To measure the degree of business cycle comovement between a given series and total hours worked, we use a statistic that is like the square of the correlation between the business cycle components in the two variables. Our statistic measures the fraction of the variance in the series that can be accounted for by the total hours worked data.[6] If this number is, say, 98 percent, this means that 98 percent of the business cycle variance in the variable can be accounted for by the business cycle in aggregate hours worked. These results are reported in the fifth column of Table 1. As expected, the results indicate that this measure of comovement is relatively low, in the sense of being below 0.50, for the mining, FIRE, and services sectors. Overall, however, the degree of comovement by this measure is high.

Going one step further in the level of disaggregation, we can get an idea about the comovement in the components of durable and nondurable manufacturing. Figure 4, panel A

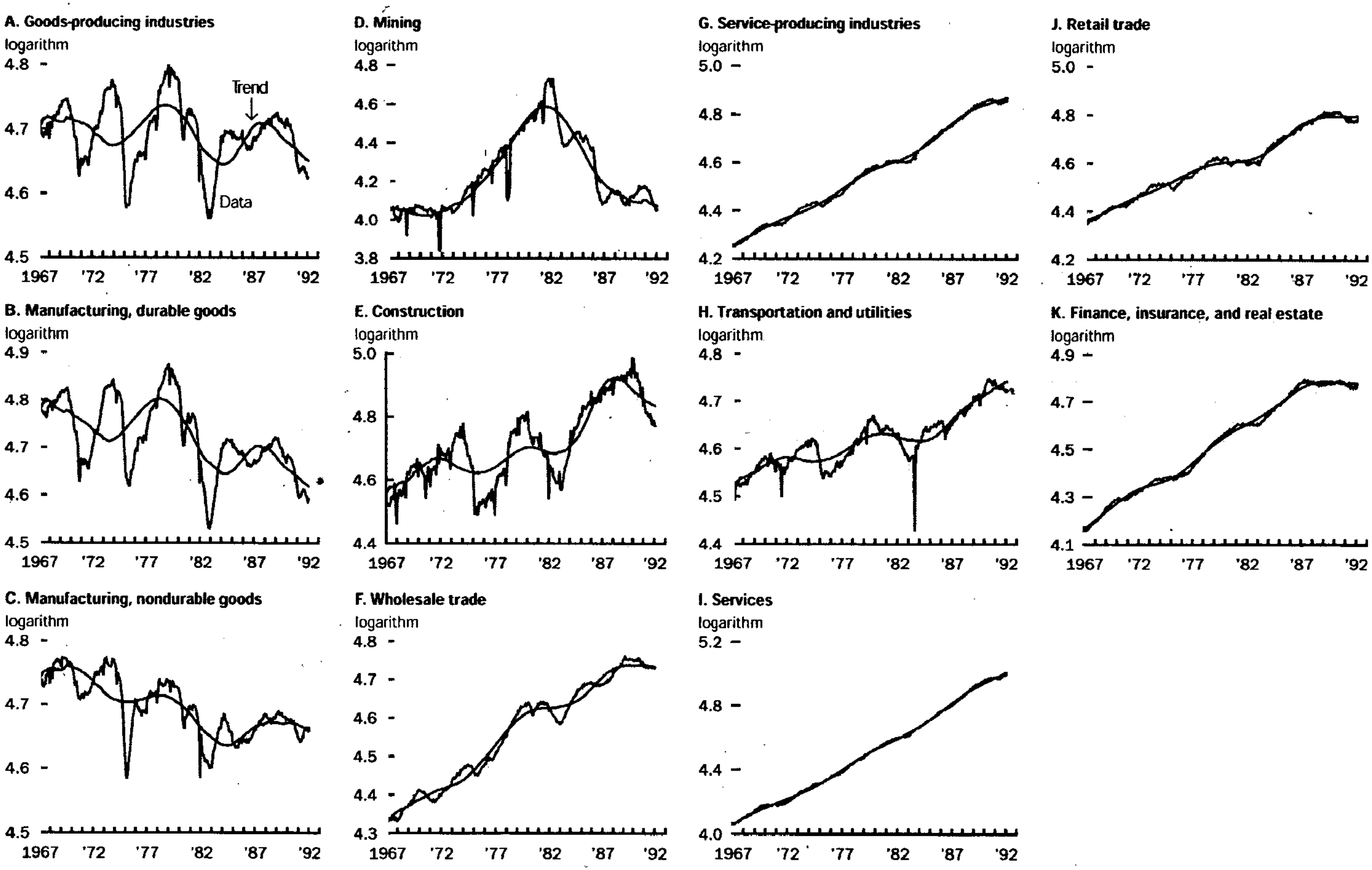


Figure 2 Hours worked in various sectors: Data and trends. Source: Authors' calculations from data of DRI Basic Economics database, 1964–96.

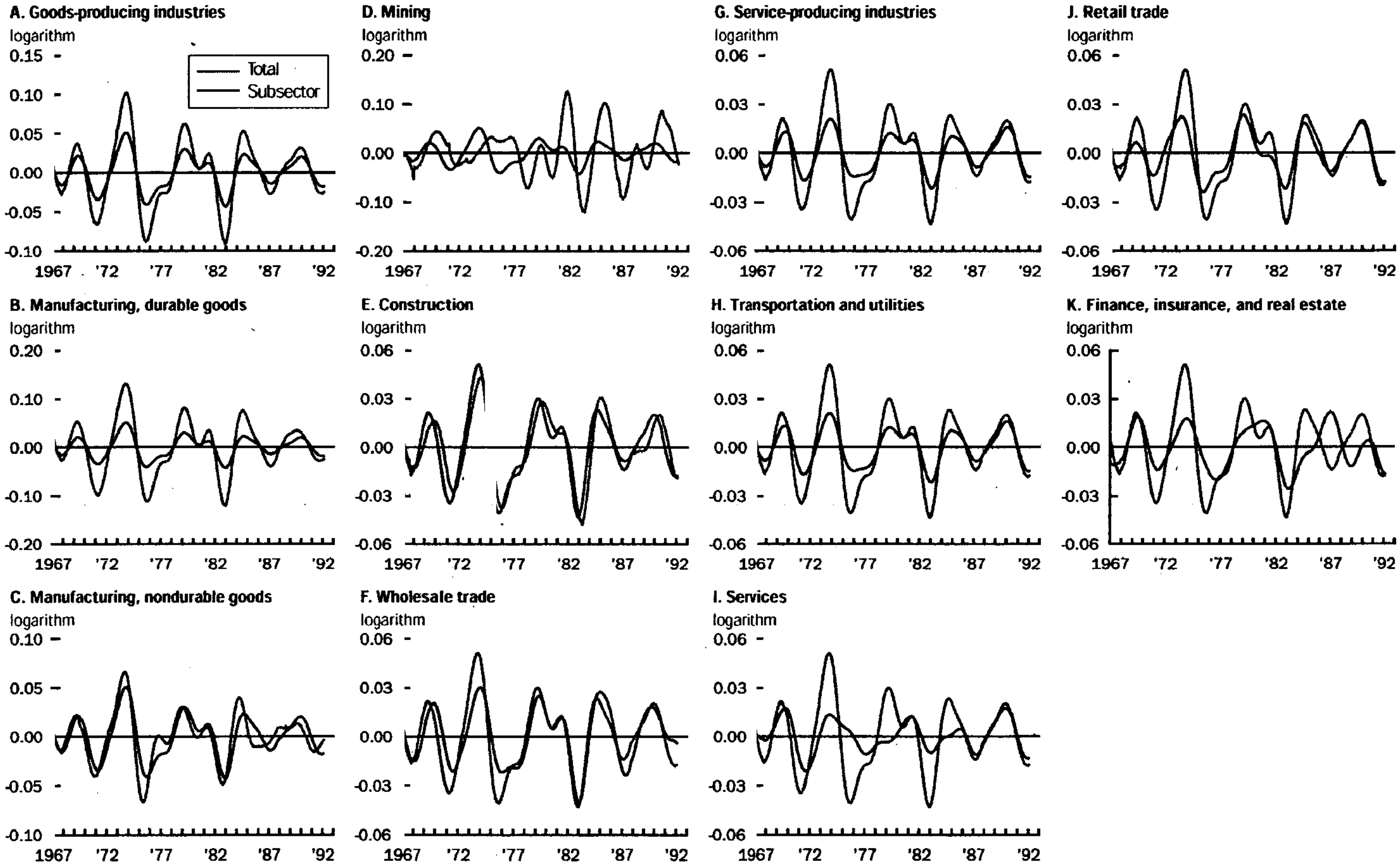


Figure 3 Business cycle component comparison: Total hours worked versus hours worked in various sectors. Note: The information displayed in the "total hours worked" line is "business cycle component," taken from Figure 1, panel B. Source: Authors' calculations from data of DRI Basic Economics database, 1964–96.

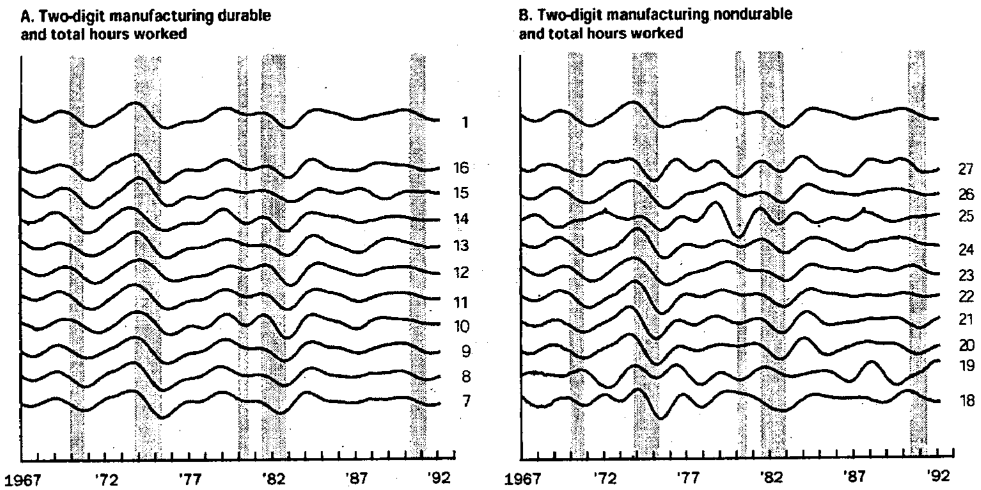


Figure 4 Business cycle component of total hours worked and hours worked in various manufacturing subsectors. Notes: The number on the right indicates the variable sector from table 1 that is being highlighted. Panel A shows: 1—total private hours; 7—lumber and wood products; 8—furniture and fixtures; 9—stone, clay, and glass products; 10—primary metal industries; 11—fabricated metal products; 12—machinery, except electrical; 13—electrical and electronic equipment; 14—transportation equipment; 15—instruments and related products; and 16—miscellaneous manufacturing. Panel B shows: 1—total private hours; 18—food and kindred products; 19—tobacco manufactures; 20—textile mill products; 21—apparel and other textile products; 22—paper and allied products; 23—printing and publishing; 24—chemicals and allied products; 25—petroleum and coal products; 26—rubber and miscellaneous plastics products; and 27—leather and leather products. Each variable has been scaled by its sample standard deviation and a constant has been added in order to spread out the data in the panels. Shaded areas indicate recessions as determined by the National Bureau of Economic Research. Source: Authors' calculations from data of DRI Basic Economics database, 1964–96.

displays the business cycle movements in the components of durable manufacturing sectors. Panel B does the same for nondurable manufacturing. In each case, the data series graphed at the top of the figure is the business cycle component of total hours worked. The series are presented so as to allow one to focus exclusively on the degree of comovement between them. Thus, we added a constant to each series to spread them out across the figure and divided each series by its sample standard deviation, so that the standard deviation of the reported data is unity in each case.[7] The number to the right of each line in the figure identifies the data series. Figure 4 also displays the NBER peak and trough dates as a convenient benchmark.

Figure 4, panel A shows that the comovement among sectors in durable manufacturing is very high. With only one minor exception, the variables move closely with each other and with aggregate employment. The exception is that instruments and related products, series 15, does not move closely with the other variables during 1987, when the other business cycle components are signaling a recession. However, overall the degree of comovement is strikingly high. Figure 4, panel B shows that the business cycle comovement in the nondurable goods manufacturing industries is lower than in the durable goods sector. Two variables that do not comove closely with the others at business cycle frequencies are tobacco manufactures, series 19, and petroleum and coal products, series 25. Both these variables are rising in the first and last NBER recession periods in our data. The comovement statistic for these variables reported in Table 1 is very low, 0.08 for tobacco manufactures and 0.16 for petroleum and coal products. The other variables in nondurable manufacturing display stronger comovement, with comovement statistics of 0.50 or higher.

Up to now, the statistics we have used to characterize comovement emphasize association with aggregate hours worked. This nicely complements the visual evidence in the graphs. However, there is a pitfall to relying exclusively on statistics like this to characterize comovement. A simple example illustrates the point. Suppose there is a variable, y_t, which is the sum of two other variables, y_{1t} and y_{2t}:

$$y_t = y_{1t} + y_{2t}.$$

Suppose further that y_{1t} and y_{2t} are uncorrelated. No one would say there is comovement between these variables. Still, each variable is strongly correlated with the aggregate. To see this, take the simple case where the variance of y_{1t} and y_{2t} is σ^2. Then, the correlation between y_{it} and y_t is 0.71, for $i = 1,2$, despite the absence of comovement between the variables.[8] This example exaggerates the point somewhat, since results are less severe when there are more than two sectors.[9] Still, this pitfall is of some concern.

With this in mind, we consider the correlation between the business cycle components of all the variables. A difficulty with this is that there are many such correlations. For example, with three variables, there are three possible correlations, with four there are six, with five there are ten, and with n there are $n(n - 1)/2$. So, with $n = 33$, there are 528 possible correlations. It is a challenge to organize and present this many correlations in a coherent way. We present the mean and histogram of the correlations for different subsectors in Figure 5. The histograms display, on the vertical axis, the fraction of correlations lying within a given interval, whose midpoint is indicated on the horizontal axis.[10]

Figure 5, panel A displays the correlations for the finest levels of aggregation for which we have data. This means hours worked in mining, construction, the 20 components of manufacturing, and the five components of the service-producing industries. Thus, we have 27 data series, with 351 correlations between them. Figure 5, panel A indicates that the mean of these correlations is 0.55. When we eliminate the three data series that we already know do not display strong business cycle comovement, the mean rises to 0.68. The histogram shows that there is a substantial fraction of high correlations in these data. We infer that the data are consistent with the impression from the preceding statistics that there is considerable evidence of comovement. Figure 5, panel B presents the results for the manufacturing durable sector. Consistent with our previous findings, the degree of comovement in this sector is very high, with a mean correlation of 0.82. Figure 5, panels C and D show the results for the nondurable manufacturing sector and the service-producing sectors, respectively. Again, the results are consistent with the notion that there is less comovement in these sectors than in manufacturing durables. Still, the degree of comovement is substantial, with mean correlations in excess of 0.6 if we consider all sectors except tobacco and petroleum and coal products.

In conclusion, we find that there is substantial business cycle comovement in the data. Only two relatively small sectors—tobacco manufactures and petroleum and coal products—exhibit little tendency to move up and down with general business conditions over the business cycle.

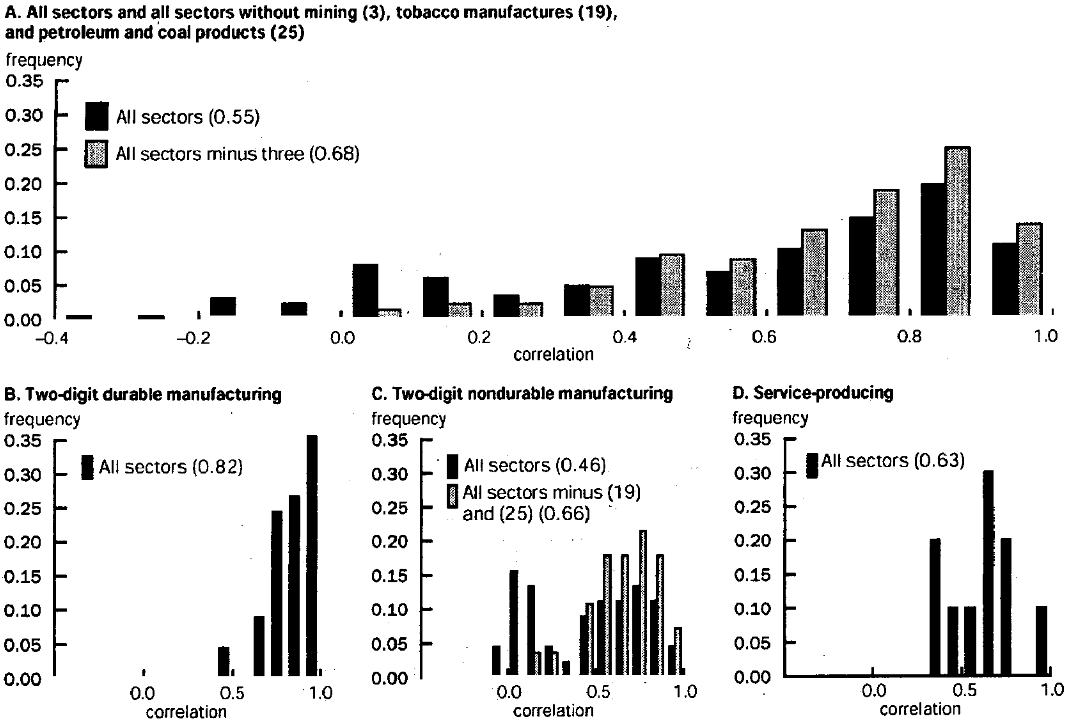


Figure 5 Distribution of correlations between business cycle components of hours worked in various sectors. Notes: Whole numbers in parentheses represent the variable sector designated in table 1. Decimals in parentheses represent mean over indicated set of correlations. Source: Authors' calculations from data of DRI Basic Economics database, 1964–96.

EXPLAINING BUSINESS CYCLE COMOVEMENT

What is it that at times pulls most sectors of the economy up, and at other times pushes them down again? This is one of the central questions in business cycle analysis. Although economists have developed a number of possible explanations, the phenomenon remains a puzzle.

In a classic article devoted to this puzzle, Robert E. Lucas, Jr., conjectures that the resolution must lie in some sort of shock that hits all sectors of the economy, a so-called aggregate shock (Lucas, 1981). Many economists today would probably agree with this conjecture. That is why, in practice, the search for the ultimate cause of business cycles often focuses on identifying an aggregate shock. However, research conducted since Lucas published his article suggests that identifying the *cause* of business cycles may not be so simple.

First, even if we do manage to identify a shock that clearly affects the whole economy, it does not necessarily follow that shock is responsible for the business cycle. A shock might well be experienced by all sectors of the economy, but they need not all *respond* in the same way. The business cycle shock, if indeed there is only one, seems to lead to a synchronized response across sectors. Second, we now know that the search for a single aggregate shock may itself be off base. Following the work of Long and Plosser

(1983), we know that, at least in theory, disturbances to individual industries, even if they are uncorrelated across industries, could result in comovement.

Currently, there is no consensus among economists as to what causes business cycles and, in particular, their key feature, comovement. At the same time, researchers are exploring a large range of possibilities. Next, we provide a selective overview of this research.

A natural starting point is what is perhaps the most thoroughly developed theory of business cycles, the *real business cycle* theory associated with Kydland and Prescott (1982), Long and Plosser (1983), and Prescott (1986).[11] We focus specifically on the *standard real business cycle model,* developed in Hansen (1985) and analyzed in Prescott (1986). Although that model posits an aggregate shock, it is inconsistent with business cycle comovement. We then explore two sets of modifications to this model. The first can be viewed as natural extensions of the model. The second departs more significantly from the model's assumptions.

STANDARD REAL BUSINESS CYCLE THEORY[12]

A key component of real business cycle theory is a *production technology.* This is a relationship that specifies how much output a firm can obtain from a given amount of capital and labor resources. This technology is subject to shocks. Sometimes a good shock occurs and more output can be produced for a given level of inputs. In this case, we say the technology has been shifted up. A good technology shock might reflect the implementation of a more efficient way to organize the work force, the acquisition of more efficient manufacturing equipment, or perhaps the discovery of a way to alter the firm's product so that it better meets customers' needs. At other times, a bad technology shock can shift a production technology down. A bad shock might reflect bad weather, a labor dispute, an accident in the workplace, a machine breakdown, or a government policy that encourages an inefficient way of organizing production. According to real business cycle theory, business cycle expansions reflect that shocks affecting firms are mostly on the positive side, while recessions reflect periods when most firms' shocks are on the negative side. Standard formulations abstract from the differences between firms and simply assume they all have the same production technology and are affected by the same shock. Thus, real business cycle theory proposes that the aggregate shock to which Lucas refers is a productivity shock.[13]

The standard real business cycle model not only assumes that all firms are affected by the same productivity shock, but also that there is just one type of good produced (and, therefore, one industry sector) in the economy. At least at first glance, this model does not seem useful for examining business cycle comovement among many sectors. However, it has recently been pointed out that this impression is misleading.[14] In fact, one *can* use the model to examine business cycle comovement. When we do so, we find that its implications are strongly counterfactual. The standard real business cycle model is at variance with the observation of business cycle comovement, despite the fact that it views the economy as being driven by a single aggregate shock. Understanding why it is incompatible with comovement is useful for gaining insight into the various lines of inquiry researchers have pursued.

The standard real business cycle model imagines that households interact with firms in competitive markets, in which they supply labor and physical capital and demand goods for consumption and to add to their stock of capital. Although there is only one type of production technology in this model, we can reinterpret the model to suggest that one type of firm produces goods for consumption (the consumption goods industry) and another type produces new investment goods for maintaining or adding to the stock of capital (the investment goods industry).

When a positive productivity shock hits, so that the real business cycle model shifts into a boom, the output of both consumption and investment goods industries increases. However, there is a relatively larger increase in the output of investment goods. This reflects a combination of two features of the model. First, a positive technology shock increases the expected return to investment, raising the opportunity cost of applying resources to the consumption sector. Second, the model assumes that households prefer not to increase consumption substantially during booms but to smooth consumption increases over a longer time horizon. The increase in the demand for investment goods relative to consumption goods that occurs in a boom implies, in the standard model, that capital and labor resources are shifted out of the production of consumption goods and into the production of investment goods. The model does predict a small rise in consumption in a boom. However, this rise is driven by the favorable technology shock, which is not fully offset by the flow of productive resources out of that sector. Thus, the model implies that hours worked in the consumption sector are countercyclical, in contrast with our empirical findings in the previous section. This is a feature of the model, despite its implication that total hours worked rise in a boom. That is, the additional hours of work all flow into the investment good sector. The standard real business cycle model also implies that investment in capital for use in the consumption sector is countercyclical. This too, is counterfactual, according to the results reported in Huffman and Wynne (1998).

So, this model is strongly at variance with comovement. Why is this so? The result may seem especially surprising to those who expect an aggregate shock to all sectors of the economy to produce comovement.

Intuitively, there are two related ways to understand the model's implication that inputs are allocated away from the sector that produces consumption goods during a boom. One is that the model overstates the value of leisure at that time. This inflates the cost of allocating labor resources to the consumption sector then. The other is that the model understates the value of the output of the sector producing consumption goods in a boom. This undercuts the incentive to allocate resources to that sector then.[15]

NATURAL EXTENSIONS OF THE STANDARD THEORY

Among the various extensions to the model that economists have pursued,[16] we focus on approaches that stress (1) factors that prevent the rise in the marginal utility of leisure in a boom and (2) factors that prevent the decline in the value of the output of the consumption sector in a boom. As in the discussion above, the work we survey here assumes two market sectors.[17]

Value of Leisure

One factor that can slow the decline of the marginal utility of leisure when the economy moves into a boom was explored in Benhabib, Rogerson, and Wright (1991) and Einarsson and Marquis (1997). Each of these papers points out that if there is a third use of time, in addition to leisure and time spent working in the market, and if that use of time declines during a boom, the marginal utility of leisure need not increase as market effort increases.[18]

Benhabib, Rogerson, and Wright (1991) suggest that the third use of time may be working in the home. For example, the amount of leisure time enjoyed by a homemaker may not change significantly if the homemaking job is exchanged for a market job. Considerations like this lead Benhabib, Rogerson, and Wright to argue that work time can be reallocated from the home to the market during a boom without substantially raising the marginal utility of leisure.[19]

Einarsson and Marquis (1997) suggest that the third use of time may be time spent accumulating human capital, such as going to school. In principle, this is an appealing idea, since it is known that time spent in educational pursuits goes down in business cycle expansions. Their work is primarily theoretical, however. A crucial issue one would have to address in pursuing this explanation at the empirical level is whether the time spent on education is sufficiently countercyclical, in a quantitative sense, to have a substantial effect in a suitably modified real business cycle model. In assessing this, one would have to confront a substantial measurement problem. In particular, time spent in educational institutions is only part of the time spent in education. Some of that time is applied in the workplace, by diverting workers from direct production. Our understanding is that there do not exist reliable measures of this use of time.

Value of the Output of the Consumption Sector

Several papers attempt to get at comovement by reducing the decline in the value of output in the consumption sector during booms. For example, Baxter (1996) adapts the standard real business cycle model by assuming that the consumption of market goods and the services of home durables are good substitutes. An example of two goods that substitute is a movie watched in a theater (a market good) and a movie watched on a home television set (a home durable good).[20]

Under Baxter's substitutability assumption, the appropriate measure of household consumption is not just market consumption, but consumption of market goods *plus* the service flow on the stock of home durables. If home durables consumption is sufficiently large, then a given jump in the consumption of market goods leads to a smaller percent drop in the marginal utility of consumption. In the extreme case where the stock of home durables is extremely large and accounts for essentially all of consumption, then a rise in market consumption would produce essentially *no* drop in the marginal utility of consumption.[21] Although Baxter shows that this mechanism does indeed produce comovement in employment across consumption and investment sectors in her model, there is a sense in which the comovement is not strong enough. That is because investment in the capital used in the two sectors is essentially uncorrelated. As noted above, the data suggest that investment across sectors comoves as well, in addition to output and employment.

One can also view the home production approach of Benhabib, Rogerson, and Wright (1991) as a strategy to generate comovement by reducing the decline in the value of output in the consumption sector during booms. In a boom, as labor is allocated away from home-produced goods toward the production of market goods, the marginal utility of the market good does not fall much because the market and home goods are assumed to be highly substitutable.[22] This allows the value of output in the consumption sector to rise sufficiently so that employment in that sector is procyclical. A shortcoming of the analysis, emphasized by Benhabib, Rogerson, and Wright (1991), is that the high substitutability between home and market goods needed for comovement of labor inputs hurts on another dimension. It has the effect that purchases of durables are countercyclical over the cycle.[23]

Christiano and Fisher (1998) take another approach. Following Boldrin, Christiano, and Fisher (1995), they modify a standard real business cycle model in two ways. First, they specify that it takes time before labor can shift between economic sectors in response to a shock. The reasons for this are not modeled explicitly, but the assumption is motivated with an informal reference to such factors as the search and training costs which inhibit real world labor mobility between industry sectors. This assumption alone is not sufficient to guarantee comovement, however. Without further changes, their model predicts that resources would be reallocated out of the consumption sector and into the investment sector as soon as labor becomes fully mobile, which they specify to occur in three months' time. As a result, this version of the model is still inconsistent with the evidence on business cycle comovement. Christiano and Fisher therefore introduce a second modification, by changing the specification of household preferences over consumption. They assume that households have a tendency to become accustomed to the level of consumption they have enjoyed in the recent past. This property of preferences is known as *habit persistence.* A household with habit persistence preferences whose consumption has recently increased is particularly unhappy if later it must return to its previous level of consumption.[24] Habit persistence preferences have the implication that when consumption rises in a boom, the marginal value of continuing to maintain consumption at a high level is increased. Christiano and Fisher show that habit persistence and limitations on the intersectoral mobility of labor are sufficient to produce comovement in hours worked and investment. To our knowledge, this is the only quantitative model in the comovement literature with this property.

The credibility of this result depends on the credibility of the underlying assumptions. The assumption that there are limitations on the speed with which productive resources can be transferred across sectors seems uncontroversial, though the model certainly takes an extreme stance. What does call for a defense is the assumption of habit persistence preferences. One defense is that these preferences help to account for observations that otherwise seem puzzling. For example, Boldrin, Christiano, and Fisher (1995) show that, consistent with results in Constantinides (1990), habit persistence and limited intersectoral mobility can account for the magnitude of the observed average premium in the return on equity over risk-free securities. The solution to this premium has eluded many researchers.[25] In addition, Christiano and Fisher (1998) show that habit persistence can help account for the so-called inverted leading indicator property of interest rates, that high interest rates tend to forecast bad economic times. King and Watson (1996) document that standard models have difficulty accounting for this observation.[26]

A third approach toward understanding comovement was recently pursued by Hornstein and Praschnik (1997). They observe that some of the output of the sector that produces consumption goods (the nondurable goods sector) is also used as intermediate goods in the production of investment goods. For example, both households and investment-good producing firms make use of the services of the transportation sector. Hornstein and Praschnik (1997) modify a real business cycle model to accommodate this feature of the economy. The modification has the effect of increasing the value of output in the consumption sector in a boom. This increased value reflects the increased need for the output of the consumption good sector during a boom for use in the investment good sector.[27] We refer to this demand channel going from investment sector to the nondurable goods sector as the *intermediate goods channel.*

There are two shortcomings of the Hornstein and Praschnik (1997) analysis. First, the model is not consistent with the observed comovement in investment across sectors. That is, the intersectoral linkages in the model are not strong enough to produce full comovement. Second, data on subsectors of the nondurable goods sector cast doubt on the notion that the intermediate good channel is the only reason there is comovement. We studied the subsectors of the nondurable goods sector and found that there is considerable variation in the fraction of total output sent to the investment goods industry. But, as documented in the previous section, most of these sectors nevertheless display strong business cycle comovement. Figure 6 is a scatter plot of the subsectors' degree of comovement, drawn (with one exception) from the fifth column of numbers in table 1, against the strength of each sector's intersectoral linkage with the investment sector, I_c. The variable I_c is the fraction of the gross output of a sector which is allocated to intermediate goods destined directly or indirectly for the production of final investment goods.[28] (Technical appendix 3 has details of how we computed this.) The numbers in parentheses in Figure 6 indicate the relative magnitude of the gross output of the sector (gross output of the sector in 1987, divided by the sum of the gross outputs across all 17 sectors). Hornstein and Praschnik's concept of the nondurable goods sector is broader than the one in Table 1. They also include agriculture; retail trade; wholesale trade; transportation, communication, and utilities; services; FIRE; and mining.[29]

Figure 6 shows that employment in most (13 of 17) nondurable good sectors is substantially procyclical (that is, the comovement statistic is 0.45 or higher), even though the strength of the intermediate good channel (the magnitude of I_c) varies from almost zero in the case of food to nearly 0.25 in the case of wholesale trade. Interestingly, although the comovement in mining is moderately weak in our data set, it is one of the sectors in which the intermediate goods channel is the strongest. Conversely, the comovement in apparel is strong, although this sector's intermediate goods channel is almost nonexistent. Based on these results, we suspect that the intermediate goods channel to the investment sector plays

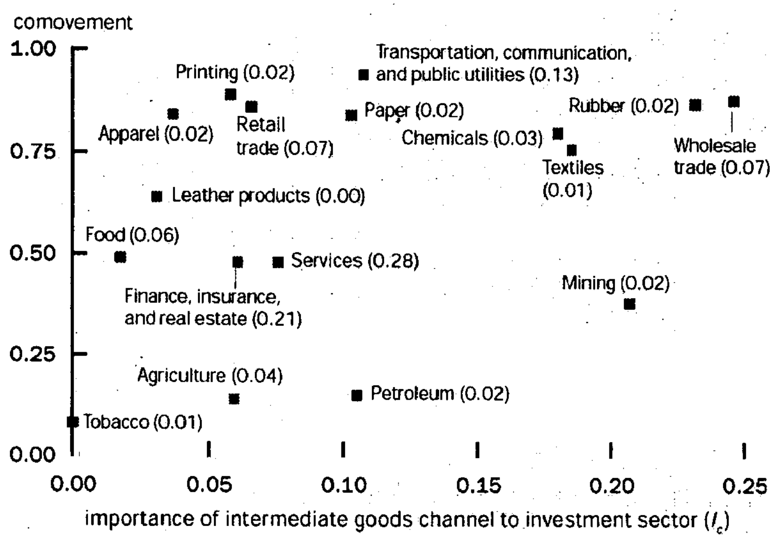


Figure 6 Business cycle comovement in nondurable goods sector. Note: Plot of sector is calculated using the comovement statistic reported in table 1 and as explained in note 6. The number in parentheses following the sector name indicates the relative size of that sector. Source: Authors' calculations from data of DRI Basic Economics database, 1964–96.

at best only a small role in accounting for comovement of employment in nondurable goods.[30] To further explore the Hornstein and Praschnik hypothesis, one would have to construct a version of their model with a disaggregated nondurable goods sector and see if it is consistent with comovement, in the sense of being able to reproduce patterns like those in Figure 6.[31]

ALTERNATIVE APPROACHES

Here, we summarize three other approaches that may ultimately lead to a satisfactory explanation of business cycle comovement—*strategic complementarities, information externalities,* and *efficiency wage theory.* The first two approaches emphasize the importance in business decisions of expectations about the future. They draw attention to the possibility that general shifts in expectations may trigger business cycle fluctuations. If so, these shifts in expectations may well constitute the aggregate shock to which Lucas (1981) refers. The third approach looks at efficiency wage theory. Although promising, the ability of these three theories to quantitatively account for the comovement aspect of business cycles is yet to be fully explored.

Strategic Complementarities

Suppose there are two people, A and B. Each has to decide on a level of work effort: high or low. Suppose that the net gain to A of exerting a high level of effort is greater if B exerts a high level of effort and that B is in a similar position. The situation is depicted in Table 2.

Table 2 has four entries, one for each possible combination of work effort. In each entry, the first number indicates the net gain to A, and the second number indicates the gain to B. Suppose A exerts high effort. Then, if B is putting out high effort too, A receives 5. But, if B exerts low effort, then A receives–5. The situation is the same for B. Table 2 captures the idea that the gain to either person from exerting high effort is high only when the other person exerts high effort. A situation like this is said to be characterized by strategic complementarity. What do we expect to happen? If the two people could sit down and reach an agreement, they would clearly both choose to exert high effort. But what if they have difficulty coordinating in this way? There are now two possibilities. One is that each expects the other to exert low effort, in which case each finds it optimal to exert low effort. This would put the two people in the bottom right box, with a low payoff going to each. They would stay there until they found a way to communicate and reach an agreement or

Table 2 Example of Strategic Complementarity

	Person B	
Person A	High effort	Low effort
High effort	(5, 5)	(–5, 0)
Low effort	(0, –5)	(2, 2)

until something happened to alter their expectation about the other's plans. Another possibility is that each expects the other to exert high effort, in which case it is in the private interest of each person to exert high effort. This situation could persist for a while, again, unless something happens to shift one person's expectations about what the other one will do.

What does this have to do with business cycles and comovement? Possibly a lot. There are aspects of business decisions that exhibit strategic complementarity. For example, suppose a firm is considering reopening a plant or starting a large capital investment project. Suppose the project involves a substantial outlay of funds, not just to hire more workers but also to purchase materials and supplies from other firms. The higher the sales the firm expects in the future, the more inclined it will be to shift to a high level of activity in this way. However, much of a firm's sales come from other firms. And those sales are greater if other firms are themselves operating at a high level of activity, for example, reopening plants or undertaking new capital investment projects. So, firm A has a greater incentive to shift to a high level of activity if it believes firm B plans to operate at a high level of activity.

What do we expect in this situation? Coordination in this setup is much more difficult than in the two person example. There are millions of firms in the economy and, even if it were technically feasible for some firms to coordinate, the antitrust laws represent another barrier. In light of these considerations, we might well expect to find results similar to those in the two person example. Thus, if firms were pessimistic about prospects for future sales, they would choose to be inactive and their pessimistic expectations would be fulfilled.[32] Optimistic expectations would be self-fulfilling in the same way. It is clear that in this setting, expectations have the potential to act as an aggregate shock driving the business cycle. Of course, that does not guarantee that they can necessarily account for comovement.[33] This is an important topic of research.[34]

Information Externalities

Another potential source of comovement is the way information about the state of the economy is transmitted to individual firms. Forecasts of the future strength of the economy are a factor in individual firms' current investment decisions. If a firm observes a series of construction projects being initiated by other firms, it may infer that those other firms have information that bodes well for the general economic outlook. When the firm combines this inference with its own information about the economic outlook, it may decide to invest too. Other firms may follow for similar reasons.

These considerations are logically distinct from the strategic complementarities discussed above. There, a firm is interested in the actions of other firms because these actions have a direct impact on the firm's profitability. Here, a firm is interested in the actions of other firms because of the associated information externality. The externality refers to the fact that a firm's action may reveal information it has on something of interest to other firms, such as the state of the economy. It is a *positive* externality, unlike the more familiar examples of externalities which tend to be negative.[35]

We present an example, taken from Banerjee (1992), to illustrate the sort of things that can happen when there are information externalities. When firms look to what other firms are doing for guidance in deciding what they should do, this can lead to what Banerjee (1992, p. 798) calls *herd behavior,* a situation with "everyone doing what everyone else is doing, even when their private information suggests doing something quite different." It hardly needs to be stated that herd behavior sounds like comovement.

Here is the example. Suppose there are 100 people trying to decide between two restaurants, A and B. Each person knows very little about the two restaurants, but thinks the odds favor A slightly. In addition, each person receives a signal about the relative quality of the two restaurants. For example, one person may read a review of the two restaurants in a travel guide. The review is several years old, however, and the signal may not be accurate. The signals received by each of the 100 persons are equally reliable. Everyone knows this, but they do not know what signal the others received. Now, suppose that 99 people get a signal that suggests B is better than A, while one person gets a signal that A is better than B. If all information were known to everyone, they would recognize that the preponderance of the evidence favors restaurant B and all 100 people would go to B. However, what actually happens is that the 99 people whose signal indicates B is better ignore their signal and flock to restaurant A, following the one person who received the signal that A is better.

This result is not due, as one might suppose, to an assumption that agents are irrational. On the contrary, the example assumes agents are completely rational. The result reflects that not all people make their decisions at the same time. Some have to be first, and as a result, the information they have has disproportionate impact, since almost everyone else is watching them. This timing assumption does not seem unrealistic. In practice, the exact timing of firms' decisions is not completely under their control.[36]

The example adopts an extreme version of the assumption that the timing of a decision is out of the agents' control, specifying that someone must choose first, then someone else must choose second after observing the choice of the first, and so on. The person choosing first happens to be the one who receives the signal that A is better than B. Since person 1's suspicion that A is better is apparently confirmed by the signal, this person rationally chooses A. The second person's signal suggests that B is better. However, person 2 knows that person 1's signal must have favored A. Since the two signals are equally reliable, they cancel in the mind of person 2. Since person 2 originally thought restaurant A was better, the rational thing for person 2 to do is to go to restaurant A. Person 3 is in precisely the same position as 2, because person 3 knows that, given person 1 went to A, person 2 would have gone to A no matter what signal she received. That is, person 3's observation that person 2 went to A provides no information at all about the relative quality of the two restaurants. Being in the same position as 2, person 3 also chooses A regardless of the signal received. In this way, all 99 people after the first ignore their own signal and go to restaurant A. Although there is a lot of information in the economy about the relative quality of the restaurants, one person acts on a small piece of it, and everyone else follows.

This example and others like it hold out some hope that a fully developed business cycle model incorporating information externalities might exhibit the synchronization of behavior across economic sectors that we observe over the business cycle. However, the above example only illustrates how information externalities can lead rational people to ignore information and synchronize on *bad* decisions. Synchronization of actions would have occurred anyway, even if there had been no information problem and all signals had been known to everyone. Another concern with this example is how heavily dependent it is upon details of the environment. For example, the outcome is very different if two people are required to choose a restaurant first. In this case, the 99 people who received the signal that B is better than A go to B.[37,38] Despite these considerations, we believe the growing literature on information externalities may eventually provide at least part of the explanation for business cycle comovement.[39]

Efficiency Wage Theory

A third strategy for understanding comovement is to make use of efficiency wage theory.

Efficiency Wage Theory: A Sketch

Under this view of labor markets, the amount of effort a worker makes (the worker's *efficiency*) depends on the wage that the worker is paid. Development economists hypothesized that in economies at a very early stage of development, a higher wage leads to greater worker efficiency because it facilitates improvements in diet and health. Efficiency wage theory holds that a higher wage also results in greater worker efficiency in a modern, developed economy, but for different reasons. Because employers cannot perfectly monitor the amount and quality of work effort expended by their employees, there is a temptation for workers to shirk. Efficiency wage theory says that a high wage rate is an effective way to combat this temptation. The higher the wage, the more a worker has to lose if caught and fired for poor job performance.

The simplest version of this idea was articulated by Robert Solow,[40] who theorized that the firm selects a wage rate, the *efficiency wage,* which maximizes worker effort per dollar paid. The firm is not willing to pay more because the resulting increase in worker effort would not be enough to warrant the extra cost. The firm is also not willing to pay less, because the resulting fall in effort would exceed the fall in cost.[41] In the Solow model, the amount of effort expended per hour by a worker is a function only of the current wage and, for example, does not depend on the general state of business conditions. As a result, the efficiency wage rate does not vary over the business cycle under Solow's efficiency wage theory.

The firm also has to decide how many workers to hire. It hires workers up to the point at which the marginal productivity of the last worker is just equal to the efficiency wage.[42] The downward sloping marginal productivity of labor curve in Figure 7 shows how the marginal productivity of labor is lower at higher levels of employment. At the level of employment, *L,* the marginal productivity of the last worker hired is equal to the efficiency wage. Since the efficiency wage in the Solow model is a constant, it follows that employ-

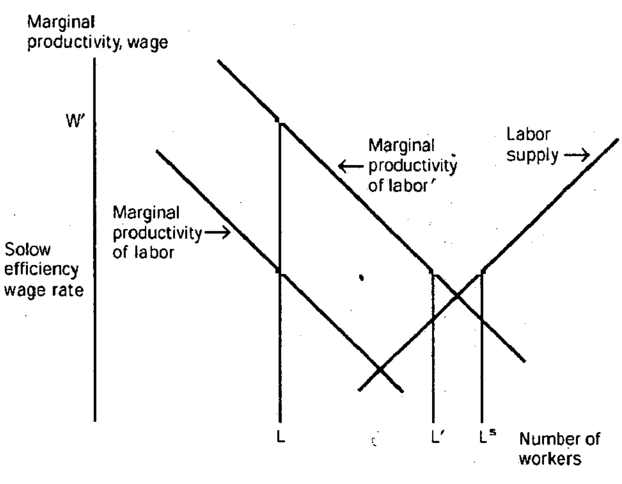


Figure 7 Efficiency wage model.

ment over the business cycle is determined by the requirement that the marginal product of labor does not change. The downward sloping curve in Figure 7, marginal productivity of labor', shows the marginal productivity curve after it has been shifted up by a positive technology shock.[43] If the firm kept employment fixed at L when technology shifted up, marginal productivity would rise to W', a point far above the efficiency wage. By increasing L to L', the firm keeps marginal productivity unchanged despite the shift up in technology.[44]

A notable feature of efficiency wage theory is that labor supply plays no role in the determination of the wage rate. The theory assumes that there are more workers willing to work than the firm is willing to hire at the efficiency wage. Still, unemployed workers cannot bid the wage down below the level of the efficiency wage. Firms are not interested in hiring workers at such a low wage because they fear it would not provide workers with enough incentive to work hard. The quantity of unemployed people is the number who are willing to work at the efficiency wage, minus the number that firms want to hire. Note how the upward sloping labor supply curve in Figure 7 is shifted to the right. At the efficiency wage, L_s, workers would like to work, but only L are hired, so unemployment is $L_s - L$. At the higher level of technology, unemployment falls to $L_s - L'$.

Efficiency Wage Theory and Business Cycle Comovement

How might efficiency wage theory help account for business cycle comovement? Suppose the business cycle is driven by an aggregate, real-business-cycle-type technology shock. As we explained earlier, in the standard real business cycle model such a shock does not lead to comovement in employment. In that model, a positive shock leads to a transfer of resources—labor and capital—away from the firms producing investment goods. Now suppose the labor market part of the real business cycle model is replaced by efficiency wage theory, which implies that firms vary the number of workers they employ to ensure that the marginal product of labor remains constant and equal to the efficiency wage rate. So, when a positive technology shock shifts up the marginal productivity of labor, employment must increase to maintain equality between the marginal product of labor and the efficiency wage.[45]

We indicated earlier that a positive real-business-cycle-type shock pushes up the production functions and the marginal labor productivity curve of all firms. According to efficiency wage theory, this results in an increase in employment by all firms, as they seek to maintain equality between marginal labor productivity and the unchanging efficiency wage rate. This is the intuition underlying the idea that efficiency wage theory may help explain business cycle comovement.[46,47]

Have we now established that efficiency wages are sufficient to account for comovement? Absolutely not. When we examine efficiency wage theory more closely, we discover that it need not necessarily work as just outlined. The relationship between how hard a worker works and the wage rate (the worker's *effort function*) is a function of the household's attitude toward risk, the resources it has available if the worker is caught shirking and fired, the probability of being caught conditional on shirking, and the precise consequences of being fired for shirking. The household effort function used in the analysis must integrate all these factors in a logically coherent way. In addition, it must be consistent with other household decisions, such as how to split income between consumption and saving. To build confidence in the idea that efficiency wage theory helps account for comovement, we must integrate all these aspects of the household into a coherent framework which also includes firms and their decisions to see if it works.

To understand why it might not work, recall the Solow model's assumption that worker effort is a function only of the wage rate. That is what led to our conclusion that the efficiency wage is a fixed number, unrelated to the state of the business cycle. But the logic of the efficiency wage argument suggests that the Solow assumption may not be consistent with rational behavior by households. According to efficiency wage theory, what motivates hard work is the fear of losing a high-wage job. Of course, the cost of that loss is not a function of the wage rate alone. It is also a function of the amount of time the worker can expect to be out of a job after being fired. This suggests that the horizontal line in Figure 7 shifts up in a boom, when the duration of unemployment is low.[48] However, if it shifts up high enough, the comovement result could disappear. This highlights the importance of integrating efficiency wage theory into a logically coherent model, before we conclude that it provides a solid foundation for understanding business cycle comovement.

Important steps have been taken in this direction, for example, Shapiro and Stiglitz (1984) and Danthine and Donaldson (1995).[49] Recent work by Gomme (1998) and Alexopoulos (1998) makes a significant further contribution toward understanding the implications of efficiency wage theory for business cycles. However, this work does not focus on the implications for business cycle comovement. We argue that doing so is a good idea.[50]

CONCLUSION

A key feature of the data is that, in a frequency range of two to eight years, output, employment and investment across a broad range of sectors move up and down together. We have documented this phenomenon—business cycle comovement—as it pertains to employment. Our survey of possible explanations for it is by no means exhaustive. Many other approaches—those based on sticky prices and wages, countercylical markups, and credit market frictions—also deserve consideration.[51] Still, we have covered a wide range of models, from straightforward modifications to standard business cycle theory to theories that suggest analogies between businesspeople and herds of animals.

Many of the approaches we have surveyed are in early stages of development, while some have been developed to the point where their implications have been quantified and compared with the data. Among these, only one has been shown to be consistent with the observed strong comovement in output, employment, and investment across sectors of the economy—the model presented in Christiano and Fisher (1998). This model incorporates a specification of household preferences, habit persistence, that is not currently standard in the macroeconomics literature. We believe that the success this model has in generating comovement warrants giving this specification of preferences further consideration.

Because comovement is such a central feature of business cycles and because we do not have a generally agreed upon theory of comovement, we conclude that the business cycle is still a puzzle.

NOTES

1. See Burns and Mitchell (1946, p. 3), Lucas (1981, p. 217), and Sargent (1979, p. 215).
2. This definition was taken from the NBER's web address, http://www.nber.org/cycles.html.
3. An important exception is Long and Plosser (1983), which does allow for multiple sectors. Their model economy is straightforward to analyze because they adopt several key simplifying

assumptions. For example, they assume the entire stock of capital in each sector wears out within three months. However, these assumptions make the model ill-suited for quantitative, empirical analysis. It took many years before economists undertook a systematic empirical analysis of versions of the Long and Plosser model without the key simplifying assumptions (see Horvath [1998a, b]).

4. Employment data are taken from DRI Basic Economics database. The hours worked data are indexes of aggregate weekly hours of production or nonsupervisory workers on private nonagricultural payrolls by industry. The data on numbers of workers are workers on nonagricultural payrolls, by industry. All data are monthly and seasonally adjusted and cover 1964:Q1–95:Q3.
5. Other studies of this question include Baxter (1996), Cooper and Haltiwanger (1990), Hornstein and Praschnik (1997), Huffman and Wynne (1998), and Murphy, Shleifer, and Vishny (1989).
6. Our statistic is the regression R^2 obtained by regressing the business cycle component of that series on the business cycle component of total hours worked, at lags 0, 1, and −1. We allow next month's employment and the previous month's employment to enter this relationship because we do not want our measure of comovement to be low just because a variable may be out of phase with total private hours worked by only one month. If we did not include these lags, our regression R^2 would coincide exactly with the square correlation referred to in the text. We construct our statistic as follows. Let y_t denote the business cycle component of a given sector's employment. Let x_t denote the corresponding measure of total hours worked. We consider the regression of y_t on x_t, x_{t-1}, and x_{t+1}, $y_t = \hat{\alpha}_0 x_t + \hat{\alpha}_1 x_{t-1} + \hat{\alpha}_2 x_{t+1} + \varepsilon_t$, where $\hat{\alpha}_t$ represents the estimated coefficients. Then, the R^2s reported in the table are $var\,(\hat{\alpha}_0 x_t + \hat{\alpha}_1 x_{t-1} + \hat{\alpha}_2 x_{t+1})/var\,(y_t)$.
7. Table 1 shows the volatility in each of these data series.
8. The correlation between y_t and y_{it} is $corr\,(y_t, y_{it}) = Cov\,(y_t, y_{it})/[Var\,(y_t)Var\,(y_{it})]^{1/2}$. But, $Cov\,(y_t, y_{it}) = \sigma^2$ and $Var\,(y_t) = 2\sigma^2$, $Var\,(y_{it}) = \sigma^2$. Substituting these results into the formula, we get $corr\,(y_t, y_{it}) = 1/\sqrt{2}. \simeq 0.71$.
9. Suppose $y_t = y_{1t} + \ldots + y_{nt}$. The logic of the previous note leads to $corr\,(y_t, y_{it}) = 1/\sqrt{n}$. With $n = 33$, this is 0.17, after rounding.
10. The midpoints are −0.35, −0.25, −0.15, . . ., 0.85, 0.95. In each case, the interval has length 0.1 and extends 0.05 above and below the midpoint.
11. Real business cycle theory has evolved considerably in recent years and now encompasses a wide variety of conceptions of the economy. The definition proposed by Prescott (1991, p. 3) reflects this: "Real business cycle theory is the application of general equilibrium theory to the quantitative analysis of business cycle fluctuations."
12. This section and the next one draw heavily on work by Christiano and Fisher (1998).
13. Some might want to dismiss the notion of a technology shock that affects all firms simultaneously as too preposterous to deserve consideration. Such a person may find it more plausible to think of technology shocks as things that are idiosyncratic to individual firms. Most of the examples of technology shocks given in the text certainly suggest this. This is the line that Lucas took when he dismissed the idea that a technology shock might be the aggregate shock needed to account for business cycles. He argued that, although technology shocks are no doubt important at the firm level, they could not be important for economy-wide aggregate output: He expected that firms affected by positive productivity shocks would be balanced by firms experiencing negative shocks. Work of Shleifer (1986) and Dupor (1998) suggests that the Lucas reason for dismissing technology shocks as an important impulse to business cycles may be premature. These researchers emphasize the distinction between the time that a new technological idea arrives in the firm, and the time the firm implements it. Consistent with Lucas's intuition, the exact timing of arrival of ideas may well be idiosyncratic at the firm level. In this case, the economy-wide average rate of arrival of new ideas would be constant: Firms discovering ideas for new products or labor-saving ways to produce output would be balanced by firms experiencing no progress or even regress. What Shleifer and Dupor emphasize, however, is that it is not the arrival of new ideas *per se* that shifts up production functions. Rather, it is the imple-

mentation of the new ideas that does this. They point out that there may well be plausible mechanisms in an economy which lead firms to implement new, technology-shifting ideas at the same time. These mechanisms involve "strategic complementarities," which are discussed further below.

14. See, for example, Benhabib, Rogerson, and Wright (1991).
15. Formally, this is what we have in mind. A standard real business cycle model, with unit elasticity of substitution in production between capital and labor, implies that the value of the output of the sector producing consumption goods, measured in utility units, is proportional to the value of the labor used in that sector, also measured in utility units. The value of the output of the consumption sector is the product of the total output of that sector, Y, and the marginal utility of consumption, u_c. The value of the labor used in the sector producing consumption goods is the product of the labor used in producing consumption goods, L_c, and the marginal utility of leisure, u_l. Thus,

$$\alpha Y u_c = u_l L_c.$$

This is just a rearrangement of the usual static efficiency condition that specifies that the marginal product of labor in producing the output of the consumption sector, $\alpha Y/L_c$, must equal the marginal rate of substitution between consumption and leisure, u_l/u_c. Note that if the term on the left of the equality falls ("the value of the output of the sector producing consumption goods falls") and u_l rises ("the marginal utility of leisure rises"), then L_c must fall.

16. The inability of the standard real business cycle model to produce comovement is surprisingly robust. Standard specifications of that model hold that the marginal rate of substitution between consumption and leisure is $\psi C/(1 - L_c - L_i)^{\xi}$, where L_c is employment in the consumption sector, L_i is employment in the investment good sector, and $1 - L_c - L_i$ is leisure. Also, ψ and ξ are non-negative constants. In Hansen's (1985) indivisible labor model, $\xi = 0$. In his divisible labor model, $\xi = 1$. The standard model assumes a Cobb–Douglas production function, so that the marginal product of labor is proportional to average labor productivity in the consumption good producing sector. Equality between the marginal product of labor and the marginal rate of substitution between consumption and leisure implies:

$$\alpha \frac{C}{L_c} = \frac{\psi C}{(1 - L_c - L_i)^{\xi}}.$$

Cancelling consumption on the two sides and rearranging, we get

$$\frac{\alpha}{\psi}(1 - L_c - L_i)^{\xi} = L_c.$$

From this it is easy to see that if, for whatever reason, L_i or L_c moves, then the other variable *must* move in the opposite direction. This demonstration summarizes a discussion in Benhabib, Rogerson, and Wright (1991) and in Murphy, Shleifer, and Vishny (1989). The result holds for the entire class of utility functions identified by King, Plosser, and Rebelo (1988) as being consistent with balanced growth. However, the same cannot be said for the entire class of production functions consistent with balanced growth. In particular, the result does *not* hold for production functions in which the elasticity of substitution between capital and labor differs from unity. We demonstrate this in technical appendix 2. We also show, however, that for plausibly parameterized versions of the standard real business cycle model, departures from unit elasticity of substitution in production do not help the model reproduce comovement.

17. One paper that is often mentioned in the comovement literature is Huffman and Wynne (1998). However, their focus is primarily on comovement in investment and output. They largely abstract from comovement in employment by making assumptions that make labor in the consumption sector essentially constant. They specify that the elasticity of substitution between labor and capital in the consumption sector is nearly unity, and that $\xi = 0$. The argument in note 16 explains why their model has the implication that L_c is essentially constant.

18. Suppose L_x is the third use of time. Then the equation in note 16 is modified as follows:

$$\frac{\alpha}{\xi}(1 - L_c - L_i - L_x)^{\xi} = L_c.$$

Evidently, now it is possible for both L_c and L_i to be procyclical, as long as L_x is sufficiently countercyclical.

19. Closely related to this is their recommendation that economists work with the following utility function in consumption and leisure: $u\,[c - \psi_0 L^{1+\psi}/(1 + \psi)]$, where $\psi, \psi_0 > 0$ and u is a concave, increasing utility function. The marginal rate of substitution between consumption and leisure with this utility function is $\psi_0 L^{\psi}$. Substituting this into the employment condition in note 16 results in

$$\alpha \frac{C}{L_c} = \psi_0 (L_c + L_i)^{\psi}.$$

The argument in that note that L_c and L_i cannot move in the same direction does not work with this utility function.

20. Baxter's model is a convenient vehicle for illustrating an issue that has to be confronted in macroeconomic models generally. The text provides an illustration of Baxter's assumption that durable goods and market goods are substitutes. However, it is just as easy to think of examples in which they are complements. Consider a car, for example. Ownership of a car makes it more attractive to go out on long road trips that require purchasing market goods like hotel and restaurant services. This suggests that cars and market goods are complements. A moment's further thought about this example suggests that most household durables actually cannot be neatly labeled as either complements or substitutes for market consumption. For example, an automobile is also a substitute for market goods because it reduces the need for market services like cab, train, and airplane rides. Similarly, consider the biggest household durable of all, the home. It substitutes for hotel and restaurant services and complements market goods such as party goods, telephone services, and food. Thus, intuition is ultimately not a good guide to assessing Baxter's assumption about the substitutability of durables and market goods. Ultimately, this must be assessed through careful econometric work to determine whether, *on average,* market goods and durables are more like substitutes or complements.

21. Consider the limiting case of perfect substitutability, so that consumption is $C + D$, where C is market consumption and D is the service flow from the stock of home durables. With log utility, the marginal utility of market consumption is $1/(C + D)$. Suppose D is fixed. Then a given jump in C reduces marginal utility by less, the larger is D.

22. Remarks in note 20 about Baxter's work are obviously relevant here too. Intuition is a very confusing guide, at best, regarding the plausibility of Benhabib, Rogerson, and Wright's assumption that the elasticity of substitution between home-produced and market-produced goods is high. The parameter must be estimated econometrically. This was done in Rupert, Rogerson, and Wright (1995), who report, based on data from the *Panel Study on Income Dynamics,* that the elasticity of substitution indeed is high.

23. Because the model predicts that consumption rises in a boom, the high degree of substitutability between home and market goods causes the marginal value of home goods to drop in a boom. This in turn causes a drop in the value of home durables, leading households to reduce their purchases of new durables. This implication is strongly counterfactual, however, since durables are in fact highly procyclical. Interestingly, Baxter's (1996) model seems to avoid this tension. In particular, her model generates comovement between employment in the consumption and investment industries and simultaneously implies that durable goods purchases are procyclical.

24. Boldrin, Christiano, and Fisher (1995) adapt the habit persistence specification of preferences proposed in Constantinides (1990) and Sundaresan (1989).

25. See Kocherlakota (1996) for a recent review. Although habit persistence helps to account for

the observed average of the premium in equity over risk-free debt, it does not account well for the volatility of these variables. For a further discussion, see Boldrin, Christiano, and Fisher (1997) and Heaton (1995).

26. See Constantinides (1990) and Sundaresan (1989) for more evidence on the plausibility of habit persistence preferences.
27. In the Hornstein and Praschnik (1997) modification, the output of the consumption sector is $C + m$, where m is intermediate goods sent to the investment good sector. Suppose the marginal utility of market consumption is $1/C$. Then, the value of the output of the consumption sector is $(C + m)/C = 1 + m/C$. Note that this jumps with a rise in C as long as m rises by a greater percentage than C. With m/C sufficiently procyclical, it is possible for employment in the investment and consumption good sectors to move up and down together over the cycle.
28. We are very grateful for instructions and advice from Mike Kouparitsas on how to analyze the input–output data.
29. We do not have an index of hours worked for this sector. Instead, we used LHAG, which is Citibase's mnemonic for number of persons employed in the agricultural industry. We obtained a measure of comovement for this variable in the same way as for the other variables.
30. The least squares regression line through the data in figure 6 is $\rho_{h,y} = 0.48 + 1.35I_c$. Thus, if a sector was not connected to the investment sector at all (that is, $I_c = 0$), employment in that sector would still exhibit substantial procyclicality (that is, $\rho_{h,y} = 0.48$).
31. Such an exercise could be pursued by building on the models in Long and Plosser (1983) and Horvath (1998a, b). To our knowledge, comovement in the sense studied in this article has not been investigated in these models.
32. A slightly different mechanism, whereby a firm's expectation that other firms will be inactive leads all firms to be inactive was analyzed by Shleifer (1986) and Dupor (1998) and summarized in note 13.
33. For example, Benhabib and Farmer (1994, 1996) incorporate strategic complementarities by way of an externality in the production function. Because their production function is of the Cobb–Douglas form, the argument in note 16 applies to these models too. In particular, in these models, employment in the production of consumption and investment goods must move in opposite directions over the business cycle.
34. The literature on the potential for expectations to be self-fulfilling is large. Influential early papers include Azariadis (1981), Bryant (1983), Cass and Shell (1983), Cooper and John (1988), Diamond (1982), Farmer and Woodford (1984), Shleifer (1986), and Woodford (1986, 1987, 1988, 1991). More recent contributions include Benhabib and Farmer (1994, 1996), Christiano and Harrison (1998), Cooper and Haltiwanger (1990, 1996), Farmer and Guo (1994), Gali (1994), and Schmitt-Grohe (1997).
35. An example of a negative externality is the pollution that is generated as a byproduct of a manufacturing process.
36. For an analysis of the case where there are information externalities and timing *is* under the control of managers, see Chamley and Gale (1994). They find, as one might expect, that there is a tendency to delay decisions under these circumstances.
37. We are grateful to Henry Siu for pointing this out to us.
38. The example is similarly sensitive to the assumption that people view the signals they receive as equally reliable to the signals received by others. It is possible that, in practice, the type of individual making investment decisions has greater confidence in her ability to interpret signals than her counterparts at other firms. This is the implication of empirical evidence that suggests that these types of people are overly confident in their own abilities. See Daniel, Hirshleifer, and Subrahmanyam (1998), and the references therein for further discussion. According to them, (p. 5–6): "Evidence of overconfidence has been found in several contexts. Examples include psychologists, physicians and nurses, engineers, attorneys, negotiators, entrepreneurs, managers, investment bankers, and market professionals such as security analysts and economic forecast-

ers. Further, some evidence suggests that experts tend to be more overconfident than relatively inexperienced individuals."

39. A small subset of the literature on information externalities includes Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1994), Caplin and Leahy (1994), and Chamley and Gale (1994).
40. See Romer (1996) for a review.
41. Let $e(w)$ be the amount of effort a worker expends per hour, given the hourly wage rate, w. The efficiency wage is the value of w that maximizes $e(w)/w$. One type of e function that guarantees that this has a maximum for $0 < w < \infty$ is one in which e, when expressed as a function of w, has an S shape: convex for w near zero and turning concave for larger values of w (see Romer, 1996). The optimal $e(w)/w$ is the slope of the ray drawn from the origin, tangent to the concave part of the e function. At the optimum, the elasticity of effort with respect to the wage is unity, that is, $e'(w)w/e(w) = 1$. Optimality requires that, when evaluated at the efficiency wage, the second derivative of e with respect to w, is negative.
42. The algebra underlying this analysis is simple. Let the production function be $f(e(w)L, K, z)$, where eL is the total amount of effort expended in L hours of work, z is a shock to technology, and K is the stock of capital. We assume that the derivative of f in its first argument is positive and strictly decreasing in eL and increasing in z. Revenues net of labor costs are $f(e(w)L, K, z) - wL$. The firm maximizes this with respect to w and L. It is convenient, however, to adopt a change of variables, $X = wl$, and let the firm choose X and w instead. Then, the revenue function is

$$f\left(\frac{e(w)}{w}X,K,z\right) - X.$$

Evidently, maximizing this with respect to w is equivalent to maximizing effort per dollar cost, $e(w)/w$ with respect to w. For a further discussion of this maximization problem, see the previous note. Maximization with respect to X implies:

$$f_1(eL,K,z)e = w,$$

that is, the marginal product of labor must equal the wage rate.

43. The marginal product of labor curve in Figure 7 graphs $f_1(e(w^*)L, K, z)e(w^*)$ as a function of L, holding K fixed. Here, w^* is the efficiency wage rate. The curve marked marginal product of labor' graphs $f_1(e(w^*)L,K,\tilde{z})e(w^*)$ for $\tilde{z} > z$.
44. These observations motivate why efficiency wage theory is sometimes viewed as a way to fix another set of counterfactual implications of the standard real business cycle model: that wages tend to fluctuate too much and employment too little over the business cycle.
45. This argument implicitly assumes that the stock of capital used by a firm, once put in place, cannot be shifted to another firm. The assumption guarantees that a positive technology shock which drives up the marginal productivity of labor curve, *must* be accompanied by a rise in employment if marginal productivity is to remain unchanged. If capital were mobile between sectors, this could even be accomplished with a *fall* in labor, as long as capital in that sector fell by an even greater percentage. The standard real business cycle model assumes that capital is freely mobile between sectors. Thus, the intuition in this article is based on *two* modifications to the real business cycle model: incorporation of efficiency wages and sectoral immobility of capital. The second of these is not sufficient to produce business cycle comovement. This is because the argument in note 16 holds even if capital is immobile between sectors.
46. In addition to verifying the logical coherence of efficiency wage theory as an explanation of comovement, there are two empirical issues to be investigated. How hard is it to monitor worker effort? If it can be monitored easily, then efficiency wage theory is irrelevant. Also, if the penalty for being fired for shirking is enormous, workers will behave as if they are being monitored continuously, and once again the theory becomes irrelevant. For a further discussion of these issues, see Alexopoulos (1998).

47. We stress that the intuition developed here relies on *two* assumptions—efficiency wages and sectorally immobile capital.
48. To be precise, suppose $e(w, D)$ is the effort supplied by workers when the wage rate is w and unemployment duration is D. At the efficiency wage, $e_{11}(w, D) < 0$. Also, we assume $e_{12}(w, D) = 0$. Totally differentiating the first order condition for the efficiency wage, $we_1(w, D)/w = 1$, with respect to w and D, and imposing the restrictions on e_{12} and e_{11} yields the result, $dw/dD < 0$.
49. In the literature, what we have called the worker's effort function, e, is referred to as the "incentive compatibility constraint."
50. Alexopoulos and Gomme have reported to us privately that their models are only partially consistent with business cycle comovement. In both cases, employment in the consumption and investment sectors is positively correlated, but investment in these two sectors is negatively correlated. However, both models assume that capital can be transferred instantaneously across sectors in response to a shock. The analysis here suggests that sectoral capital immobility may be important for obtaining comovement.
51. For an introduction to the literature on sticky prices and wages, see Romer (1996). To see why countercyclical markups might help, recall the key equation in note 16, used to show why hours worked making consumption goods and hours worked making investment goods in a standard real business cycle model must move in opposite directions. A version of that model with market power, for example, the model of Rotemberg and Woodford (1992), implies that it is the ratio of the marginal product of labor to the markup that must equal the marginal rate of substitution between consumption and leisure. That is, that equation must be modified as follows:

$$\frac{\alpha}{\mu}\frac{C}{L_c} = \frac{\psi C}{(1 - L_c - L_i)^{\xi}},$$

where μ is the markup of price over marginal cost. Cancelling consumption on the two sides and rearranging, we get

$$\frac{\alpha}{\psi}(1 - L_c - L_i)^{\xi} = \mu L_c.$$

Suppose a boom occurs, driving up L_i. If μ falls, as in the Rotemberg and Woodford model, then it is possible for L_c to rise too. (For another model with countercyclical markups see Gali [1994]). See Murphy, Shleifer and Vishny (1989) for a conjecture about how limited intersectoral labor mobility, together with credit market restrictions, may help account for comovement.

TECHNICAL APPENDIX 1 EXTRACTING THE BUSINESS CYCLE COMPONENT OF A TIME SERIES

In casual discussions of economic time series, we often think of the data as being the sum of components that have different frequencies of oscillation: the business cycle component lasting two to eight years, components lasting shorter periods, etc. The theory of the spectral analysis of time series makes this intuition rigorous. It clarifies how one can think of data as being composed of components that fluctuate at different frequencies. The method we use to extract the business cycle component of economic time series builds on this theory. For this reason, we begin with a brief section which attempts to convey the basic intuition of spectral analysis. The second section uses this intuition to describe and motivate our method for extracting the business cycle component of a time series.

Decomposing a Time Series into Frequency Components

At the core of the spectral analysis of time series is the view that the data can be thought of as the sum of periodic functions. The purpose of this section is to explain this. We begin by reviewing the basic periodic function used in spectral analysis, which is composed of a sine and a cosine function.

Consider the following cosine function of time, t:

$$\cos(t\omega),\ t = 0, 1, 2, \ldots.$$

A graph of this, with cos $(t\omega)$ on the vertical axis and t on the horizontal axis, exhibits the oscillations between 1 and–1 familiar from high school trigonometry. Recall too, that the period of the cosine function is 2π. That is, $\cos(y) = \cos(y + 2\pi h)$, for $h = 1, 2, \ldots$. Thus, after the argument of the cosine function increases by 2π, the function repeats itself in a periodic fashion.

What is of interest here is the period of oscillation of $\cos(t\omega)$, expressed in units of time. This is the amount by which t must increase so that $t\omega$ increases by 2π. Thus, suppose t_1 is a given point in time. We want to know what is the later point in time, $t_2 > t_1$, when the cosine function begins to repeat itself. This is just t_2 such that $t_2\omega - t_1\omega = 2\pi$, or, $t_2 - t_1 = 2\pi/\omega$. Thus, the period of oscillation of $\cos(t\omega)$, in units of time, is $2\pi/\omega$. The parameter ω is referred to as the frequency of oscillation.

The function, $\sin(t\omega)$, behaves similarly. It fluctuates between 1 and–1, and has a period of oscillation of $2\pi/\omega$. Thus, the two functions have the same *amplitude* (magnitude of vertical variation) and period. However, the sine function has a different *phase* than $\cos(t\omega)$. For example, a graph of the two functions together shows that one looks like the other, apart from a horizontal shift. The phase difference between the two functions is a measure of the magnitude of this horizontal shift. Figure A1 displays sine and cosine functions for $t = 0, 1, \ldots, 200$. The period of oscillation is 100, so that the frequency is $\omega = 2\pi/100$. Thus, the figure displays the graphs of $\cos(t2\pi/100)$ and $\sin(t2\pi/100)$.

We can now describe the central periodic function in spectral analysis, namely the linear combination of a sine and a cosine function.

$$a\cos(t\omega) + b\sin(t\omega) \tag{1}$$

where a and b are parameters. This function obviously has a period, in units of time, equal to $2\pi/\omega$. But, its amplitude and phase depend on the values of a and b. If a and b are both

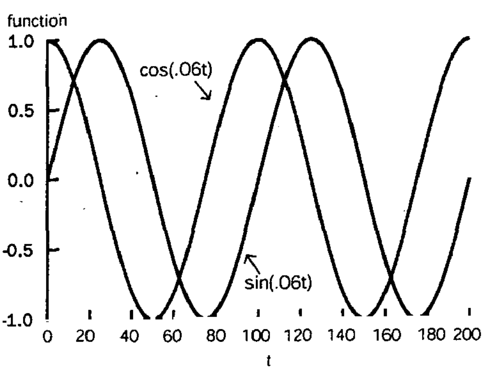


Figure A1 Trigonometric functions.

very small, the resulting function has very small amplitude and if a and b are both very large it has a large amplitude. Also, as the size of a is increased and the size of b is decreased, the phase of the function shifts, as more weight is allocated to the cosine and less to the sine.

It turns out that sums of periodic functions like equation 1 look very much like actual data. Thus, suppose we have a time series of data, x_t, $t = 1, \ldots, T$. To see that x_t can be expressed as a sum of periodic functions, suppose we specify $T/2$ (suppose T is even) such functions, each with a different frequency of oscillation ω. For concreteness, let $\omega_j = 2\pi j/T$, for $j = 1, \ldots, T/2$. To distinguish the parameters associated with each of these functions, we denote them by a_j and b_j for $j = 1, \ldots, T/2$. It should not be surprising that, in general, a time series, $x_1, \ldots, x_T$ can be written as the sum of these $T/2$ functions

$$x_t = a_1 \cos(t\omega_1) + b_1 \sin(t\omega_1) + \ldots + a_{T/2} \cos(t\omega_{T/2}) + b_{T/2} \sin(t\omega_{T/2}), \tag{2}$$

for $t = 1, \ldots, T$. That is, we can always find values for the T parameters, $(a_j, b_j; j = 1, \ldots, T/2)$, so that the T equations, equation 2 for $t = 1, \ldots, T$, are satisfied. To see this, consider the following regression. Let the explanatory variables be:

$$X = \begin{bmatrix} \cos(\omega_1) & \sin(\omega_1) & \ldots & \cos(\omega_{T/2}) & \sin(\omega_{T/2}) \\ \cos(2\omega_1) & \sin(2\omega_1) & \ldots & \cos(2\omega_{T/2}) & \sin(2\omega_{T/2}) \\ \vdots & \vdots & & \vdots & \vdots \\ \cos(T\omega_1) & \sin(T\omega_1) & \ldots & \cos(T\omega_{T/2}) & \sin(T\omega_{T/2}) \end{bmatrix}.$$

Let the $T \times 1$ vector of "independent variables," Y, and the $T \times 1$ vector of regression coefficients, β, be

$$Y = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}, \qquad \beta = \begin{bmatrix} a_1 \\ b_1 \\ \vdots \\ a_{T/2} \\ b_{T/2} \end{bmatrix}.$$

Then, the regression is

$$Y = X\beta + u.$$

Note, however, that since the number of explanatory variables is T, the error term is exactly zero, and β is computed from $(X' X)^{-1} X' Y = X^{-1}Y^1$. Thus a time series of length T can be expressed exactly as the sum of $T/2$ simple processes like equation 1, each having a different frequency of oscillation.

Unfortunately, taken literally, equation 2 is *not* a very sensible way to think of the data. With T observations, one has only to compute the a_is and the b_is and then the $T + 1$st observation can be predicted *exactly*. No one believes that there is *any* way to use T observations on any economic data series and predict the next observation exactly. Imagine, for example, that you could do this with the Dow Jones Industrial Average. If you could, then after one minute of reading this appendix, you would have the information needed to go out and become fabulously wealthy.

Of course, one could instead suppose that the data are a realization from an expression like equation 2, in which the number of periodic functions exceeds the number of observations by, say, 10. In this case, there is no longer the implication that one can perfectly predict next period's value of x_t. However, there *is* the implication that after 20 more periods, the data series will *then* become perfectly predictable. No one would think this is a sensible way to view economic data either. The theory of spectral analysis assumes, sensibly, that no matter how many observations on x_t one accumulates, the data *never* become perfectly predictable. That is, it in effect assumes that the number of periodic functions in equation 2 is infinitely large by comparison with the size of the available data set. When this is so, equation 2 is written in the form of an integral, as follows.[2]

$$x_t = \int_0^{\pi} [a(\omega)\cos(\omega t) + b(\omega)\sin(\omega t)]d\omega, \tag{3}$$

where $a(\omega)$ and $b(\omega)$ are functions of ω. In view of these observations, it is perhaps not surprising that any covariance stationary time series process, x_t, can be expressed in the form of equation 3 (Koopmans, 1974).

Extracting the Business Cycle Component

Equation 3 allows us to make precise the notion of extracting the business cycle component of x_t. That representation views the time series process, x_t, as the sum of components with periods of oscillation $2\pi/\omega$ for ω lying in the interval 0 to π. In monthly data, the business cycle corresponds to components with period greater than 24 months and less than 96 months. In terms of frequencies of oscillation, this corresponds to ω belonging in the interval $\underline{\omega} = 2\pi/96$ to $\bar{\omega} = 2\pi/24$. Thus, we seek the business cycle component of x_t, y_t, such that

$$y_t = \int_{\underline{\omega}}^{\bar{\omega}} [a(\omega)\cos(\omega t) + b(\omega)\sin(\omega t)]d\omega. \tag{4}$$

It is well known that y_t can be computed as a particular centered moving average of observations on the observed data, x_t

$$y_t = B_0 x_t + B_1(x_{t-1} + x_{t+1}) + B_2(x_{t-2} + x_{t+2}) + \ldots, \tag{5}$$

where

$$B_j = B_{-j} = \frac{\sin(j\bar{\omega}) - \sin(j\underline{\omega})}{\pi j}, j = 1, 2, \ldots$$
$$B_0 = \frac{\bar{\omega} - \underline{\omega}}{\pi} \tag{6}$$

There is a major practical stumbling block to using equation 5 for extracting the business cycle component of x_t. It requires an infinite amount of data! Some sort of approximation is needed, if one is to estimate y_t given only the available data, $x_1, \ldots, x_T$.

An extensive analysis of this problem appears in Christiano and Fitzgerald (1998), which also provides a review of the related literature. We provide only the briefest review of that discussion here, just enough to enable us to describe exactly how we isolated the business cycle component of the data.

We denote our approximation of y_t by $\hat{y}_t$. Here, we focus on the approximations of the following form:

$$\hat{y}_t = \hat{B}_0 x_t + \hat{B}_1(x_{t-1} + x_{t+1}) + \ldots + \hat{B}_K(x_{t-K} + x_{t+K}). \tag{7}$$

That is, we approximate y_t by a finite ordered, centered, symmetric moving average. But, how should we choose the weights? The natural way is to choose them so that $\hat{y}_t$ is as close to y_t as possible, that is, so that they solve

$$\min_{\hat{B}_i,\, i=0,1,\ldots,K} E(y_t - \hat{y}_t)^2. \tag{8}$$

The solution to this problem is a function of the details of the time series representation of x_t. For example, if we suppose that x_t is a random walk, that is, $x_t - x_{t-1}$ is a process that is uncorrelated over time, then the solution is:[3]

$$\begin{aligned} \hat{B}_j &= B_j, j = 0, \ldots, K - 1 \\ \hat{B}_K &= -[\tfrac{1}{2} B_0 + B_1 + B_2 + \ldots + B_{K-1}]. \end{aligned} \tag{9}$$

Suppose the data at hand are $x_1, \ldots, x_T$, so that the objects of interest are, $y_1, \ldots, y_T$. We computed $\hat{y}_{36}, \ldots, \hat{y}_{T-36}$ as follows. For $\hat{y}_{36}$ we applied equation 7 with $K = 35$, for $\hat{y}_{37}$ we applied equation 7 with $K = 36$, and so on. For each $\hat{y}_t$ that we computed, we used the largest value of K possible. Christiano and Fitzgerald (1998) argue that this procedure for estimating y_t works well in terms of optimizing equation 8, even if the true-time series representation of x_t is not a random walk. They show that an even better approach uses an asymmetric set of weights, so that the estimate of $\hat{y}_t$ for each t uses *all* the available data on x_t.

NOTES

1. Note that $\sin(t\,\omega_{T/2}) = 0$ for all integers t. Since the right column of X is zero in this case, X is singular and so cannot be inverted. In practice, the last column of X is replaced by a column of ones, to accommodate a non-zero sample mean in x_t. Under these conditions, the columns of X are orthogonal, so that $X^{-1}Y$ is trivial to compute. In particular, for $j = 1, \ldots, T/2 - 1$:

$$a_j = \frac{2}{T}\sum_{t=1}^{T} \cos(\omega_j t)x_t,$$

$$b_j = \frac{2}{T}\sum_{t=1}^{T} \sin(\omega_j t)x_t.$$

Also,

$$a_{T/2} = [\sum_{t=1}^{T} \cos(\omega_{T/2} t)x_t]/T$$

$$b_{T/2} = [\sum_{t=1}^{T} x_t]/T.$$

2. To gain further intuition into the relationship between equations 2 and 3, it is useful to recall the simplest definition of an integral, the Riemann integral. Thus, let $f(y)$ be a function, with domain $\underline{y} \leq y \leq \bar{y}$. Let $y_t, j = 1, \ldots, M$ be a set of numbers that divide the domain into M equally spaced parts. That is, $y_1 = \underline{y} + \Delta_M, y_2 = y_1 + \Delta_M, \ldots, y_M = y_{M-1} + \Delta_M$, where $\Delta_M = (\bar{y} - \underline{y})/M$. Note that $y_M = \bar{y}$. The integral of f over its domain is written,

$$\int_{\underline{y}}^{\bar{y}} f(y)dy.$$

This is approximated by the sum of the areas of the M $f(y_j)$ by Δ_M rectangles:

$$\sum_{j=1}^{M} f(y_j)\Delta_M.$$

The Riemann interpretation of the integral is that it is the limit of the above sums, as $M \to \infty$. The relationship between the above finite sum and the integral resembles that between equations 2 and 3 if we adopt $y_j = \omega_j = 2\pi j/T$, $\Delta_M = 2\pi/T$, $M = T/2$, $f(y_j) = a(\omega_j)\cos(\omega_j t) + b(\omega_j)\sin(\omega_j t)$, $a(\omega_j) = a_j T/2\pi$, $b(\omega_j) = b_j T/2\pi$.

3. Actually, the theory as we summarized it here technically does not accommodate nonstationary processes like random walks. Christiano and Fitzgerald (1998) discuss standard ways of extending the theory to this case. Also, optimizing the mean square error criterion, equation 7, requires a constant term in equation 8. See Christiano and Fitzgerald (1998) for more details.

TECHNICAL APPENDIX 2 COMOVEMENT AND THE ELASTICITY OF SUBSTITUTION

The standard real business cycle model assumes that the elasticity of substitution between capital and labor in production is unity. In the text, we discussed a result due to Benhabib, Rogerson, and Wright (1991): With this kind of production function and with utility functions consistent with balanced growth, comovement in employment is impossible (see note 16). Here, we show that comovement is a technical possibility when the elasticity of substitution is different from unity. However, we find that comovement does not occur for plausible parameter values. These results suggest that attempts to account for comovement by adjusting the elasticity of substitution in production in a standard real business cycle model are unlikely to be successful.

We begin by describing a version of the standard real business cycle model. We assume that households have identical preferences of the following form:

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{[C_t(1-L_t)^{\psi}]^{1-\sigma}}{1-\sigma},$$

where $\sigma, \psi > 0$ satisfy the various conditions required for utility to be strictly concave. Also, $C_t > 0$ denotes per capita consumption, and L_t denotes per capita hours worked. We require $0 \leq L_t \leq 1$. The resource constraint is

$$c_t + k_{t+1} - (1-\delta)k_t \leq \left[(1-\alpha)k_t^{\frac{v-1}{v}} + \alpha L_t^{\frac{v-1}{v}}\right]^{\frac{v}{v-1}} z_t.$$

Here, $0 < \delta < 1$ is the rate of depreciation on capital and $0 < \alpha < 1$ is a parameter. The elasticity of substitution between capital and labor is $v > 0$. Also,

$$\log(z_t) = \rho\log(z_{t-1}) + \varepsilon_t, 0 < \rho < 1,$$

where ε_t is a zero mean random variable, that is independently distributed over time. Finally, $k_t > 0$ denotes the beginning of period t stock of capital, which is a given quantity at time t.

As noted in the body of the article, it is possible to interpret this as a two sector model: one in which consumption goods, c_t, and investment goods, $k_{t+1}-(1-\delta)k_t$, are produced in different sectors by different firms. It is assumed that both sectors use the same production function, the one stated above, and that capital and labor can move freely between the two sectors, subject only to the obvious constraint that the sum of capital and labor in the two sectors equals k_t and L_t, respectively. Thus, letting L_{ct}, k_{ct} and L_{it}, k_{it} denote the amount of labor and capital, respectively, used in the consumption and investment sectors, we require

$$L_{ct} + L_{it} = L_t, k_{ct} + k_{it} = k_t.$$

As mentioned in note 16, the marginal product of labor in the sector producing the consumption good equals households' marginal rate of substitution between consumption and leisure.

$$\alpha\left(\frac{C}{L_c}\right)^{\frac{1}{v}} z^{\frac{v-1}{v}} = \frac{\psi C}{(1 - L_c - L_i)}.$$

We drop the time subscripts to simplify the notation. Rearranging this equation, we obtain

$$\left(\frac{\alpha}{\psi}\right)^{v} \frac{C}{z}^{1-v} (1 - L_c - L_i)^{v} = L_c.$$

Note first that when $v = 1$, we reproduce the result in note 16, which indicates that L_c and L_i *cannot* move in the same direction. When $v \neq 1$, this reasoning no longer holds. We can see intuitively that employment in the two sectors might move together with $v >$ 1. In particular, consider the case $v = 1$. In this case, we have found in many numerical examples that C/z falls with a rise in z due to a positive shock in ε. Continuity suggests that this also happens when v is a little above unity. We conclude that if the resulting rise in $(C/z)^{1-v}$ is sufficiently large, then it is possible for both L_c and L_i to increase in response to a positive shock in ε (see note 16 for the sort of reasoning used here).

We approximated the solution to this model using the undetermined coefficient method in Christiano (1991). We assigned parameter values in the following way. For our baseline parameterization, we set $\beta = 0.99$, $\delta = 0.025$, $\rho = 0.95$, and $\sigma = 1$. We chose ψ and α to guarantee that, in the model's steady state, labor's share of income is 0.64 and steady state hours worked is 0.30. An empirical defense for the choice of values for β, δ, ρ, labor's share, and steady state hours worked may be found in the real business cycle literature. For the calculations reported below, we set $\sigma = 0.01$, and drew 1,000 observations on ε_t from a normal distribution with mean zero and standard deviation, σ_ε.

In our first experiment, we considered values of v on a grid between 0.7 and 20. For each value of v, 1,000 observations on L_{ct} and L_{it}, and

$$I_{c,t} = k_{c,t+1} - (1 - \delta)k_{c,t},\ I_{i,t} = k_{i,t+1} - (1 - \delta)k_{i,t}$$

were generated using our approximation to the model's solution. The 1,000 observations were then used to compute the correlations, r^l_{ic} between L_{ct} and L_{it} and the correlations, ρ^i_{ic}, between $I_{c,t}$ and $I_{i,t}$. A model exhibits comovement in employment and investment if both ρ^i_{ic}, $\rho^l_{ic} > 0$. We found ρ^i_{ic}, $\rho^l_{ic} < 0$ for each value of n using the benchmark parameterization.

We repeated these calculations several times, each time perturbing one, and only one, of the parameters in the benchmark parameterization. We considered the following alternatives: $\sigma = 5$; $\rho = 0.99$; $\rho = 0.0$; steady state hours equal to 0.1; steady state labor's share equal to 0.3; $\delta = 0.05$, $\delta = 0.01$; and $\beta = 0.97$, $\beta = 0.995$. The perturbations in σ, ρ, β, and δ did not produce a parameterization exhibiting comovement. The reduction in labor's share resulted in comovement in employment, but not investment, for values of v between about 3 and 13. Lowering steady state hours to 0.10 also resulted in comovement in employment but not investment. Hence, we conclude that altering the elasticity of factor substitution in production does not improve the standard real business cycle model's ability to reproduce full comovement for reasonable parameter values.

TECHNICAL APPENDIX 3 ANALYSIS OF THE INPUT-OUTPUT TABLES

Our analysis of the input–output tables is based on the 1987 benchmark, 95 variable input–output table for the U.S. economy. Our main objective here is to define the fraction of a sector's final output which is used directly or indirectly in the production of final investment goods. Let Y denote the vector of gross outputs for the production sectors of the economy. Let $A = [a_{ij}]$ be the matrix of input–output coefficients. That is, a_{ij} is the quantity of the ith industry's output used to produce one unit of the jth industry's output. Let I^f, C, G, O denote the vectors of gross private fixed investment, personal consumption expenditures, government (federal, state, and local) purchases, and "other" for each sector. Here, "other" is essentially exports minus imports. Total output, Y, is broken down into a part allocated to intermediate inputs, AY, and a part allocated to final output, $I^f + C + G + O$ as follows:

$$AY + I^f + C + G + O = Y.$$

Solving this for Y, we get

$$Y = Y_{If} + Y_C + Y_G + Y_O,$$
$$Y_i = [I - A]^{-1}\, i,\; i = I^f, C, G, O.$$

For convenience, we report Y_i, $i = I^f$, C, G, O for the 95 sectors of the U.S. economy which are included in the input–output table underlying the analysis reported in Figure 6. Table A1 reports results for the 17 sectors of the nondurable goods industry, as defined in the Hornstein and Praschnik (1997) analysis. That table reports the input–output table industry numbers that make up the industries whose name is in the middle column. Table A2 reports the numbers for the other sectors. The sum of the numbers in a row must be unity.

Table A1 Results for Consumption

I-O industry number	I-O industry title	Y_{I^f}	Y_C	Y_G	Y_O
1+2+3+4	Agriculture, forestry, and fisheries	0.060	0.894	0.052	–0.006
5+6+7+8+9+10	Mining	0.207	0.893	0.181	–0.282
14	Food and kindred products	0.018	0.962	0.041	–0.021
15	Tobacco products	0.000	0.914	0.000	0.086
16+17	Textile mill products	0.185	0.995	0.072	–0.252
18+19	Apparel and other textile products	0.037	1.284	0.041	–0.362
24+25	Paper and allied products	0.103	0.833	0.112	–0.047
26A+26B	Printing and publishing	0.058	0.795	0.121	0.026
27A+27B+28	Chemicals and allied products	0.180	0.698	0.138	–0.016
31	Petroleum refining and related products	0.105	0.782	0.145	–0.032
32	Rubber and miscellaneous plastics products	0.246	0.761	0.127	–0.134
33+34	Footwear, leather, and leather products	0.031	2.154	0.037	–1.222
65A+ … +68C	Transportation, communications, and utilities	0.107	0.740	0.123	0.029
69A	Wholesale trade	0.232	0.589	0.098	0.082
69B	Retail trade	0.066	0.919	0.015	0.000
70A+70B+71A+71B	Finance, insurance, and real estate	0.061	0.877	0.043	0.020
72A+ … +77B	Services	0.076	0.879	0.044	0.002

Notes: Y_{I^f} measures amount of gross output of industry in indicated row sent directly or indirectly to industry i, where $i = I^f$, C, G, O. Row numbers are scaled so they sum to unity.

Source: Authors' calculations based on data from U.S. Department of Commerce. Bureau of Economic Analysis, 1992. *Survey of Current Business*. Volume 12, Number 4, April.

Table A2 Results for Nonconsumption

I-O industry number	I-O industry title	Y_i	Y_p	Y_o	Y_a
11	New construction	0.805	0.000	0.195	0.000
12	Maintenance and repair construction	0.180	0.574	0.243	0.002
13	Ordnance and accessories	0.011	0.051	0.838	0.100
20+21	Lumber and wood products	0.542	0.340	0.169	−0.051
22+23	Furniture and fixtures	0.477	0.585	0.067	−0.128
29A	Drugs	0.018	0.963	0.125	−0.107
29B	Cleaning and toilet preparations	0.018	0.949	0.033	0.000
30	Paints and allied products	0.422	0.445	0.168	−0.036
35	Glass and glass products	0.202	0.798	0.116	−0.116
36	Stone and clay products	0.575	0.314	0.206	−0.095
37	Primary Iron and steel manufacturing	0.515	0.501	0.207	−0.223
38	Primary nonferrous metals manufacturing	0.400	0.485	0.247	−0.132
39	Metal containers	0.057	0.891	0.064	−0.012
40	Heating, plumbing, and fabricated structural metal products	0.604	0.201	0.198	−0.002
41	Screw machine products and stampings	0.358	0.641	0.132	−0.131
42	Other fabricated metal products	0.377	0.573	0.175	−0.124
43	Engines and turbines	0.377	0.362	0.246	0.015
44+45	Farm, construction, and mining machinery	0.731	0.151	0.097	0.021
46	Materials handling machinery and equipment	0.876	0.134	0.105	−0.115
47	Metalworking machinery and equipment	0.779	0.261	0.108	−0.147
48	Special industry machinery and equipment	0.962	0.154	0.028	−0.145
49	General industrial machinery and equipment	0.729	0.305	0.130	−0.164
50	Miscellaneous machinery, except electrical	0.309	0.438	0.258	−0.006
51	Computer and office equipment	0.786	0.148	0.156	−0.090
52	Service industry machinery	0.636	0.289	0.120	−0.045
53	Electrical industrial equipment and apparatus	0.639	0.308	0.168	−0.114
54	Household appliances	0.242	0.842	0.045	−0.129
55	Electric lighting and wiring equipment	0.471	0.447	0.185	−0.104
56	Audio, video, and communication equipment	0.626	0.564	0.206	−0.396
57	Electronic components and accessories	0.338	0.437	0.322	−0.097
58	Miscellaneous electrical machinery and supplies	0.321	0.687	0.148	−0.156
59A	Motor vehicles (passenger cars and trucks)	0.478	0.776	0.051	−0.304
59B	Truck and bus bodies, trailers, and motor vehicles parts	0.437	0.746	0.080	−0.263
60	Aircraft and parts	0.134	0.049	0.546	0.270
61	Other transportation equipment	0.145	0.543	0.336	−0.024
62	Scientific and controlling instruments	0.442	0.166	0.372	0.020
63	Ophthalmic and photographic equipment	0.347	0.590	0.228	−0.165
64	Miscellaneous manufacturing	0.175	1.121	0.071	−0.368
78	Federal government enterprises	0.079	0.814	0.104	0.003
79	State and local government enterprises	0.033	0.928	0.043	−0.003
80	Noncomparable imports	0.000	0.000	0.000	0.000
81	Scrap, used, and secondhand goods	−9.699	7.493	1.830	1.377
82	General government industry	0.000	0.000	1.000	0.000
83	Rest of the world adjustment to final uses	0.000	0.000	0.000	0.000
84	Household industry	0.000	1.000	0.000	0.000
85	Inventory valuation adjustment	0.000	0.000	0.000	1.000

Notes: Y_i measures amount of gross output of industry in indicated row sent directly or indirectly to industry *y i*, where *i* = *P. C. G. O.* Row numbers are scaled so they sum to unity

Source: Authors' calculations based on U.S. Department of Commerce. Bureau of Economic Analysis, 1992 *Survey of Current Business*. Volume 12, Number 4, April.

REFERENCES

Alexopoulos, M. 1998. Efficiency wages, unemployment, and the business cycle. Northwestern University, manuscript.

Alexopoulos, M. 1997. Shirking in a monetary business cycle model. Northwestern University, manuscript.

Azariadis, C. 1981. Self-fulfilling prophesies. *Journal of Economic Theory,* Vol. 25, pp. 380–396.

Banerjee, A. 1992. A simple model of herd behavior. *Quarterly Journal of Economics,* Vol. 107, pp. 797–817.

Baxter, M. 1996. Are consumer durables important for business cycles? *Review of Economics and Statistics,* Vol. 78, No. 1, February, pp. 147–155.

Benhabib, J., Rogerson, R. and Wright, R. 1991. Homework in macroeconomics: Household production and aggregate fluctuations. *Journal of Political Economy,* Vol. 99, pp. 1166–1187.

Benhabib, J., and Farmer, R. E. A. 1996. Indeterminacy and sector-specific externalities. *Journal of Monetary Economics,* Vol. 37, No. 3, June, pp. 421–443.

Benhabib, J., and Farmer, R. E. A. 1994. Indeterminacy and growth. *Journal of Economic Theory,* Vol. 63, pp. 19–41.

Bikhchandani, S., Hirshleifer, D., and Welch, I. 1992. A theory of fads, fashion, custom, and cultural change as information cascades. *Journal of Political Economy,* Vol. 100, No. 5, October, pp. 992–1026.

Boldrin, M., Christiano, L. J., and Fisher, J. D. M. 1997. Habit persistence and asset returns in an exchange economy. *Macroeconomic Dynamics,* Vol. 1, No. 2, pp. 312–332.

Boldrin, M., Christiano, L. J., and Fisher, J. D. M. 1995. Asset pricing lessons for modeling business cycles. National Bureau of Economic Research, working paper, No. 5262.

Burns, A. and Mitchell, W. C. 1946. *Measuring Business Cycles, Studies in Business Cycles, No. 2.* New York: National Bureau of Economic Research.

Bryant, J. 1983. A simple rational expectations Keynes-type model. *Quarterly Journal of Economics,* Vol. 98, No. 3, pp. 525–528.

Caplin, A. and Leahy, J. 1994. Business as usual, market crashes, and wisdom after the fact. *American Economic Review,* Vol. 84, No. 3, pp. 548–565.

Cass, D. and Shell, K. 1983. Do sunspots matter? *Journal of Political Economy,* Vol. 91, pp. 193–227.

Chamley, C. and Gale, D. 1994. Information revelation and strategic delay in a model of investment. *Econometrica,* Vol. 62, No. 5, September, pp. 1065–1085.

Chari, V. V. and Kehoe, P. J. 1997. Hot money. Federal Reserve Bank of Minneapolis, staff report, No. 228.

Christiano, L. J. and Fisher, J. D. M. 1998. Stock market and investment good prices: Implications for macroeconomics. Federal Reserve Bank of Chicago, working paper, July, available on the Internet at www.econ.nwu.edu/faculty.html.

Christiano, L. J. and Fisher, J. D. M. 1995. Tobin's q and asset returns: Implications for business cycle analysis. National Bureau of Economic Research, working paper, No. 5292.

Christiano, L. J. and Fitzgerald, T. 1998. The band pass filter: Optimal approximations. Northwestern University, manuscript in process, available on the Internet at www.econ.nwu.edu/faculty.html.

Christiano, L. J. and Harrison, S. 1999. Chaos, sunspots, and automatic stabilizers. *Journal of Monetary Economics,* Vol. 44, No. 1, August, pp. 3–31.

Constantinides, G. 1990. Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy,* Vol. 98, pp. 519–542.

Cooper, R. and John, A. 1988. Coordinating coordination failures in Keynesian models. *Quarterly Journal of Economics,* Vol. 103, August, pp. 441–463.

Cooper, R. and Haltiwanger, J. 1996. Evidence on macroeconomic complementarities. *Review of Economics and Statistics,* Vol. 78, No. 1, February, pp. 78–93.

Cooper, R. and Haltiwanger, J. 1990. Inventories and the propagation of sectoral shocks. *American Economic Review,* Vol. 80, No. 1, March, pp. 170–190.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. 1998. Investor psychology and security market under- and overreactions. Northwestern University, Kellogg Graduate School, manuscript, October 12.

Danthine, J.-P. and Donaldson, J. B. 1995. Non-Walrasian economies. In *Frontiers of Business Cycle Research,* Thomas F. Cooley (ed.). Princeton, NJ: Princeton University Press, chapter 8.

Diamond, P. 1982. Aggregate demand management in search equilibrium. *Journal of Political Economy,* Vol. 90, No. 5, pp. 881–894.

Dupor, W. 1998. Aggregate fluctuations and production complementarities. University of Pennsylvania, Wharton School, manuscript, July.

Einarsson, T. and Marquis, M. M. 1997. Home production with endogenous growth. *Journal of Monetary Economics,* Vol. 39, pp. 551–569.

Farmer, R. E. A. and Woodford, M. 1997. Self-fulfilling prophecies and the business cycle. *Macroeconomic Dynamics,* Vol. 1, No. 4, pp. 740–769.

Farmer, R. E. A. and Guo, J.-T. 1995. The econometrics of indeterminacy: An applied study. UCLA Working Paper Series: 720 October 1994.

Farmer, R. E. A. and Guo, J.-T. 1994. Real business cycles and the animal spirits hypothesis. *Journal of Economic Theory,* Vol. 63, pp. 42–73.

Gali, J. 1994. Monopolistic competition, business cycles, and the composition of aggregate demand. *Journal of Economic Theory,* Vol. 63, pp. 73–96.

Gomme, P. 1999. Shirking, unemployment, and aggregate fluctuations. *International Economic Review,* Vol. 40, No. 1, February, pp. 3–21.

Hansen, G. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics,* Vol. 16, pp. 309–327.

Heaton, J. 1995. An empirical investigation of asset pricing with temporally dependent preference specifications. *Econometrica,* Vol. 63, pp. 681–717.

Hornstein, A. and Praschnik, J. 1997. Intermediate inputs and sectoral comovement in the business cycle. *Journal of Monetary Economics,* Vol. 40, No. 3, pp. 573–595.

Horvath, M. 1998a. Sectoral shocks and aggregate fluctuations. Stanford University, Department of Economics, manuscript, February.

Horvath, M. 1998b. Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. Stanford University, Department of Economics, manuscript, April.

Huffman, G. and Wynne, M. 1999. The role of intratemporal adjustment costs in a multisector economy. *Journal of Monetary Economics,* Vol. 43, No. 2, April, pp. 317–350.

King, R. Plosser, C. and Rebelo, S. 1988. Production, growth, and business cycles: I. The basic neoclassical model. *Journal of Monetary Economics,* Vol. 21, No. 213, pp. 195–232.

King, R. G. and Watson, M. W. 1996. Money, prices, interest rates, and the business cycle. *The Review of Economics and Statistics,* February, pp. 35–53.

Kocherlakota, N. 1996. The equity premium: It's still a puzzle. *Journal of Economic Literature,* Vol. 34, No. 1, March, pp. 42–71.

Koopmans, L. 1974. *The Spectral Analysis of Time Series.* New York: Academic Press.

Kydland, F. E. and Prescott, E. C. 1982. Time to build and aggregate fluctuations. *Econometrica,* Vol. 15, No. 6, pp. 1345–1370.

Long, J. and Plosser, C. 1983. Real business cycles. *Journal of Political Economy,* Vol. 91, pp. 39–69.

Lucas, R. E. Jr. 1981. Understanding business cycles. In *Studies in Business-Cycle Theory,* R. E. Lucas, Jr. (ed.). Boston: MIT Press. Reprinted from K. Brunner and A. H. Meltzer (eds.). 1977. *Stabilization of the Domestic and International Economy,* Vol. 5, *Carnegie-Rochester Series on Public Policy.* Amsterdam: North-Holland Publishing Company.

Murphy, K. M., Shleifer, A. and Vishny, R. W. 1989. Building blocks of market clearing business cycle models. *NBER Macroeconomics Annual 1989,* Olivier J. Blanchard and Stanley Fischer (eds.). Boston: MIT Press.

Prescott, E. 1991. Real business cycle theory: What have we learned? *Revista de Analisis Economico,* Vol. 6, No. 2, November, pp. 3–19.

Prescott, E. 1986. Theory ahead of business cycle measurement. *Carnegie-Rochester Conference Series on Public Policy,* Vol. 25, No. 0, Autumn, pp. 11–44.

Romer, D. 1996. *Advanced Macroeconomics.* New York: McGraw-Hill.

Rotemberg, J. and Woodford, M. 1992. Oligopolistic pricing and the effects of aggregate demand on economic activity. *Journal of Political Economy,* Vol. 100, December, pp. 1153–1207.

Rupert, P., Rogerson, R. and Wright, R. 1995. Estimating substitution elasticities in household production costs. *Economic Theory,* Vol. 6, No. 1, June, pp. 179–193.

Sargent, T. 1979. *Macroeconomic theory.* New York: Academic Press.

Schmitt-Grohe, S. 1997. Comparing four models of aggregate fluctuations due to self-fulfilling expectations. *Journal of Economic-Theory,* Vol. 72, No. 1, January, pp. 96–147.

Shapiro, C. and Stigliz, J. E. 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review,* Vol. 74, June, pp. 443–444.

Shleifer, A. 1986. Implementation cycles. *Journal of Political Economy,* Vol. 94, pp. 1163–1190.

Sundaresan, S. M. 1989. Intertemporally dependent preferences and the volatility of consumption and wealth. *Review of Financial Studies,* Vol. 2, pp. 73–89.

Woodford, M. 1991. Self-fulfilling expectations and fluctuations in aggregate demand. In *New Keynesian Economics, Vol. 2, Coordination Failures and Real Rigidities,* N. Gregory Mankiw and David Romer (eds.). Boston: MIT Press, pp. 77–110.

Woodford, M. 1988. Expectations, finance, and aggregate instability. In *Finance Constraints, Expectations, and Macroeconomics,* M. Kohn and S-C. Tsiang (eds.). New York: Oxford University Press, pp. 230–261.

Woodford, M. 1987. Three questions about sunspot equilibria as an explanation of economic fluctuations. *American Economic Review Papers and Proceedings,* Vol. 77, No. 2, May, pp. 93–98.

Woodford, M. 1986. Stationary sunspot equilibria in a finance constrained economy. *Journal of Economic Theory,* Vol. 40, No. 1, October.

3

Changes in the Business Cycle

Carl E. Walsh

Professor of Economics, University of California, Santa Cruz, California

In December 1998, the current expansion reached a milestone—it became the longest peacetime expansion in post–World War II U.S. economic history, surpassing the record previously held by the 1982–1990 expansion. In fact, if the expansion continues through January 2000, it will tie the expansion associated with the Vietnam War as the longest expansion since our records of such things start in 1854.

The experience of the United States during the last twenty years has been quite remarkable. The long economic expansion of the 1980s was followed by a relatively short recession in 1990–91, and the economy has been expanding ever since. The United States has experienced only 8 months of recession in the last 16 years. The most visible sign of the continued expansion is provided by the unemployment rate. For the past year, it has remained below 4.5 percent, hovering at levels not seen since the early 1970s.

Not surprisingly, the long expansion has raised questions about the whole notion of the business cycle. Extended periods of expansion always lead a few commentators to speculate that the conventional business cycle is dead. In 1969, for example, a conference volume titled "Is the Business Cycle Obsolete?" was published just as the 1961–69 expansion came to an end and the economy entered a recession. With two record-setting expansions in a row, and the current one still going, it is to be expected that the notion of regular business cycles is again being questioned. The current favorite hypothesis is that a "new economy" has emerged in which our old understanding of business cycle forces is no longer relevant.

While few economists believe we have seen the end of business cycles (just look at Asia and Latin America!), the views of economists about business cycles have changed. These changes reflect real changes in the U.S. economy, changes in our ability to measure economic developments, and changes in economic theory.

DATING BUSINESS CYCLES

Although virtually all data used to analyze the U.S. economy are produced by some agency of the federal government, the standard dates identifying business cycle peaks and troughs

Reprinted from: *FRBSF Economic Letter,* No. 99-16 (May 14, 1999).

are determined by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER). The NBER is a private, non-profit research organization whose research affiliates include many of the world's most influential economists.

The NBER defines a recession as "a recurring period of decline in total output, income, employment, and trade, usually lasting from six months to a year, and marked by widespread contractions in many sectors of the economy." Recessions are, therefore, *macroeconomic* in nature. A severe decline in an important industry or sector of the economy may involve great hardships for the workers and firms in that industry, but a recession is more than that. It is a period in which many sectors of the economy experience declines. Recessions are sometimes said to occur if total output declines for two consecutive quarters. However, this is not the formal definition used by the NBER.

Business cycle peaks and troughs cannot be identified immediately when they occur for two reasons. First, recessions and expansions are, by definition, recurring periods of either decline or growth. One quarter of declining GDP would not necessarily indicate that the economy had entered a recession, just as one quarter of positive growth need not signal that a recession had ended. The recession of 1981–82 provides a good example. Real GDP declined from the third quarter of 1981 to the fourth quarter, and then again from the fourth quarter to the first quarter of 1982. It then grew in the second quarter of 1982. The recession was not over, however, as GDP again declined in the third quarter of 1982. Only beginning with the fourth quarter did real output begin a sustained period of growth.

Second, the information that is needed to determine whether the economy has entered a recession or moved into an expansion phase is only available with a time lag. Delays in data collection and revisions in the preliminary estimates of economic activity mean the NBER must wait some time before a clear picture of the economy's behavior is available. For example, it was not until December 1992 that the NBER announced that the trough ending the last recession had occurred in March 1991, a delay of 20 months.

EXPANSIONS AND CONTRACTIONS SINCE 1854

U.S. business cycle peaks and troughs going back to the trough in December 1854 have been dated by the NBER. Based on their dates, we can ask whether basic business cycle facts have changed over time.

One important aspect of a recession or an expansion is its duration. The lengths of recessions since 1854 are shown in Figure 1. Several interesting facts are apparent from the figure. First, measured solely by duration, the Great Depression of 1929–1933 pales in comparison with the 1873–1879 depression that lasted over five years. And the 1882–1885 recession lasted nearly as long as the Great Depression. Some lasting images of American history survive from this period, including the great debate over silver coinage.

Second, while the Great Depression was not the longest period of economic decline, it does appear to represent a watershed; no recession since has lasted even half as long as the 1929–1933 contraction.

Third, it is not just that recessions have been shorter on average in the post–World War II era, they have *all* been much shorter. Of the 19 recessions before the Great Depression, only three lasted less than a year; of the 11 recessions since the Great Depression, only three have lasted more than a year.

Figure 2 shows the duration of economic expansions since 1854. Darker bars mark wartime expansions. Based on duration, the changing nature of expansions is not quite as

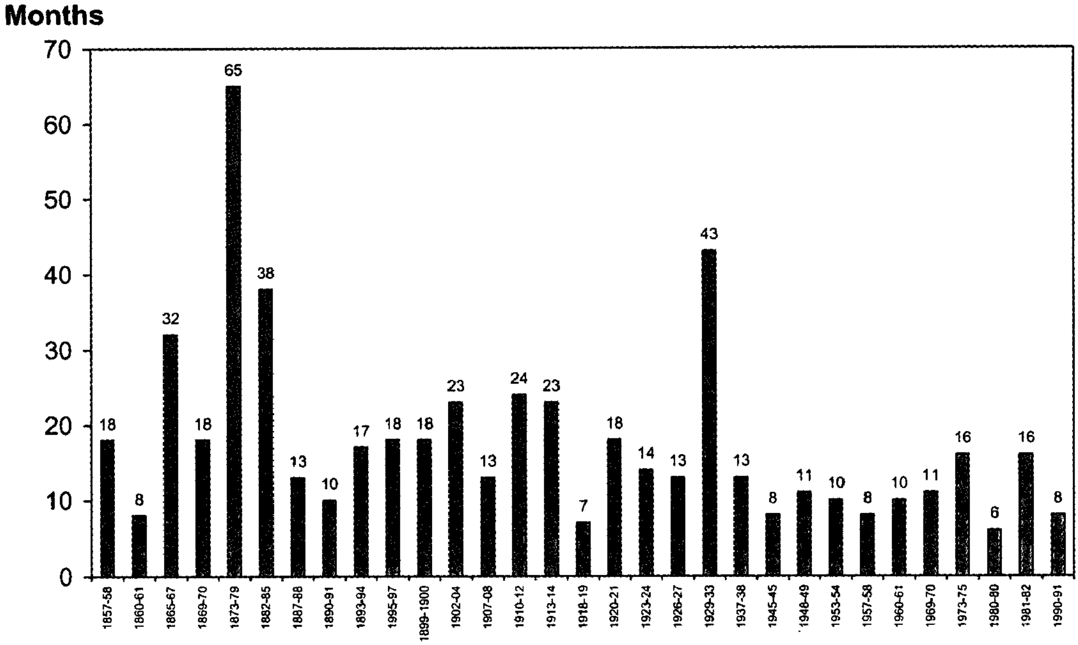


Figure 1 U.S. contractions, 1857–1999.

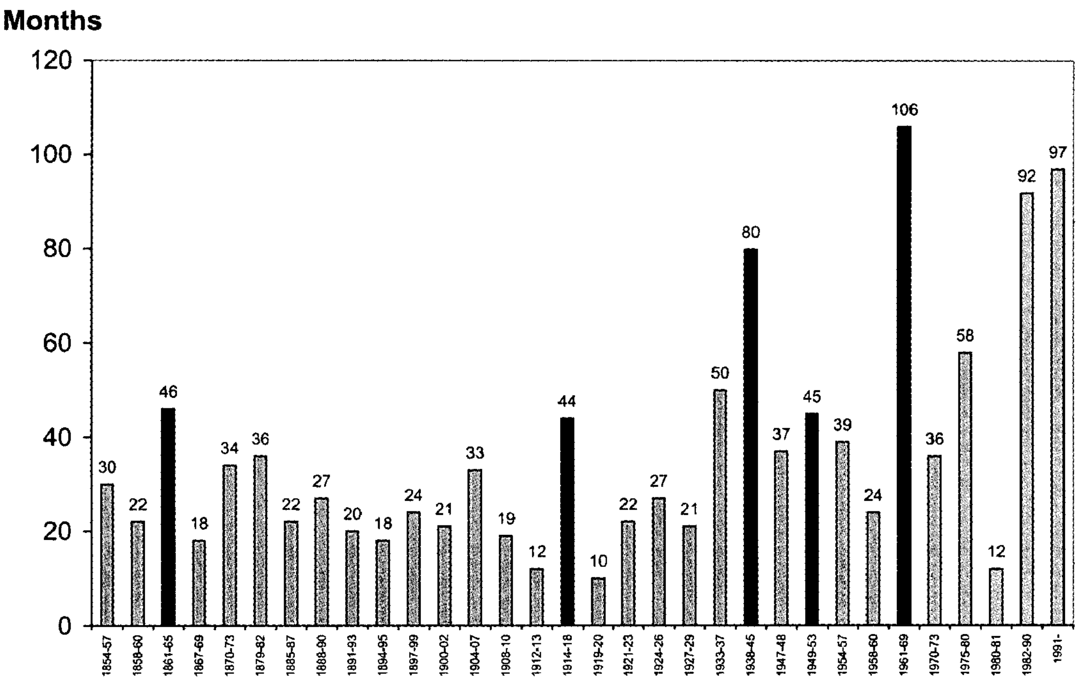


Figure 2 U.S. economic expansions, 1854–1999.

evident as for contractions. But of the 21 expansions prior to World War II, only three lasted more than three years. In contrast, of the 10 expansions since, only three have lasted less than three years. Even if the wartime expansions associated with Korea and Vietnam are ignored, post–World War II expansions have averaged 49 months, compared to an average of only 24 months for pre–World War II peacetime expansions.

IS THE ECONOMY MORE STABLE?

A simple comparison of the duration of expansions and contractions does suggest the U.S. economy has performed better in the post–World War II era. Recessions are shorter, expansions are longer. These changes strongly suggest that business cycles have changed over time. However, a simple comparison of duration cannot tell us about the severity of recessions or the strength of expansions. This would be better measured by the decline in output that occurs in a recession or the growth that occurs in an expansion. However, most studies that examine how volatile economic activity has been do conclude that output has been somewhat more stable in the post–World War II era.

This conclusion, however, is not universally accepted. There are three reasons that comparing the business cycle over time is difficult.

First, the quality of economic data has improved tremendously over the past 100 years. If the earlier data on the U.S. economy contained more measurement error because the quality of our statistics was lower, the measured path of the economy may show some fluctuations that simply reflect random errors in output data. This will make the earlier period look more unstable. In addition, earlier data on economic output tended to provide only a partial coverage of the economy. For example, better statistics were available on industrial output than on services. Since services tend to fluctuate less over a business cycle, the earlier data undoubtedly exaggerated the extent of fluctuations in the aggregate economy.

Second, NBER dating methods have not remained consistent. Romer (1994) argues that the dating of pre–World War II business cycles was done in a manner that tended to date peaks earlier and troughs later than the post–World War II methods would have done. This contributes to the impression that prewar recessions were longer and expansions shorter.

Third, the economy is increasingly becoming a producer of services, and productivity in the service sector is often difficult to measure. In general, the tremendous changes experienced in recent years associated with the information revolution are likely to affect the cyclical behavior of the economy in ways not yet fully understood.

IMPLICATIONS FOR MACROECONOMIC POLICY

Understanding changes in the nature of the business cycle is important for policymakers. Most central banks view contributing to a stable economy as one of their responsibilities. Promoting stable growth has important benefits, and reducing the frequency or severity of recessions is desirable as part of a policy to ensure employment opportunities for all workers. Preventing expansions from generating inflation is also important since once inflation gets started, high unemployment is usually necessary to bring it back down.

One might think, then, that policy designed to stabilize the economy should attempt to eliminate fluctuations entirely. This is not the case, for a very important reason. A business cycle represents fluctuations in the economy around full-employment output, but an economy's full-employment output, often called potential GDP, can also change. It grows over time due to population growth, growth in the economy's capital stock, and technological change. Developments in economic theory have led to a better understanding of how an economy adjusts to various disturbances. These adjustments can cause potential GDP to fluctuate, and it would be inappropriate for policy to attempt to offset these fluctuations. Identifying fluctuations in potential GDP from cyclical fluctuations can be difficult, however, as the current economic expansion illustrates. Is the economy in danger of overheating, risking a revival of inflation? Or have changes in the economy increased potential GDP?

While the U.S. economy has enjoyed two consecutive record expansions, a longer historical perspective does help to remind us that business cycles are unlikely to be gone for good. Despite talk of the "new economy," all economies experience ups and downs that are reflected in swings in unemployment, capacity utilization, and overall economic output. Though changes in the structure of the economy may alter the extent of these fluctuations, they are unlikely to eliminate them.

In addition, the business cycle record is not independent of policy decisions. The economy may not have changed fundamentally; perhaps we have simply benefited from good economic policy (see Taylor 1998 for a discussion along these lines). With less successful policies, recessions could become more frequent and longer again. The Great Depression, for example, was prolonged by, among other things, poor economic and monetary policy decisions, and the recessions of the early 1980s were the price of policy mistakes in the 1970s that allowed inflation to rise significantly (Romer 1999). Thus, one reason business cycles can change, even if the underlying economy or source of disturbances haven't, is because policymakers do a better (or worse) job of stabilizing the economy.

REFERENCES

Romer, C. 1999. Changes in Business Cycles: Evidence and Explanations. NBER Working Paper 6948.

Romer, C. 1994. "Remeasuring Business Cycles." *Journal of Economic History* 54, pp. 573–609.

Taylor, J. B. 1998. Monetary Policy and the Long Boom. Federal Reserve Bank of St. Louis *Review* (November/December).

This Page Intentionally Left Blank

4
Why Is Financial Stability a Goal of Public Policy?

Andrew Crockett
General Manager, Bank for International Settlements, Basel, Switzerland

A number of developments in recent years have combined to put the issue of financial stability at the top of the agenda, not just of supervisory authorities, but of public policymakers more generally. These developments include: the explosive growth in the volume of financial transactions, the increased complexity of new instruments, costly crises in national financial systems, and several high profile mishaps at individual institutions.

The growth in the volume of financial transactions and the increasing integration of capital markets have made institutions in the financial sector more interdependent and have brought to the fore the issue of systemic risk. International capital flows, though generally beneficial for the efficient allocation of savings and investment, now have the power in unstable conditions to undermine national economic policies and destabilize financial systems.

The increased complexity of new instruments makes it harder for senior management in financial firms, let alone supervisory authorities, to understand intuitively the risks to which the institutions concerned are exposed. There are fears that the models underlying the pricing of the new instruments may not be sufficiently robust, that the mathematics of the models may have become disconnected from the realities of the marketplace, or that the operational controls within financial institutions may be inadequate to control the resultant risks.

The crises in financial systems that have occurred have demonstrated the close linkages between financial stability and the health of the real economy. In Mexico, for example, what began as a currency crisis led to a serious recession and created huge strains in the banking system, further deepening the recession. The consequences of the Mexican crisis destabilized several other Latin American countries, notably Argentina, and threatened for a while to have even wider repercussions. In industrial countries, financial strains in Scandinavia and Japan, among others, had adverse consequences for the real economy.

Reprinted from: *Federal Reserve Bank of Kansas City Economic Review* (Fourth Quarter 1997) 5–22. The views expressed are those of the author and do not necessarily reflect those of the BIS or the Federal Reserve System.

Lastly, there have been a number of well-publicized losses at individual institutions, due to the breakdown of operational or other controls. Episodes such as Drexel Burnham, Procter & Gamble, Orange County, Metallgesellschaft, Barings, Daiwa, and Sumitomo, though reasonably well contained, demonstrate how quickly losses can mount, and illustrate the systemic risks that would be inherent in a larger scale mishap.

The central case for making the health of the financial system a public policy concern rests on two propositions: firstly, that, left to itself, the financial system is prone to bouts of instability; and secondly, that instability can generate sizable negative spillover effects (externalities). It will be the purpose of this paper to examine these propositions more closely, and in the light of this examination, to consider what forms public policy intervention in the financial sector might take. More specifically, I will address the following questions: what do we mean by financial stability? Why should official intervention (as opposed to reliance on market forces) be required to promote stability? And what concrete approaches can be employed?

WHAT IS FINANCIAL STABILITY?

A distinction is commonly made nowadays between monetary stability and financial stability (interestingly, this distinction would not have been so easily recognized a generation ago, either by economists or public officials). Monetary stability refers to the stability of the general price level; financial stability to the stability of the key institutions and markets that go to make up the financial system. While these are conceptually separate objectives of policy, the linkages between the two are now increasingly recognized.[1]

The debate on monetary stability has progressed further and its definition has reached a greater degree of consensus than is the case with financial stability. Nobody disputes that the avoidance of excessive inflation is an appropriate objective. And nobody doubts that it is public policy (specifically, monetary policy) that ultimately determines the inflation rate. Remaining debates surround issues such as how to accurately measure inflation; what, within a relatively narrow range (usually 1–3 percent), should be considered an optimal inflation rate; whether the objective should be expressed in terms of the inflation rate or the price level, and how quickly one should return to price stability after having been forced away from it.[2]

No such general consensus applies in the case of the definition of financial stability. For the time being, at least, each writer can supply his own. In my case, I will take financial stability to apply to both institutions and markets. In other words, stability requires (i) that the key *institutions* in the financial system are stable, in that there is a high degree of confidence that they can continue to meet their contractual obligations without interruption or outside assistance; and (ii) that the key *markets* are stable, in that participants can confidently transact in them at prices that reflect fundamental forces and that do not vary substantially over short periods when there have been no changes in fundamentals.

This does not, however, provide a full definition. Which are the "key institutions" whose stability is important? And what is the degree of price stability in financial markets that is required?

Stability in financial institutions means the absence of stresses that have the potential to cause measurable economic harm beyond a strictly limited group of customers and counterparties. Occasional failures of smaller institutions, and occasional substantial losses at larger institutions, are part and parcel of the normal functioning of the financial system. Indeed, they serve a positive function by reminding market participants of their obligation to exercise discipline over the activities of the intermediaries with whom they do business.

Similarly, stability in financial markets means the absence of price movements that cause wider economic damage. Prices can and should move to reflect changes in economic fundamentals. And the prices of assets can often move quite abruptly when something happens to cause a reassessment of the future stream of income associated with the asset, or the price at which this income stream should be discounted. It is only when prices in financial markets move by amounts that are much greater than can be accounted for by fundamentals, and do so in a way that has damaging economic consequences, that one is justified in talking about "instability" or "crisis" in the financial system.

A practical issue that is worth addressing at this point is whether all financial institutions and all markets should be treated similarly. Are problems in the banking sector to be considered in the same light as problems at nonbank financial institutions? Is the failure of a big bank the same as that of a small bank? And should central banks be as concerned about excessive volatility in asset prices as they are about instability among financial institutions? These are issues that have been, and remain, controversial.

Consider first the question of which institutions are important for financial stability. This raises two further issues: are banks special? And are some institutions "too big to fail"? Two reasons are usually given for believing that banks warrant special treatment in the preservation of financial stability.[3] The first is that banks' liabilities are repayable at par on demand, while their assets are typically comparatively illiquid. This makes them more liable to runs that cause illiquidity and even insolvency. The second is that banks remain responsible for the operation of the payment system. This means that difficulties at one institution are transmitted, semi-automatically, to the rest of the financial system, with the risk, at the extreme, that the payments system could seize up.

Both of these reasons continue to have force, though perhaps not to the same extent as previously. While illiquid loans remain a disproportionate share of banks' assets, holdings of marketable securities have tended to increase. And the "moneyness" of banks' liabilities may have become less of a distinguishing characteristic, as banks increase their reliance on marketable claims to meet funding requirements, and nonbank institutions issue liabilities that are repayable on demand. Banks continue to dominate the payments system, and the failure of one bank immediately generates losses to those banks exposed to it in the settlement system. Cascading losses through these arrangements have the potential to undermine the payments systems, which is the basis for monetary exchange in all economies. But interlocking claims and settlement exposures among other entities at the core of the financial system have grown sizably as nonbank financial intermediaries have come to greater prominence. These have increased the potential for knock-on effects among them.

The conclusion is that banks remain "special," in that instability in the banking system has a greater capacity to generate systemic contagion than difficulties elsewhere in the financial sector. But the distinctions are becoming more blurred, with problems at key nonbank institutions having growing potential for significant spillover consequences.

In many respects size has become more important than an institution's formal character in determining its systemic significance. Regulators frequently deny that there is a "too big to fail" doctrine. One can understand why they do, since to make it explicit would court moral hazard. Still, it is only realistic to recognize that certain institutions are so central to the financial system that their failure would constitute a systemic crisis. Their obligations to counterparties are so large that failure to discharge them would cause widespread contagion. This group of institutions includes both banks and nonbanks.

Next, what about price volatility in asset markets? How much price movement can take place before we should classify markets as being "unstable"? And which markets are of particular concern for the health of the financial system and the economy more generally?

There are obviously no hard-and-fast answers to these questions. Any price movements that exceed what can be justified on grounds of changing fundamentals have the potential to result in resource misallocation. Sustained price volatility that generates uncertainty, leading to an unwillingness to enter into long-term contracts, hampers economic performance through discouraging the mobilization and allocation of savings through the financial system. And sudden or sharp price movements that place the liquidity or solvency of prudently run financial institutions at risk have more immediate dangers.

As to which markets should be the focus of concern, once again the criterion should be the capacity to cause wider economic damage. Financial and other asset markets, because of their broad linkages to saving and investment decisions, obviously have a greater potential impact on other macroeconomic variables than do developments in markets for goods and services. This impact can occur through wealth effects, as the prices of financial assets change; through changing the expected returns on savings and investment; or through generalized effects on consumer and business confidence.

A further point concerns the capacity for contagion among financial markets. Just as difficulties at one financial intermediary appear to have the effect of undermining confidence more generally, so experience suggests that sharp movements in one market can destabilize others. Examples of this phenomenon include the broadly similar movements in international equity prices in 1987, following the price break on Wall Street; the general upward movement in bond yields in 1994 (see Table 1); and the spread of exchange rate difficulties in Europe in 1992–93 and in Southeast Asia in 1997.

In conclusion, there is still no clear-cut definition of what constitutes financial instability. What may distinguish the financial system from other areas of economic activity,

Table 1 Equity Prices in 1987 and Bond Yields in 1994

	Equity price movements in 2 weeks of October 1987*	Bond yield rise end-January through end-July 1994 (basis points)†
United States	–20.2	142
Japan	–12.2	89
Germany	–14.2	142
France	–16.7	159
United Kingdom	–24.8	236
Italy	–11.3	235
Canada	–18.5	297
Netherlands	–18.9	124
Belgium	–10.7	156

* 9th to 23rd October 1987.
† Ten-year benchmark.
Sources: National sources.

however, is the potential for healthy flexibility to develop—in a short period of time—into more troublesome instability and eventually, in extreme circumstances, into crisis. This is because precautionary action taken by individuals in the face of asymmetric information can in certain circumstances have the effect of amplifying, rather than dampening, natural volatility. This potential brings us closer to an understanding of why the maintenance of stability is often considered to be a natural responsibility of public authorities.

Assessing the point at which movements in asset prices, or in the financial position of intermediaries, risk becoming self-perpetuating is obviously a matter of judgment. Because the costs of mistakes are so high, it is of key importance to understand the dynamics of the process. It is also important to come to an assessment of the ways in which the financial instability interacts with the real economy to intensify (or moderate) an initial shock. It is for this reason that, whatever the specific arrangements in place in any country to monitor or underwrite the health of individual institutions, there needs to be close cooperation between the authorities responsible for the supervision of individual institutions, those responsible for broader systemic stability, and those concerned with stability in prices and the real economy.

WHY IS OFFICIAL INTERVENTION REQUIRED TO PROMOTE STABILITY?

There can be little doubt that financial stability, properly defined, is a "good thing." It creates a more favorable environment for savers and investors to make intertemporal contracts, enhances the efficiency of financial intermediation and helps improve allocation of real resources. It provides a better environment for the implementation of macroeconomic policy. Instability, on the other hand, can have damaging consequences, from the fiscal costs of bailing out troubled institutions to the real GNP losses associated with banking and currency crises.

The only qualification to be made is that stability must not be confused with rigidity. Market prices must be allowed to move as supply and demand conditions change. And financial institutions should not be prevented from going out of existence when they are unable to make a profit. The trick is to permit the necessary flexibility in market prices and structures, without generating instability that has damaging consequences on confidence and real economic activity.

Financial stability is a public good in that its "consumers" (i.e., users of financial services) do not deprive others of the possibility of also benefiting from it. In this sense, public authorities have an interest in seeing that it is "supplied" in an appropriate quantity. This does not mean, however, that public authorities should necessarily intervene in financial markets so as to promote stability. There is no public agency directly concerned with stability in the market for foodstuffs or automobiles (although governments generally accept a responsibility for health and safety and for competition). Is finance any different?

It cannot be denied that all financial instability has costs for someone. The collapse of a financial firm imposes direct costs on shareholders, who lose their investment; on employees, who lose their jobs; and on depositors and unsecured creditors, whose claims may be forfeit. Instability in asset prices creates losses for those whose investments prove unsuccessful. In this (i.e., the direct or "private" costs of instability), financial firms and markets are not qualitatively different from other sectors of the economy. And while there are always pressures to compensate private losses, it is generally assumed that the public interest is served best by allowing market disciplines to work—unless there is evidence of market failure.

In what follows, I will examine the argument that the financial system is particularly subject to market failure, and that the consequences of such failure justify public policy intervention. It will be convenient to divide this discussion into two parts: that concerned with the potential for instability at financial *institutions;* and that concerned with excessive volatility in prices in financial *markets.*

Instability at Financial Institutions

The reasons why difficulties at a financial firm may give rise to public policy concerns may be grouped under several (overlapping) heads:

(a) Losses to depositors and other creditors may be exacerbated because of the unique vulnerability of financial institutions to "runs."
(b) The scope for losses to spread to other financial institutions through "contagion" or direct exposures is high.
(c) There may be budgetary costs from the perceived need to protect depositors or bailout troubled institutions.
(d) There may be more widespread macroeconomic consequences from instability in the financial sector.
(e) A loss of confidence in financial intermediation may lead to financial "repression" resulting in suboptimal levels of savings and misallocation of investment.

The first two of these points concern the potential for an "instability bias" in the financial system; the last three to the external costs generated by such instability. Let us now consider them in slightly more detail.

"Runs" and the Protection of Individual Institutions

There are two broad reasons why the authorities may wish to be involved with the stability of individual institutions (other than contagion risk, which is dealt with below). One rests on the vulnerability of banks to runs; the other on economies of scale in monitoring the behavior of complex firms.

A well-known feature of banks is that they issue liabilities that are redeemable on demand at par, while they hold longer term assets that are less readily marketable and have an uncertain value. Under normal circumstances, this does not pose a major problem, since deposit withdrawals are subject to the law of large numbers and well-managed loans that are held to maturity are mostly repaid at face value. A bank's holding of capital covers the risk of loan loss, and a cushion of liquid assets is sufficient to preserve confidence in its ability to meet withdrawals.

If, however, something happens to disturb confidence, the situation can be destabilized. Depositors perceive that those who withdraw their funds first will be able to do so without loss or penalty; those who delay may find that the bank's capital has been eroded by a "fire-sale" of less marketable assets. What this means is, firstly, the value of a bank (like other firms) is greater as a going concern than it is in a forced liquidation. Secondly, because of the leverage inherent in banks' operations, forced liquidation is more likely than in the case of nonfinancial firms. This argues in favor of an outside agent to preserve potentially solvent institutions as going concerns, or else to intervene to gradually wind down firms that have become insolvent.

A slightly different argument for intervention to protect depositors is that they have inadequate information to protect themselves. Monitoring financial institutions is costly, and pooled monitoring may be more efficient than individual monitoring. (Note that this argument may apply to all firms, not just those, like banks, whose liabilities are repayable at par on demand.) In this view, the public authorities are performing a service (like that of a rating agency) that it would be too difficult or too costly for individual depositors to perform for themselves. This argument can be given a political slant by recognizing that, to be realistic, certain depositors will always act foolishly when faced with the incentive of high returns. Since political pressure to provide compensation for losses is bound to ensue, it is better for the authorities to step in to avert losses, or rationalize the process by which compensation is provided.

"Contagion" Effects at Other Financial Institutions

Potentially more serious than the losses that accrue to individual depositors at a failed institution is the danger that difficulties may be propagated more widely. Such contagion can take place through two main channels: firstly, the pattern of interlocking claims among financial institutions; and secondly, the potential for difficulties at one institution to provoke a loss of confidence in others thought to be similarly placed.

There can be little doubt that the exposure of financial firms to other financial intermediaries has grown dramatically in recent years. A major factor has been the increase in trading activities. Daily foreign exchange trading has increased threefold over the last decade and stood at $1.25 trillion in 1995. Well in excess of 80 percent of these trades are between dealing counterparties. Derivatives and securities trading has grown even faster and is also dominated by interdealer activity. The place where the resulting interintermediary exposures get concentrated is the interlocking network of payments and securities settlement systems. Although individual exposures are of short duration, at any point in time they are very large in size. In many cases, the unsecured exposure of financial institutions to a single counterparty exceeds capital. It is this fact that has led some observers to conclude that a disruption transmitted through the payment system is the largest single threat to the stability of the financial system.[4]

Contagion can also occur indirectly, when strains at one financial institution provoke a loss of deposits from, or an unwillingness to enter into transactions with, other firms that are also thought to be vulnerable. Following the Barings collapse, for example, a number of small to medium-sized investment banks in London and elsewhere were reported to have suffered deposit withdrawals, even though there was nothing to suggest that they had incurred losses similar to Barings'. In other words, contagion can be indirectly as well as directly induced.

Contagion is one of the basic reasons why public authorities are concerned with the health and survival of individual financial institutions. This relates to the "public good" aspect of financial stability. Confidence in the financial system benefits individual participants without imposing costs on others. If the failure of one institution causes a contagious loss of confidence elsewhere, the adverse consequences to the system as a whole may be much greater than those resulting from the initial disruption.

Resolution Costs

Turning now to the spillover consequences of instability, the transfer costs of resolving financial crises are the most readily quantifiable, and in many ways the most striking. To public policy officials, the costs that fall on the public budget surely provide the most persuasive evidence of the need to do whatever is necessary to strengthen financial systems.

The U.S. public is acutely aware of the savings and loan debacle of the 1980s, the resolution costs of which are estimated at anywhere between 2 percent and 4 percent of GDP. These costs, however, pale in comparison with the fiscal costs incurred in a number of other countries.[5] In France, the losses incurred by a *single* bank, Credit Lyonnais, are now put at some $30 billion, or over 2 percent of GNP. Honohan estimates the fiscal costs of resolving crises in developing countries alone as being as much as $250 billion.[6] A World Bank Study estimates that 14 countries had to devote more than 10 percent of GNP to the resolution of banking sector crises (Table 2).[7] And a by now well-known IMF study concludes that almost three-quarters of IMF member countries encountered "significant" banking sector problems during the period 1980–96; of these as many as one-third warrant the designation "crisis."[8] Part of the resolution costs of these crises fall on the banking system and its clients. More frequently, however, the government budget is left to pick up the lion's share.

GNP Costs of Financial Instability

The resolution costs of financial sector crises are, of course, transfer costs. They cannot be taken as an accurate guide to losses in economic welfare, which could be either greater or smaller. They could be smaller than the transfer costs if the real assets financed by failed banks remained in existence and continued to yield productive services. On the other hand,

Table 2 World Bank Study of Countries Who Devoted More than 10 Percent of GNP to Resolve Banking Sector Crises

Country (time period of crisis)	Estimate of total losses/costs (percentage of GDP)
Latin America	
Argentina (1980–82)	55
Chile (1981–83)	41[a]
Venezuela (1994–95)	18
Mexico (1995)	12–15[b]
Africa	
Benin (1988–90)	17
Cote d'Ivoire (1988–91)	25
Mauritania (1984–93)	15
Senegal (1988–91)	17
Tanzania (1987–95)	10[c]
Middle East	
Israel (1977–83)	30[d]
Transition countries	
Bulgaria (1990s)	14
Hungary (1995)	10
Industrial countries	
Spain (1977–85)	17
Japan (1990s)	10[e]

[a] 1982–85.
[b] Accumulated losses to date.
[c] In 1987.
[d] In 1983.
[e] Estimate of potential losses.
Source: Goldstein (1997) based on Caprio and Klingebiel (1996a).

the cumulative misallocation of financial resources represented by bad loans suggests that the overall loss to society from inefficient financial intermediation may have been even larger than the losses that eventually fell on the budget or on the shareholders and other claimants of banks. How can one go about assessing the macroeconomic costs of instability?

Even if instability does not lead to crisis, they can make it harder for the authorities to gauge the appropriateness of a given policy stance. Financial fragility complicates the interpretation of the indicators used to guide monetary policy decisions. Somewhat more seriously, weaknesses at financial institutions can limit the willingness to lend, thus creating "head winds" for the expansion of demand. Overall economic performance suffers as a result.

Where financial difficulties are more serious, the impact on GNP can be larger and more direct, whether or not the authorities decide to support the financial system. In Mexico, for example, the interaction of financial sector difficulties and a currency crisis led to a sharp setback to GNP. By mid-1995 industrial output in Mexico had fallen 12 percent from its level two quarters earlier. Even in Argentina, which successfully defended its exchange rate, GDP is estimated to have temporarily fallen some 7 percent below trend as a result of the "tequila effect." The banking crisis of the 1980s in Chile saw output growth drop from 8 percent in the five years preceding the crisis to only 1 percent in the five years after it.

Among industrial countries, it is harder to detect cause-and-effect relationship between financial instability and GDP. In the United States, the savings and loan crisis had little measurable impact on growth, costly though it was to the budget. In Nordic countries and in Japan, the consequences are more readily apparent. Growth in Finland averaged 4.5 percent in the years preceding the outbreak of the banking crisis, and was *minus* 4.0 percent in the three succeeding years (though doubtless not all of the difference is attributable to financial difficulties). In Sweden and Norway, there were economic downturns following the strains in the banking system, though again other factors were also at work. And in Japan, the "head winds" caused by financial sector weaknesses held growth in the mid-1990s below the underlying potential of the economy.

It bears repeating here that the relationship between financial instability and macroeconomic instability is two-way. Macroeconomic instability is usually a major factor in financial difficulties, often because an unsustainable expansion induces unwise lending. Credit-fueled "bubbles" in financial asset and property prices frequently play a contributory role, especially when a large share of lending is used to finance the acquisition of real estate or financial assets whose price is, for a time, rising rapidly.[9] A recession then reveals serious weakness in lending portfolios. When the financial system encounters difficulties, problems can quickly worsen macroeconomic performance. Weakened intermediaries cease to lend, losses in the financial sector create negative wealth or income effects, generalized uncertainty inhibits investment, and the public sector is often forced to rein in real expenditure to help offset the budgetary cost of increased transfers.

Instability and the Development of the Financial Sector

Beyond the direct effects of financial instability on real economic activity, there can be indirect adverse consequences for longer run growth potential if financial intermediation is stunted. As Akerlof has shown, in any market where participants have asymmetric information, moral hazard and adverse selection reduce exchange below levels that could be beneficial if market participants had better information (the market for lemons). The market for intertemporal exchange is characterized by extreme asymmetry of information between providers of funds and potential borrowers. The potential negative consequences are, however, offset by the existence of specialized intermediaries. Financial intermediaries perform the role of agents for lenders, screening out uncreditworthy borrowers, monitoring borrower's performance after a loan is made, adding creditworthiness through the commitment of their own capital, and creating liquidity through providing for the ready marketability of claims.

All of this, however, depends upon the preservation of confidence in the stability of the network of financial intermediaries: if lenders lose confidence in the continued stability of the institutions to whom they have entrusted their funds, or in the integrity of the markets in which they have invested, they will seek to reduce their exposure and place their assets elsewhere. In the limit, they may choose consumption over saving, or may place their savings in nonproductive but "safe" forms (such as precious metals). If this happens, the contribution of the financial sector in providing improved methods of risk pricing and management, and in adding liquidity and creditworthiness, will be much diminished. Mishkin indeed defines a financial crisis as "a disruption to financial markets in which adverse selection and moral hazard problems become much worse, so that financial markets are unable to channel funds efficiently to those who have the most productive investment opportunities."[10]

Instability in Financial Markets

While there is broad (though not universal) acceptance that the stability of financial institutions should be an objective of public policy, this is much less true with regard to financial asset prices or financial flows. The majority view is that free markets are the best guarantors of equilibrium in prices, and that official intervention should be limited to removing market imperfections, e.g., by promoting the disclosure of relevant information and preventing the emergence of monopoly practices. Yet financial markets can, in principle, be subject to the same kind of "instability bias" and adverse spillovers that affect financial institutions.

Instability bias arises if a disturbance affecting prices generates forces creating further moves in the same direction. These are generally based on extrapolative expectations, which can result from asymmetric information, reinforced by herd instincts. Certain technical features of markets, such as margin requirements, can also play a role. In a rising market, those who invest on margin find their net worth rising, and are thereby enabled to make further leveraged purchases, pushing prices still higher. The opposite effects come into play in a falling market, with margin calls forcing liquidation of holdings and exacerbating price declines.

The importance of such instability biases are very hard to assess on a priori grounds. The sudden drop in equity prices in 1987 suggests that they can sometimes be significant, though the relative infrequency of such occurrences provides some reassurance. Swings in exchange rates could be taken as evidence that similar pressures work in currency markets; though full-blown currency crises are more apt to be result of attempts to defend a fixed rate at an unsustainable level.

Volatility in financial asset prices has the capacity to create "spillover" effects of various kinds. Firstly (and perhaps least troublesome), is the added difficulty it creates for the authorities in formulating macroeconomic policies. Movements in asset prices influence all of the channels by which monetary policy traditionally affects the real economy: the interest rate channel, the wealth channel, the exchange rate channel. Moreover, they can, if severe, have pervasive effects on confidence. There is at present a lively debate about whether and how monetary policy should respond to asset price movements. The fact that the debate is still unresolved is evidence of the uncertainties created for policymakers when financial markets are unstable.

Another type of spillover effect occurs when asset price movements undermine the stability of financial institutions. This can happen if intermediaries are heavily exposed to certain categories of assets (e.g., equities or real estate), or if their lending is secured on such assets. It can also occur if financial institutions have mismatched foreign currency or interest rate books, or if higher volatility suddenly increases the costs of hedging options positions.

Lastly, asset price volatility can create real economic costs if the authorities are led to take extreme measures to restore stability. Perhaps the most prominent examples of such costs occur in currency crises. Instability in foreign exchange markets is almost invariably accompanied by sharply higher interest rates in the country whose currency is under downward pressure. And higher interest rates usually provoke a downturn in economic activity, whether accompanied and exacerbated by a financial sector crisis or not.

What are the specific markets that are particularly vulnerable to instability, and what is the nature of the spillover effects? Let us briefly consider four.

Firstly, the foreign exchange market. Two types of instability should be distinguished: the turmoil that surrounds speculation against a pegged exchange rate; and the volatility that seems to characterize floating rates. The defense of pegged rates, especially when it is ultimately unsuccessful, is most likely to be classified as a currency "crisis." In such a case, it can be argued that the problem is as much one of policy as of market instability. Should the authorities have selected a fixed rate regime? Should they have changed the peg (or the regime) earlier? Should they have pursued a different mix of policies? Some have argued, however, that attacks on a fixed peg can also be speculatively induced.[11] Where there are dual or multiple equilibria in exchange rate relationships, the movement from one to another may owe more to market dynamics than to fundamentals.

Where exchange rates are floating, volatility is harder to explain, especially when movements in fundamentals are modest. Swings in relative real values among the U.S. dollar, Deutsche Mark and Japanese Yen have approached 50 percent or more in the past decade and a half. Such swings complicate macroeconomic policies; generate the potential for resource misallocation and give rise to protectionist pressures. While it can be argued that exchange markets are responding to policy divergences (actual and expected), the link is often not at all clear.

Secondly, instability in equity markets can also have external consequences. Stock market volatility can undermine the stability of financial institutions who are directly or indirectly exposed to equity prices; exacerbate the investment cycle (via Tobin's "q"); and, if prices fall sharply, have adverse effects on confidence. However, although stock market crashes have a fascination for lay opinion, the impact of equity price instability has for most of the time been relatively mild. This may be because there are nonlinearities at work. Modest movements in equity values do little if any harm; but a larger movement has a disproportionately greater potential both to set up self-perpetuating forces and to do real economic damage.

Thirdly, much the same can be said of price fluctuations in bond markets. Despite the generalized run-up in bond yields in 1994, adverse spillovers were rather well contained.

So long as the central bank is thought able to stabilize inflation, the scope for extreme movements in bond prices is limited.

Fourthly and finally, real estate, though not strictly speaking a financial asset, can be subject to "bubble" phenomena. A real estate bubble complicates the formulation of monetary policy while it is being created, and can leave a string of failures in its wake when it bursts. Some of the difficulties faced in mid-1997 by Southeast Asian economies can be traced, in part, to real estate bubbles.

What should be concluded from the foregoing brief survey? If there are disequilibrium tendencies in financial and other asset markets, and if price volatility has had adverse spillover consequences, does this argue for making the stability of asset prices a focus of public policy concern in the same way as the stability of financial institutions?

Here the answer is, at best, not clear-cut. Few economists would be confident that governments could be better at determining equilibrium prices than markets. Even when prices move by an amount that is clearly greater than "fundamentals" justify, it can rarely be said that the price was more appropriate before the move than after it. And frequently, the blame for price volatility is due to unstable policies just as much as to unstable markets. So the broad consensus among economists (with which I agree) is that official policy to stabilize financial asset prices should be focused more on sustainable policies and removing market imperfections, than on direct actions to limit price movements.

One should recognize that there can, occasionally, be exceptions to this general rule. When currencies become substantially misaligned (as in 1985, say), governments may try to give a lead to markets (albeit through statements concerning policies). And if domestic asset prices were to fall to an extent that threatened financial stability, it would not be surprising to see a policy response aimed at stabilizing prices. In fact, central banks responded to the 1987 stock market crash by easing the provision of liquidity to financial markets. In general, however, official responses to extreme price movements tend to be ad hoc, rather than part of a cohesive "policy" on financial market stability.

APPROACHES TO ENSURING FINANCIAL STABILITY

The foregoing section has listed a number of reasons why financial instability has negative externalities. These are probably sufficient to make achieving and maintaining stability a public policy goal. It is of less help, however, in determining how public authorities should promote stability. This section reviews several broad approaches to promoting stability, implying varying degrees of intervention by the authorities. The principal focus is on policies to promote stability at financial institutions, since these have been the subject of more coherent analysis. At the end of the section, however, there are a few observations on preventing instability in key market prices.

Reliance on Market Forces

With the possible exception of New Zealand, where certain special circumstances apply, no countries have adopted the position that market forces can be relied on as the sole guarantor of stability at financial institutions. But while official support for the pure market solution is limited, there is a stronger academic tradition in this vein, going back to the free banking school, and finding recent expression in the writing of Dowd.[12] Other academics have questioned whether the contagion effect that lies behind official concern with systemic stability is in reality all that significant.[13]

The case for the market solution is, to simplify, as follows: when all actors, including depositors, counterparties, managers, and shareholders of financial institutions realize they are "on their own," they will exercise a much higher degree of care, and financial institutions will thereby be forced to operate in a sounder and more prudent fashion. The failure of an individual institution will become less likely, and the risk of systemic contagion will be almost nonexistent. The moral hazard implied by official intervention will be removed, with favorable consequences for the efficiency of resource allocation.

The case against can be put on several levels. Most fundamentally, it is argued that there are events that may occur very infrequently, that cannot be predicted, and that have the capacity to destabilize the financial system if not resisted. These could include political events such as the outbreak of war or the election of radical governments; economic events, such as the 1929 stock market crash; or natural disasters such as a major earthquake in a large metropolitan center. If governments were to stand aside from helping the financial system under such extraordinary circumstances, financial institutions would have to carry such a large cushion of capital as to greatly reduce their capacity to contribute to economic welfare in normal times.

More prosaically, it is pointed out by Goodhart and others that political pressures make it very hard for elected authorities to refuse assistance to institutions whose depositors have powerful electoral influence.[14] Since most market participants know this, any ex ante announcement by governments not to support the financial system lacks credibility. Moral hazard is not, therefore, avoided. Thus, despite the attraction of reliance on market forces, most observers accept that it is insufficient, by itself, to guarantee stability in all circumstances.

Safety Nets

The most effective way of ensuring continued confidence in financial institutions is to provide their users with some sort of explicit safety net. The main types of safety net are deposit insurance schemes, and the presence of a lender of last resort. The primary drawback of safety nets is moral hazard, which appears in a particularly overt form with deposit insurance. Insured depositors have no incentive to monitor the institutions with whom they place their funds. Borrowing institutions are therefore able to pursue risky strategies and, at the limit, to "gamble for resurrection" when their capital has been eroded. The potential for imprudent behavior is exemplified by the savings and loan episode in the United States.

Various means have been suggested to address the moral hazard issue. These include limiting the coverage of deposit insurance, charging risk-based insurance premia, and limiting insurance coverage to a specific category of institutions (100 percent reserve banks). None is entirely satisfactory. Limiting the coverage of insurance schemes means that uninsured depositors can still precipitate a "run" when they fear for bank solvency. Risk-based insurance premia are difficult to calculate on a formulaic basis. And 100 percent reserve banking, despite impressive academic support from Henry Simons to Milton Friedman and James Tobin, has never gained much support.[15] Probably most observers conclude that 100 percent reserve banks would not be successful in winning a major share of the market during normal times, and therefore the issue of how to safeguard stability at other institutions would not go away.

Lender-of-last-resort support has been a recognized role of central banks since Bagehot. The object is to provide support to solvent but illiquid institutions to avoid the possibility that they would have to liquidate assets in a "fire sale" that would generate losses and lead to an avoidable insolvency. Aside from the practical difficulty of distinguishing between insolvency and illiquidity, the lender-of-last-resort role does not avoid the problem of moral hazard.[16] One answer to this is "constructive ambiguity"—a phrase made popular by Jerry Corrigan meaning that central banks reserve the right to intervene to preserve stability but give no assurances, explicit or implicit, to individual institutions. Such an approach is intended to make institutions act more prudently by making them uncertain whether they would be rescued in a crisis. In some circumstances, however, "constructive ambiguity" may turn out to be a cloak for "too big to fail," if the lender of last resort is more willing to take the risk of allowing a small institution to go under than a large one.

Regulation

If there were no safety net, regulation would be justified by the need to protect the interests of depositors and other creditors. With a safety net, the justification shifts to one of protecting the deposit insurance fund (often taxpayers) and avoiding moral hazard. In practice the focus of regulation has shifted significantly over time, and may now be in the process of a further shift. Three different focuses for regulation can be distinguished.

Regulation to Protect Franchise Values

Until about 20 years ago, regulation in most countries had the effect of limiting competition in the financial industry. Entry to the industry was controlled, there were restrictions on interest rate competition, and cartel-type practices were tolerated. In a number of countries, including the United States and Japan, there was strict segregation between commercial and investment banking activities. Since franchise values were high as a result, losses were less likely and, when they did occur, more often led to industry-sponsored takeover or rescue than to outright failure.

Several developments in the 1970s and 1980s undermined this form of regulation. The growing dominance of the free market philosophy made protective practices less acceptable. Liberalization and deregulation increased competition which in turn eroded banks' profitability and diminished franchise values. With relatively thin capital cushions, this made banks more vulnerable to adverse external shocks. As a result regulation to limit competition and bolster the profitability of financial institutions was no longer a practicable or acceptable means of ensuring systemic stability.

Risk-Based Capital Adequacy

In recent years the dominant form of regulation to promote systemic stability has been risk-based capital adequacy. Instead of limiting banks' activities, regulators have sought to ensure that banks are adequately capitalized against the risks they run. This is the philosophy behind a series of documents issued by the Basle Committee on Banking Supervision. Supervisors have divided assets into a number of "risk classes" and specified the amount of capital to be held against each.

Such an approach has several advantages. The notion of relating capital to risk is in conformity with the reason financial institutions hold capital in the first place. And the increased capitalization of the banking system that has followed from the decisions of the Basle Committee has undoubtedly improved systemic resiliency. Nevertheless, certain aspects of the way the approach has been implemented have drawbacks, which are becoming increasingly recognized.

Firstly, and most importantly, there is the potential for a discrepancy between risk, as calculated by the financial institution itself, and risk as measured by regulatory criteria. To take two obvious examples: the Basle Committee risk weights make no distinction between high and low quality credits within the same risk weight category (e.g., between a AAA borrower and a junk bond issuer); nor do they take account of the possibility of risk reduction through diversification. Most financial firms now find that there is a significant discrepancy between the "economic capital" they consider appropriate to cover the total risk of their portfolio and the "regulatory capital" they are required to hold under the Basle ratios.

This would not matter much from the viewpoint of stability if the only problem were an excess of prudence on the part of supervisors. Indeed, it could well have advantages, since the additional capital cushion required by supervisors could be considered the "price" to be paid for the safety net provided by the lender of last resort. As some writers have pointed out, however, this is not the only implication. Even adjusting for supervisory caution, a portfolio's riskiness may appear significantly different when internal risk models are used than when the Basle risk weights are applied. It is possible for banks with higher risk appetites to deliberately add risk to their portfolios (e.g., through the use of credit derivatives) without having an effect on the regulatory capital required to meet the Basle ratios.[17]

A second problem with the current approach is that it focuses only on certain categories of risk. One gap in the original Basle Accord has now been plugged with the extension of capital requirements to market risk as well as credit risk. But several of the most recent examples of serious losses in the financial sector have come from operational risk (Barings, Daiwa), legal risk (swaps with UK local authorities) and model risk (Metallgesellschaft). As a result of these perceived shortcomings, growing attention is now being given to using regulation to better harness market incentives in support of stability.

Regulation to Support Market Forces

In any market, self-regulation is a powerful force. The strongest incentive to act with prudence and integrity comes from those with most to lose when they fail to do so. Recent thinking has therefore focused on ways of strengthening the incentives on individual institutions to manage their own affairs prudently and on their counterparties to exercise appropriate discipline: in the jargon, "incentive-compatible financial regulation."[18]

Consider the Assessment of Risk

The managers of a financial institution have a strong incentive to monitor accurately their risk exposure. It therefore seems likely that an internal assessment of risk will be a better measure than a simplified external formula. This philosophy has been accepted by the Basle Committee and incorporated in the market risk amendment to the Capital Accord. The market risk amendment allows firms to use their own models (subject to external validation) to measure the risk in their trading portfolio then prescribes a "multiplication" factor which translates value at risk into required capital holding.

It seems, therefore, as though the debate is moving towards a distinction between the *measurement* of risk, which is best done by those who are closest to the portfolio, and have the tools to do it; and the *capitalization* of risk, decisions on which raise public policy issues. Since the authorities, by underwriting the stability of the financial system are in essence providing financial institutions with catastrophic risk insurance, it is legitimate for them to limit the potential recourse to such insurance by requiring a minimum *level* of capital holding.

Conceivably, one could go even further and assign responsibility for decisions on capital holding to the private sector as well. This is the philosophy behind the so-called "pre-commitment" approach. An institution would itself choose how much capital it would assign to cover the value at risk in its portfolio. If losses exceeded the calculated probability, then the institution would be subject to some kind of penalty. This is an intriguing idea, though it would present a number of complex practical issues. Moreover, it is not clear that it would lead to an appropriate pricing of the safety net.

The idea of harnessing self-disciplining forces is also behind the proposal of the Group of Thirty to develop industry-led standards for risk management, internal operating controls and public disclosure.[19] The proposal would call for major international institutions to commit to standards that they would undertake to meet themselves and to require of their counterparties. When endorsed by supervisors, these would then presumably spread, through market pressures, to all institutions. Being developed by practitioners, these standards, it is argued, are more likely to provide an appropriate balance between benefits and costs. In particular, by allowing the industry to propose more efficient ways of reducing risk, they would reduce the danger that firms would cut corners in an effort to avoid burdensome official regulation.

Before ending this section, a word should be added on policies to preserve stability in financial market prices. Theory provides much less help in addressing this issue than that of stability in financial institutions. Certain approaches to providing a more stable market environment would not be controversial. These include the encouragement of stable and sustainable macroeconomic policies; fuller disclosure and dissemination of relevant financial data; and the outlawing of anticompetitive practices in financial markets. Other measures have also attracted a measure of support, such as the use of "circuit-breakers" when prices move by more than a certain threshold amount.

What to do when a significant "bubble" is thought to be developing, or when a bubble bursts, is a matter on which there is little agreement. Public authorities can warn about "irrational exuberance," but central bankers are in general unwilling to adjust macroeconomic policy to stabilize financial asset prices. If prices were to fall, the reaction might be different, if only because experience suggests that price falls tend to be more rapid and disorderly than price rises.

CONCLUDING COMMENTS

There is persuasive evidence that financial stability provides a favorable environment for efficient resource allocation and more rapid economic growth.[20] Instability has been associated with lower levels of saving and investment, fiscal costs, and setbacks to GNP. It is, therefore, unavoidable that securing stability should be a concern of public policy authorities.

What is less clear, however, is whether the maintenance of stability requires an activist approach on the part of the authorities, or alternatively whether it can best be achieved by reliance on market forces. Arguments against a pure laissez-faire approach include the following: that there are disequilibrium tendencies within the financial system that can, via contagion, turn instability into crisis; and that the costs of a financial crisis for economic welfare are so great that it is irresponsible to take chances. On the other hand, too great a level of support for the financial system, or support in inappropriate ways, can lead to inefficiency and moral hazard.

A consensus therefore seems to be developing among central bankers that regulation should, as far as possible, be directed at reinforcing the self-disciplining tendencies of the market. This probably means less detailed or prescriptive regulation, and a greater reliance on the internal controls of market participants, supported by mechanisms that sharpen the incentive for prudent behavior.

It may be worth ending with a few observations on regulatory structure. A tendency has developed in recent years to draw a distinction between the function of institutional supervision; responsibility for systemic stability; and responsibility for price stability. These are indeed separate functions, and there may be cases in which the pursuit of any one of them is handifcapped by the simultaneous pursuit of the others.

There are also powerful linkages, however. Systemic stability is linked to the health of the individual institutions that comprise the system; and instability in the financial system can both cause and be caused by instability in the real economy. What this means is that there must be close collaboration between those responsible for monetary and financial stability, respectively, and that both must be aware of the financial condition of the key institutions. Moreover, in order not to stifle innovation, all concerned need to have a healthy respect for market forces and recognize the need, in a market economy, for bankruptcy as an ultimate sanction for unsuccessful enterprises.

This does not lead to any universally applicable conclusions concerning regulatory structure. It should, however, give pause to those who believe that separating functions is a straightforward and costless measure to tackle perceived shortcomings in present arrangements.

NOTES

1. *BIS Annual Report,* 1996–97.
2. Federal Reserve Bank of Kansas City's 1996 symposium, "Achieving Price Stability," held in Jackson Hole, Wyoming, August 29–31.
3. Goodhart, C. A. E., Hartmann, P., Llewellyn, D. T., Rojas-Suares, L., and Weisbrod, S. R. 1997. Financial Regulation: Why, How and Where Now? Monograph for the Central Bank Governors' Meeting at the Bank of England, June 6.
4. Corrigan, E. G. 1996. Remarks at the Symposium on Risk Reduction in Payments Clearance and Settlement Systems, New York, Goldman Sachs and Co., January 25.
5. Goldstein, M. 1997. The Case for an International Banking Standard. Institute for International Economics, Washington, D.C., April.
6. Honohan, P. 1996. Financial System Failures in Developing Countries: Diagnosis and Prescriptions. Unpublished manuscript. Washington: International Monetary Fund.
7. Caprio, G. and Klingebiel, D. 1996. Bank Insolvencies; Cross-Country Experience. Unpublished manuscript. Washington: World Bank.
8. Lindgren, C.-J., Garcia, G., and Seal, M. 1995. Bank Soundness and Macroeconomic Policy. Washington: International Monetary Fund.
9. Borio, C. E. V., Kennedy, N. and Prowse, S. D. 1996. Exploring Aggregate Asset Price Fluctuations Across Countries: Measurements, Determinants and Monetary Policy Implications. *BIS Economic Papers,* no. 40, April.
10. Mishkin, F. S. 1991. Anatomy of a Financial Crisis. National Bureau of Economic Research Working Paper no. 3934.
11. Eichengreen, B. and Wyplosz, C. 1993. The Unstable EMS. Brookings Papers on Economic Activity, no. 1, pp. 51–143.
12. Dowd, K. 1988. Private Money. Institute of Economic Affairs, Hobart Paper, no. 112.
13. Kaufman, G. 1994. Bank Contagion: A Review of the Theory and Evidence. *Journal of Financial Services Research,* no. 8, April, pp. 123–50.
14. Goodhart, C. A. E., Hartmann. P., Llewellyn, D. T., Rojas-Suares, L., and Weisbrod, S. R. 1997.

Financial Regulation: Why, How and Where Now? Monograph for the Central Bank Governors' Meeting at the Bank of England, June 6.

15. Simons, H. 1948. *Economic Policy for a Free Society.* Chicago: University of Chicago Press. Also, Friedman, M. 1959. *A Program for Monetary Stability.* New York: Fordham University Press. Also, Tobin, J. 1985. Financial Innovation and Deregulation in Perspective. *Bank of Japan Monetary and Economic Studies,* no. 3, September, pp. 19–29.
16. Of course, it is usually accepted that the management of a failed institution should forfeit their positions and the shareholders should lose their money: the debate surrounds the extent to which uninsured depositors should be protected.
17. Yellen, J. L. 1996. The 'New' Science of Credit Risk Management at U.S. Financial Institutions, presented at the Conference on Recent Developments in the Financial System, Bard College, April 11, and reproduced in BIS Review.
18. Greenspan. A. 1996. Remarks to 32nd Annual Conference on Bank Structure and Competition. Federal Reserve Bank of Chicago, May, and reproduced in BIS Review, No. 58.
19. Group of Thirty. 1997. Global Institutions, National Supervision and System Risk. A Study Group Report, Washington.
20. King, R. G., and Levine, R. 1993. Finance and Growth: Schumpeter Might Be Right. *Quarterly Journal of Economics,* August, pp. 717–37.

5

The New View of Growth and Business Cycles

Jonas D. M. Fisher
Senior Economist, Federal Reserve Bank of Chicago, Chicago, Illinois

INTRODUCTION AND SUMMARY

Two central concerns of economic policy are growth and business cycle stabilization. There is considerable interest in devising government policies and institutions to influence prospects for economic growth and mitigate the distress associated with economic downturns. Proper evaluation of the benefits and costs of a given policy proposal requires knowledge of the determinants of growth and business cycles. This is one reason for the considerable body of research aimed at understanding these phenomena.

The last two decades have seen considerable advances in this research. Recent empirical evidence, however, brings into question two of its basic assumptions—first, that technological change is homogeneous in nature, in that it affects our ability to produce all goods symmetrically, including consumption and investment goods; and second, that business cycles are driven by shocks which affect the demand for investment goods.

In this article, I document the key evidence that challenges the conventional views of growth and business cycles. I then discuss the plausibility of alternative theories that have been advanced to meet the challenge. To date, the evidence seems to support a new view of growth and business cycles, one that is based on technical change biased toward new investment goods like capital equipment.

The key evidence involves two observations on the behavior of the *relative price* of business equipment over the last 40 years. First, in almost every year since the end of the 1950s, business equipment has become cheaper than the previous year in terms of its value in consumption goods. This means that if one had to trade restaurant meals for a piece of equipment that makes the same number and quality of, say, bicycles, one would forgo fewer meals in 1998 than in 1958. Second, this relative price tends to fall the most when the economy, and investment expenditures in particular, are growing at relatively high rates, that is, it is *countercyclical.*

The first piece of evidence is striking because it suggests that much of post-WWII economic growth can be attributed to technological change embodied in new capital equip-

---

Reprinted from: *Federal Reserve Bank of Chicago Economic Perspectives* (Fourth Quarter 1999) 35–56.

ment. This conflicts with conventional views on what drives economic growth. A piece of capital equipment is a good that is used to produce another good, such as a crane or a computer. An improvement in *capital-embodied technology* is the invention of equipment that takes the same amount of labor and preexisting equipment to produce as the old equipment but that produces more goods when combined with the same amount of labor as before. If a new production process yields the same units of capital equipment with less factor inputs, then this has the same economic implications as if the capital equipment produced were itself more efficient. Hence, an *equivalent* interpretation of what constitutes capital-embodied technical change is that it involves an improvement in the technology that *produces* capital equipment.

To understand the relationship between capital-embodied technical change and the trend in the equipment price, suppose the technology for producing consumption goods is fixed. With improvements in technology embodied in equipment, the supply of (quality-adjusted) investment goods increases relative to consumption goods, so the equipment price falls. Greenwood et al. (1997) build on this insight to show that a large fraction of economic growth can be attributed to *capital-embodied technical change.* This conflicts with the conventional view that most growth is due to disembodied technical change, or *multifactor productivity.* Improvements in disembodied technology, usually measured as the Solow (1957) residual, make it possible to produce all kinds of goods, not just capital goods, with less capital and labor.[1] If this were the dominant source of growth, then we should not have seen such a large drop in the price of equipment over the last 40 years.

The second piece of evidence runs counter to standard views of the business cycle. Standard theories hold that the business cycle is driven by shocks which affect the demand for investment goods. For example, consider the *IS–LM* model, which summarizes much of what is often called *Keynesian* macroeconomics. This model is the focus of most textbooks on macroeconomics and underlies much of the discussion of macroeconomic policy in the media.[2] In this model, business cycles are due to shocks to aggregate demand, such as monetary and fiscal disturbances. For example, expansionary monetary policy stimulates demand for investment goods through lower interest rates. If there is an upward sloping supply schedule for investment goods, we would expect the relative price of investment goods to *rise.* The same holds for expansionary fiscal policy, if government spending does not fully crowd out investment. Another view of business cycles, often attributed to Keynes, is that they are primarily investment cycles driven by variation in *animal spirits,* that is, changes in confidence about future growth prospects.[3] With the same assumptions on investment supply, we would expect investment prices to be high when investment is high. In summary, traditional Keynesian views of business cycles imply that investment good prices should be *procyclical,* that is, be high when overall economic activity is relatively high.

In recent years, an alternative view of business cycles, based on "fundamentals" that influence aggregate supply, has gained credence. This *real business cycle* view says that business cycles are driven in large part by disturbances to multifactor productivity. Just as the shocks to aggregate demand which are central to Keynesian theories, these disturbances influence business cycles through their effect on the demand for investment goods.[4] Hence, if there are costs in terms of forgone consumption of expanding investment good production, that is, if the supply schedule of capital is upward sloping, these models also predict the relative price of investment goods to be procyclical (Greenwood and Hercowitz, 1988).

Since the relative price evidence contradicts the major schools of business cycle thought, it poses a challenge to our understanding of business cycles. There are two leading hypotheses that could reconcile the theory and evidence. One, the *embodied technology* view, is built from the real business cycle tradition and takes into account the trend evidence on equipment prices. Falling equipment prices are compelling evidence of capital-embodied technological progress over long horizons. Perhaps changes in the rate of such technological progress occur over shorter horizons as well. Suppose the business cycle were driven, to a large extent, by these disturbances. An increase in the rate of capital-embodied technical change would lead to an outward shift in the supply schedule for investment goods. With stable investment demand, investment would rise and equipment prices would fall. This new view of business cycles, which complements the new view of growth suggested by the longrun evidence on equipment prices, has been explored by Christiano and Fisher (1998), Fisher (1997), and Greenwood et al. (1998).

The other leading theory is more easily understood in the context of traditional Keynesian views of the business cycle. If shocks to aggregate demand occur with a downward sloping investment supply curve, then the price of equipment could fall in a boom. A downward sloping investment supply curve would arise if increasing returns to scale played an important part in the production of capital equipment, so this is called the *increasing returns* view. This view has been advanced by Murphy, Shleifer, and Vishny (1989).

Below, I document the trend and business cycle evidence on equipment prices. There is no reason to expect that capital-embodied technological change is unique to equipment. Equipment is one of many investment good aggregates, that is, types of capital. Moreover, for simplicity most economic models assume only one or two types of capital. Therefore, in addition to equipment prices, I analyze other investment good aggregates. Next, I discuss research that sheds light on the plausibility of the alternative views, including some new evidence. To date, the evidence seems to support the new view of growth and business cycles based on capital-embodied technical change.

If growth and business cycles are originating from changes in capital-embodied technology, then the models we use for policy analysis have to incorporate this and, consequently, policy recommendations could change. For example, to the extent that technological change is embodied in capital equipment, government policies that affect equipment investment could have a dual impact on growth via the quality and quantity of capital goods. This could mean, for example, that investment tax credits directed toward improvements in the efficiency of capital equipment could have a significant impact on growth.

The implications for stabilization policy of the embodied technology view are less obvious. The fact that it seems to supplant the increasing returns view means that the arguments for interventionist stabilization policy that this view lends support to are less compelling. For example, increasing returns could provide scope for policy intervention, as it either involves externalities or is inconsistent with perfect competition. Moreover, it makes models based on animal spirits more plausible, which also has implications for stabilization policy (see Christiano and Harrison, 1999). The embodied technology view is more in line with the real business cycle tradition, in which policy interventions are counterproductive.

EVIDENCE ON INVESTMENT GOOD PRICES

To study the trend and business cycle properties of investment good prices, we need two things—a way to extract real prices and quantities from data on nominal investment ex-

Box 1 Measuring Real Quantities and Prices from Nominal Expenditure Data

The U.S. Bureau of Economic Analysis (BEA) uses the chain-type Fisher index to measure real output and prices. For a thorough discussion of the procedures the BEA uses, see Landefeld and Parker (1997), which this box draws on. This index, developed by Irving Fisher, is a geometric mean of the conventional fixed-weighed Laspeyres index (which uses weights of the first period in a two-period example) and a Paassche index (which uses the weights of the second period). The Laspeyres price index for period t constructed using base year $t - 1$, L_t is given by

$$L_t = \frac{\sum_{i=1}^{N} P_t^i \times q_{t-1}^i}{\sum_{i=1}^{N} P_{t-1}^i \times q_{t-1}^i}.$$

The Paassche price index for period t constructed using base year t, S, is given by

$$S_t = \frac{\sum_{i=1}^{N} P_t^i \times q_t^i}{\sum_{i=1}^{N} P_{t=1}^i \times q_t^i}$$

Here N is the number of goods whose prices are being summarized by the index, P_t^i is the date t dollar price of the ith quality-adjusted good, and q_t^i is the quality-adjusted quantity of good t purchased at date t. The Fisher price index at date t, F_t is

$$F_t = \sqrt{L_t \times S_t}$$

From this definition we see that changes in F, are calculated using the "weights" of adjacent years. These period to period changes are "chained" (multiplied) together to form a time series that allows for the effects of changes in relative prices and in the composition of output over time. Notice that a quantity index can be computed in a manner analogous to the price index. A nice feature of the Fisher index is that the product of these two indexes equals nominal expenditures. Landefeld and Parker (1997) discuss several advantages of this index over previously used fixed weight indexes.

To measure relative prices we need to choose a numeraire. In the introduction the term "value in consumption goods" was used. Implicit in this statement is the assumption that consumption goods, specifically nondurable and services consumption, is the numeraire. Define the price deflator for nondurable and services consumption as P_t^c. Then the relative price of the good i at the time t, p_t^i is defined as

$$p_t^i = \frac{P_t^i}{P_t^c} = \frac{\text{time } t \text{ dollars/good } i}{\text{time } t \text{ dollars/consumption good}}$$

$$= \frac{\text{consumption goods}}{\text{good } i}$$

Notice that the units of the price are what we require. The BEA does not provide a measure of price deflator for nondurable and services consumption. To construct the consumption deflator used in this article, I applied the chain-weighing methodology outlined above, treating the NIPA quantity and price indexes for nondurable consumption and service consumption as the prices and quantities in the formulas.

penditures; and a precise definition of what we mean by the business cycle component of the data. Below, I address these issues. Then, I introduce the data and present the results characterizing the trend and cycle behavior of investment good prices.

Measuring Prices and Quantities

This section describes how relative prices and real quantities of investment goods are measured. My measures of prices and quantities are based on measures published in the "National income and product accounts" (NIPA) of the U.S. Bureau of Economic Analysis (BEA).

The basis of the BEA procedure is to construct a *price deflator.* To be concrete, a given nominal quantity of expenditures on some good i, X_t^i, is decomposed into a price deflator, P_t^i, (which measures the *nominal* price of the good) multiplied by a quality-adjusted index of the *real* quantity of the good, q_t^i.

The BEA measures P_t^i and q_t^i for different goods using a so-called chain-weighting procedure, which is summarized in Box 1. My measure of quantity is simply q_t^i, measured in units of 1992 dollars. My measure of the *real price,* alternatively the *relative price,* of good i at date t, p_t^i, is the real quantity of consumption goods that would need to be sold in order to purchase one unit of good i at time t. It is defined as the price deflator for good i divided by the price deflator for consumption of nondurables and services. The rationale for this measure is described in Box 1.

Measuring the Business Cycle Component of the Data[5]

In the introduction I described how the price of producer durable equipment (PDE) varies over the business cycle. Below, I provide a brief description of how I measure the business cycle component of the data. A detailed discussion of the procedure is given in Christiano and Fitzgerald (1998).

Figure 1 illustrates the basic idea behind the procedure. Panel A of Figure 1 displays real 1992 dollar chain-weighted gross domestic product (GDP). The reported data are the logarithm of the raw data. The advantage of using the logarithm is that the resulting movements correspond to percent changes in the underlying data. The deviations between the data and the trend line (graphed in panel B) contain the rapidly varying, erratic component, inherited from the choppy portion of the data that is evident in panel A. Panel B shows my measure of the business cycle component of real GDP. This measure excludes both the trend part of the data and the rapidly varying, erratic component. It includes only the com-

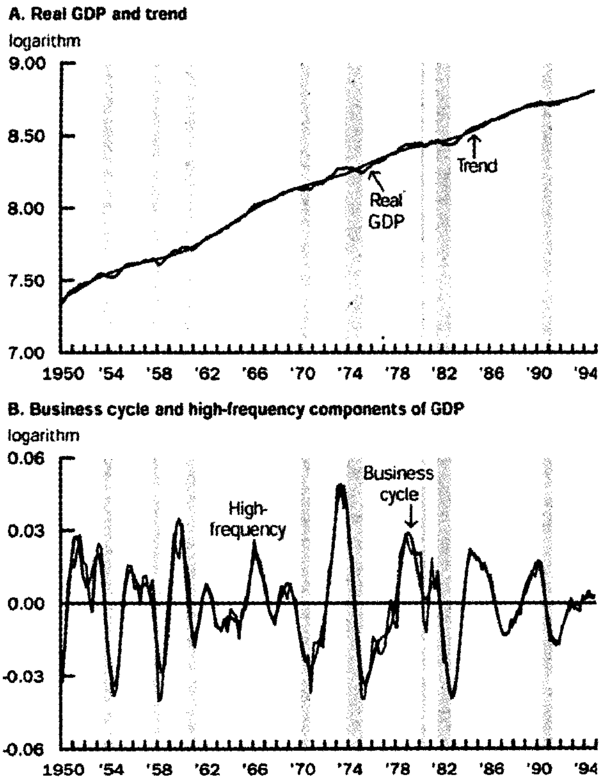


Figure 1 The trend and business cycle components of real GDP. Notes: Real GDP in panel A is real 1992 dollar chain-weighted GDP. The reported data are the logarithm of the real data. Shaded areas indicate recessions as determined by the National Bureau of Economic Research. Source: Author's calculations from U.S. Department of Commerce. Bureau of Economic Analysis, "National income and product accounts." *Survey of Current Business,* extracted from DRI Basic Economics database. 1947–98.

ponent of the data that contains fluctuations in the range of two to eight years. According to this approach, the economy is in recession when the business cycle measure is negative and in prosperity when it is positive.

Figure 1 also compares this measure of the business cycle with the one produced by the National Bureau of Economic Research (NBER). This organization decides, based on an informal examination of many data series by a panel of experts, when the economy has reached a business cycle peak or trough. The start of each shaded area indicates the date when, according to the NBER, the economy reached a business cycle peak. The end of each shaded area indicates a business cycle trough. Note how real GDP falls from peak to trough and then generally grows from trough to peak. An obvious difference in the two business cycle measures is that the measure used in this article is a continuous variable, while the NBER's takes the form of peak and trough dates. As a result, my measure not only indicates when a recession occurs, but also the intensity of the recession. Apart from these differences, the two measures appear reasonably consistent. For example, near the trough of every NBER recession, my measure of the business cycle is always negative. However, the two measures do not always agree. According to my measure, the economy was in recession in 1967 and 1987, while the NBER did not declare a recession then. In part, this is because there must be several quarters of negative GDP growth before the NBER declares a recession. The procedure I use only requires a temporary slowdown.

The Data

I consider a broad variety of investment goods, as outlined in Table 1. The broadest measure of investment is total private investment (TPI). This measure includes all private expenditures on capital goods and consumer goods designed to last more than three years.[6] This is a broader measure of investment than the conventional NIPA measure of investment, private fixed investment (PFI), which excludes expenditures on consumer goods. Within TPI, I define two main components, nonresidential and residential. Nonresidential has two main subcomponents, structures (NRS, for example, factory buildings and office buildings) and producer durable equipment (PDE, for example, auto-assembly robots and personal computers). Similarly, residential is broken down into residential structures and

Table 1 Measures of Investment Used in the Analysis

	Nominal share			Real share			Business cycle volatility	
	1958	1978	1998	1958	1978	1998	$\sigma_{q^i}/\sigma_{q^Y}$	$\sigma_{p^i}/\sigma_{q^Y}$
Total private investment	0.2184	0.2641	0.2378	0.1558	0.2154	0.2627	2.97	0.55
Nonresidential	0.4165	0.4496	0.4656	0.4254	0.4356	0.4815	2.83	0.98
Structures	0.1730	0.1510	0.1226	0.2613	0.1674	0.1032	2.66	0.90
Nonresidential buildings	0.0976	0.0821	0.0906	0.1435[a]	0.0982	0.0764	3.63	0.49
Utilities	0.0418[a]	0.0399	0.0238[b]	0.0552[b]	0.0412	0.0226[b]	2.67	0.64
Mining exploration, shafts, & wells	0.0233	0.0255	0.0115	0.0302[a]	0.0207	0.0089	5.49	3.19
Producer durable equipment	0.2438	0.2985	0.3430	0.1981	0.2676	0.3841	3.18	0.85
Information & related	0.0355	0.0690	0.1153	0.0080[c]	0.0327	0.1300[a]	3.05	0.95
Industrial	0.0796	0.0783	0.0734	0.1134[c]	0.0958	0.0747[a]	3.63	0.91
Transportation & related	0.0598	0.0782	0.0871	0.0797[c]	0.0900	0.0762[b]	5.25	0.63
Residential	0.5830	0.5504	0.5344	0.5730	0.5646	0.5200	3.98	0.43
Residential structures & equipment	0.2188	0.2175	0.1782	0.3259	0.2526	0.1546	6.24	0.57
Single family structures	0.1289	0.1203	0.0897	0.2118[a]	0.1350	0.0754	8.89	0.81
Multifamily structures	0.0228	0.0212	0.0122	0.0406[a]	0.0257	0.0109	10.80	0.81
Other structures	0.0627	0.0715	0.0721	0.1028[b]	0.0860	0.0645	3.18	0.34
Consumer durables	0.3643	0.3329	0.3562	0.2835	0.3184	0.3663	2.99	0.61
Motor vehicles & parts	0.1453	0.1539	0.1414	0.1361	0.1629	0.1290	5.16	0.94
Furniture & household equipment	0.1659	0.1223	0.1443	0.0918	0.0933	0.1717	1.94	0.54
Other	0.0534	0.0567	0.0705	0.0551	0.0670	0.0697	1.52	0.59
Private fixed investment	0.6357	0.6671	0.6438	0.7294	0.6812	0.6333	3.09	0.58

[a] 1959 data.
[b] 1995 data.
[c] 1960 data.

Notes: See box 1 for a description of the notation. For total private investment and gross domestic product. Y, nominal shares in the first row are $P_t^{TPI} q_t^{TPI}/(P_t^Y q_t^Y)$. Nominal and real shares for investment good i in the other rows are given by $P_t^Y q_t^Y/(P_t^{TPI} q_t^{TPI})$. Real shares for total private investment and gross domestic product are q_t^{TPI}/q_t^Y. Real shares for investment good i in the other rows are given by q_t^Y/q_t^{TPI}.

Source: U.S. Department of Commerce. Bureau of Economic Analysis. 1947–98, "National income and product accounts." *Survey of Current Business*, and author's calculations of the business cycle component of the data.

equipment (RSE, for example, single family homes and refrigerators) and consumer durables (CD, for example, televisions and vacuum cleaners). These four major subcomponents of TPI are then broken down further.[7]

The "nominal share" and "real share" data provide information on the relative magnitudes of expenditures on the different measures of investment, as well as a preliminary indication of interesting trends in relative prices. The nominal and real shares for TPI are calculated as the ratio of nominal and real TPI relative to nominal and real GDP, respectively. For example, in 1958 nominal TPI expenditures were 22 percent of nominal GDP and real TPI expenditures were 16 percent of real GDP. The remaining shares are calculated using TPI as the base for the share calculations. For example, PDE expenditures accounted for 24 percent of nominal TPI and 20 percent of real TPI in 1958.[8] (I explain the last two columns in Table 1 in the section on prices of investment goods over the business cycle.)

Table 1 reveals several interesting facts about how expenditures on investment have changed since 1958 and underlying trends in relative prices. First, nominal TPI expenditures have been roughly stable (abstracting from short-run movements) as a fraction of nominal GDP since the late 1950s. Yet, the real quantity of this broadest measure has been growing as a fraction of real GDP. In 1958, TPI was 16 percent of 1992 chain-weighted GDP, compared with 26 percent in 1998. The fact that nominal and real shares behave in this way is an indication that the relative price of this bundle of investment goods fell between 1958 and 1998. Notice that there are differences between real and nominal shares for many of the components of investment listed in Table 1, suggesting that trends in relative prices are exhibited by many of the subcomponents of TPI. Second, the difference between the real shares of TPI and PFI (the former is a fraction of GDP, while the latter is a fraction of the former) is seen to be due to the increasing quantities of consumer durables being purchased. Third, the much talked about "information age" manifests itself here as the huge increase in the fraction of TPI that has been due to expenditures on information and related equipment since 1960. In 1960 this type of investment accounted for less than 1 percent of real TPI. By 1995, its share had grown to 13 percent. Finally, note that both residential and nonresidential structures account for less of TPI in 1998 than in 1958.

Trends in Investment Good Prices

In this section, I explain two main findings relating to the long-run behavior of relative prices for the various components of investment listed in Table 1. First, the relative price of TPI has fallen consistently since the mid-1950s. Second, there is considerable heterogeneity in the long-run behavior of the prices of the subcomponents of TPI. Generally, the behavior of the price of TPI is dominated by dramatic drops in the prices of PDE and CD, which are also evident in the prices of most of the main subcomponents of these investment aggregates. The prices of RSE and NRS and their subcomponents, while exhibiting trends over subsamples of the period studied, have not fallen as consistently and their changes over time are much smaller than those of PDE and CD.

Figure 2 displays the relative price trend evidence. The black lines in Figure 2 are measures of the (natural logarithm of the) relative price of each of the investment components listed in Table 1 over the period for which data are available.[9] The gray lines are the trends calculated in the same way as the trend of real GDP displayed in Figure 1. The first column of panels in Figure 2 displays prices and trend lines for the main aggregates. The remaining columns display prices and trends for the four broad categories of TPI and their main subcomponents.

Figure 2 shows that the relative prices of different components of investment have behaved quite differently in the postwar era. The price of the broadest investment measure, TPI, has been falling consistently since the early 1950s. Since the plot of the relative price of TPI is in natural logarithms, one can take the difference between the prices for two years to calculate the percentage change. This procedure indicates that the price of TPI in terms of consumption goods fell about 42 percent between 1958 and 1998.

Studying the other plots in Figure 2, we see that this large drop in the price of TPI can be attributed to strong downward trends in PDE (particularly information and related and transportation equipment) and CD (all three types). The drop in the relative price of information equipment is particularly dramatic, at almost 200 percent since 1961. The prices of NRS and its components were generally rising until the late 1970s, were falling for most of the rest of the sample period, and have started to rise again in the 1990s. RSE and its components display a similar pattern. Generally, the long-run changes in structures prices have been much smaller than in PDE and CD prices. When the investment components are aggregated into nonresidential and residential, the strong downward trends in PDE and CD prices dominate the changing trends in structures.[10]

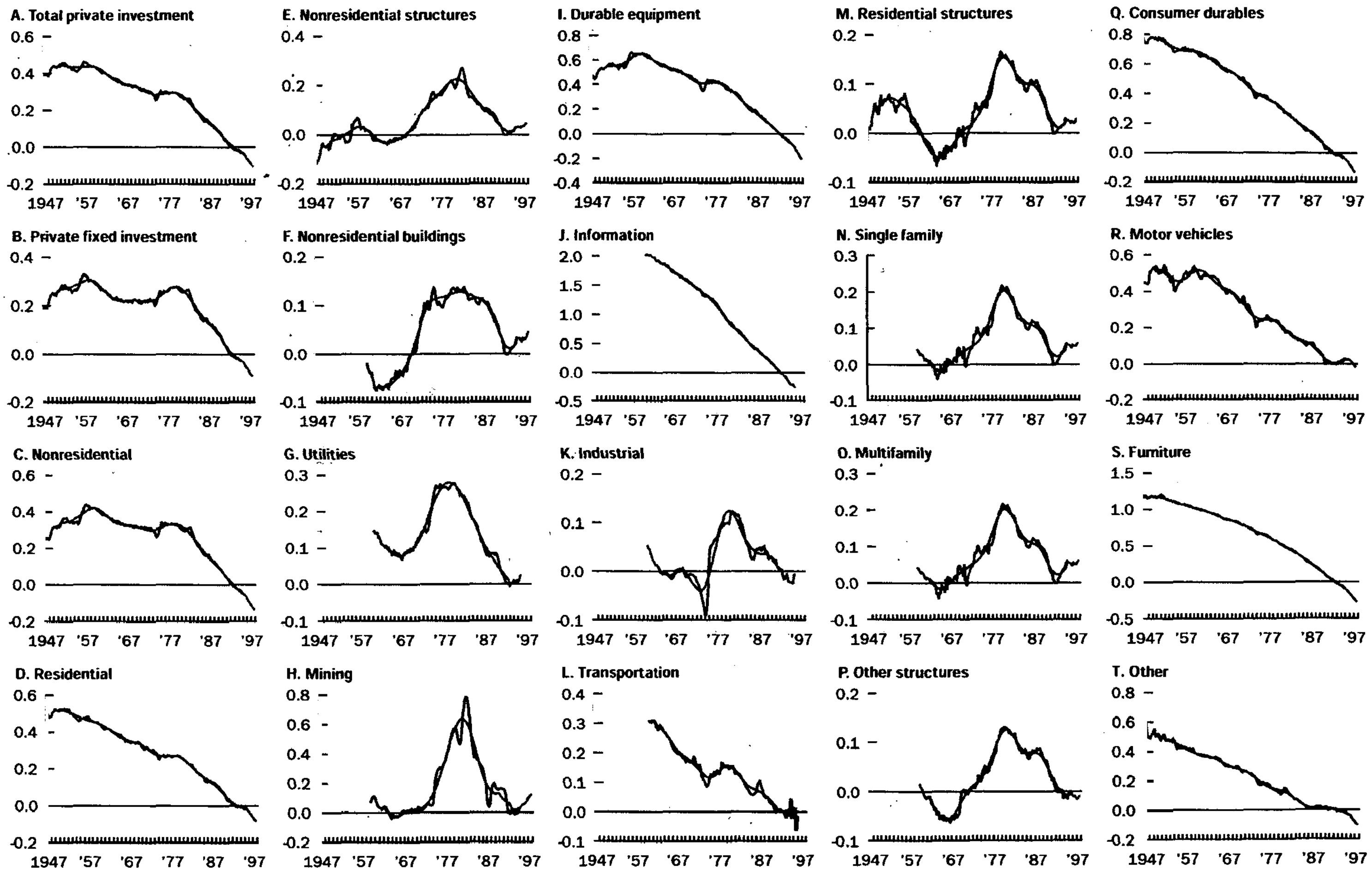
Prices of Investment Goods over the Business Cycle

My objective here is to determine the extent to which investment good prices are generally procyclical, countercyclical, or *acyclical* (do not display any distinctive pattern over the business cycle). I find that, generally speaking, prices of PDE, NRS, and their components are countercyclical, prices of RSE and its components are procyclical, and prices of CD and its components are acyclical. There is some sample period sensitivity, as outlined below.

In Table 1, the column headed σ_{qi}/σ_{qy}, indicates the relative volatility of the different investment components over the business cycle. This is the standard deviation of the business cycle component of the indicated real quantity series divided by the standard deviation of the business cycle component of real GDP. We see that TPI varies almost three times as much as GDP. The most volatile components of investment are single family structures, multifamily structures, and consumer expenditures on motor vehicles and parts. The least volatile components are NRS, furniture and household equipment, and the "other" component of CD. The column headed σ_{pi}/σ_{qy} indicates the relative volatility of the prices of different investment components over the business cycle. This is the standard deviation of the business cycle component of the indicated relative price series divided by the standard deviation of the business cycle component of real GDP. The prices are much less volatile than the quantities. With one exception (mining exploration, shafts, and wells), all the prices are less volatile than real GDP over the business cycle.

As a preliminary look at the cyclicality of investment good prices, Figure 3 displays the business cycle components of the prices (gray lines) and quantities (black lines) of seven of the broadest measures listed in Table 1, along with the business cycle component of the deflator for consumption of nondurables and services. The latter price is used in the denominator of all the investment relative prices, so its business cycle dynamics will influence all the relative price measures discussed here.[11]

Notice first that the consumption deflator rises in all but one recession, 1981:Q3–82:Q4 (see shaded areas in Figure 3). This is a force for procyclicality of investment good prices. For example, if the price deflator for an investment good were constant,


A. Total private investment
E. Nonresidential structures
I. Durable equipment
M. Residential structures
Q. Consumer durables
B. Private fixed investment
F. Nonresidential buildings
J. Information
N. Single family
R. Motor vehicles
C. Nonresidential
G. Utilities
K. Industrial
O. Multifamily
S. Furniture
D. Residential
H. Mining
L. Transportation
P. Other structures
T. Other
1947 '57 '67 '77 '87 '97

then the real price of that good would be procyclical. As expected, the quantities are generally procyclical, although the peaks and troughs do not exactly coincide with the NBER dates. The prices do not display as consistent a pattern as the quantities. For example, sometimes the price of TPI moves with the quantity of TPI (1950s, 1960s, and 1990s) and sometimes it moves in the opposite direction (1970s and 1980s). More distinct patterns emerge when TPI is decomposed into nonresidential and residential. In the 1950s and 1990s, the prices and quantities of nonresidential appear to move closely together. In the 1960s, 1970s, and 1980s, prices and quantities of this investment measure generally move in opposite directions. Prices and quantities of residential show more evidence of moving together. The most striking pattern to emerge among the subcomponents of nonresidential and residential is in PDE. With the exception of the 1950s, almost every time the quantity of PDE moves up, the price of PDE moves down. This suggests countercyclical behavior in the real price of PDE.

For a more formal examination of how the prices of investment goods vary with the business cycle, I use a *cross-correlogram.* A cross-correlogram is a diagrammatic device for describing how two variables are related dynamically. For example, it provides a measure of whether, say, movements in one variable tend to occur at the same time and in the same direction as movements in another variable. It can also be used to measure whether, for example, positive movements in a variable tend to occur several quarters ahead of positive movements in another variable.

The basis for the cross-correlogram is the *correlation coefficient,* or correlation. A correlation is a measure of the degree to which two variables move together and always takes on values between –1 and 1. If a correlation is positive, then the two variables are said to be positively correlated. Similarly, if a correlation is negative, the variables are said to be negatively correlated. Larger absolute values in a correlation indicate a stronger pattern of moving together. A correlation for two variables measured contemporaneously is a measure of how much two variables move together *at the same time.* A correlation can be computed for two variables measured at different times. For example, we can measure the correlation between variable x at time t and variable y at time $t - k$, where k is a positive integer. This would measure the degree to which variations in y occur before movements in x. A cross-correlogram plots these correlations for various values of k.

Figure 4 displays cross-correlograms (along with a two-standard-deviation confidence interval, a measure of how precisely the correlations are estimated) for various business cycle components of real investment and GDP, $-6 \leq k \leq 6$. For example, panel A of figure 4 displays the correlations of real nonresidential investment at date t and real GDP at date $t - k$ for the various values of k. The fact that the correlation for $k = 0$ is positive and close to 1 for all the plots in figure 4 shows that all the components of investment displayed are strongly positively correlated with GDP contemporaneously. This confirms the impression given by Figure 3 that real expenditures on these investment goods are

◀

Figure 2 Trends in investment good prices, 1947–98 (logarithm). Notes: Relative price (black line) is a measure of the (natural logarithm of the) relative price of each of the investment components listed in Table 1 over the period for which data are available. The trend (colored line) is calculated in the same way as the trend of real GDP displayed in Figure 1. Panels A through D show prices and trends for the main aggregates. Panels E through T show prices and trends for the four broad categories of total private investment, along with their main subcomponents. Source: See Figure 1.

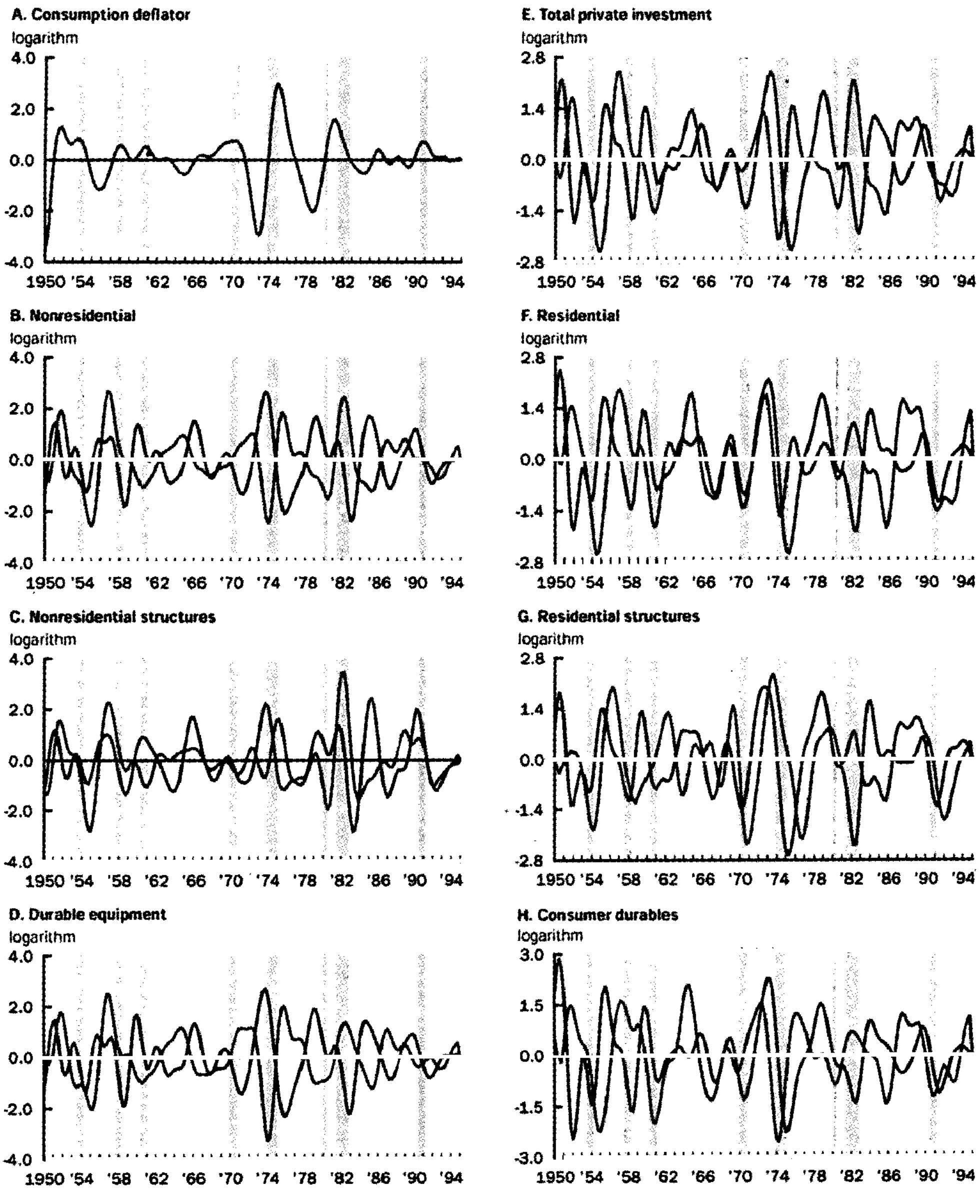


Figure 3 Business cycle components of investment good prices and quantities. Notes: Each business cycle component has been scaled by its standard deviation, and all data are quarterly. The colored lines represent the business cycle component of the price series for the indicated variable and the black lines represent the business cycle component of the quantity series for the indicated variable. Shaded areas indicate recessions as determined by the National Bureau of Economic Research. Source: See figure 1.

strongly procyclical. Notice that the largest correlations for nonresidential and its two main subcomponents, NRS and PDE, are for $k > 0$. This says that these components of investment tend to *lag* GDP over the business cycle. Another way of saying this is that movements above trend in GDP tend to occur *before* movements above trend in these measures of investment. On the other hand, the largest correlations for residential and its main subcomponents, RSE and CD, are all for $k < 0$. This says that these components of investment *lead* output over the business cycle. Because the correlations in Figure 4 are mostly positive, this figure shows that the main components of investment are generally procyclical. (If they had been mostly negative, then this would have been evidence of countercyclicality. If the correlations were mostly close to zero, this would have been evidence of acyclicality.)

Figure 5 displays cross-correlograms (with standard errors) for the prices of the broadest measures of investment and real GDP. The plots in Figure 3 indicate that there may be some sample period sensitivity in the estimation of the underlying correlations, so Figure 5 displays cross-correlograms based on two sample periods. The first column of panels in Figure 5 is based on the sample period 1947:Q1–98:Q3 and the second column is based on 1959:Q1–98:Q3. Notice that none of the correlations for the TPI price based on the longer sample are significantly different from zero. This means that the price of the broadest measure of investment is essentially acyclical. There is some evi-

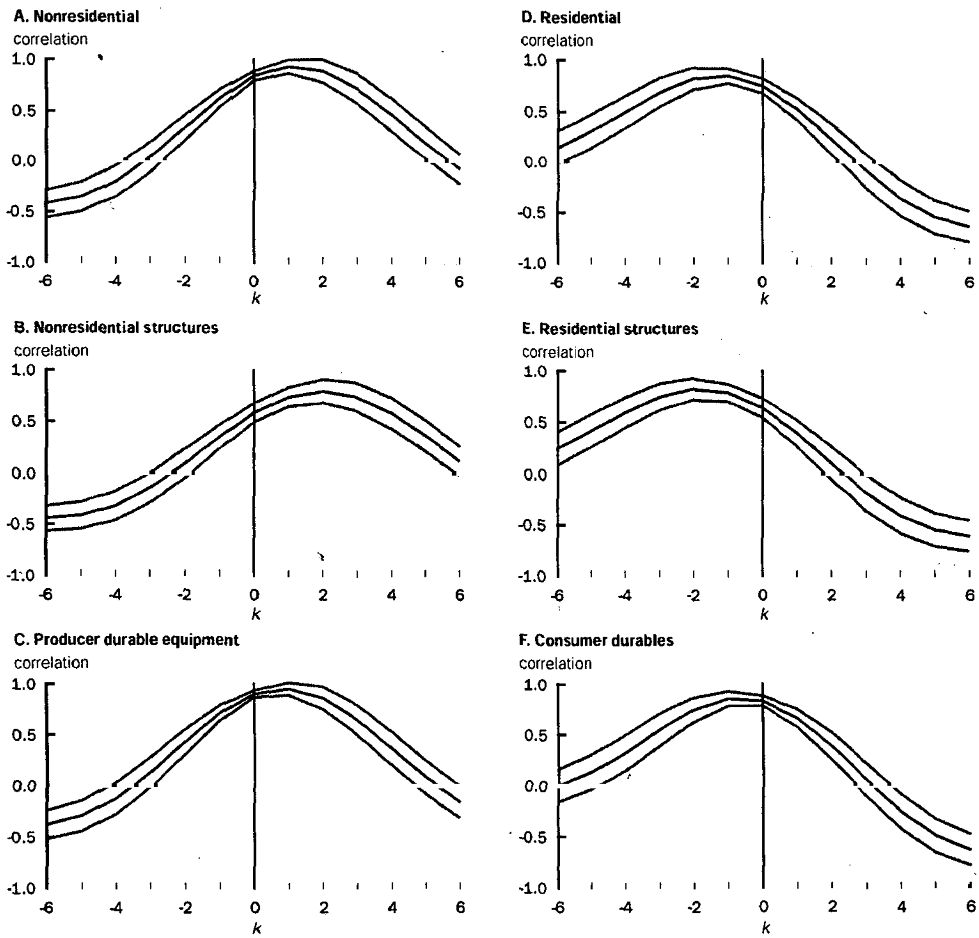


Figure 4 Business cycle correlations between investment quantities (t) and output (t-k). Note: Black lines are point estimates of correlations for the indicated series: colored lines are a two-standard-error confidence band. Source: See figure 1.

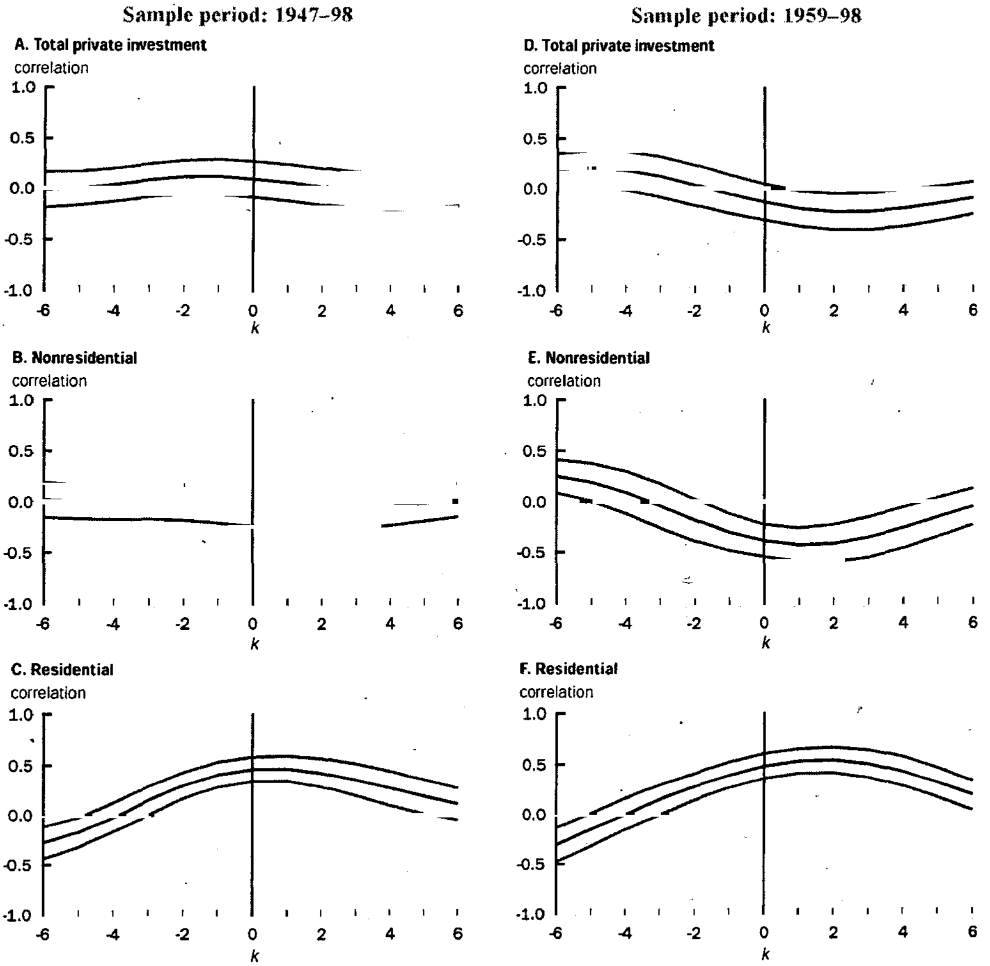


Figure 5 Business cycle correlations between investment prices (t) and output (t-k). Note: Black lines are point estimates of correlations for the indicated series; colored lines are a two-standard-error confidence band. Source: See figure 1.

dence of countercyclical movements in this price for the shorter sample, although the correlations in this case are generally not very large in absolute value or statistically significant.

The cyclical behavior of prices for the narrower investment aggregates displayed in Figure 5 reveals that the lack of any distinct cyclical pattern for the price of TPI masks interesting differences between the prices of nonresidential and residential goods. Over the longer sample, the nonresidential price is estimated to be essentially acyclical, but the residential price is clearly procyclical. Over the shorter sample the nonresidential investment price is clearly countercyclical and the residential price remains procyclical. The difference in the estimated cross-correlogram for nonresidential over the two sample periods turns out to be due to differences in the behavior of the price of PDE in the 1950s compared with the later sample period (see Figure 3).

The evidence in Figure 5 suggests two things. First, the cyclical behavior of investment good prices depends to some extent on the sample period examined. Second, considering a broad investment aggregate masks potentially interesting cyclical characteristics of more narrowly defined investment good prices. Figures 6 and 7 try to uncover whether the cyclical behavior of nonresidential and residential prices also masks different cyclical behavior among the subcomponents of these broad investment aggregates. These figures display price–output cross-correlograms for the main subcomponents of nonresi-

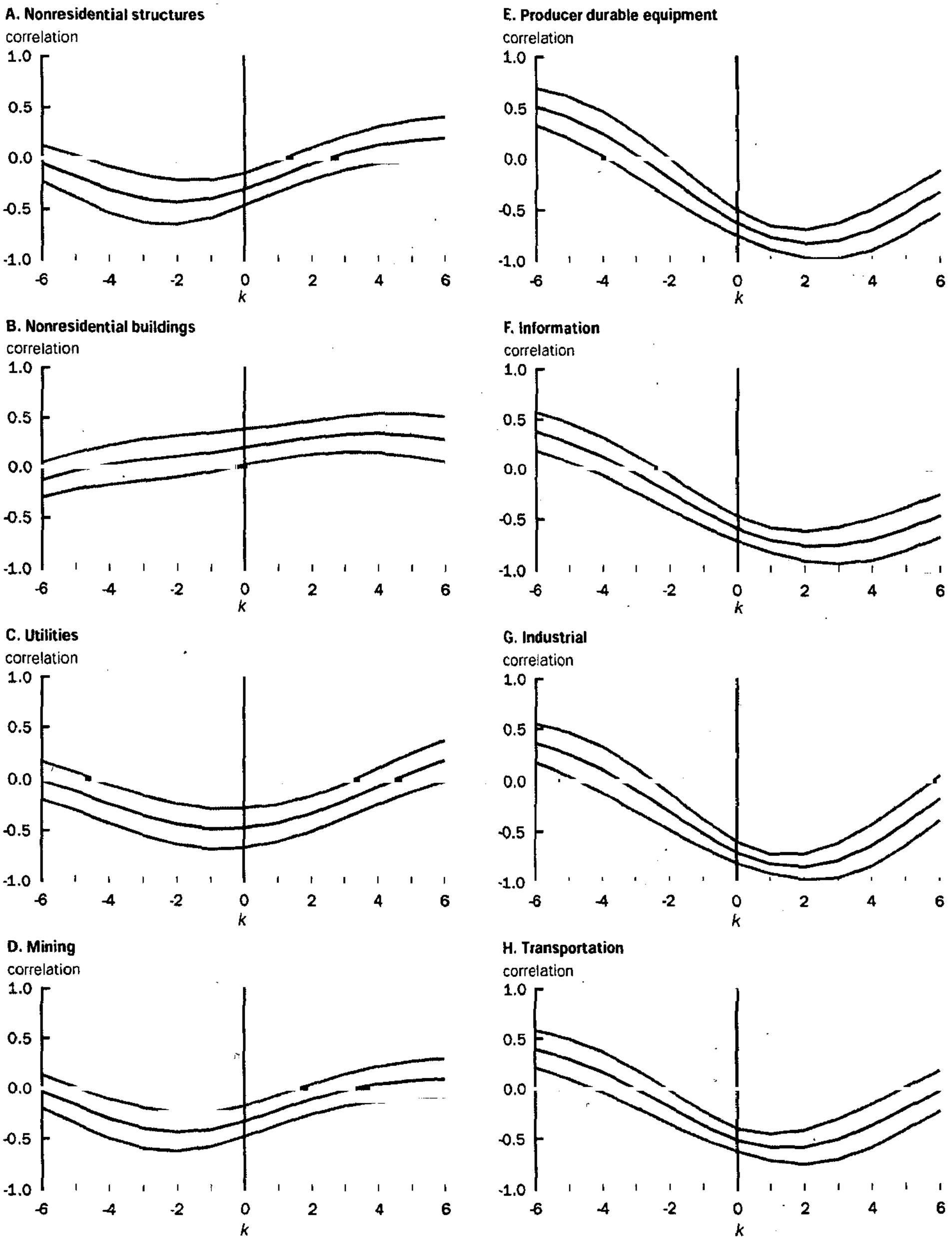


Figure 6 Business cycle correlations between nonresidential prices (t) and output (t-k). Note: Black lines are point estimates of correlations for the indicated series; colored lines are a two-standard-error confidence band. Source: See Figure 1.

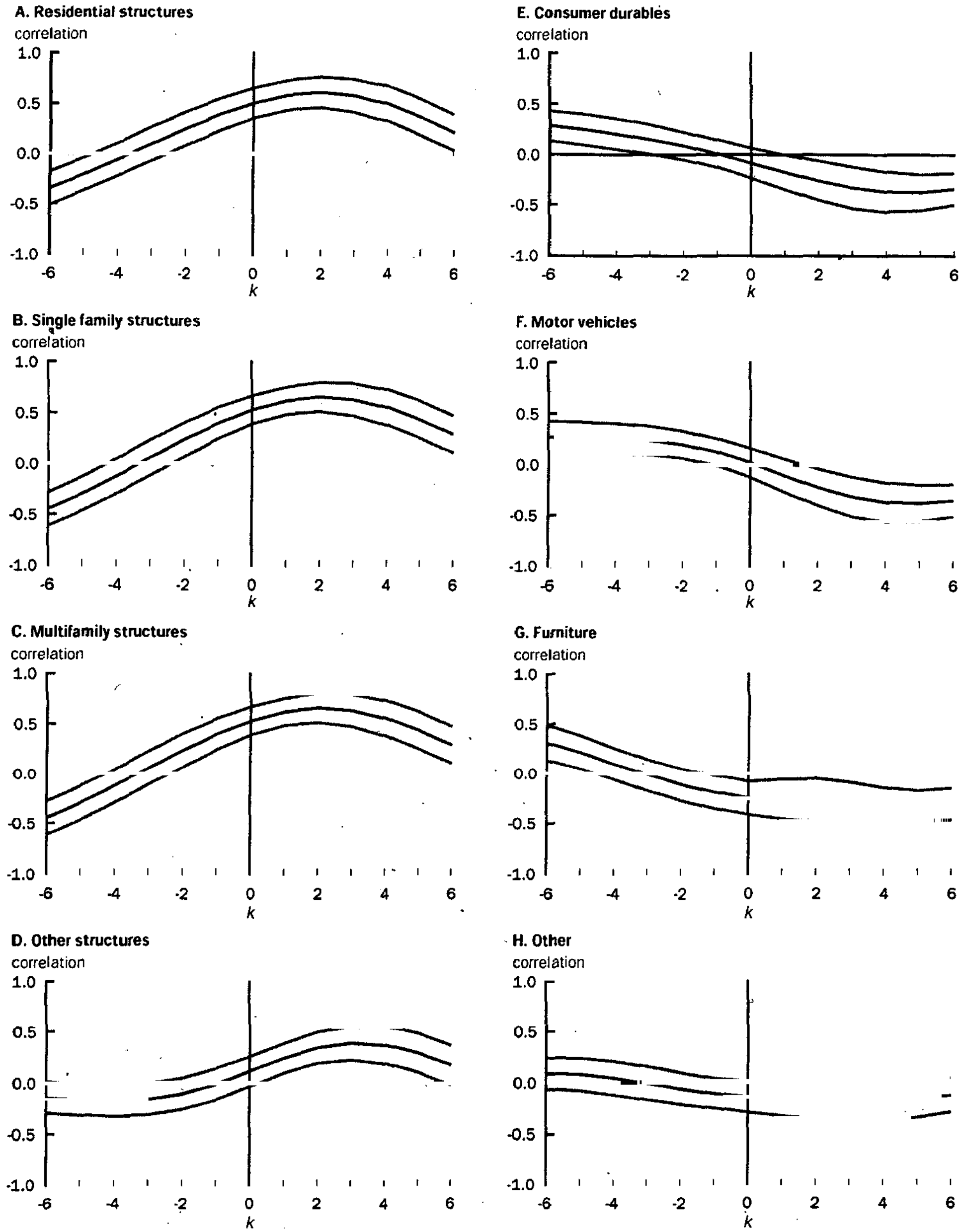


Figure 7 Business cycle correlations between residential prices (t) and output (t-k). Note: Black lines are point estimates of correlations for the indicated series; colored lines are a two-standard-error confidence band. Source: See Figure 1.

dential and residential. Due to data availability, the sample period for estimating the correlations is 1959:Q1–98:Q3.

The first column in Figure 6 pertains to NRS and its main subcomponents, nonresidential buildings, utilities, and mining. The price of NRS is significantly countercyclical. This appears to be mainly driven by the price of utilities and mining. The second column of Figure 6 pertains to PDE and its main subcomponents, information and related equipment, industrial equipment, and transportation equipment. There are two observations to make here. First, the price of PDE is strongly and significantly countercyclical. The contemporaneous ($k = 0$) correlation is–0.63 with a standard error of 0.03. The largest correlation in absolute value is for $k = 2$, indicating that this price lags output by about two quarters, about the same as the quantity of PDE (see Figure 4). The second observation is that the prices of the main components of PDE behave almost identically: They are strongly and significantly negatively correlated with output and lag output by about two quarters. The behavior of the industrial equipment price is particularly striking, given that the long-run behavior of this price is so different from that of the other two subcomponents of PDE (see Figure 2).

Figure 7 is constructed similarly to Figure 6, with RSE and its subcomponents in the first column and CD and its subcomponents in the second column. This figure shows that prices of RSE are generally procyclical and prices of CD goods are mostly acyclical. The behavior of RSE is driven mostly by the cyclicality of single and multifamily structures. Interestingly, despite the fact that investment in RSE tends to lead output over the business cycle, the real price of RSE and its components lags output. The real price of CD is driven mostly by motor vehicles and other. Of the subcomponents of CD, only the furniture price displays significant countercyclicality.

Summary of the Evidence

The key features of the evidence presented in this section can be summarized as follows. First, there is strong evidence of a downward trend in the price of investment goods in terms of consumption goods. This downward trend is concentrated among components of PDE and CD. Second, the broadest category of investment, TPI, displays little distinct cyclical variation over the sample period 1947:Q1–98:Q3, but is moderately countercyclical in the later period, 1959:Q1–98:Q3. If we are willing to abstract from the 1950s, say because of the dominating influence of the Korean war, then it seems reasonable to say that the price of the broadest component of investment is weakly countercyclical. Certainly it is difficult to make the case that this price is procyclical, regardless of the sample period considered.

Many components of TPI display distinct cyclical characteristics, even if we include the 1950s. The prices of the two main components, nonresidential and residential, behave differently. The former is significantly countercyclical and the latter is significantly procyclical. The behavior of the nonresidential price is dominated by the PDE price. The PDE price is strongly countercyclical, as are the prices of all its subcomponents. The price of NRS is mildly countercyclical, but this pattern is not shared by all its subcomponents. The behavior of the residential price is dominated by RSE prices, which are strongly procyclical. CD prices are acyclical or weakly countercyclical.

IMPLICATIONS FOR GROWTH AND THE BUSINESS CYCLE

How does the trend and cycle behavior of investment goods prices presented above challenge conventional views about growth and business cycles? Next, I discuss various attempts to reconcile theory with the evidence and some empirical work that sheds light on the plausibility of competing theories.

Growth Theory

Recent years have seen an explosion of theoretical and empirical research into economic growth.[12] On the theoretical side, two leading classes of models of the determinants of economic growth have emerged. The first is based on the accumulation of *human capital* and follows from the work of Lucas (1988). Human capital consists of the abilities, skills, and knowledge of particular workers. The basic idea behind this view of economic growth is that it is fundamentally based on improvements in the stock of human capital of workers over time. This view of growth holds that, other things being equal, the larger is the stock of human capital of workers, the more productive they are. This means that one expects an improvement in the stock of human capital to increase the amount of output of *any* good that can be produced for a fixed quantity of workers and capital. In this sense, growth due to the accumulation of human capital has a homogeneous impact on the economy's ability to produce goods.

The second leading class of models focuses on research and development. Pioneering work along these lines includes Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992). One of the key insights of this literature is that growth can emerge if there are nondecreasing returns to produced factors of production (such as knowledge or capital, but not labor).[13] The bottom line of this theory is similar to that of the human capital models. Improvements in technology due to research and development usually increase the productivity of all factors of production. Consequently, if there is such an improvement in technology, more of all goods can be produced with a fixed quantity of capital and labor. Again, technological change is assumed to have a homogeneous impact on produced goods.

The evidence on trends in investment good prices, particularly the trend in the price of PDE, challenges these views of growth, because it strongly suggests that there have been substantial improvements in technology that have affected one kind of good but not another. Specifically, the data suggest that the quality and technology of capital goods production have advanced almost nonstop since the end of World War II. Why do the data suggest this? Assuming that the prices and quantities of PDE are correctly measured, the real price of PDE measures how many (constant quality) consumption goods need to be sold in order to raise the funds to purchase one (constant quality) unit of PDE. If this price has been falling, then fewer and fewer consumption goods are needed to buy a unit of PDE. This suggests that the supply of PDE has grown relative to the supply of consumption goods. One way the supply of PDE can rise in this way is if the technology for producing capital goods improves at a faster rate than that for producing consumption goods. In this case, the same amount of capital and labor applied to producing PDE or consumption goods will yield more PDE than consumption as time passes. That is, the supply of PDE will grow relative to consumption goods. The basic logic of supply and demand then dictates that the price of PDE in terms of consumption goods must fall. Greenwood et al. (1997) build on this intuition to show how the trend in the relative price of PDE and the associated increase in the share of PDE in aggregate output (see Table 1) can be accounted for in a growth model in which most growth is due to capital-embodied technical change. In addition, the authors argue that other potential explanations for the price and quantity trends are implausible or boil down to essentially the same explanation.[14]

Greenwood et al. (1997) apply their model of growth to reevaluate conventional estimates of the importance of technological change in improving standards of living. This line of research is called growth accounting. The effects of technical change using standard models, like the ones briefly described above, can be summarized by multifactor productivity, which is also called the Solow residual. Multifactor productivity is an index of the quantity of aggregate output that can be produced using a fixed quantity of (quality-adjusted) capital and labor. The higher the multifactor productivity, the more output can be produced. Traditionally, most of growth is viewed as being due to improvements in multifactor productivity. Greenwood et al. (1997) use their model to show that approximately 60 percent of all improvements in productivity can be attributed to capital-embodied technical change, while the multifactor productivity index accounts for the rest. This says that capital-embodied technical change is a fundamental part of growth.

Business Cycle Theory

To assess the cyclical evidence on relative prices, we need to understand how various shocks to the economy might influence the cost of investment goods compared with consumption goods. Figure 8 displays a production possibilities frontier (PPF) for consumption and investment goods. The PPF depicts the various quantities of consumption and investment goods that can be produced if capital and labor are fully employed and used efficiently. The shape of the frontier reflects the fact that, holding fixed the quantity of labor and capital employed in producing goods, it is costly to shift production toward either producing more consumption goods or more investment goods.[15] This is reflected in the figure by the increase in the (absolute value of the) slope of the frontier as one moves from the upper left to the lower right. In a competitive equilibrium, the slope of the frontier equals the relative price of the goods. Hence, as more investment goods are produced, the relative price of investment goods rises.

The PPF summarizes the supply side of the economy. The actual price in a competitive equilibrium is determined by the interaction of the demand for consumption and investment goods with the supply. Suppose that the demand for consumption and investment goods dictates that the quantity of consumption goods and investment goods actually produced is given by C_0 and I_0 in Figure 8. Now, suppose a Keynesian demand shock—for example, an increase in the money supply which lowers interest rates—increases the demand for investment goods relative to consumption goods. Since this is a demand shock,

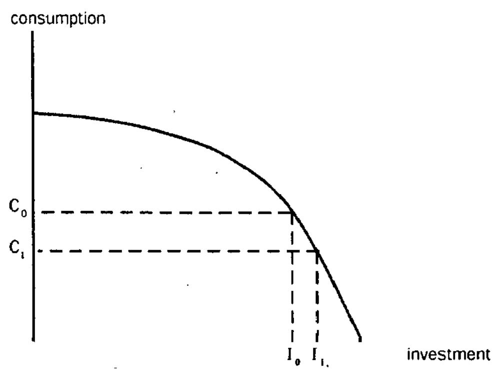


Figure 8 Effect of increase in demand for investment goods.

the PPF in Figure 8 does not change. The change in demand leads to a movement down the frontier, say to a point where consumption and investment are given by C_1 and I_1. Since the slope of the frontier is steeper at this point, the relative price of investment goods must rise. If aggregate output is driven by shocks to investment demand, then the price of investment goods is predicted to be procyclical.

An aggregate supply shock has a similar implication. The conventional assumption about these kinds of shocks is that they raise multifactor productivity and influence all produced goods symmetrically. This is shown in Figure 9 as a proportional shift out in the solid line PPF to the dashed line PPF. The dashed line PPF has been drawn so that its slope is identical to the slope of the solid line PPF along a straight line from the origin. This means that if the ratio of consumption to investment goods produced before and after the technology shock is constant, then the relative price of investment goods will be unchanged. However, this is not what is predicted in standard models. These models say that when a good technology shock arrives, which raises the productivity of all factors of production, the optimal response of individuals is to smooth consumption. That is, not have consumption change too much in the short run. The result of this is that investment rises more than consumption. In Figure 9, this is represented by consumption and investment changing from C_0 and I_0 before the productivity shock to C_1 and I_1 after the shock. It follows that the price of investment goods must rise in this case as well. Since output also rises with a positive technology shock, the price of investment goods is predicted to be procyclical.[16]

In view of the cyclical evidence presented earlier, these model predictions are problematic. They are consistent with the behavior of residential investment, but inconsistent with the behavior of the other major components of investment and the broadest measure, TPI. Why are investment goods prices not procyclical? The two leading explanations involve assumptions about the technology for producing investment goods. One is based on *increasing returns* to scale in the production of investment goods (but not consumption goods). The other is based on a variation in the rate of capital-embodied technical change. The increasing returns view assumes that the more investment goods that are produced, the less costly it is to produce a unit of investment goods. One way to represent this is shown in Figure 10, which displays a pseudo-PPF.[17] Notice that the shape is different from Figures 8 and 9. Now when more investment goods are produced relative to consumption

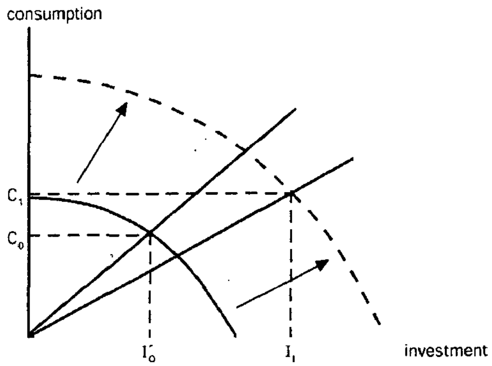


Figure 9 Effect of increase in multifactor productivity.

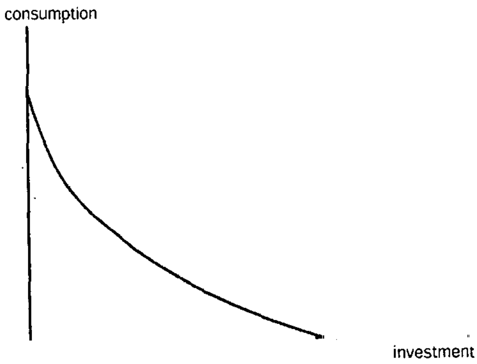


Figure 10 Increasing returns to scale case.

goods, the price of investment goods falls. In this case, both aggregate technology shocks and Keynesian demand shocks can lead to countercyclical relative prices.

To understand the embodied technology view, consider an increase in the productivity of producing investment goods that has no direct impact on the production of consumption goods. This could take the form of improvements in the efficiency of producing investment goods. It could also take the form of an improvement in the quality of investment goods produced so that a given quantity of capital and labor can produce a higher quantity of quality-adjusted goods. Either way, we can represent the change in technology as in Figure 11. The improvement in technology is shown by the shift from the solid to the dashed frontier. Along the dashed frontier, for each quantity of consumption goods produced, more investment goods can be produced. Moreover, along any straight line from the origin, the slope of the dashed frontier is flatter than the solid frontier. That is, for any fixed ratio of consumption to investment goods, the investment goods are cheaper in terms of consumption goods after the change in technology. Now, after the increase in technology, there will be a shift in favor of the production of investment goods. If this shift is strong enough, the movement along the dashed frontier could in principle raise the investment good price. In practice, this does not happen. Since aggregate output rises after this kind of technology shock, if business cycles are in part driven by this kind of disturbance, then investment good prices could be countercyclical.

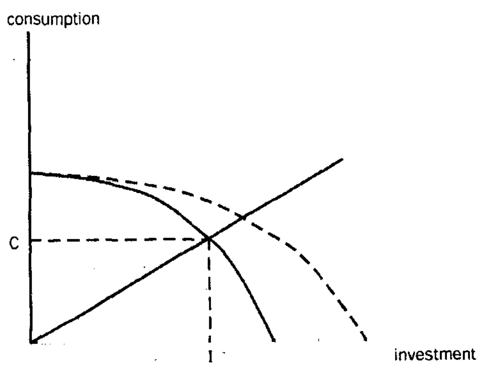


Figure 11 Effect of increase in capital-embodied technology.

EVALUATING THE THEORIES

Beyond the work of Greenwood et al. (1997), little has been done to evaluate the plausibility of the capital-embodied technological change theory of the trend evidence on investment prices. However, more work has been done to evaluate the differing views on the cyclicality of investment good prices.

Generally, the empirical evidence seems to go against the increasing returns interpretation of the cyclical evidence on prices. Harrison (1998) examines annual data on capital, labor, and value added in various industries in the consumption good sector and the investment good sector. She finds some empirical support for increasing returns associated with capital and labor in the production of investment goods. However, she does not find a sufficient degree of increasing returns to generate increasing returns in the factor of production, labor, that is variable in the short run. Consequently, the work does not support the increasing returns view. Other research on measuring increasing returns focuses on the manufacturing sector. Basu and Fernald (1997), Burnside (1996), and Burnside, Eichenbaum, and Rebelo (1995) have overturned previous empirical claims of increasing returns in the manufacturing sector, including capital equipment industries.

Other empirical work attempts to address a key implication of the increasing returns view—that the supply curve for investment goods slopes down. That is, holding other things constant, the cost of investment goods is diminishing in the quantity of investment goods produced. Shea (1993), in a study of many sectors of the economy, uses instrumental variables econometric techniques to distinguish *supply* shocks from *demand* shocks to trace out the slope of supply curves. The author's main conclusion is that, broadly speaking, supply curves slope up. Goolsbee (1998) focuses specifically on the supply of capital goods and uses a series of "natural experiments" (involving periodic changes in federal laws providing for investment tax credits) to identify a disturbance that affects the demand for investment goods but not the supply. He finds clear evidence of an upward sloping investment supply curve. To summarize, empirical work on the sign of the slope of the investment good supply schedule finds that it is positive.

Other research assesses the plausibility of the embodied technology view. Christiano and Fisher (1998) and Greenwood et al. (1998) evaluate business cycle models in which a major driving force for fluctuations is variations in capital-embodied technical change. They test the embodied technology view by examining the ability of their models to account for various business cycle phenomena. Both studies find that their models do about as well as other business cycle models in accounting for business cycle phenomena. As a measure of the importance of capital-embodied technical change as a driving force for business cycles, Greenwood et al. (1998) find that about 30 percent of business cycle variation in output can be attributed to this kind of shock. Christiano and Fisher (1998), in a very different model, find that about three-quarters of output fluctuations are due to this shock. Either way, the evidence suggests that variation in the rate of technical change embodied in capital equipment accounts for a significant proportion of business cycle variation in output.

New Evidence

Some new research attempts to distinguish the increasing returns view from the embodied technology view of the cyclical behavior of investment good prices. This evidence is based on two econometric procedures designed to identify disturbances to the aggregate economy that influence the demand for investment goods, but leave supply unchanged. The specific shocks considered are an exogenous increase in government purchases (that is an increase in government purchases that is unrelated to developments in the economy) and an exogenous monetary contraction.

In the government spending case, the idea is to investigate how particular investment quantities and prices respond to an exogenous increase in government purchases. The exogenous increase in government spending takes the form of a large military buildup (specifically the Korean war, the Vietnam war, and the Carter–Reagan buildup). The methodology is identical to that employed by Eichenbaum and Fisher (1998).[18] Figure 12 displays the estimates, which are based on quarterly data for 1947:Q1–98:Q3. The first row of Figure 12 plots the response to an exogenous increase in government purchases of real investment in PDE and RSE (solid lines) along with a 68 percent confidence band (gray lines). The second row plots the corresponding relative price responses. Interestingly, PDE investment rises and RSE investment falls.[19] Under the increasing returns view, we would expect the PDE price to fall and the RSE price to rise. The second row of plots indicates that the RSE price response is inconsistent with the increasing returns view, while the PDE price response seems to confirm it.

The monetary shocks case examines how quantities and prices of PDE and RSE respond to an estimate of a contractionary monetary disturbance. The methodology is standard[20] and has been summarized by Christiano (1996) (see also Christiano, Eichen-

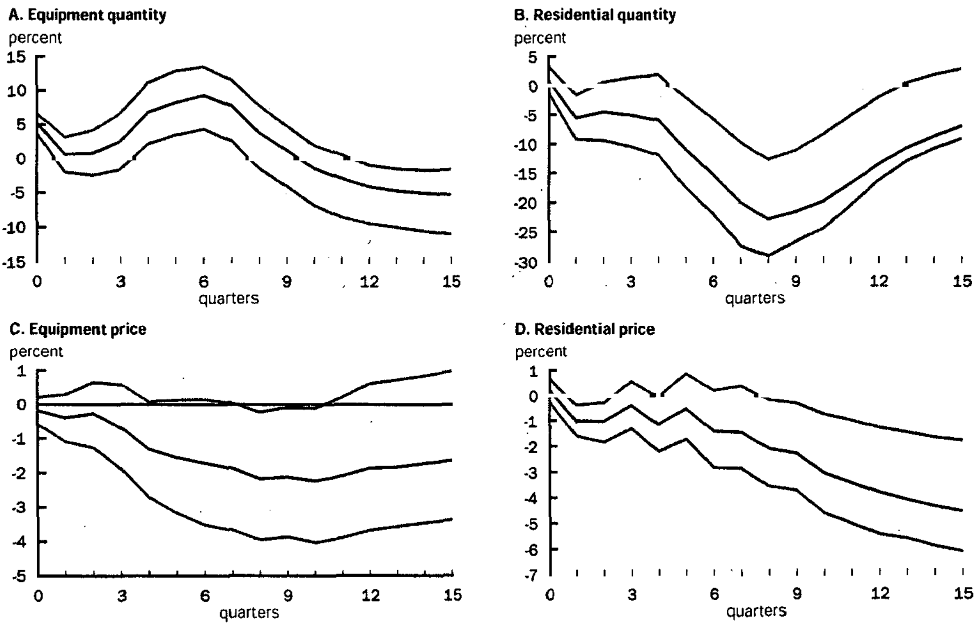


Figure 12 Response of investment quantities and prices to exogenous increase in government purchases. Notes: Black lines are point estimates of the responses; the colored lines are a symmetric 68 percent confidence band. Source: Author's calculations from data extracted from DRI Basic Economics database, 1959–98.

baum, and Evans, 1999). The estimated responses (along with a 95 percent confidence interval) are presented in Figure 13. Looking at the quantities in the first row of plots, notice that both PDE and RSE fall after an exogenous monetary contraction. Under the increasing returns view, one would expect the prices of both investment goods to rise. Studying the second row of plots, we see that the PDE price response is not significantly different from zero and the RSE price drops significantly.

Taken together, the evidence on the responses of RSE prices and quantities to government spending and monetary shocks goes against the increasing returns view. It conforms to a standard neoclassical view of investment, in the sense that it is consistent with the discussion of the production possibilities frontier in Figure 8. Of course, the increasing returns view is really intended to apply to PDE investment. The responses of PDE prices and quantities provide mixed signals. The responses to a monetary shock provide evidence neither for nor against increasing returns, since the quantity falls but the price response is not very precisely estimated and could be either positive, negative, or zero. The responses to a government spending shock might be viewed as evidence in favor of increasing returns. However, one interpretation of the PDE price response in this case is that it is dominated by the Korean war military buildup. This occurred just after World War II, when military spending had fallen from very high levels. The increasing returns that could support a lower price with higher investment might conceivably be due to the resumption of large-scale production at facilities that had been operating far below minimum efficient scale. If this is true, it seems more like a special case than an enduring feature of the U.S. economy.

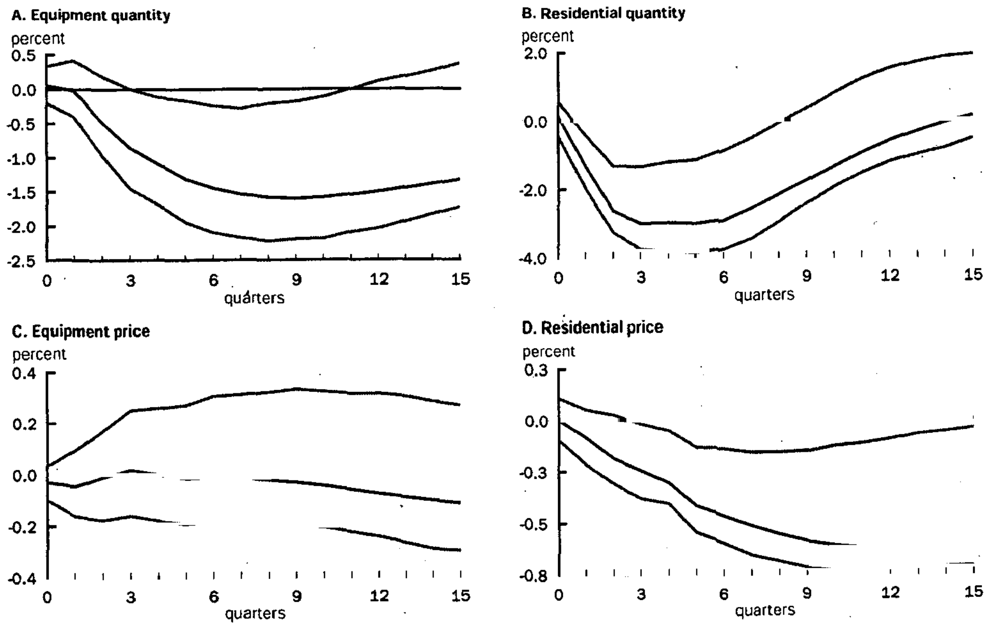


Figure 13 Response of investment quantities and prices to an exogenous monetary contraction. Notes: Black lines are point estimates of the responses; the colored lines are a symmetric 95 percent confidence band. Source: See figure 1.

CONCLUSION

In this article, I have presented evidence on trends and business cycle variation in the prices of investment goods relative to nondurables and services consumption. This evidence seems to go against conventional views of both business cycles and growth. How can one reconcile theory with the evidence? The leading views include one based on increasing returns to scale in the production of investment goods and another based on capital-embodied technical change. While some of the evidence I presented could be viewed as supporting the increasing returns view, generally, there is little empirical support for increasing returns. At this point, then, the leading candidate to reconcile theory with the data appears to be the one based on capital-embodied technical change, that is, the embodied technology view.

This conclusion has implications for our understanding of growth and business cycles, future research on these subjects, and policy. The prospect of a comprehensive theory of growth and business cycles is appealing because of its simplicity. Disembodied technical change has gained credence for its supposed ability to account for growth and business cycles. Yet, the theory of business cycles based on disembodied technology has always been problematic because the shocks are hard to interpret. The growth accounting results of Greenwood et al. (1997) bring into question the growth implications of this theory as well. In the search for a comprehensive theory of growth and business cycles, then, advances in capital-embodied technology seem to offer a promising alternative. In addition, they provide a much more tangible notion of growth. These considerations suggest that future research on growth and business cycles that emphasizes capital-embodied technical change may be fruitful.

If growth and business cycles are originating from changes in capital-embodied technology, then the models we use for policy analysis have to incorporate this and, consequently, policy recommendations could change. To the extent that technological change is embodied in capital equipment, government policies that affect equipment investment could have a dual impact on growth via the quality and the quantity of capital goods. This could mean, for example, that investment tax credits directed toward improvements in the efficiency of capital equipment could have a significant impact on growth. More research is required to uncover the full implications of this.

The implications for stabilization policy of the embodied technology view are less obvious. The fact that it seems to supplant the increasing returns view means that the arguments for interventionist stabilization policy that this view supports are less compelling. For example, increasing returns could provide scope for policy intervention, because it either involves externalities or is inconsistent with perfect competition. Moreover, it makes animal spirits models more plausible, which also has implications for stabilization policy (see, for example, Christiano and Harrison, 1999). The embodied technology view is more in line with the real business cycle tradition, in which policy interventions are counterproductive. Real business cycle theory says that the business cycle is largely the result of optimal behavior by individuals in the economy interacting, for the most part, in perfectly competitive markets. Any policy interventions in such an environment tend to reduce overall welfare. To the extent that the embodied technology view is more compelling than previous incarnations of real business cycle models, it lends greater support to the argument that interventionist stabilization policy cannot improve the well-being of any individual in the U.S. economy without hurting some other individual. Of course, this still leaves open the possibility that equity considerations might be used to defend interventionist stabilization policy.

NOTES

1. Equivalently, higher quality goods of all kinds can be produced with the same amount of capital and labor. As described in more detail below, new models of endogenous growth have reduced forms, which have similar implications for growth accounting to those of models written in terms of exogenous disembodied technical change.
2. Examples of textbooks that emphasize the IS–LM model are Abel and Bernanke (1997), Gordon (1998), Hall and Taylor (1997), and Mankiw (1997).
3. For a survey of theories based on animal spirits, see Farmer (1993).
4. A good summary of this view is Prescott (1986). For a discussion of how this view can be used to explain the 1990–91 recession, see Hansen and Prescott (1993).
5. This section relies heavily on Christiano and Fitzgerald (1998, pp. 58–59).
6. This is the empirical counterpart to investment as it is usually defined in the real business cycle literature.
7. The aggregation in this table is identical to the aggregation used by the BEA, except for "residential," which is calculated as the chain-weighted aggregate of "residential structures and equipment" and "consumer durable." See Box 1 for the chain-weighting procedure.
8. For TPI and GDP, y, the nominal shares in the first row are $P_t^{TPI} q_t^{TPI}/(P_t^y q_t^y)$ and the real shares are q_t^{TPI}/q_t^y. Nominal and real shares for investment good i in the other rows are given by $P_t^i q_t^i/(P_t^{TPI} q_t^{TPI})$ and the real shares are q_t^i/q_t^{TPI}.
9. In the notation used above, the black lines are (the natural logarithm of) p_t^i for i corresponding to the 20 types of investment listed in Table 1 over the period for which data are available.
10. Many of the trends evident in Figure 2 are not apparent in the NIPA fixed-weighted constant 1982 dollar and earlier NIPA data. In a very influential book, Gordon (1989) argued that the conventional BEA treatment of investment good quality severely underestimated the degree of quality change in investment goods. His analysis was the first to show that there is a substantial downward trend in the prices of PDE and CD. The BEA now incorporates many of the adjustments for quality change advocated by Gordon (1989).
11. The procedure used to extract the business cycle component of the relative price data involves the application of a linear filter. This, combined with the fact that this filter is applied to the natural logarithm of the relative prices, implies that the business cycle component of each relative price is the business cycle component of the relevant investment deflator minus the business cycle component of the consumption deflator.
12. For a comprehensive review of this literature, see Barro and Sala-i-Martin (1995).
13. The assumption of constant returns to scale is usually based on a replication argument. A fixed quantity of capital and labor applied to produce x amount of some good can always be applied again to produce another x of the good. That is, increasing the quantity of factors of production by some proportion changes the amount produced by the same proportion. This argument seems harder to apply in the case of technology. For example, suppose a group of researchers have discovered a new process for making steel. If another group of researchers make the same discovery, there is no net improvement in knowledge. In this case, there would be decreasing returns. On the other hand, fixed costs or advantages to having many researchers working on similar projects may mean that increasing returns to scale are important in the process of knowledge creation.
14. Greenwood et al. (1997) show how the research and development and human capital classes of models can be used to account for the evidence, if these activities have a disproportionate impact on the production of equipment compared with consumption goods. Two explanations they consider differ fundamentally from their basic story. They both involve a two-sector

interpretation of the evidence, in which equipment and consumption goods are produced in separate sectors (using separate production functions). In one case, the production functions have different factor shares, that is, the different goods require capital and labor in different proportions to produce a unit of the good. The authors conclude that the "prospect for explaining the relative price decline with a two-sector model based on differences in share parameters looks bleak, given the implausibly large differences required in the structure of production across sectors (p. 358)." The other explanation involves an externality in the production of investment goods. Specifically, the productivity of factors in the investment good sector is increasing in the quantity of investment goods along the lines described in Romer (1986). Greenwood et al. (1997) show that this explanation can, in principle, account for the trend evidence. However, this theory relies on an externality which is difficult to identify empirically. Some evidence on increasing returns to scale, which the production externality implies, is discussed below. Generally, there is little empirical support for this view.

15. The shape of the frontier can be justified by standard neoclassical assumptions about how goods are produced, in particular that they are produced using constant returns to scale production functions in labor and capital and that it is costly to transfer labor and/or capital across sectors producing consumption goods and sectors producing investment goods. Note that adjustment costs in the installation of investment goods affect the relative price of installed capacity, *not* the relative price of investment goods.
16. This discussion assumes that the shares of factors in production are identical in producing consumption and investment goods and/or that there are costs of adjusting factors of production across sectors. It is possible for the price of investment goods to be countercyclical in this type of model if the share of labor in production is greater in the consumption sector than in the investment goods sector. As long as factors of production are perfectly mobile across sectors (that is, there are no costs to shifting factors across sectors), an increase in technology lowers the price of investment goods in this case. Factor shares are difficult to measure, so assessing the plausibility of this possibility is difficult. However, the Greenwood et al. (1997) results for long-run trends suggest that the differences in factor shares required to reconcile the empirical evidence on prices with this explanation may be implausible. Also, it is implausible to assume that there are no costs of shifting factors of production across sectors.
17. This frontier does not necessarily reflect true technological possibilities, but takes into account the restrictions on individual decisionmaking, such as individuals not internalizing a production externality, such that the points on the frontier are consistent with optimizing behavior of producers.
18. The methodology is identical to that employed by Eichenbaum and Fisher (1998). This methodology uses four variables, in addition to the investment good quantity and price variables, in a vector autoregression, along with a dummy variable which takes on the value zero at all dates except 1950:Q3, 1965:Q1 and 1980:Q1, in which cases the variable equals unity. These dates correspond to the beginning of three large military buildups. The key identifying assumption is that these buildups were exogenous events. For further discussion, see Edelberg, Eichenbaum, and Fisher (1999). The four variables are the log level of time t real *GDP*, the net three-month Treasury bill rate, the log of the Producer Price Index of crude fuel, and the log level of real defense purchases, g_t. Six lags were used. The plotted responses in Figure 12 correspond to the average response of the indicated variable across the three military buildup episodes, taking into account the endogenous variation in the variable.
19. See Edelberg, Eichenbaum, and Fisher (1999) for a discussion of how this evidence can be explained within the context of a standard neoclassical model.
20. Technically, I estimate a vector autoregression in the deflator for nondurables and services, real GDP, an index of changes in sensitive materials prices, the federal funds rate, plus the investment price and quantity I am interested in. All variables except the federal funds rate are first logged. The impulse response functions in Figure 13 correspond to an orthoganalized innovation in the federal funds rate. The orthoganalization procedure assumes the order of the

vector autoregression is the same as listed in the text and a triangular decomposition. Ordering is not important for the investment responses as long as standard assumptions are made about the variables that precede the federal funds rate in the ordering (see Christiano, Eichenbaum, and Evans, 1999). Finally, the standard errors are computed using the procedure described by Christiano, Eichenbaum, and Evans (1999).

REFERENCES

Abel, A. and Bernanke, B. 1997. *Macroeconomics,* 3rd edition. New York: Addison-Wesley.

Aghion, P. and Howitt, P. 1992. A model of growth through creative destruction. *Econometrica,* Vol. 60, No. 2, March, pp. 323–351.

Barro, R. and Sala-i-Martin, 1995. *Economic Growth.* New York: McGraw-Hill.

Basu, S. and Fernald, J. 1997. Returns to scale in U.S. production: Estimates and implications *Journal of Political Economy,* Vol. 105, No. 2, April, pp. 249–283.

Burnside, C. 1996. Production function regressions, returns to scale, and externalities. *Journal of Monetary Economics,* Vol. 37, No. 2, April, pp. 177–201.

Burnside, C. Eichenbaum, M. and Rebelo, S. 1995. Capacity utilization and returns to scale. In *NBER Macroeconomics Annual 1995,* Vol. 10, Ben Bernanke and Julio Rotemberg (eds.). Cambridge, MA: MIT Press.

Christiano, L. J. 1996. Identification and the liquidity effect: A case study. *Economic Perspectives,* Federal Reserve Bank of Chicago, Vol. 20, No. 3, May/June, pp. 2–13.

Christiano, L. J., Eichenbaum, M. and Evans, C. 1999. Monetary policy shocks: What have we learned and to what end? In *Handbook of Macroeconomics,* John Taylor and Michael Woodford (eds.). Amsterdam: Elsevier Science, forthcoming.

Christiano, L. J. and Fisher, J. D. M. 1998. Stock market and investment good prices: Implications for macroeconomics. Federal Reserve Bank of Chicago, working paper.

Christiano, L. J. and Fitzgerald, T. 1998. The business cycle: It's still a puzzle. *Economic Perspectives,* Federal Reserve Bank of Chicago, Vol. 22, No. 4, Fourth Quarter, pp. 56–83.

Christiano, L. J. and Harrison, S. 1999. Chaos, sunspots, and automatic stabilizers. *Journal of Monetary Economics,* Vol. 44, No. 1, pp. 3–31.

Edelberg, W. Eichenbaum, M. and Fisher, J. D. M. 1999. Understanding the effects of a shock to government purchases. *Review of Economic Dynamics,* Vol. 2, forthcoming.

Eichenbaum, M. and Fisher, J. D. M. 1998. How does an increase in government purchases effect the economy? *Economic Perspectives,* Federal Reserve Bank of Chicago, Vol. 22, No. 3, Third Quarter, pp. 29–43.

Farmer, R. 1993. *The Macroeconomics of Self-Fulfilling Prophecies.* Cambridge: MIT Press.

Fisher, J. D. M. 1997. Relative prices, complementarities, and comovement among components of aggregate expenditures. *Journal of Monetary Economics,* Vol. 39, No. 3, August, pp. 449–474.

Goolsbee, A. 1998. Investment tax incentives, prices, and the supply of capital goods. *Quarterly Journal of Economics,* Vol. 113, No. 1, February, pp. 121–148.

Gordon, R. 1998. *Macroeconomics,* 7th edition. New York: Addison-Wesley.

Gordon, R. 1989. *The Measurement of Durable Goods Prices.* Chicago: University of Chicago Press.

Greenwood, J. and Hercowitz, Z. 1988. Investment, capacity utilization, and the real business cycle. *American Economic Review,* Vol. 78, No. 3, June, pp. 402–417.

Greenwood, J., Hercowitz, Z. and Krusell, P. 1998. The role of investment-specific technological change in the business cycle. University of Rochester, manuscript, January.

Greenwood, J., Hercowitz, Z., and Krusell, P. 1997. Long-run implications of investment-specific technological change. *American Economic Review,* Vol. 87, No. 3, June, pp. 342–362.

Grossman, G. and Helpman, E. 1991. *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press.

Hall, R. and Taylor, J. 1997. *Macroeconomics,* 5th edition. New York: W. W. Norton.

Hansen, G. and Prescott, E. 1993. Did technology shocks cause the 1990–91 recession? *American Economic Review,* Vol. 83, No. 2, May, pp. 280–286.

Harrison, S. G. 1998. Evidence on the empirical plausibility of externalities and indeterminacy in a two sector model. Barnard College, Columbia University, unpublished manuscript.

Hornstein, A. and Praschnik, J. 1997. Intermediate inputs and sectoral comovement in the business cycle. *Journal of Monetary Economics,* Vol. 40, No. 3, December, pp. 573–595.

Huffman, G. and Wynne, M. 1999. The role of intratemporal adjustment costs in a multisector economy. *Journal of Monetary Economics,* forthcoming.

Krusell, P., Ohanian, L., Rios-Rull, V., and Violante, G. 1995. Capital-skill complementarity and inequality. University of Rochester, manuscript.

Landefeld, J. S. and Parker, R. P. 1997. BEA's chain indexes, time series, and measures of long-term economic growth. *Survey of Current Business,* pp. 58–68.

Lucas, R. E. 1988. On the mechanics of economic development. *Journal of Monetary Economics,* Vol. 22, No. 1, July, pp. 3–42.

Mankiw, G. 1997. *Macroeconomics,* 3rd edition. New York: Worth.

Murphy, K. M., Shleifer, A. and Vishny, R. W. 1989. Building blocks of market clearing business cycle models. In *NBER Macroeconomics Annual 1989,* Olivier J. Blanchard and Stanley Fischer (eds.). Cambridge, MA: MIT Press.

Prescott, E. 1986. Theory ahead of business cycle measurement. *Quarterly Review,* Federal Reserve Bank of Minneapolis, Fall, pp. 9–22.

Romer, P. 1990. Endogenous technical change. *Journal of Political Economy,* Vol. 98, No. 5, Part 2, October, pp. S71–S102.

Romer, P. 1986. Increasing returns and long-run growth. *Journal of Political Economy,* Vol. 94, No. 5, October, pp. 1002–1037.

Shea, J. 1993. Do supply curves slope up? *Quarterly Journal of Economics,* Vol. 108, February, pp. 1–32.

Solow, R. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics,* Vol. 39, pp. 312–320.

This Page Intentionally Left Blank

6
Social Norms and Economic Theory

Jon Elster
Professor of Political Science and Philosophy, University of Chicago, Chicago, Illinois

One of the most persistent cleavages in the social sciences is the opposition between two lines of thought conveniently associated with Adam Smith and Emile Durkheim, between *homo economicus* and *homo sociologicus.* Of these, the former is supposed to be guided by instrumental rationality, while the behavior of the latter is dictated by social norms. The former is "pulled" by the prospect of future rewards, whereas the latter is "pushed" from behind by quasi-inertial forces (Gambetta, 1987). The former adapts to changing circumstances, always on the lookout for improvements. The latter is insensitive to circumstances, sticking to the prescribed behavior even if new and apparently better options become available. The former is easily caricatured as a self-contained, asocial atom, and the latter as the mindless plaything of social forces. In this chapter I characterize this contrast more fully, and discuss attempts by economists to reduce norm-oriented action to some type of optimizing behavior.[1]

Rational action is concerned with outcomes. Rationality says: If you want to achieve *Y*, do *X*. By contrast, I define social norms by the feature that they are *not outcome-oriented.* The simplest social norms are of the type: Do *X*, or: Don't do *X*. More complex norms say: If you do *Y*, then do *X*, or: If others do *Y*, then do *X*. More complex norms still might say: Do *X* if it would be good if everyone did *X*. Rationality is essentially conditional and future-oriented. Social norms are either unconditional or, if conditional, are not future-oriented. For norms to be *social,* they must be shared by other people and partly sustained by their approval and disapproval. They are also sustained by the feelings of embarrassment, anxiety, guilt and shame that a person suffers at the prospect of violating them. A person obeying a norm may also be propelled by positive emotions, like anger and indignation. Djilas (1958, p. 107) refers to the feeling of a person enacting the norms of vengeance in Montenegro as "the wildest, sweetest kind of drunkenness." Social norms have a grip on the mind that is due to the strong emotions they can trigger.

This initial statement somewhat exaggerates the mechanical, unreflective character of norm-guided behavior. Social norms offer considerable scope for skill, choice, interpretation and manipulation. For that reason, rational actors often deploy norms to achieve their ends. Yet there are limits to the flexibility of norms, otherwise there would be nothing to manipulate.

Reprinted from: *Journal of Economic Perspectives,* v. 3, n. 4 (Fall 1989) 99–117.

Social norms must be distinguished from a number of other, related phenomena. First, social norms differ from moral norms. Some moral norms, like those derived from utilitarian ethics, are consequentialist. Secondly, social norms differ from legal norms. Legal norms are enforced by specialists who do so out of self-interest: they will lose their job if they don't. By contrast, social norms are enforced by members of the general community, and not always out of self-interest (see below). Thirdly, social norms are more than the convention equilibria described by Robert Sugden. As Sugden explains, the evolution of a convention equilibrium is guided by whether the conventions lead to a substantively better outcome. I argue below, however, that many social norms do not benefit anyone. Fourthly, social norms differ from private norms, the self-imposed rules that people construct to overcome weakness of will (Ainslie 1982, 1984, 1986). Private norms, like social norms, are non-outcome-oriented and sustained by feelings of anxiety and guilt. They are not, however, sustained by the approval and disapproval of others since they are not, or not necessarily, shared with others. Finally, norm-guided behavior must be distinguished from habits and compulsive neuroses. Unlike social norms, habits are private. Unlike private norms, their violation does not generate self-blame or guilt. Unlike neuroses and private norms, habits are not compulsive. Unlike social norms, compulsive neuroses are highly idiosyncratic. Yet what in one culture looks like a compulsive neurosis may, in another society, be an established social norm (Fenichel 1945, p. 586). Compulsive revenge behavior could be an example (Djilas, 1958).

To fix our ideas, let me give some examples of social norms.

Consumption norms regulate manners of dress, manners of table and the like. As shown by Proust's masterful account of life in the Guermantes circle, conformity with such norms can be vitally important to people, in spite of the fact that nothing of substance seems to be at stake. Pierre Bourdieu (1979) has extended the notion of consumption norms to cover cultural behavior: which syntax, vocabulary and pronunciation do you adopt? which movies do you see? which books do you read? which sports do you practice? what kind of furniture do you buy?

Norms against behavior "contrary to nature" include rules against incest, cannibalism, homosexuality and sodomy. The rule against cannibalism allows, however, for exceptions in case of *force majeure* (Edgerton, 1985, p. 51). The point obtains quite generally: Whenever there is a norm, there are often a set of adjunct norms defining legitimate exceptions. Often, these are less explicit than the main norm, and rely heavily on judgment and discretion.

Norms regulating the use of money often become legal, like the law against buying and selling votes. Often, however, they remain informal, like the norm against buying into a bus queue or the norm against asking one's neighbor to mow one's lawn for money. I discuss both of these cases later.

Norms of reciprocity enjoin us to return favors done to us by others (Gouldner, 1960). Gift-giving is often regulated by these norms. There may not be an unconditional norm of giving Christmas presents to a first cousin, but once the cousin begins to give me a gift I am under an obligation to return it.

Norms of retribution enjoin us to return harms done to us by others. Rules regulating revenge are often highly elaborate (Hasluck, 1954; Boehm, 1984; Miller, 1990). Nevertheless, revenge often seems to be contrary to self-interest: "Who sees not that vengeance, from the force alone of passion, may be so eagerly pursued as to make us knowingly neglect every consideration of ease, interest, or safety?" (Hume, 1751, Appendix II).

Work norms. The workplace is a hotbed for norm-guided action. There is a social norm against living off other people and a corresponding normative pressure to earn one's income from work (Elster, 1988). At the workplace one often finds informal norms among the workers that regulate their work effort. Typically, these set lower as well as upper limits on what is perceived as a proper effort: neither a chiseler nor a ratebuster be (Roethlisberger and Dickson, 1939, p. 522). Akerlof (1980) argues that employed workers have a "code of honor" that forbids them to train new workers who are hired to do the same job for lower wages.[2]

Norms of cooperation. There are many outcome-oriented maxims of cooperation. A utilitarian, for instance, would cooperate if and only if his contribution increases the average utility of the members in the group. There are also, however, non-outcome-oriented norms of cooperation. One is what one may call "everyday Kantianism:" cooperate if and only if it would be better for all if all cooperated than if nobody did. Another is a "norm of fairness:" cooperate if and only if most other people cooperate. Among the phenomena based on norms of cooperation one may cite voting (Barry, 1979) and tax compliance (Laurin, 1986).

Norms of distribution regulate what is seen as a fair allocation of income or other goods. In democratic societies, the norm of equality is especially strong. As Tocqueville (1969, p. 505) wrote: "the passion for equality seeps into every corner of the human heart, expands and fills the whole. It is no use telling them that by this blind surrender to an exclusive passion they are compromising their dearest interests; they are deaf." People may be willing to take a loss rather than accept a distribution they find unfair (Kahneman, Knetsch and Thaler, 1986). The solution concept for cooperative bargaining proposed by Kalai and Smorodinsky (1975) embodies a norm of fair distribution (McDonald and Solow, 1981, pp. 905–6).

Drawing on these examples, I shall consider a number of arguments that have been made to the effect that social norms are "nothing but" instruments of individual, collective or genetic optimization. First, however, I want to make two brief remarks.

To accept social norms as a motivational mechanism is not to violate methodological individualism. True, many sociologists who have stressed the importance of social norms have also advocated methodological holism (e.g. Durkheim, 1958), but there is no logical connection between these views. Social norms, as I understand them here, are emotional and behavioral propensities of individuals.

To accept social norms as a motivational mechanism is not to deny the importance of rational choice. One eclectic view is that some actions are rational, others are norm-guided. A more general and more adequate formulation would be that actions typically are influenced both by rationality and by norms. Sometimes, the outcome is a compromise between what the norm prescribes and what rationality dictates. The subjects in the experiment of Kahneman, Knetsch and Thaler (1986) who rejected very unfair distributions, preferring to take nothing rather than to be exploited by others, did accept mildly skewed distributions. At other times, rationality acts as a constraint on social norms. Many people vote out of civic duty, except when the costs become very high. Conversely, social norms can act as a constraint on rationality. Cutthroat competitiveness in the market can go together with strict adherence to norms of honesty (Coleman, 1982).

ARE NORMS RATIONALIZATIONS OF SELF-INTEREST?

Is it true, as argued by early generations of anthropologists and sociologists, that norms are in the saddle and people merely their supports? Or is it true, as argued by more recent generations, that rules and norms are just the raw material for strategic manipulation or, perhaps, for unconscious rationalization?

Sometimes, people will invoke a social norm to rationalize self-interest. Suppose my wife and I are having a dinner party for eight, and that four persons have already been invited. We discuss whether to invite a particular couple for the last two places, and find ourselves in disagreement, for somewhat murky reasons. I like the woman of the couple, and my wife doesn't like it that I like her. But we don't want to state these reasons. (Perhaps there is a social norm against doing so.) Instead we appeal to social norms. I invoke the norm of reciprocity, saying, "Since they had us over for dinner, it is our turn to invite them now." My wife invokes another norm: "Since we have already invited two single men, we must invite two women, to create a balance."

In wage negotiations, sheer bargaining power counts for much. Appeals to accepted social norms can also have some efficacy, however. There is a norm of fair division of the surplus between capital and labor. Employers will appeal to this norm when the firm does badly, workers when it does well. There is a norm of equal pay for equal work. Workers will appeal to this norm when they earn less than workers in similar firms, but not when they earn more. The norm of preservation of status, or wage differences, can also be exploited for bargaining purposes.

Social psychologists have studied norms of distribution to see whether there is any correlation between who subscribes to a norm and who benefits from it. Some findings point to the existence of a "norm of modesty:" high achievers prefer the norm of absolute equality of rewards, whereas low achievers prefer the norm of equity, or reward proportionally to achievement (Mikula, 1972; Kahn, Lamm and Nelson, 1977; Yaari and Bar-Hillel, 1988). More robust, however, are the findings which suggest that people prefer the distributive norms which favor them (Deutsch, 1985, Ch. 11; Messick and Sentis, 1983). This corresponds to a pattern frequently observed in wage discussions. Low-income groups invoke a norm of equality, whereas high-income groups advocate pay according to productivity.

Conditional norms lend themselves easily to manipulation. There is, for instance, a general norm that whoever first proposes that something be done has a special responsibility for making sure that it is carried out. This can prevent the proposal from ever being made, even if all would benefit from it. A couple may share the desire to have a child and yet neither may want to be the first to lance the idea, fearing that he or she would then get special child-caring responsibility.[3] The member of a seminar who suggests a possible topic for discussion is often saddled with the task of introducing it. The person in a courtship who first proposes a date is at a disadvantage (Waller, 1937). The fine art of inducing others to make the first move, and of resisting such inducements, provides instances of instrumentally rational exploitation of a social norm.

Some have said that this is all there is to norms: they are tools of manipulation, used to dress up self-interest in more acceptable garb. But this cannot be true. Some norms, like the norm of vengeance, obviously override self-interest. In fact, the cynical view of norms is self-defeating. "Unless rules were considered important and were taken seriously and followed, it would make no sense to manipulate them for personal benefit. If many people did not believe that rules were legitimate and compelling, how could anyone use these rules for personal advantage?" (Edgerton, 1985, p. 3). Or again, "if the justice arguments are such transparent frauds, why are they advanced in the first place and why are they given serious attention?" (Zajac, 1985, p. 120). If some people successfully exploit norms for self-interested purposes, it can only be because others are willing to let norms take precedence over self-interest. Moreover, even those who appeal to the norm usually believe in it, or else the appeal might not have much power (Veyne, 1976).

The would-be manipulator of norms is also constrained by the need—in fact, the social norm—to be consistent. Even if the norm has no grip on his mind, he must act as if it had. Having invoked the norm of reciprocity on one occasion, I cannot just dismiss it when my wife appeals to it another time. An employer may successfully appeal to the workers and get them to share the burdens in a bad year. The cost he pays is that in a good year he may also have to share the benefits. By making the earlier appeal, he committed himself to the norm of a fair division of the surplus (Mitchell, 1986, p. 69). The Swedish metal workers in the 1930s successfully invoked a norm of equality to bring about parity of wages with workers in the construction industry. Later, when they found themselves in a stronger bargaining position, their previous appeal to equality forced them to pull their punches (Swenson, 1989, p. 60). Finally, the manipulator is constrained by the fact that the repertoire of norms on which he can draw is, after all, limited. Even if unconstrained by earlier appeals to norms, there may not be any norm available that coincides neatly with his self-interest.

When I say that manipulation of social norms presupposes that they have some kind of grip on the mind since otherwise there would be nothing to manipulate, I am not suggesting that society is made up of two sorts of people: those who believe in the norms and those who manipulate the believers. Rather, I believe that most norms are shared by most people—manipulators as well as manipulated. Rather than manipulation in a direct sense, we are dealing here with an amalgam of belief, deception and self-deception. At any given time we believe in many different norms, which may have contradictory implications for the situation at hand. A norm that happens to coincide with narrowly defined self-interest easily acquires special salience. If there is no norm handy to rationalize self-interest, or if I have invoked a different norm in the recent past, or if there is another norm which overrides it, I may have to act against my self-interest. My self-image as someone who is bound by the norms of society does not allow me to pick and choose indiscriminately from the large menu of norms to justify my actions, since I have to justify them to myself no less than to others. At the very least, norms are soft constraints on action. The existence of norms of revenge shows that sometimes they are much more than that.

ARE NORMS FOLLOWED OUT OF SELF-INTEREST?

When people obey norms, they often have a particular outcome in mind: they want to avoid the disapproval—ranging from raised eyebrows to social ostracism—of other people. Suppose I face the choice between taking revenge for the murder of my cousin and not doing anything. The cost of revenge is that I might in turn be the target of a counter-vengeance. At worst, the cost of not doing anything is that my family and friends desert me, leaving me out on my own, defenselessly exposed to predators. At best, I will lose their esteem and my ability to act as an autonomous agent among them. A cost-benefit analysis is likely to tell me that revenge (or exile) is the rational choice. More generally, norm-guided behavior is supported by the threat of social sanctions that make it rational to obey the norms. Akerlof (1976) argues, along these lines, that in India it is rational to adhere to the caste system, even assuming that "tastes" are neutral.

In response to this argument, we can first observe that norms do not need external sanctions to be effective. When norms are internalized, they are followed even when violation would be unobserved and not exposed to sanctions. Shame or anticipation of it is a sufficient internal sanction. I don't pick my nose when I can be observed by people on a train passing by, even if I am confident that they are all perfect strangers whom I shall never see again and who have no power to impose sanctions on me. I don't throw litter in the park, even when there is nobody around to observe me. If punishment was merely the price tag attached to crime, nobody would feel shame when caught. People have an internal gyroscope that keeps them adhering steadily to norms, independently of the current reactions of others.

A second answer to the claim that people obey norms because of the sanctions attached to violations of norms emerges if we ask why people would sanction others for violating norms. What's in it for them? One reply could be that if they do not express their disapproval of the violation, they will themselves be the target of disapproval by third parties. When there is a norm to do *X*, there is usually a "meta-norm" (Axelrod, 1986) to sanction people who fail to do *X*, perhaps even a norm to sanction people who fail to sanction people who fail to do *X*. As long as the cost of expressing disapproval is less than the cost of receiving disapproval for not expressing it, it is in one's rational self-interest to express it. Now, expressing disapproval is always costly, whatever the target behavior. At the very least it requires energy and attention that might have been used for other purposes. One may alienate or provoke the target individual, at some cost or risk to oneself. Opportunities for mutually beneficial transactions are lost when one is forbidden to deal with an ostracized person. By contrast, when one moves upwards in the chain of actions, beginning with the original violation, the cost of receiving disapproval falls rapidly to zero. People do not usually frown upon others when they fail to sanction people who fail to sanction people who fail to sanction people who fail to sanction a norm violation.[4] Consequently, some sanctions must be performed for other motives than the fear of being sanctioned.

DO NORMS EXIST TO PROMOTE SELF-INTEREST?

I believe that for many economists an instinctive reaction to the claim that people are motivated by irrational norms would be that on closer inspection the norms will turn out to be disguised, ultrasubtle expressions or vehicles of self-interest. Gary Becker (1976, pp. 5, 14) argues, for example, that the "combined assumptions of maximizing behavior, market equilibrium and stable preferences, used relentlessly and unflinchingly . . . provides a valuable unified framework for understanding *all* human behavior." This view suggests that norms exist because they promote self-interest, over and above the avoidance of sanctions.

Some social norms can be individually useful, such as the norm against drinking or overeating. Moreover, people who have imposed private norms on their own behavior may join each other for mutual sanctioning, each in effect asking the others to punish him if he deviates, while being prepared to punish them if they do not punish him. Alcoholics Anonymous provide the best-known example (Kurtz, 1979, p. 215): "Each recovering alcoholic member of Alcoholics Anonymous is kept constantly aware, at every meeting, that he has *both* something to give *and* something to receive from his fellow alcoholics." Most norms, however, are not social contracts of this kind.

It might also be argued that social norms are individually useful in that they help people to economize on decision costs. A simple mechanical decision rule may, on the whole and in the long run, have better consequences for the individual than a fine-tuned search for the optimal decision. This argument, however, confuses social norms and habits. Habits certainly are useful in the respect just mentioned, but they are not enforced by other people, nor does their violation give rise to feelings of guilt or anxiety.

A further argument for the view that it is individually rational to follow norms is that they lend credibility to threats that otherwise would not be believable. They help, as it were, to solve the problem of time inconsistency. Vendettas are not guided by the prospect of future gain but triggered by an earlier offense. Although the propensity to take revenge is not guided by consequences, it can have good consequences. If other people believe that I invariably take revenge for an offense, even at great risk to myself, they will take care not to offend me. If they believe that I will react to offense only when it is in my interest to react, they need not be as careful. From the rational point of view, a threat is not credible unless it will be in the interest of the threatener to carry it out when the time comes. The threat to kill oneself, for instance, is not rationally credible. Threats backed by a code of honor are very effective, since they will be executed even if it is in the interest of the threatener not to do so.

This observation, while true, does not amount to an explanation of the norm of vengeance. When a person guided by a code of honor has a quarrel with one who is exclusively motivated by rational considerations, the first will often have his way. But in a quarrel between two persons guided by the code, both may do worse than if they had agreed to let the legal system resolve their conflict. (Mafiosi seem to do better for themselves in the United States than in Sicily.) Since we are talking about codes of honor that are shared social norms, the latter case is the typical one. The rationality of following the code then reduces to the desire to avoid sanctions, discussed above.

In any case, one cannot rationally decide to behave irrationally, even when one knows it would be in one's interest to do so. To paraphrase Max Weber, a social norm is not like a taxi from which one can disembark at will. Followers of a social norm abide by it even when it is not in their interest to do so. In a given situation, following the norm may be useful, but that is not to say that it is always useful to follow it. Moreover, there is no presumption that its occasional usefulness can explain why it exists.

The distinction between the usefulness of norms and their rationality can also be brought out by considering Akerlof's explanation of why workers refuse to train new workers who are hired at lower wages. In an analysis of wage rigidity, Assar Lindbeck and Dennis Snower (1986) argue that the explanation is to be sought in the self-interest of the employed workers. By keeping potential entrants out, they can capture a greater deal of the benefits of monopoly power. The weapons at their disposal for keeping the unemployed at bay include the following:

> First, by being unfriendly and uncooperative to the entrants, the insiders are able to make the entrants' work more unpleasant than it otherwise would have been and thereby raise the wage at which the latter are willing to work. In practice, outsiders are commonly wary of underbidding the insiders. This behavior pattern is often given an *ad hoc* sociological explanation: "social mores" keep outsiders from "stealing" the jobs from their employed comrades. Our line of argument, however, suggests that these mores may be traced to the entrants' anticipation of hostile insider reaction and that this reaction may follow from optimisation behavior of insiders. Second, insiders are usually responsible for training the entrants and thereby influence their productivity. Thus insiders may be able to raise their wage demands by threatening to conduct the firm's training programs inefficiently or even to disrupt them. . . . In sum, to raise his wage, an insider may find it worthwhile to threaten to become a thoroughly disagreeable creature.

The insider may, to be sure, make this threat, *but is it credible?* If an outsider *is* hired, would it then still be in the insider's interest to be unfriendly and uncooperative? Since Lindbeck and Snower (1988, p. 171) believe that "harassment activities are disagreeable to the harassers," they ought also to assume that outsiders will recognize this fact and, in consequence, will not be deterred by fear of harassment. I believe Akerlof is right in arguing that it takes something like a social norm to sustain this behavior. While useful, the ostracism is not rational.

DO NORMS EXIST TO PROMOTE COMMON INTERESTS?

Among economists, those who do not subscribe to the individual rationality of norms will mostly argue for their collective rationality, claiming that social norms have collectively good consequences for those who live by them and that, moreover, these consequences explain why the norms exist. Most writers on the topic probably use the term "socially useful" to mean that a society with the norm is at least as good for almost everybody and substantially better for many than a society in which the norm is lacking, perhaps with an implied clause that no other norm could bring further Pareto-improvements.

Among those who have argued for the collective optimality of norms, Kenneth Arrow (1971, p. 22) is perhaps the most articulate and explicit:

> It is a mistake to limit collective action to state action. . . . I want to [call] attention to a less visible form of social action: norms of social behavior, including ethical and moral codes. I suggest as one possible interpretation that they are reactions of society to compensate for market failure. It is useful for individuals to have some trust in each other's word. In the absence of trust, it would become very costly to arrange for alternative sanctions and guarantees, and many opportunities for mutually beneficial cooperation would have to be foregone. Banfield has argued that the lack of trust is indeed one of the causes of economic underdevelopment.
>
> It is difficult to conceive of buying trust in any direct way (though it can happen indirectly, e.g. a trusted employee will be paid more as being more valuable); indeed, there seems to be some inconsistency in the very concept. Non-market action might take the form of a mutual agreement. But the arrangement of these agreements and especially their continued extension to new individuals entering the social fabric can be costly. As an alternative, society may proceed by internalization of these norms to the achievement of the desired agreement on an unconscious level.
>
> There is a whole set of customs and norms which might be similarly interpreted as agreements to improve the efficiency of the economic system (in the broad sense of satisfaction of individual values) by providing commodities to which the price system is inapplicable.[5]

I shall adduce three arguments against this view. First, not all norms are Pareto-improvements. Some norms make everybody *worse* off, or, at the very least, they do not make almost everybody better off. Secondly, some norms that would make everybody better off are not in fact observed. Thirdly, even if a norm does make everybody better off, this does not explain why it exists, unless we are also shown the feedback mechanism that specifies how the good consequences of the norm contribute to its maintenance.

To support the first argument I shall consider a number of norms that do not appear to be socially useful in the sense defined. The social sciences being what they are, no conclusive proof can be given, but I hope the overall impact of the counterexamples will be persuasive.

Consumption norms do not appear to have any useful consequences. If anything, norms of etiquette seem to make everybody worse off, by requiring wasteful investments in pointless behaviors. Let me, nevertheless, mention three possible arguments for the social usefulness of these norms, together with corresponding objections.

First, there is the argument that norms of etiquette serve the useful function of confirming one's identity or membership in a social group. Since the notion of social identity is elusive, the argument is hard to evaluate, but one weakness is that it does not explain why these rules are as complicated as they often are. To signal or confirm one's membership in a group one sign should be sufficient, like wearing a badge or a tie. Instead, there is often vast redundancy. The manner of speaking of an Oxford-educated person differs from standard English in many more ways than what is required to single him out as an Oxford graduate.

Secondly, there is the argument that the complexity of the rules serves an additional function, that of keeping outsiders out and upstarts down (Bourdieu, 1979). It is easy to imitate one particular behavior, but hard to learn a thousand subtly different rules. But that argument flounders on the fact that working-class life is no less norm-regulated than that of the upper classes. Whereas many middle-class persons would like to pass themselves off as members of the upper class, few try to pass themselves off as workers.

Thirdly, one might combine the first and the second position, and argue that norms simultaneously serve functions of inclusion and exclusion. Evans-Pritchard's (1940, p. 120) classical argument about the Nuer can help us here. "A man of one tribe sees the people of another tribe as an undifferentiated group to whom he has an undifferentiated pattern of behavior, while he sees himself as a member of a segment of his own group." Fine-tuned distinction and gamesmanship within a group is consistent with "negative solidarity" towards outsiders. This view is more plausible, but it does not really point to social benefits of norm following. It is not clear why the working-class as a whole would benefit from the fact that it contains an infinite variety of local subcultures, all of them recognizably working-class and yet subtly different from each other in ways that only insiders can understand. Nor is it clear that the local varieties provide collective benefits to members of the subculture. One might say, perhaps, that norms are useful in limiting the number of potential interaction partners to a small and manageable subset, thus making for greater focus and consistency in social life. A community of norms would then be a bit like a convention equilibrium, since it is important that one's partners limit *their* partners by the same device. This explanation, however, fails to account for the emotional tonality of norms and for their capacity to induce self-destructive behavior.

Consider, as a second example, the social norms against behavior "contrary to nature." Some of these norms like those against cannibalism and incest, are good candidates for collectively beneficial norms. Everybody benefits from a norm that forces people to look elsewhere than to other people for food.[6] Norms against incest may well be optimal from a number of perspectives: individual, collective or genetic. Norms against sodomy, by contrast, involve only harmful restrictions of freedom and no benefits. They make everybody worse off. Norms against homosexuality might also, under conditions of overpopulation, make everybody worse off.

Many social norms against various uses of money do not appear to be collectively rational either. Consider the norm against walking up to a person in a bus queue and asking to buy his place. Nobody would be harmed by this action. Other people in the queue would not lose their place. The person asked to sell his place is free to refuse. If the forbidden practice were allowed, some would certainly gain: the norm does not create a

Pareto-improvement. Yet I cannot assert that it makes everybody worse off, since some individuals could lose from its abolition. That question can only be answered in a general-equilibrium model which, to my knowledge, does not exist.

The norm that prevents us from accepting or making offers to mow other people's lawn for money seems more promising. Consider a suburban community where all houses have small lawns of the same size.[7] Suppose a houseowner is willing to pay his neighbor's son ten dollars to mow his lawn, but not more. He would rather spend half an hour mowing the lawn himself than pay eleven dollars to have someone else do it. Imagine now that the same person is offered twenty dollars to mow the lawn of another neighbor. It is easy to imagine that he would refuse, probably with some indignation. But why is mowing one lawn worth $10 or less, while mowing an identical lawn is worth $20 or more?

Thaler (1980) has suggested, as one possible explanation, that people evaluate losses and gains foregone differently. (Credit card companies exploit this difference when they insist that stores advertise cash discounts rather than credit card surcharges.) The houseowner is more affected by the out-of-pocket expenses that he would incur by paying someone to mow his lawn, than by the loss of a windfall income. But this cannot be the full story, because it does not explain why the houseowner should be indignant at the proposal. Part of the explanation must be that he doesn't think of himself as the kind of person who mows other people's lawns for money. It *isn't done*, to use a revealing phrase that often accompanies social norms.

One may argue that the norm serves an ulterior purpose. Social relations among neighbors would be disturbed if wealth differences were too blatantly displayed, and if some treated others as salaried employees. An unintended consequence of many monetary deals among neighbors could be the loss of the spontaneous self-help behavior that is a main benefit from living in a community. By preventing deals, the norm preserves the community.

The norm could also have a more disreputable aspect, however. The norm against flaunting one's wealth may just be a special case of a higher-order norm: *Don't stick your neck out.* "Don't think you are better than us, and above all don't behave in ways that make us think that you think you are better than us" (Sandemose, 1936). This norm, which prevails in many small communities, can have very bad consequences. It can discourage the gifted from using their talents, and may lead to their being branded as witches if nevertheless they go ahead and use them (Thomas, 1973, p. 643–44). By preserving the community, the norm stifles progress.

It is plausible that norms of reciprocity do, on the whole, have good consequences. Even in this case, however, there are counterexamples, since these norms can become the object of strategic manipulation. An extreme example of such ambiguous altruism is found in Colin Turnbull's description of gift and sacrifice in this society among the miserable Ik of Uganda:

> These are not expressions of the foolish belief that altruism is both possible and desirable: they are weapons, sharp and aggressive, which can be put to divers uses. But the purpose for which the gift is designed can be thwarted by the non-acceptance of it, and much Icien ingenuity goes into thwarting the would-be thwarter. The object, of course, is to build up a whole series of obligations so that in times of crisis you have a number of debts you can recall, and with luck one of them may be repaid. To this end, in the circumstances of Ik life, considerable sacrifice would be justified, to the very limits of the minimal survival level. But a sacrifice that can be rejected is useless, and so you have the odd phenomenon of these otherwise singularly self-interested people going out

> of their way to "help" each other. In point of fact they are helping themselves and their help may very well be resented in the extreme, but it is done in such a way that it cannot be refused, for it has already been given. Someone, quite unasked, may hoe another's field in his absence, or rebuild his stockade, or join in the building of a house that could easily be done by the man and his wife alone. At one time I have seen so many men thatching a roof that the whole roof was in serious danger of collapsing, and the protests of the owner were of no avail. The work done was a debt incurred. It was another good reason for being wary of one's neighbors. Lokeléa always made himself unpopular by accepting such help and by paying for it on the spot with food (which the cunning old fox knew they could not resist), which immediately negated the debt.[8]

Similarly, I may try to benefit from the conditional norm that if I give something to a friend for Christmas, he has an obligation to reciprocate. Suppose the friend is wealthy and that there is a norm that wealthier people should give more in absolute terms (although allowed to give less in relative terms). I can then exploit the situation to my advantage by making the initial gift.

Norms of retribution are often said to serve the social function of resolving conflicts and reducing the level of violence below what it would otherwise have been. There will be fewer quarrels in societies regulated by codes of honor, since everybody knows that they can have disastrous consequences (Boehm, 1984, p. 88). But it is not clear that this is a good thing. One could probably get rid of almost all criminal behavior if all crimes carried the death penalty, but the costs of creating this terror regime would be prohibitive. Also, it is not clear that there is less violence in a vendetta-ridden society than in an unregulated state of nature. In the state of nature, people are supposed to be rational. Hence there would be less violence because people would not harm others just to get even. Also, codes of honor generate quarrels, because honor is attained by brinkmanship and demonstrated willingness to run the risk of initiating a feud (Boehm, 1984, p. 146). On the other hand, the state of nature could be more violent, since people need not fear that others might retaliate just to get even. The net effect is anybody's guess, since the state of nature is not really a well-defined notion.

Consider next Akerlof's analysis of the norm against two-tiered wage systems. This norm does not seem to benefit the employed workers, while harming both employers and the unemployed who have a common interest in such systems. If the employed workers have good reasons to think that the new workers would drive their wages down, the code of honor makes good collective sense, at least with respect to the short-run interests of the local group of workers. Society as a whole might, however, suffer because of the unemployment generated by the practice. In that case, codes of honor would embody solutions to local collective action problems while also creating a higher-order problem.

Somewhat similar arguments apply to the norm against rate-busting. It has been argued that this norm is due to sheer conformism (Jones, 1984) or to envy (Schoeck, 1987, pp. 31, 310). The obvious alternative explanation is that the norm is a collectively optimal response to the constant pressure of management to change piece-rates. Workers often express the view that any increase in effort will induce management to reduce rates. It remains to be shown, however, that this argument is more than rationalization of envy. In the words of one notorious rate-buster: "There are three classes of men: (1) Those who can and will; (2) those who can't and are envious; (3) those who can and won't—they're nuts!" (Dalton, 1948, p. 74). The third category, presumably, are moved by solidarity and norms of justice.

The question cannot be treated separately from the behavior of management. On the one hand, management has a clear incentive to make it clear that they will never cut rates as a result of increased efforts. "Changes in piece rates at the Western Electric Company . . . are not based upon the earnings of the worker. The company's policy is that piece rates will not be changed unless there is a change in the manufacturing process" (Roethlisberger and Dickson, 1939, p. 534).

On the other hand, how can management make this promise credible? They cannot commit themselves to never introducing new methods of production, nor easily prove that a new method is not just a subterfuge for changing rates. A knowledgeable engineer wrote, "I was visiting the Western Electric Company, which had a reputation of never cutting a piece rate. It never did; if some manufacturing process was found to pay more than seemed right for the class of labor employed on it—if, in other words, the rate-setters had misjudged—that particular part was referred to the engineers for redesign, and then a new rate was set for the new part" (Mills, 1946, p. 9, cited after Roy, 1952). Knowing that management has the capability of taking actions of this kind, workers have good reasons to be skeptical.

Three conclusions emerge. First, both management and workers would benefit if a way was found to distinguish "good" from "bad" changes in the piece rates. Second, the worker collective as a whole may well benefit from the norm against rate-busting, given that management cannot credibly commit itself to maintain rates. Third, however, the norm may work against the interest of society as a whole, including the working-class as a whole, if the loss of productivity caused by the norm is sufficiently serious.[9] Even granting that the norm represents the successful solution of a collective action problem within the enterprise, it might create a new problem among enterprises.

At the very least, I believe these examples demonstrate that the social usefulness of social norms cannot be taken for granted. In fact, I think I have shown more than that. Even though each of my claims about non-optimality could be contested and the facts be represented and explained in different ways, I believe that the cumulative impact of the claims is very difficult to refute.

A second strategy for attacking the claim that social norms spring from collective rationality is to imagine some socially useful norms that do not, in fact, exist. If public transportation was widely chosen over private driving, the roads would be less congested and everyone would spend so much less time commuting that the loss of comfort would be offset. Yet there is no social norm to use public transportation in crowded cities. In many developing countries private insurance motives create an incentive to have large families, although the aggregate effect is overpopulation and pressure on resources. Yet there is no social norm against having many children. Japan has apparently imposed the norm "Buy Japanese," but other countries have been less successful. The small Italian village described by Edward Banfield (1958) would certainly have benefited from a social norm against corruption. Instead it had what appears to have been a norm against public-spirited behavior. Nobody would frequent a person stupid enough not to violate the law when he would get away with it. Criminals could benefit from a minimum of solidarity among themselves. A book about the Brooklyn wiseguys suggests, however, that as soon as you're in trouble, you're forgotten: there is no honesty among thieves (Pileggi, 1986). The reader is encouraged to think of other examples.

A third strategy is to criticize the explanatory impact of the collective benefits of social norms. In the absence of a mechanism linking the benefits to the emergence or perpetuation of the norm we cannot know if they obtain by accident. Social scientists should be suspicious of theories of society that deny the possibility of accidental benefits. Moreover, and perhaps more importantly, the beneficial or optimal nature of the norm is often controversial. It is only a slight exaggeration to say that any economist worth his salt could tell a story—produce a model, that is, resting on various simplifying assumptions—which proves the individual or collective benefits derived from the norm. The very ease with which such "just-so stories" can be told suggests that we should be skeptical about them. We would be much more confident about the benefits if a mechanism could be demonstrated.

There are not many plausible candidates for a feedback mechanism. Individual reinforcement could not work here, since the benefits are collective rather than individual. Chance variation and social selection might seem a better alternative.[10] On this account, social norms arise by accident. Societies which happen to have useful norms thrive, flourish and expand; those which do not disappear or imitate the norms of their more successful competitors. Whether the successful societies proceed by military conquest or economic competition, the end result is the same. The argument is popular, but weak. The norms of the strong are not as a rule taken over by the weak, nor do the weak always disappear in competition with the strong. Greece was conquered by Rome, but Rome assimilated more Greek norms than the other way around. When China was conquered by the barbarians, the latter ended up assimilating and defending the culture they had conquered. Today, few developing countries are taking over the norms and work habits that were a precondition for Western economic growth, nor is there any sign of these countries going out of existence.

These arguments do not add up to a strong claim that the social usefulness of norms is irrelevant for their explanation. I find it as hard as the next man to believe that the existence of norms of reciprocity and cooperation has *nothing* to do with the fact that without them civilization as we know it would not exist. Yet it is at least a useful intellectual exercise to take the more austere view, and to entertain the idea that civilization owes its existence to a fortunate coincidence. On this view, social norms spring from psychological propensities and dispositions that, taken separately, cannot be presumed to be useful, yet happen to interact in such a way that useful effects are produced.

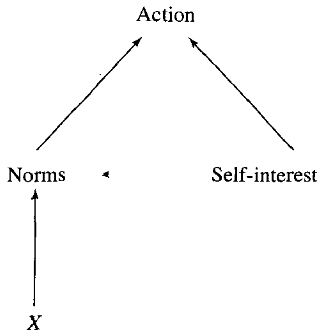
DO NORMS EXIST TO PROMOTE GENETIC FITNESS?

The final argument against the autonomy of norms is that they owe their existence to their contribution to genetic fitness. I do not know of explicit statements of this view. Several writers have, however, taken this position on the closely related issue of the emotions of guilt and shame that sustain norm-guided behavior (Trivers, 1971; Hirschleifer, 1987; Frank, 1988). Chagnon (1988) argues that revenge can be explained as fitness-maximizing behavior, but he does not explicitly consider norms of revenge. I know too little about evolutionary biology to evaluate these claims. I would like, nevertheless, to record my skepticism and make a few general remarks, largely inspired by Kitcher (1985).

Evolutionary explanations do not take the narrow form "Feature *X* exists because it maximizes the genetic fitness of the organism." Rather, their general form is "*X* exists because it is part of a package solution that at some time maximized the genetic fitness of the organism." The latter form allows for two facts that the former excludes. First, there is the omnipresent phenomenon of *pleiotropy.* A tendency to conform to a social norm might detract from genetic fitness and yet be retained by natural selection if it is the by-product of a gene whose main product is highly beneficial. Secondly, the general form allows for time lags. A social norm may be maladaptive today and yet have been adaptive at the stage in history when the human genome evolved and, for practical purposes, was fixed.

When I said that norms might owe their existence to "psychological propensities and dispositions," a natural reply would be to say that these in turn must be explicable in terms of genetic fitness. Let me concede the point, provided that the explanation is allowed to take this general form. Advocates of evolutionary explanations, however, usually have the narrower form in mind. I am not saying that in doing so they are always wrong, only that they cannot take it for granted that an explanation of the narrow form always exists. What is true, is that a plausible story of the narrow form can almost always be told. Again, however, the very ease with which just-so stories are forthcoming should make us wary of them.

Let me summarize the discussion in a diagram:



I believe that both norms and self-interest enter into the proximate explanations of action. To some extent, the selection of the norm to which one subscribes can also be explained by self-interest. Even if the belief in the norm is sincere, the choice of one norm among the many that could be relevant may be an unconscious act dictated by self-interest. Or one might follow the norm out of fear of the sanctions that would be triggered by violation. But I do not believe that self-interest provides the full explanation for adherence to norms. There must be some further explanation, *X,* of why norms exist. I have discussed various candidates for *X,* and found them wanting. I have no positive account of my own to offer. In particular, I have no suggestion as to how norms emerge and disappear. I suggest, however, that a good research strategy might be to investigate the role of emotions in maintaining social norms. Also, the often-ignored phenomena of envy and honor might repay further study. Finally, the psychological theory of conformism should be brought to bear on the subject.

NOTES

1. A fuller account of norms, with applications to collective action and bargaining problems, is found in Elster (1989).
2. This was written before the introduction of two-tiered wage systems in several American airlines.
3. I am indebted to Ottar Brox for this example.
4. The argument in Akerlof (1976, p. 610) seems to rest on the assumption that sanctions can go on forever, without losing any of their force. Anyone who violates any rule of caste, including anyone who fails to enforce the rules, automatically becomes an outcaste. Abreu (1988) offers a formal analysis built on a similar assumption. I know too little about the caste system to assess the validity of the assumption in this case, but I am confident that it is false in the cases

about which I have some knowledge. Sanctions tend to run out of steam at two or three removes from the original violation.
5. See also Ullmann-Margalit (1977), p. 60.
6. Note that the norm cannot be justified by individual "Tit for Tat" rationality: if I eat someone I have no reason to fear that he may eat me on a later occasion.
7. I am indebted to Amos Tversky for suggesting this to me as an example of social norms.
8. Turnbull (1972), p. 146. These strategies are universally employed. As I was completing this paper, I came across a passage in a crime novel (Engel, 1986, p. 155) making the same point: "I decided to make a fast getaway. I had done Pete a favour and it didn't pay to let him thank me for doing it. It was more negotiable the other way. I heard him calling after me but I kept going."
9. As participant-observer in a machine shop Roy (1952) found substantial losses due to deliberately suboptimal efforts.
10. Faia (1986) has a good discussion of the (severely limited) range of cases in which social selection arguments make good sense.

REFERENCES

Abreu, D. On the Theory of Informally Repeated Games with Discounting. *Econometrica,* 1988, *56,* 383–396.

Ainslie, G. A Behavioral Economic Approach to the Defense Mechanisms: Freud's Energy Theory Revisited. *Social Science Information,* 1982, *21,* 735–79.

Ainslie, G. Behavioral Economics II: Motivated Involuntary Behavior. *Social Science Information,* 1984, *23,* 247–74.

Ainslie, G. Beyond Microeconomics. In Elster, J., ed., *The Multiple Self.* Cambridge: Cambridge University Press, 1986, pp. 133–76.

Akerlof, G. The Economics of Caste and of the Rat Race and Other Woeful Tales. *Quarterly Journal of Economics,* 1976, *90,* 599–617.

Akerlof, G. A Theory of Social Custom, of Which Unemployment May be One Consequence. *Quarterly Journal of Economics,* 1980, *94,* 749–75.

Arrow, K. Political and Economic Evaluation of Social Effects and Externalities. In Intriligator, M., ed., *Frontiers of Quantitative Economics.* Amsterdam: North-Holland, 1971, pp. 3–25.

Axelrod, R. An Evolutionary Approach to Norms. *American Political Science Review,* 1986, *80,* 1095–1111.

Banfield, E. G. *The Moral Basis of a Backward Society.* New York: The Free Press, 1958.

Barry, B. *Sociologists, Economists and Democracy,* 2nd Edition. Chicago: University of Chicago Press, 1979.

Becker, G. *The Economic Approach to Human Behavior.* Chicago: University of Chicago Press, 1976.

Boehm, C. *Blood Revenge: The Anthropology of Feuding in Montenegro and Other Tribal Societies.* Lawrence, KS: University of Kansas Press, 1984.

Bourdieu, P. *La Distinction.* Paris: Editions de Minuit, 1970.

Chagnon, N. Life Histories, Blood Revenge, and Warfare in a Tribal Population. *Science,* 1988, *239,* 985–92.

Coleman, J. S. Systems of Trust. *Angewandte Sozialforschung,* 1982, *10,* 277–300.

Dalton, M. The Industrial "Rate-Buster:" A Characterization. *Applied Anthropology,* Winter 1948, 5–18.

Deutsch, M. *Distributive Justice.* New Haven: Yale University Press, 1985.

Djilas, M. *Land Without Justice.* London: Methuen, 1958.

Durkheim, E. *The Rules of Sociological Method.* Glencoe, Ill.: The Free Press, 1958.

Edgerton, R. *Rules, Exceptions and the Social Order.* Berkeley: University of California Press, 1985.

Elster, J. Is There (Or Should There Be) A Right to Work? In Guttman, A., ed., *Democracy and the Welfare State.* Princeton: Princeton University Press, 1988, pp. 53–78.

Elster, J. *The Cement of Society.* Cambridge: Cambridge University Press, 1989.

Engel, H. *A City Called July.* New York: Penguin Books, 1986.

Evans-Pritchard, E. *The Nuer.* Oxford: Oxford University Press 1940.

Faia, M. A. *Dynamic Functionalism.* Cambridge: Cambridge University Press, 1986.

Fenichel, O. *The Psychoanalytic Theory of Neurosis.* New York: Norton, 1945.

Frank, R. K. *Passions within Reason.* New York: Norton, 1988.

Gambetta, D. *Did They Jump or Were They Pushed?* Cambridge: Cambridge University Press, 1987.

Gouldner, A. The Norm of Reciprocity. *American Sociological Review,* 1960, *25,* 161–78.

Hasluck, M. *The Unwritten Law in Albania.* Cambridge: Cambridge University Press, 1954.

Hirschleifer, J. On the Emotions as Guarantors of Threats and Promises. In Dupre, J., ed., *The Latest on the Best.* Cambridge: MIT Press, 1987, pp. 307–26.

Hume, D. *An Enquiry Concerning the Principles of Morals,* New York: Oxford University Press, 1999.

Jones, S. F. *The Economics of Conformism.* Oxford: Blackwell, 1984.

Kahn, A., Lamm, H. and Nelson, R. Preferences for an Equal or Equitable Allocator. *Journal of Personality and Social Psychology,* 1977, *35,* 837–44.

Kahneman, D., Knetsch, J. and Thaler, R. Fairness and the Assumptions of Economics. *Journal of Business,* 1986, *59,* 5285–5300.

Kitcher, P. *Vaulting Ambition.* Cambridge: MIT Press, 1985.

Kurtz, E. *Not-God: A History of Alcoholics Anonymous.* Center City, Minnesota: Hazelden Educational Services, 1979.

Laurin, U. *Pa Heder och Samvete.* Stockholm: Norstedts, 1986.

Lindbeck, A., and Snower, D. J. Wage Rigidity, Union Activity and Unemployment. In Beckerman, W., ed., *Wage Rigidity and Unemployment.* London: Duckworth, 1986, pp. 97–126.

Lindbeck, A., and Snower, D. J. Cooperation Harassment and Involuntary Unemployment. *American Economic Review,* 1988, *78,* 167–88.

McDonald, I. M., and Solow, R. Wage Bargaining and Employment. *American Economic Review,* 1981, *71,* 896–908.

Messick, D. M., and Sentis, K. Fairness, Preference and Fairness Biases. In Messick, D. M., and K. Cook, eds., *Equity Theory.* New York: Praeger, 1983, pp. 61–94.

Mikula, G. Gewinnaufteilung in Dyaden bei variiertem Leistungsverhaltnis. *Zeitschrift für Sozialpsychologie,* 1972, *3,* 126–33.

Miller, W. *Bloodtaking and Peacemaking: Society and the Disputing Process in Medieval Iceland.* Chicago: University of Chicago Press, 1990.

Mills, J. *The Engineer in Society.* New York: Nostrand, 1946.

Mitchell, D. J. Explanations of Wage Inflexibility. In Beckerman, W., ed., *Wage Rigidity and Unemployment.* London: Duckworth, 1986, pp. 43–76.

Pileggi, N. *Wiseguy.* New York: Pocket Books, 1986.

Roethlisberger, F. J., and Dickson, W. J. *Management and the Worker.* Cambridge: Harvard University Press, 1939.

Roy, D. Quota Restriction and Goldbricking in a Machine Shop. *American Journal of Sociology,* 1952, *67,* 427–42.

Sandemose, A. *A Fugitive Crosses His Track.* New York: Knopf, 1936.

Schoeck, H. *Envy.* Indianapolis: Liberty Press, 1987.

Swenson, P. *Fair Shares.* Ithaca: Cornell University Press, 1989.

Thaler, R. Towards a Positive Theory of Consumer Behavior. *Journal of Economic Behavior and Organization,* 1980, *1,* 39–60.

Thomas, K. *Religion and the Decline of Magic.* Harmondsworth: Penguin, 1973.

Tocqueville, de A. *Democracy in America.* New York: Anchor Books, 1969.

Trivers, R. E. The Evolution of Reciprocal Altruism. *Quarterly Review of Biology,* 1971, *46,* 35–57.

Turnbull, C. *The Mountain People.* New York: Simon and Schuster, 1972.
Tversky, A., and Kahneman, D. The Psychology of Choice and the Framing of Preferences. *Science,* 1981, *211,* 4353–58.
Ullmann-Margalit, E. *The Emergence of Norms.* Oxford: Oxford University Press, 1977.
Veyne, P. *Le Pain et le Cirque.* Paris: Editions du Seuil, 1976.
Waller, W. The Rating and Dating Complex. *American Sociological Review,* 1937, *2,* 727–34.
Yaari, M., and Bar-Hillel, M. Judgments of Justice. unpublished manuscript, 1988.
Zajac, E. P. Perceived Economic Justice: The Example of Public Utility Regulation. In Young, H. P., ed., *Cost Allocation.* Amsterdam: North Holland, 1985, pp. 119–53.

This Page Intentionally Left Blank

7

Okun's Law Revisited
Should We Worry About Low Unemployment?

David Altig
Vice President and Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Terry Fitzgerald
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Peter Rupert
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

> *The Commerce Department yesterday raised its estimate of first-quarter growth to a 5.8 percent annual rate—the highest in more than a decade. Coming at a time of low unemployment, the unusually rapid growth sparked concerns that it could cause an increase in inflation.*
>
> —John M. Berry, *Washington Post Online*, May 31, 1997

The quotation above expresses a common, if not dominant, view of the genesis of inflationary pressure in an economy. The story goes something like this: High GDP growth eventually places excessive strain on a nation's resources. This strain can become particularly acute in labor markets, where it is manifested as low unemployment. The labor market tightness associated with this low unemployment ultimately leads to higher prices.

The connection between unemployment and GDP growth is often formally summarized by the statistical relationship known as "Okun's law." As developed by the late economist Arthur Okun in 1962, the "law" related decreases in the unemployment rate to increases in output growth. Over time, the exact quantitative form of this relationship has changed somewhat (a point we will return to below). However, the negative correlation between changes in the unemployment rate and changes in GDP growth is viewed as one of the most consistent relationships in macroeconomics.

The widely accepted connection of Okun's law to inflation can be better understood by noting that the unemployment/output relationship was more precisely considered by Okun as relating percentage deviations of output from its *potential* level to deviations of the unemployment rate from its "natural" level. Potential output was—and is—understood to answer the question, in Professor Okun's words, "How much output can the economy

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (May 15, 1997).

produce under conditions of full employment?"[1] Because "full employment" is defined as the state in which labor markets are neither tight nor slack, inflationary pressures are presumed to arise when output growth pushes beyond its normal levels, which in turn is related to declines in the unemployment rate below its normal levels.

A key aspect of this perspective is the implicit, but critical, role of the potential-output and full-employment concepts in determining whether a particular growth rate or unemployment rate is inherently "inflationary."[2] The meanings and implications of these concepts are the subject of considerable debate among economists. We are ourselves skeptical that there exists a definitive notion of labor market tightness associated with above-trend (or above-potential) real GDP growth that is reliably related to price pressures. Nonetheless, since the conventional wisdom holds that "appropriate" monetary policy should consider this relationship, we will address it on its own terms.

This *Economic Commentary* reviews the connection between labor resource utilization and the growth/unemployment correlation summarized by Okun's law. In general, we emphasize the same caveats offered in Okun's original presentation of the statistical relationship that came to bear his name. The essence of our argument is that recognizing the instability in the relationship between GDP growth and changes in the unemployment rate can help us understand how we have found ourselves in the happy circumstance of an economy operating with low (and falling) unemployment, robust growth, and stable inflation.

OKUN'S LAW UPDATED

In his original research, Okun found that a 1-percentage-point decline in the unemployment rate was, on average, associated with additional output growth of about 3 percentage points. Okun's law is now widely accepted as stating that a 1-percentage-point decrease in the unemployment rate is associated with additional output growth of about 2 *percent.* Figure 1, which plots annual changes in real GDP against annual changes in the unemployment rate, illustrates the basis for the current version of Okun's law.[3] The line drawn through the scatter of points indicates that the percentage change in output is roughly 3.2 minus two times the change in the unemployment rate.[4]

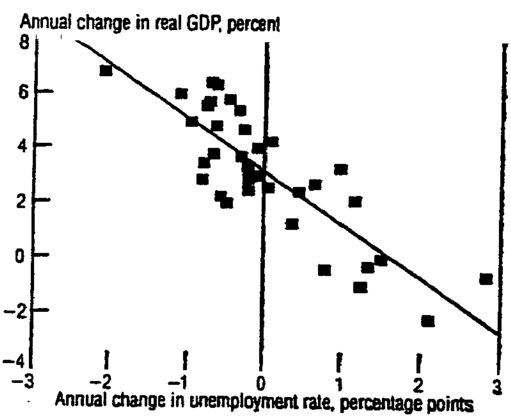


Figure 1 Unemployment and output growth, 1960–1996.

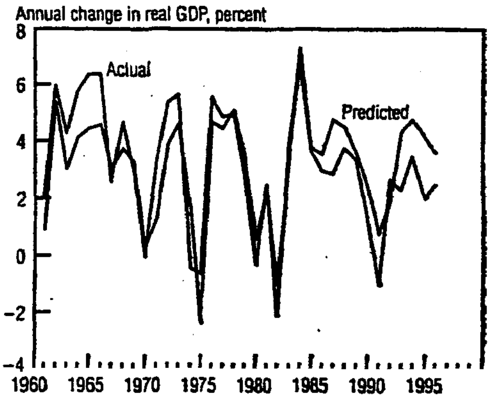


Figure 2 Output growth and Okun's law.

This relationship says that every percentage point of output growth in excess of 3.2 percent per year is associated with a drop in the unemployment rate of half a percentage point. For example, output growth of 4.2 percent would coincide with a 0.5-percentage-point decline in the jobless rate.[5]

The 3.2 percent value in this discussion has a natural interpretation as "potential GDP growth" in the Okun's law formulation. Seen in this light, the numbers reported above may seem startling. It has been a while since most Americans believed that the long-run growth rate of the economy is as high as 3.2 percent. In fact, as shown in Figure 2, the Okun equation predicts output growth very well over most of the 1970s, but less well from about the mid-1980s on (or in the 1960s, for that matter).

In effect, the relationship between the jobless rate and a given amount of economic growth has shifted over time (and with it, the implied value of potential output). This can be clearly seen in Figure 3, which shows the rates of output growth obtained from Okun's law models estimated for three distinct subsamples over the past three decades: 1961–70, 1971–80, and 1981–90. Although the predicted pattern of GDP growth is similar across all

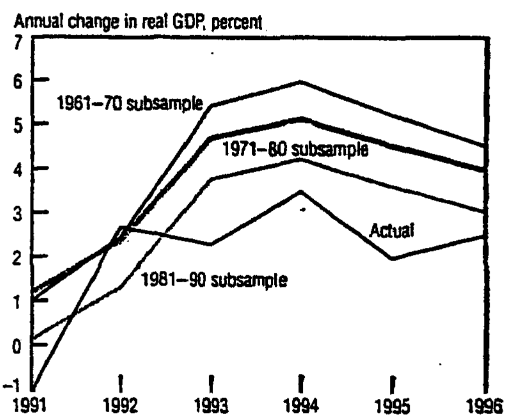


Figure 3 Okun's law predictions. NOTE: Subsamples in figure 3 are estimates. SOURCES: U.S. Department of Commerce. Bureau of Economic Analysis; U.S. Department of Labor. Bureau of Labor Statistics; and authors' calculations.

three variants, the projected levels are quite different, with the "fit" of the model improving when more recent data are used in the estimation.

Note that although Okun's law does generally capture the shape of the time series of output, there are several instances in which the direction of GDP growth is inconsistent with the model's predictions. In 1993, for instance, Okun's law would have had GDP growth increasing substantially, whereas it in fact fell relative to 1992. The reverse occurred in 1996: GDP growth was higher than in the prior year, despite the decline predicted by the Okun equation.

These short-run "mistakes" in the model occur for precisely the same reasons that the relationship between growth and unemployment does not remain constant over long periods. GDP growth depends in a fundamental way on the level and rate of change of labor resource utilization. This is the case in both the short and the long run, and such changes are only imperfectly captured by changes in the unemployment rate.

LABOR RESOURCE UTILIZATION AND GDP

Although Okun's law expresses a relationship between changes in the unemployment rate and output growth, it is more appropriate to think of it as a "rule of thumb," as it was intended, rather than as an immutable law derived from theory. To understand why the rule of thumb holds, it is necessary to understand the relationship between a nation's output and its inputs to production.

Basic economic theory tells us that output depends on both the amount of inputs used and the level of technology. In a very general categorization, the inputs to production are the labor services provided by a nation's citizens and the services provided by the current capital stock. It follows, then, that changes in output can result from changes in overall productivity, in the flow of capital services, and/or in the quantity of labor services. When observed over a relatively short horizon, such as a quarter or a year, shifts in the aggregate capital stock are likely to be limited because of the difficulty of quickly adjusting the size of this stock. Therefore, changes in output will largely reflect changes in productivity (output per hour) and in the level of labor services.[6]

The output of an economy does not depend directly on the unemployment rate. However, the labor services provided are related to the unemployment rate, and it is through this channel that unemployment is related to GDP growth. As shown in Box 1,

Box 1 Output, Labor Services, and the Unemployment Rate

$$\text{Percentage change in output} \cong \text{Percentage change in productivity} + \text{A} * \text{Percentage change in labor services}$$

$$\text{Percentage change in labor services} \cong \text{Percentage change in hours per worker} + \text{Percentage change in population} + \text{Percentage change in labor force participation} + \text{Percentage change in other factors} - \text{Absolute change in the unemployment rate}$$

NOTE: Given typical assumptions about the technology that describes the macroeconomy, the parameter A is the first expression can be interpreted as the share of total national income earned by labor. This, of course, would imply that $0 < A < 1$.

a simple accounting exercise suggests that in the short run (holding capital services fixed), the percentage change in output can be written as the sum of the percentage change in productivity plus a constant times the percentage change in labor services. The percentage change in labor services, in turn, can be written as the sum of the percentage change in hours per worker, population growth, labor force participation, and other relevant factors (such as worker efficiency levels), minus the absolute change in the unemployment rate.

From this perspective, it may seem surprising that there is much association at all between changes in the unemployment rate and changes in output. After all, a change in the jobless rate is but one of several factors that contribute to a change in GDP. Also surprising is the fact that a 1-percentage-point decrease in unemployment would, on average, be connected with such a large (2-percentage-point) increase in output growth. The simple accounting intuition above suggests that a 1-percentage-point drop in the unemployment rate would translate into roughly a 1-percentage-point rise in employment, which itself translates into an increase in GDP growth of something less than 1 percentage point.

The reason that the association between the unemployment rate and output is relatively strong is that changes in the unemployment rate are related to changes in the other factors that affect output growth. For example, using average annual data, a rising unemployment rate is strongly associated with decreases in both hours per worker and labor force participation. Since each of these factors contributes to falling output, it is clear how small upticks in the unemployment rate could be associated with larger declines in GDP.

WHEN AND WHY THE OKUN'S LAW RELATIONSHIP CHANGES

The preceding discussion sheds light on why and when the association between the unemployment rate and output may change or fail to hold altogether. As noted, a stable and relatively predictable relationship between these two measures requires that the relationship between unemployment and the full set of factors determining labor resource utilization itself be relatively predictable and constant over time.

The association between changes in the unemployment rate and output growth can become less reliable for a variety of reasons, including breaks in the historical magnitude (or even the direction) of the correlation between the unemployment rate and the population growth rate, labor force participation rates, or average hours per worker. However, the predominant factor that has tended to undermine specific representations of Okun's law (in both the short and the long run) has been changes in productivity.

Productivity changes are only slightly correlated with changes in the unemployment rate, and the variability of these changes is large—roughly two-thirds of the variability of output. Since 1960, annual changes in productivity, measured as total output per hour, have varied from as low as–1.6 percent in 1974 to as high as 4.5 percent in 1962.

Consider the situation in 1993, when the unemployment rate fell 0.6 percent, compared to an increase of 0.6 percent in the previous year. Contrary to what Okun's law would suggest, output growth was lower in 1993 than in 1992. Why? Using the accounting method described above, we find that the difference is due to a relatively large, 2.7 percent increase in productivity that occurred in 1992, versus a slight decline of 0.4 percent in 1993.

Changes in productivity are difficult to predict, even in richer models of the interaction between unemployment and output. Furthermore, it is widely accepted that the trend growth of productivity has tended to shift over time. As the trend path of productivity changes, so must the conventional measures of potential GDP embedded in the standard Okun's law model. Considering this, and the variability of productivity changes generally, it is no surprise that productivity changes are not reliably captured by the simple rule-of-thumb relationship represented by Okun's law.

OKUN'S LAW AND MONETARY POLICY

Our discussion can be summarized by the simple observation that the relationship between the unemployment rate and GDP growth changes through time or, in Okun's language, that potential GDP growth is not constant over time. Although this is widely understood to be true over extended periods—decade to decade, say—it is equally true over the much shorter horizons that characterize the business cycle. If these changes are substantial, Okun's rule of thumb can send very misleading signals about the rate of economic growth associated with any given change in the unemployment rate.

From the standpoint of conventional wisdom, the implications of this observation should be clear. Even accepting the proposition that higher output growth begets lower unemployment rates, which in turn beget inflationary pressure, this chain of effects is useful for monetary policy only to the extent that these crucial relationships are stable and predictable. On this score, both theory and evidence provide ample ammunition for the skeptic.

Perhaps the recent downticks in the unemployment rate do indeed signal strains on labor resource utilization that threaten the promotion of price stability, as suggested by the opening quotation. On the other hand, we could be witnessing a fundamental shift in the level of potential GDP, as traditionally defined, and a transition to a new version of Okun's law associated with permanently higher GDP growth. The uncertainty about which of these alternatives is true raises serious questions about the usefulness of Okun's-law-type relationships when conducting monetary policy in real time.

NOTES

1. See Arthur M. Okun, "Potential GNP: Its Measurement and Significance," American Statistical Association, *Proceedings of the Business and Economics Section,* 1962, pp. 98–103. Gross National Product (GNP) was the typical gauge of total output at the time Okun was writing. It measures the total output of U.S citizens, independent of the country in which production occurs. Gross Domestic Product (GDP), the measure we use today, represents total output produced in the United States, independent of what country's citizens own the resources used in production.
2. There is a sense in which the terminology used here is misleading and unfortunate. It is widely accepted that in the long run, inflation is a purely monetary phenomenon. In other words, inflation that persists when output is at its potential and unemployment is at its natural rate is solely attributable to monetary growth in excess of demand. Seen in this context, the trend inflation rate is wholly unconnected to the level of resource utilization in labor markets. Nonetheless, for purposes of this discussion, we will proceed using the conventional, though imprecise, language.
3. In his original work (footnote 1), Okun offered changes in GDP growth and the unemployment rate as empirical proxies for deviations from potential, or "natural," levels.
4. The actual estimated linear regression equation is 3.17 for the intercept and–1.93 for the slope.

5. This follows from a simple algebraic rearrangement of the Okun's law equation: Letting u be the unemployment rate and g be the growth rate of GDP, the equation can be expressed as

$$g = 3.2 - 2\Delta u$$
$$\Rightarrow g - 3.2 = 2\Delta u$$
$$\Rightarrow \Delta u = \frac{g - 3.2}{2}.$$

For $g = 4.2$, this implies $\Delta u = 1/2$.

6. This ignores the important possibility that capital services can adjust in the short run via fluctuations in capital's utilization rate. As a measurement issue, failure to control for variable capital utilization would lead us to ascribe too large a fraction of total output to factor productivity.

This Page Intentionally Left Blank

8

Nobel Laureate Robert E. Lucas, Jr. Architect of Modern Macroeconomics

V. V. Chari

Professor of Economics, University of Minnesota, and Advisor, Research Department, Federal Reserve Bank of Minneapolis, Minneapolis, Minnesota

In the late 1960s and early 1970s, Robert E. Lucas, Jr., wrote a number of papers which have rightly been revered as modern classics. For this body of work, Lucas received the Nobel Prize in Economic Sciences in the fall of 1995. The purpose of this review is to place Lucas' work in a historical context and to evaluate the effect of this work on the economics profession. In writing this review, I have benefited greatly from Lucas' (1996) Nobel lecture and from the essay of Thomas Sargent (1996) which was written to kick off a conference held at the Federal Reserve Bank of Minneapolis to celebrate the 25th anniversary of the publication of Lucas' (1972) seminal paper, "Expectations and the Neutrality of Money."

Lucas' work is sometimes heralded as revolutionary, marking the beginning of the end of Keynesian economics and the birth of rational expectations economics. This tendency to mark all key developments in economics as revolutionary is popular enough, but in my view, it is a misreading of the history of economic thought. My thesis is that Lucas' work is very much a part of the natural progress of economics as a science. Scientific progress arises from the interaction between theory and data and the desire to have one unified theory to account for the observations at hand. The search for such a theory proceeds by developing specific abstractions, or *models,* to understand specific observations. These abstractions then lead to the development of a more general theory, which in turn leads to discarding models which are inconsistent with data and to the development of better models. Lucas' central contribution was to develop and apply economic theory to specific questions in macroeconomics and to make obsolete one class of models. With trenchant vigor and uncommon grace, Lucas argued that economic theory could be used to illuminate old and puzzling substantive questions.

Lucas' contributions are both methodological and substantive. The methodological contribution is to illustrate how one goes about constructing dynamic, stochastic general equilibrium models to shed light on questions of substantive economic interest. The substantive contribution is to develop and analyze a specific mechanism by which monetary instability leads to fluctuations in output and inflation. It is hard to overemphasize the

Reprinted from: *Journal of Economic Perspectives,* vol. 12, no. 1 (Winter 1998) 171–186.

contribution to method. Economists today routinely analyze systems in which agents operate in complex probabilistic environments in order to understand interactions about which the great theorists of an earlier generation could only speculate. This sea change is due in substantial part to Lucas.

THE THEORETICAL FOUNDATIONS OF MACROECONOMICS

By the 1960s, the models used in macroeconomics described the aggregate economy as consisting of a system of equations: one equation to describe consumption, one to describe investment, one to describe money demand, and so on. Each of these equations was loosely thought of as arising from a deeper formulation of individual or firm decision making. This approach was attractive because the models were mathematically explicit and the parameters of the equations could be estimated using the powerful econometric procedures that had been developed in the postwar era under the influence of the Cowles Commission. These macroeconometric models were widely used for answering questions such as, How does the conduct of monetary policy affect output, inflation, and unemployment? A growing consensus in economics viewed these models as fitting the behavior of the U.S. economy and as suitable for generating answers to policy questions; for an expression of this confidence, see Franco Modigliani's (1977) presidential address to the American Economic Association. At the same time, the desirability of making specific the relationship between macroeconometric models and microeconomic theory was widely recognized. That is, macroeconomics needed theoretical foundations.

The chief difficulty in developing these foundations was that macroeconomic questions necessarily involve dealing with dynamics and uncertainty. An individual choosing how much to spend today is necessarily making a choice of how much to consume in the future. Investment decisions are based on the expectations of future returns. These and other decisions are fraught with risk. Furthermore, they are based on anticipations of prices that will prevail in the future. How does one model this decision making and the way in which anticipations are made and revised?

Economic theory is about developing frameworks that can be used to analyze such situations. The theory has at its base two fundamental postulates. First, individuals act purposefully to achieve the ends they seek, and this feature can best be captured in models where agents maximize a well-defined objective function. Second, since outcomes depend upon the actions of everyone in society, agents must form expectations about the actions of others and, indeed, expectations about the expectations of others, and so on. This feature can be captured by the notion of *equilibrium.*

The equilibrium postulate is a convenient and powerful way of summarizing these expectations and ensuring consistency in decision making. As the name suggests, *equilibrium* is the rest point of a system, and it was conventional to think of this rest point in terms of quantities and prices. However, this conventional view is not particularly helpful in thinking about a world which is continually buffeted by shocks. In such a world, the sensible way to think about decision problems is as formulating decision rules or contingency plans for choosing actions which depend upon agents information. The central theoretical breakthrough of the last 50 years is that economists now think of equilibrium as a rest point in the space of decision rules. This breakthrough appeared in the most theoretical and abstract reaches of the discipline in the work of John Nash (1950) in game theory and the work of Kenneth Arrow (1951) and Gerard Debreu (1959) in the theory of competitive equilibrium. Lucas is perhaps the foremost recent developer and expositor of this view. Thinking of equilibrium as a rest point in the space of decision rules has given economists the conceptual framework to analyze a bewildering variety of environments in which dynamics and uncertainty play central roles.

The contrast between the theoretical foundations of the 1960s-style macroeconometric models and those of modern models is stark: the book edited by Thomas Cooley (1995) is a collection of papers which illustrate the style of modern macroeconomic modeling. The earlier generation of macroeconometric models was frequently rationalized as representing the equilibria of static general equilibrium models together with tacked on dynamics representing slow wage and price adjustment to shocks. The parameters describing the speed of adjustment were not derived from maximizing behavior. The notion that people setting wages and prices will not react rationally to the expected future state of the economy, or will react in a mechanistic way, is fundamentally at odds with the maximization postulate. It was well understood that this was not a happy state of affairs. Resistance to conventional economic theory came in substantial part because equilibrium models were thought to be inconsistent with high rates of unemployment.

Thus, the macroeconomics of the 1960s and early 1970s needed firm theoretical foundations, and the great contribution of Lucas and others (including Robert Barro, Edward Prescott, Sargent, and Neil Wallace) lies in the attempt to reformulate old questions in the language of economic theory. In doing so, these theorists clarified the questions for which macroeconometric models could provide reliable answers and the questions for which such models could not provide reliable answers. More important, these theorists laid out a research program for studying substantive questions in macroeconomics. Modern economic models apply economic theory consistently. These models also have a surprising ability to reproduce observations that were thought to be inconsistent with equilibrium, including unemployment, underutilization of capital, and fluctuations in economic aggregates.

In the next sections of this review, I focus on two papers by Lucas: his 1972 paper "Expectations and the Neutrality of Money" and his 1976 paper "Econometric Policy Evaluation: A Critique." These papers were explicitly cited by the Royal Swedish Academy of Sciences in awarding Lucas the Nobel prize.

EXPECTATIONS AND THE NEUTRALITY OF MONEY

The Setting

By the late 1960s, there was a consensus among macroeconomists that the Phillips curve was a central feature of business cycles. A. W. Phillips (1958) plotted the rate of growth of nominal wages against the unemployment rate for the United Kingdom and showed that these variables were negatively associated. Subsequent analyses focused on the relationship between the rate of change of a broad index of prices of goods and services—that is, the inflation rate—and the deviations of gross national product, or output, from a trend. A stable relationship of this kind has immediate policy implications. It suggests that monetary authorities can lower unemployment at the cost of a somewhat higher inflation rate and can reduce the inflation rate only by incurring the cost of higher unemployment.

However, Milton Friedman (1968) and Edmund Phelps (1968) soon mounted powerful theoretical arguments against these policy recommendations. They argued that economic theory suggests that sustained inflation can have no effect because people care about real quantities, not nominal ones. Once people anticipate sustained inflation, they will adjust their pricing, employment, and job search decisions in ways that take inflation into account, rendering the inflation irrelevant to real economic decisions. These considerations suggest that sustained inflation cannot lead to a permanent reduction in unemployment. Friedman emphasized that expectations adjust slowly to permanent changes in the inflation rate. This slow adjustment implies that unemployment can be temporarily low when the economy is stimulated by, say, expansionary monetary policy. But eventually the monetary expansion will filter through to the economy in higher prices, and unemployment will return to the level determined by underlying real forces. Phelps (1970), in his introductory essay to a marvelous volume, sketched out a formulation in which informational imperfections lead people to believe that overall price changes reflecting monetary fluctuations are instead relative price changes favoring the industry or sector in which they are employed. The stage was set for Lucas to flesh out this sketch in the language of modern economic theory.

The Question

In "Expectations and the Neutrality of Money," Lucas (1972) asks one of the oldest questions in economics: How do changes in the conduct of monetary policy affect inflation, output, and unemployment? At least since David Hume in 1752 (in Rotwein 1970), economists have struggled with this question, and it continues to occupy center stage two and a half centuries later. The evidence is unambiguous in one respect: Business cycle booms are times in which the growth rate of monetary aggregates is higher than average, and contractions are times in which the growth rate of monetary aggregates is lower than average. A central question in macroeconomics is whether monetary policy can and should be used to moderate business cycle fluctuations. It is the kind of question that the data alone cannot answer. Models are needed.

Lucas made a substantive and a methodological contribution in his 1972 paper. The substantive contribution is to develop and analyze a specific mechanism by which monetary instability leads to fluctuations in output and inflation. In this mechanism, people with limited information confuse monetary disturbances with relative price movements, so that monetary fluctuations lead to aggregate output fluctuations. The methodological contribution is to illustrate how one goes about constructing dynamic, stochastic general equilibrium models to shed light on questions of substantive economic interest.

Lucas set his argument in a framework originally introduced by Paul Samuelson (1958). In this overlapping generations framework, people live for two periods, so that in any period the economy always has people of two age groups, the young and the old. At the end of each period, the old die, the young become old, and a new generation is born. There is only one good. Only the young can work and produce the good, but both young and old people like to consume it. The good cannot be stored. In this highly stylized economy, current and future generations can all be made better off if they could devise some mechanism to transfer part of production in every period to those who are old. One obvious such institution is social security, and indeed, modern analyses of social security programs use the overlapping generations model as a point of departure.

Samuelson (1958) noted that other institutions could also perform much the same function. In particular, intrinsically useless pieces of paper, called *money,* could provide the old with a claim to part of the output produced by the young. Each generation of young people willingly gives up part of what they produce for pieces of paper, because they think that future generations will exchange these pieces of paper for goods. Suppose that the number of pieces of paper, or the *stock of money,* is fixed and held by the initial generation of old people and that each generation of young people is identical. The simplest way of thinking about this kind of economy is that people behave competitively; that is, they take the price of goods in terms of money as unaffected by their individual decisions on how much to produce and consume. Old people supply all the money they possess and consume what the market provides. Young people have a more interesting problem. In choosing how much of their production to supply to the market, they need to forecast the value of money when they are old. The value of money, of course, depends upon the decisions of the next generation and therefore upon the forecasts that will be made by the next generation. Rational decision making by today's young requires forecasting the forecasts of others.

It is here that the notion of equilibrium allows analysis of an apparently intractable problem. In this unchanging world, the notion of equilibrium requires that expectations of future prices, or forecasts, be the same as the prices that actually prevail. An equilibrium, then, is a price in each period and a choice by young people in each period of how much to sell to the market, given the price when they are young and the price when they are old, such that the amount of money brought by old people into the market is the same as the amount of money young people want to carry into the future. This last requirement is sometimes described as a *market-clearing condition.*

This kind of monetary economy shares a feature with all sensibly formulated economies. The units in which prices are quoted have no effect on the outcomes people care about. If we split up dollar bills into pennies and quote prices in pennies rather than dollars, it is obvious that all that happens is that prices are multiplied by a factor of 100. This property is called *zero-degree homogeneity* of prices. An implication of zero-degree homogeneity is that if we double the number of dollar bills once and for all in the hands of the initial generation of old people, all that happens is that prices double in all periods. Monetary economies with this feature are said to display *neutrality.* More generally, money is said to be *neutral* if a proportionate change in all nominal, or dollar-denominated, quantities in all periods is associated with a proportionate change in all prices and no change in real quantities. In Samuelson's (1958) economy, a one-time change in the number of dollars held by the initial generation of old people leads to a proportionate change in all nominal quantities and in all prices, so that monetary injections of this kind are neutral.

Monetary injections of other kinds may or may not be neutral. Suppose, for example, that a monetary authority expands the quantity of money at a constant rate and does so by continually handing out money to old people in a lump-sum fashion, that is, independently of the amount of money any particular old person may have. One would expect this kind of injection to lead to a constant increase in the price of goods, that is, to *inflation.* In this kind of a world, young people see that the purchasing power of money will be diminished by the time they are old, and this reduced reward to work today leads to a fall in their willingness to work and in output. That is, inflation induced by ongoing monetary expansions of this kind acts much as a tax does. The inflation tax is an important feature of actual economies, but in this context, it leads to the implication that anticipated expansions in the stock of money depress current economic activity. Lucas, remember, is seeking to create a framework in which a current expansion in the stock of money first creates a surge in real activity, as the evidence suggests actually occurs, but is neutral in the long run. For this purpose, the inflation tax argument works in exactly the wrong way.

However, there are other ways of injecting money into the economy which lead to continuing inflation but do not alter real decisions about how much to produce and consume. Consider, for example, handing out money to old people in exact proportion to the amount of money they have carried over from the past. Monetary injections of this kind are neutral because they do not change the rate of return to holding money. With injections of this kind, the negative effect of inflation on willingness to work is exactly undone by the higher return associated with the proportionate transfer.

The Answer

To make the informational mechanism play a central role, Lucas assumes that transfers are proportional. Consider a situation in which these transfers are random. The interesting feature of this economy is that even if the young do not know the size of the monetary transfer in the current period, money is still neutral. The reason is that in competitive markets, the young generation can observe the price of goods before making their production decision. Therefore, in equilibrium, the prices reveal the size of the monetary transfer. As a result, in equilibrium, prices simply rise in each period by the amount of the transfer, and real allocations are completely unaffected.

The central economic idea that Lucas wants to formalize is that monetary disturbances lead to movements in prices that people interpret as meaning that the present is a favorable time to produce. The elegant formulation Lucas chooses is one in which trade occurs in "two physically separated markets" (1972, p. 103). Specifically, think of the economy as two islands, each with an equal number of old people. The overall number of young people is fixed, but they are divided randomly between the two islands in a given time period. Suppose for the moment that the stock of money is fixed for all time. Young people who find themselves on an island with few young people will find that the price of the good they sell is high, since there are few producers. This temporarily high price signals to them that they should produce a relatively large amount. Young people assigned to the other island find a low price and choose to produce little. In this economy, output on one island is higher than average, and output on the other island is lower than average. There is no particular reason that total output should exactly be equal to its average value, so in this sense output will fluctuate over time, depending on the exact assignment of young people. However, these fluctuations seem to have little to do with business cycles, since a key feature of the business cycle is that essentially all sectors of the economy move together.

Now consider adding monetary disturbances to this economy. A higher than average transfer induces prices to rise on both islands. Consider the problem facing a typical young person. Prices could be high because of the monetary disturbance, in which case the best thing to do is not to respond in terms of production decisions, or prices could be high because there are relatively few people on the island, in which case the best thing to do is to produce more. If a producer does not know why the price is high, the optimal decision is a mix of these extremes, so that output in both islands rises relative to the case when there was no monetary disturbance. Thus, in this economy, prices are higher than average precisely when output is higher than average—and this is precisely when the rate of growth of the money supply is higher than average. Prices and output are lower than average when the rate of growth of money is lower than average. Notice, however, that if the size of the monetary disturbance is known, there is no scope for confusion about the source of the price increase, and monetary disturbances are neutral. The model requires that we draw a sharp distinction between *anticipated* monetary fluctuations, which are neutral, and *unanticipated* fluctuations, which induce output movements.

Lucas (1972, p. 119) also uses the model to argue for a particular sense in which the best monetary policy is one in which the monetary authority follows a "k-percent rule," in which the rate of growth of the quantity of money is constant. At this point it is best to quote from the conclusion to the paper (Lucas 1972, pp. 121–22):

> This paper has been an attempt to resolve the paradox posed by Gurley (1961), in his mild but accurate parody of Friedmanian monetary theory: "Money is a veil, but when the veil flutters, real output sputters." This resolution has been effected by postulating economic agents free of money illusion, so that the Ricardian hypothetical experiment of a fully announced, proportional monetary expansion will have no real consequences (that is, so that money *is* a veil). These rational agents are then placed in a setting in which the information conveyed to traders by market prices is inadequate to permit them to distinguish real from monetary disturbances. In this setting, monetary fluctuations lead to real output movements in the same direction.
>
> In order for this resolution to carry any conviction, it has been necessary to adopt a framework simple enough to permit a precise specification of the information available to each trader at each point in time, and to facilitate verification of the rationality of each trader's behavior. To obtain this simplicity, most of the interesting features of the observed business cycle have been abstracted from, with one notable exception: the Phillips curve emerges not as an unexplained empirical fact, but as a central feature of the solution to a general equilibrium system.

The Legacy

The demonstration that a Phillips curve could emerge in an economic model with rational agents is at one level an impressive display of technical wizardry. The key to the technical contribution is that prices are thought of as *functions* of the state of the economy, where the state is the stock of money and the distribution of young people across islands. This notion has its antecedents in the work of Arrow (1951), Debreu (1959), John Muth (1961), and Lucas and Prescott (1971). Muth advanced the principle of rational expectations as a model-building device: the idea is that the expectations attributed to economic agents in a model should be the same as those implied by the model. More generally, the rational expectations hypothesis is that agents use available information in the best way.

It took some time before this principle was widely used in economics. Once it began to be used, however, it took the field by storm for three reasons. First, unlike the alternatives, the notion of rational expectations adds no free parameters but, instead, imposes restrictions across equations. In contrast, for example, the notion of adaptive expectations involves adding free parameters to describe how expectations are formed and revised. Second, rational expectations is consistent with individual maximization, since it rules out the existence of obvious profit opportunities. Third, the equilibrium point of view practically forces one to use rational expectations. Once prices and choices are thought of as functions of the state of the economy, one is forced to impart beliefs to economic agents about how the state evolves and therefore beliefs about the model of the economy held by agents in our models. Today, it seems hard to imagine starting anywhere else.

Some of the most interesting recent theoretical work involves studying how agents learn; a good introduction to this literature is Sargent 1993. Specifically, one question that many authors have attempted to study is whether agents who start off with beliefs other than those implied by rational expectations will eventually come to hold rational expectations. Another line of research assumes that people are boundedly rational and asks whether such economies will eventually look like economies with fully rational agents. The starting point for both literatures is a rational expectations equilibrium.

With the model in "Expectations and the Neutrality of Money," Lucas emphasizes the distinction between anticipated and unanticipated changes in the stock of money. In this sense, the approach represents a difference, and I think an advance, over the distinction between the long run and the short run which both Friedman and the Keynesian literature emphasized. The specific formulation led to a long and misdirected debate over whether rational expectations implies that anticipated monetary policy could have no real effects. It is abundantly clear from the model that neutrality of anticipated monetary policy depends critically upon the manner in which money is injected. Other ways of injecting money have effects on output. For example, if monetary injections were made in a lump-sum manner, the inflation tax would affect the behavior of output. However, we have good reason to believe that these effects are likely to be small. In any event, the economy will respond quite differently to anticipated and unanticipated changes in the stock of money.

This contribution of "Expectations and the Neutrality of Money" led to an extensive empirical literature. (See, for example, Sargent 1976 and Barro 1977.) In "Some International Evidence on Output-Inflation Tradeoffs," Lucas (1973) noted that a key implication of the 1972 paper is that when monetary fluctuations become very volatile, agents will pay no attention to the price signal when making their decisions. This immediately suggests that countries with volatile inflation rates should have less volatile output. The international evidence lent some support to this view. The distinction between the effect on output of anticipated and unanticipated changes in the stock of money was also tested for U.S. time series data by many economists, notably Sargent and Barro. The evidence here is mixed, and it is fair to say that the effects of price surprises appear to be weak.

The idea that informational limitations play a central role in how monetary policy affects output in the real world has largely fallen by the wayside. In part, this is because of the evidence from U.S. time series. The main reason, however, is that it seems quite difficult to use this mechanism to generate persistent effects of monetary shocks on output. In developed economies like the United States, information about economywide outcomes is readily and quickly available. It may be reasonable to suppose that people are confused about the sources of price changes for perhaps two or three months, but it seems difficult to see how people could continue to be misinformed for two or three years. Since business cycle fluctuations last at least that long, this mechanism is not persuasive as a model of business cycles.

One especially interesting logical descendant of "Expectations and the Neutrality of Money" is Finn Kydland and Prescott's (1982) "Time to Build and Aggregate Fluctuations." Both papers take seriously the ideas of focusing on one key driving force behind business cycle fluctuations, of using the best economic theory available, and of taking the implications of the theory seriously. Substantively, however, it is difficult to imagine two papers more at odds with each other. In 1972, Lucas modeled business cycles as an avoidable consequence of erratic monetary policy. In 1982, Kydland and Prescott modeled business cycles as the efficient response of the economy to technology disturbances. Neither paper's substantive message has been accepted by the profession at large, but the methodological contributions are overwhelming. It is hard to pick up a recently published paper in macroeconomics that does not routinely use the notion of rational expectations equilibrium, and dynamic, stochastic general equilibrium models in the style of these papers have become the workhorses of modern macroeconomics.

What, then, is the legacy of "Expectations and the Neutrality of Money"? The paper is a contribution to method. It led to a simple reduced-form model of output fluctuations which continues to be widely used in the time-consistency literature and in positive models of central bank policy. Along with the work of Friedman and Phelps, the paper contributed to the demise of the belief that there was a long-run trade-off between unemployment and inflation for policymakers to exploit. The great inflation of the 1970s was surely due in part to the economics profession's acceptance of the Phillips curve, just as the great disinflation of the 1980s and 1990s was due in part to the profession's acceptance that the long-run Phillips curve is vertical (or perhaps even slopes upward). In this, as in so much else, ideas have profound consequences.

ECONOMETRIC POLICY EVALUATION: A CRITIQUE

The process of integrating economic theory into macroeconomics has fundamentally altered the profession's perspective on a variety of questions. Most notably, it has discredited the usefulness of 1960s-style macroeconometric models for answering a variety of policy questions. For example, suppose we want to ask how the behavior of the U.S. economy would change if the Federal Reserve System were to adopt a policy of maintaining the growth rate of the money supply at 4 percent per annum. Using parameter estimates in a macroeconometric model generated from a time period when the Federal Reserve was pursuing a completely different policy makes sense if one believes these parameters would not change under a different policy regime. The problem is that economists have every reason to believe that parameters in such models are a mongrel of the way in which people's expectations are formed and of underlying features of the economy, such as preferences and technology. Expectations depend upon the nature of the policy regime in place and therefore are likely to change systematically with the regime. This is the substance of the 1976 *Lucas critique.*

In some ways, the Lucas critique has had a more substantial impact than did "Expectations and the Neutrality of Money" (1972). In part, this is due to the simplicity of the examples Lucas used in the critique to make his point. But the greater impact of the critique stems from the fact that it uses entirely conventional theoretical formulations to criticize the use of macroeconometric models in policy evaluation. Economists have long understood that economic models cannot sensibly be used for policy evaluation unless one has confidence that the structure of the model and its parameters are likely not to change under alternative policies. The typical macroeconometric model is a system of equations which are interpreted as describing the behavior of the people, the firms, and the government in the economy. When such models are used for evaluating alternative policies, they implicitly presume that the parameters of the equations will be invariant with respect to alternative policies. However, as Lucas (1976, p. 25) wrote in a later paper: *"Everything we know about dynamic economic theory indicates that this presumption is unjustified"* (emphasis in original). The argument behind this bold claim is that the equations in macroeconometric models are implicitly based on decision rules which specify what people will do, given the state of the economy. However, these decision rules depend on their expectations of future policies, which in turn surely depend on the kinds of policies chosen in the past. If policymakers choose policies in a new manner, surely people's expectations about future policies will change, and their decision rules will also.

The distinction between structural and reduced-form parameters and warnings about using reduced-form models for policymaking were well known in economics far before Lucas. (See, for example, the work of Jan Tinbergen, 1952, and Jacob Marschak, 1953.) The value of the critique lies in its use of graphic examples to illustrate the argument and the alternative program it advocated. The first example Lucas uses is one based on Friedman's permanent income hypothesis. Friedman (1957) hypothesized that consumption is a function of *permanent income,* which is defined as that constant flow which yields the same present value as an individual's expected present value of actual income. Friedman also posited that permanent income is a weighted average of past incomes. Muth (1960) then showed more rigorously that a particular stochastic behavior of income over time, together with optimal forecasting by agents, implies that the best estimate of permanent income is an exponentially weighted average of past incomes. This stochastic process for income is given by the sum of a highly *persistent* part (a random walk) and a very *transitory* part (an independent random variable). Muth showed that the weights on past incomes depend on the relative variabilities of the two components; for example, if the transitory part has large variance, rational individuals attribute income fluctuations to the transitory part, and thus, the weight on current income is low. In terms of the relationship between consumption and income, this theory gives consumption as a function of current and past incomes, where the weights depend upon the relative variabilities. From an econometric point of view, one can obtain the relationship between consumption and income from historical data by running a regression.

Lucas used this framework to make his point that this kind of regression relation cannot be used to uncover parameter values which are invariant across some interesting policy experiments. Consider a policy which supplements the individual's income by a constant amount forever. If this policy is known to the individual, it is clear that permanent income rises by the amount of the supplement and consumption rises in proportion to permanent income. Traditional uses of macroeconometric models regard the relationship between consumption and income as given by the historical data and use the estimated relationship to forecast the implied time path of the expected change in consumption. This relationship implies that expected consumption will gradually rise. The theory, however, says that consumption should rise immediately and that expected consumption should be permanently higher immediately. This apparent conflict between the implications of a widely accepted theory and conventional procedure has had a lasting effect on the profession.

Lucas used other examples to make the point that conflicts of this variety are pervasive. One example concerns the effect of a temporary investment tax credit to stimulate economic activity in recessions. It makes the point that anticipations of an investment tax credit, while the proposed credit is moving through the political process, may induce firms to postpone investments and, thereby, may accentuate the very recession the policy is designed to eliminate.

The real value of the critique lies in the clearly articulated research program it envisions. This research program involves specifying a structural model as well as the policy regime under which the economy is thought to operate. A *policy regime* is simply a function which prescribes the policies for each state of the economy. Economic agents in the model are thought of as knowing the policy regime. Data can then be used to uncover the regime as well as the details of the model. Policy evaluation, then, consists of evaluating the properties of the model under alternative policy regimes. This contribution has led directly to a vast literature on rational expectations econometrics (for example, Lucas and Sargent 1981).

The research program has also had a profound impact on the old argument over rules versus discretion in economic policymaking. Friedman (1968) has been perhaps the most prominent proponent of the view that economic and especially monetary policy should be constrained by rules that specify policy as an explicit function of the state of the economy. His arguments are primarily based on the practical view that discretionary policymaking has led to bad outcomes and that economists and policymakers do not know enough about the structure of the economy for discretionary policy to work well. Lucas' argument is based on the view that economists simply have no hope of understanding the effect of policies unless we think of policies as choosing among alternative rules.

Consider, for example, the question, Should the Federal Reserve raise interest rates next quarter? Answering this question requires that we know how future expectations will change in response to this action. If the current policy regime prescribes that the Federal Reserve should raise interest rates, then it is clear that we can forecast the effects of this action. If it does not, then we need to know how private agents will react to an apparent change in the regime. The problem is that we have no way of knowing what the new regime is. Private agents may even view this action as simply erratic monetary policy, and it is not clear that policy exercises which involve introducing noise are desirable.

Economists can, however, offer sensible policy advice when it comes to choosing among alternative policy rules, which are ways in which actions should be chosen depending upon the state of the economy. From this perspective, the question economists can answer should be posed as, Is raising interest rates next quarter part of a rule for the conduct of monetary policy that will lead to good outcomes on average? It should be emphasized that the point that economists can offer sensible advice only when it comes to choosing among alternative rules in no way implies that proposals such as Friedman's, that the stock of money should grow at 4 percent a year, are necessarily optimal. For example, John Taylor (1979) developed a model with staggered wage-setting in which a monetary policy rule which reacts to the state of the economy is better than a fixed money growth rate rule.

The perspective that policies should be thought of as rules has also led to an influential research program which uses game-theoretic techniques to understand the relative advantages of rules and discretion in policymaking. Beginning with the seminal contributions of Kydland and Prescott (1977) and Guillermo Calvo (1978), this literature thinks of discretionary policy as a situation in which the actions in each period are required to be optimal for the policymaker relative to other possible actions. This criterion generates policies as rules. It turns out, however, that the rules for policymaking implied by this procedure can be dominated in an average sense by other policies. Put differently, the policymaker can be made better off by committing to follow future policies. The simplest example is the payment of ransom to hostages. If a government could credibly commit never to pay ransom, it is possible that kidnappers would choose never to take prisoners hostage. The problem, of course, is that once hostages are taken, it may well be optimal to pay the ransom to save the hostages' lives. This issue shows up in economic situations as well. Consider, for example, the problem of default on government debt. Since revenues to pay interest on such debt typically must be raised from taxes that distort private decisions, it is optimal to default on government debt and promise never to do so again. Obviously, nobody would buy such debt if the promises were not believed. This example, then, illustrates the importance of being able to commit to an action (not to default on the debt) even though one would like to deviate from the committed action later. The models used in the rules versus discretion literature do not provide simple answers. However, taking economic theory seriously, as Lucas did, has led to an enormously influential and rich research agenda.

OTHER CONTRIBUTIONS

Lucas has made significant contributions to a number of fields in economics including financial economics (1978), monetary theory (1980a, Lucas and Stokey 1987), public finance (Lucas and Stokey 1983), international economics (1982), and, most recently, economic growth (1988). In every area, his work has set new standards and generated a large new literature. Here, let me discuss only a few of my favorites.

The work of Lucas and Leonard Rapping (1969) is, quite simply, a classic. Lucas and Rapping tried to understand why employment fell so dramatically during the Great Depression and rose so dramatically during World War II. Central to their argument is the idea that households work more hours when wages are temporarily high and fewer hours when wages are temporarily low. In the jargon of economics, the intertemporal elasticity of labor supply is high. Labor economists and macroeconomists to this day continue to argue over the size of this elasticity. It plays a central role in any model which attempts to understand the fluctuations of employment over the business cycle. Lucas and Rapping used a form of adaptive expectations in their model, but emphasis on intertemporal labor supply substitution continues in Lucas' own work and in much other work on business cycles.

Lucas 1978 is one of the most influential papers in financial economics. Here Lucas showed how asset prices could be expressed as a function of the economy's state variables and that this function is the solution to a functional equation that arises from individual optimization and market-clearing. This elegant characterization is now routinely used in the asset-pricing literature.

The field of economic growth has been a growth industry in the last decade. In the so-called new growth literature, the long-run growth rate is determined by the accumulation of physical capital, human capital, and technological know-how. In this area, Lucas (1988) has made powerful arguments that human capital accumulation has important external effects and that learning by doing plays an important role in the process of human capital accumulation.

In many ways, my personal favorite of Lucas' work is "Methods and Problems in Business Cycle Theory" (1980b), which is a piece on methodology in economics. In general, I am hostile to methodological pieces; I prefer to read about work that has been done rather than be preached at about how to do it. However, the basic premise of this engaging article is that, as scientists, economists are limited by the tools at their disposal rather than by their ability to make verbal conjectures about how the world works. Lucas argued that improvements in economic theory and computational abilities have been driving forces in the postwar transformation of economics into a quantitative science. Lucas (1980b, pp. 709–10) wrote that "Our task as I see it . . . is to write a FORTRAN program that will accept specific economic policy rules as 'input' and will generate as 'output' statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies." And how are we to build this FORTRAN program? "Progress in economic thinking means getting better and better abstract, analogue economic models, not better verbal observations about the world" (Lucas 1980b, p. 700).

SUMMING UP

Robert E. Lucas, Jr., is the preeminent macroeconomist of the last 25 years. Even when academic macroeconomists disagree over substantive questions, most work today under a common set of standards that define high-quality work. We use similar equilibrium concepts, econometric techniques, and models of policymaking. This agreement over method is due in substantial part to Lucas. The logical structure of his arguments has been central in this methodological victory, although the flair and grace of his writing and his ability to craft persuasive examples to make telling points have played important supporting roles.

Sargent (1996, p. 536) has written that "the late 1960s were good times to be a young macroeconomist." Ideas and controversies were in the air. There was a general feeling that economic science was on the verge of making sharp quantitative statements about a host of issues. The late 1990s seem to me to be even better times. The controversies are just as pronounced, but the sophistication of our theoretical tools and our abilities to make quantitative assessments are now vastly greater. Progress has by some measure been slow over the last three decades, but it is sobering to think how much slower it would have been without Lucas' contributions.

ACKNOWLEDGMENTS

I first got to know Bob Lucas when I, as a graduate student at Carnegie-Mellon, was fortunate enough to spend a year at Chicago. I still have not gotten over being treated as an equal. I am pleased to have this opportunity to acknowledge my intellectual debt. Bob is a charming and delightful person, but you do want to be thoroughly armed in any debate with him. His rhetorical skills are formidable beyond belief, and since he reads widely and majored in history, it is tough to win a debate with him. I haven't, as yet, but I keep trying. A marvelous autobiography is available through the home page of the Nobel Foundation (http://www.nobel.se). I highly recommend it.

Comments from Brad De Long, Narayana Kocherlakota, Alan Krueger, and Timothy Taylor were enormously helpful in writing this essay. The views expressed here are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

REFERENCES

Arrow, K. J. 1951. An extension of the basic theorems of classical welfare economics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability,* ed. Jerzy Neyman, pp. 507–32. Berkeley: University of California Press.

Barro, R. J. 1977. Unanticipated money growth and unemployment in the United States. *American Economic Review* 67 (March): 101–15.

Calvo, G. 1978. On the time consistency of optimal policy in a monetary economy. *Econometrica* 46 (November): 1411–28.

Cooley, T. F., ed. 1995. *Frontiers of business cycle research.* Princeton, N.J.: Princeton University Press.

Debreu, G. 1959. *Theory of value: An axiomatic analysis of economic equilibrium.* New Haven. Conn.: Yale University Press.

Friedman, M. 1957. *A theory of the consumption function.* Princeton, N.J.: Princeton University Press.

Friedman, M. 1968. The role of monetary policy. *American Economic Review* 58 (March): 1–17.

Gurley, J. G. 1961. Review of *A program for monetary stability* by Milton Friedman. *Review of Economics and Statistics* 43 (August): 307–8.

Kydland, F. E., and Prescott, E. C. 1977. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85 (June): 473–91.

Kydland, F. E., and Prescott, E. C. 1982. Time to build and aggregate fluctuations. *Econometrica* 50 (November): 1345–70.

Lucas, R. E., Jr. 1972. Expectations and the neutrality of money. *Journal of Economic Theory* 4 (April): 103–24.

Lucas, R. E., Jr. 1973. Some international evidence on output-inflation tradeoffs. *American Economic Review* 63 (June): 326–34.

Lucas, R. E., Jr. 1976. Econometric policy evaluation: A critique. In *The Phillips curve and labor markets,* ed. Karl Brunner and Allan H. Meltzer. Carnegie-Rochester Conference Series on Public Policy 1: 19–46. Amsterdam: North-Holland.

Lucas, R. E., Jr. 1978. Asset prices in an exchange economy. *Econometrica* 46 (November): 1429–45.

Lucas, R. E., Jr. 1980a. Equilibrium in a pure currency economy. *Economic Inquiry* 18 (April): 203–20.

Lucas, R. E., Jr. 1980b. Methods and problems in business cycle theory. *Journal of Money, Credit, and Banking* 12 (November, part 2): 696–715.

Lucas, R. E., Jr. 1981. *Studies in business-cycle theory.* Cambridge, Mass.: MIT Press.

Lucas, R. E., Jr. 1982. Interest rates and currency prices in a two-currency world. *Journal of Monetary Economics* 10 (November): 335–59.

Lucas, R. E., Jr. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22 (July): 3–42.

Lucas, R. E., Jr. 1996. Nobel lecture: Monetary neutrality. *Journal of Political Economy* 104 (August): 661–82.

Lucas, R. E., Jr., and Prescott, E. C. 1971. Investment under uncertainty. *Econometrica* 39 (September): 659–81.

Lucas, R. E., Jr., and Rapping, L. A. 1969. Real wages, employment, and inflation. *Journal of Political Economy* 77 (September/October): 721–54.

Lucas, R. E., Jr., and Sargent, T. J., eds. 1981. *Rational expectations and econometric practice.* 2 vols. Minneapolis: University of Minnesota Press.

Lucas, R. E., Jr., and Stokey, N. L. 1983. Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics* 12 (July): 55–93.

Lucas, R. E., Jr., and Stokey, N. L. 1987. Money and interest in a cash-in-advance economy. *Econometrica* 55 (May): 491–513.

Marschak, J. 1953. Economic measurements for policy and prediction. In *Studies in econometric method,* ed. William C. Hood and Tjalling G. Koopmans, pp. 1–26. Cowles Commission Monograph 14. New York: Wiley.

Modigliani, F. 1977. The monetarist controversy or, should we forsake stabilization policies? *American Economic Review* 67 (March): 1–19.

Muth, J. F. 1960. Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association* 55 (June): 299–306.

Muth, J. 1961. Rational expectations and the theory of price movements. *Econometrica* 29 (July): 315–35.

Nash, J. F. 1950. Equilibrium points in *n*-person games. *Proceedings of the National Academy of Sciences of the United States of America* 36 (January): 48–49.

Phelps, E. S. 1968. Money-wage dynamics and labor-market equilibrium. *Journal of Political Economy* 76 (July/August): 678–711.

Phelps, E. S., ed. 1970. *Microeconomic foundations of employment and inflation theory.* New York: Norton.

Phillips, A. W. 1958. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Economica* 25 (November): 283–99.

Rotwein, E., ed. 1970. *Writings on economics* by David Hume, 1711–1776. Madison, University of Wisconsin Press.

Samuelson, P. A. 1958. An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy* 66 (December): 467–82.

Sargent, T. J. 1976. A classical macroeconometric model for the United States. *Journal of Political Economy* 84 (April): 207–37.

Sargent, T. J. 1993. *Bounded rationality in macroeconomics: The Arne Ryde memorial lectures.* Oxford: Clarendon Press.

Sargent, T. J. 1996. Expectations and the nonneutrality of Lucas. *Journal of Monetary Economics* 37 (June): 535–48.

Taylor, J. B. 1979. Estimation and control of a macroeconomic model with rational expectations. *Econometrica* 47 (September): 1267–86.

Tinbergen, J. 1952. *On the theory of economic policy.* Amsterdam: North-Holland.

This Page Intentionally Left Blank

9
The Great Depression in the United States from a Neoclassical Perspective

Harold L. Cole
Senior Economist, Federal Reserve Bank of Minneapolis, Minneapolis, Minnesota

Lee E. Ohanian
Economist, Federal Reserve Bank of Minneapolis, Minneapolis, Minnesota

Between 1929 and 1933, employment fell about 25 percent and output fell about 30 percent in the United States. By 1939, employment and output remained well below their 1929 levels. Why did employment and output fall so much in the early 1930s? Why did they remain so low a decade later?

In this article, we address these two questions by evaluating macroeconomic performance in the United States from 1929 to 1939. This period consists of a *decline* in economic activity (1929–33) followed by a *recovery* (1934–39). Our definition of the *Great Depression* as a 10-year event differs from the standard definition of the Great Depression, which is the 1929–33 decline. We define the Depression this way because employment and output remained well below their 1929 levels in 1939.

We examine the Depression from the perspective of neoclassical growth theory. By *neoclassical growth theory,* we mean the optimal growth model in Cass 1965 and Koopmans 1965 augmented with various shocks that cause employment and output to deviate from their deterministic steady-state paths as in Kydland and Prescott 1982.[1]

We use neoclassical growth theory to study macroeconomic performance during the 1930s the way other economists have used the theory to study postwar business cycles. We first identify a set of shocks considered important in postwar economic declines: technology shocks, fiscal policy shocks, trade shocks, and monetary shocks. We then ask whether those shocks, within the neoclassical framework, can account for the decline and the recovery in the 1930s. This method allows us to understand which data from the 1930s are consistent with neoclassical theory and, especially, which observations are puzzling from the neoclassical perspective.

Reprinted from: *Federal Reserve Bank of Minneapolis Quarterly Review*, vol. 23, no. 1, (Winter 1999) 2–24. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

In our analysis, we treat the 1929–33 decline as a long and severe recession.[2] But the neoclassical approach to analyzing business cycles is not just to assess declines in economic activity, but to assess recoveries as well. When we compare the decline and recovery during the Depression to a typical postwar business cycle, we see striking differences in duration and scale. The decline, as well as the recovery, during the Depression lasted about four times as long as the postwar business cycle average. Moreover, the size of the decline in output in the 1930s was about 10 times the size of the average decline. (See Table 1.)

What factors were responsible for these large differences in the duration and scale of the Depression? One possibility is that the *shocks*—the unexpected changes in technology, preferences, endowments, or government policies that lead output to deviate from its existing steady state growth path—were different in the 1930s. One view is that the shocks responsible for the 1929–33 decline were much larger and more persistent versions of the same shocks that are important in shorter and milder declines. Another view is that the types of shocks responsible for the 1929–33 decline were fundamentally different from those considered to be the driving factors behind typical cyclical declines.

To evaluate these two distinct views, we analyze data from the 1930s using the neoclassical growth model. Our main finding differs from the standard view that the most puzzling aspect of the Depression is the large decline between 1929 and 1933. We find that while it may be possible to account for the 1929–33 decline on the basis of the shocks we consider, none of those shocks can account for the 1934–39 recovery. Theory predicts large increases in employment and output beginning in 1934 that return real economic activity rapidly to trend. This prediction stands in sharp contrast to the data, suggesting to us that we need a new shock to account for the weak recovery.

We begin our study by examining deviations in output and inputs from the trend growth that theory predicts in the absence of any shocks to the economy. This examination not only highlights the severity of the economic decline between 1929 and 1933, but also raises questions about the recovery that began in 1934. In 1939, real per capita output remained 11 percent below its 1929 level: output increases an average of 21 percent during a typical 10-year period. This contrast identifies two challenges for theory: accounting for the large decline in economic activity that occurred between 1929 and 1933 and accounting for the weak recovery between 1934 and 1939.

We first evaluate the importance of *real shocks*—technology shocks, fiscal policy shocks, and trade shocks—for this decade-long period. We find that technology shocks may have contributed to the 1929–33 decline. However, we find that the real shocks predict a very robust recovery beginning in 1934. Theory suggests that real shocks should have led employment and output to return to trend by 1939.

Table 1 Duration and Scale of the Depression and Postwar Business Cycles: Measured by the Decline and Recovery of Output

	Length of Decline	Size of Decline	Length of Recovery
Great Depression	4 years	–31.0%	7 years
Postwar Cycle Average	1 year	–2.9%	1.5 years

Sources: National Bureau of Economic Research: U.S. Department of Commerce. Bureau of Economic Analysis

We next analyze whether *monetary shocks* can account for the decline and recovery. Some economists, such as Friedman and Schwartz (1963), argue that monetary shocks were a key factor in the 1929–33 decline. To analyze the monetary shock view, we use the well-known model of Lucas and Rapping (1969), which connects changes in the money supply to changes in output through intertemporal substitution of leisure and unexpected changes in wages. The Lucas-Rapping model predicts that monetary shocks reduced output in the early 1930s, but the model also predicts that employment and output should have been back near trend by the mid-1930s.

Both real shocks and monetary shocks predict that employment and output should have quickly returned to trend levels. These predictions are difficult to reconcile with the weak 1934–39 recovery. If the factors considered important in postwar fluctuations can't fully account for macroeconomic performance in the 1930s, are there other factors that can? We go on to analyze two other factors that some economists consider important in understanding the Depression: *financial intermediation shocks* and *inflexible nominal wages.* One type of financial intermediation shock is the bank failures that occurred during the early 1930s. Some researchers argue that these failures reduced output by disrupting financial intermediation. While bank failures perhaps deepened the decline, we argue that their impact would have been short-lived and, consequently, that bank failures were not responsible for the weak recovery. Another type of financial intermediation shock is the increases in reserve requirements that occurred in late 1936 and early 1937. While this change may have led to a small decline in output in 1937, it cannot account for the weak recovery prior to 1937 and cannot account for the significant drop in activity in 1939 relative to 1929.

The other alternative factor is inflexible nominal wages. The view of this factor holds that nominal wages were not as flexible as prices and that the fall in the price level raised real wages and reduced employment. We present data showing that manufacturing real wages rose consistently during the 1930s, but that nonmanufacturing wages fell. The 10-year increase in manufacturing wages is difficult to reconcile with nominal wage inflexibility, which typically assumes that inflexibility is due to either money illusion or explicit nominal contracts. The long duration of the Depression casts doubt on both of these determinants of inflexible nominal wages.

The weak recovery is a puzzle from the perspective of neoclassical growth theory. Our inability to account for the recovery with these shocks suggests to us that an alternative shock is important for understanding macroeconomic performance after 1933. We conclude our study by conjecturing that government policies toward monopoly and the distribution of income are a good candidate for this shock. The National Industrial Recovery Act (NIRA) of 1933 allowed much of the economy to cartelize. This policy change would have depressed employment and output in those sectors covered by the act and, consequently, have led to a weak recovery. Whether the NIRA can quantitatively account for the weak recovery is an open question for future research.

THE DATA THROUGH THE LENS OF THE THEORY

Neoclassical growth theory has two cornerstones: the aggregate production technology, which describes how labor and capital services are combined to create output, and the willingness and ability of households to substitute commodities over time, which govern how households allocate their time between market and nonmarket activities and how households allocate their income between consumption and savings. Viewed through the lens of this theory, the following variables are keys to understanding macroeconomic performance: the allocation of output between consumption and investment, the allocation of time (labor input) between market and nonmarket activities, and productivity.[3]

Output

In Table 2, we compare levels of output during the Depression to peak levels in 1929. To do this, we present data on consumption and investment and the other components of real gross national product (GNP) for the 1929–39 period.[4] Data are from the national income and product accounts published by the Bureau of Economic Analysis of the U.S. Department of Commerce. All data are divided by the *working-age* (16 years and older) population. Since neoclassical growth theory indicates that these variables can be expected to grow, on average, at the trend rate of technology, they are also *detrended,* that is, adjusted for trend growth.[5] With these adjustments, the data can be directly compared to their peak values in 1929.

As we can see in Table 2, all the components of *real output* (GNP in base-year prices), except government purchases of goods and services, fell considerably during the 1930s. The general pattern for the declining series is a very large drop between 1929 and 1933 followed by only a moderate rise from the 1933 trough. Output fell more than 38 percent between 1929 and 1933. By 1939, output remained nearly 27 percent below its 1929 detrended level. This detrended decline of 27 percent consists of a raw 11 percent drop in per capita output and a further 16 percent drop representing trend growth that would have normally occurred over the 1929–39 period.[6]

The largest decline in economic activity occurred in business investment, which fell nearly 80 percent between 1929 and 1933. Consumer durables, which represent household, as opposed to business, investment, followed a similar pattern, declining more than 55 percent between 1929 and 1933. Consumption of nondurables and services declined almost 29 percent between 1929 and 1933. Foreign trade (exports and imports) also fell considerably

Table 2 Detrended Levels of Output and Its Components in 1929–39* (Index, 1929 = 100)

		Consumption				Foreign trade	
Year	Real output	Nondurables and services	Consumer durables	Business investment	Government purchases	Exports	Imports
1930	87.3	90.8	76.2	69.2	105.1	85.2	84.9
1931	78.0	85.2	63.3	46.1	105.3	70.5	72.4
1932	65.1	75.8	46.6	22.2	97.2	54.4	58.0
1933	61.7	71.9	44.4	21.8	91.5	52.7	60.7
1934	64.4	71.9	48.8	27.9	100.8	52.7	58.1
1935	67.9	72.9	68.7	41.7	99.8	53.6	69.1
1936	74.7	76.7	70.5	52.6	113.5	55.0	71.7
1937	75.7	76.9	71.9	59.5	105.8	64.1	78.0
1938	70.2	73.9	56.1	38.6	111.5	62.5	58.3
1939	73.2	74.6	64.0	49.0	112.3	61.4	61.3

* Data are divided by the working-age (16 years and older) population
Source: U.S. Department of Commerce. Bureau of Economic Analysis

Table 3 Changes in the Composition of Output in 1929–39

Year	Shares of Output				
				Foreign trade	
	Government consumption	Investment	Purchases	Exports	Imports
1929	.62	.25	.13	.05	.04
1930	.64	.19	.16	.05	.04
1931	.67	.15	.18	.05	.04
1932	.72	.08	.19	.04	.04
1933	.72	.09	.19	.04	.04
1934	.69	.11	.20	.04	.04
1935	.66	.15	.19	.04	.04
1936	.63	.17	.20	.04	.04
1937	.63	.19	.18	.04	.04
1938	.65	.14	.21	.04	.04
1939	.63	.16	.20	.04	.04
Postwar Average	.59	.20	.23	.06	.07

Source: U.S. Department of Commerce. Bureau of Economic Analysis

between 1929 and 1933. The impact of the decline between 1929 and 1933 on government purchases was relatively mild, and government spending even rose above its trend level in 1930 and 1931.

Table 2 also makes clear that the economy did not recover much from the 1929–33 decline. Although investment improved relative to its 1933 trough level, investment remained 51 percent below its 1929 (detrended) level in 1939. Consumer durables remained 36 percent below their 1929 level in 1939. Relative to trend, consumption of nondurables and services increased very little during the recovery. In 1933, consumption was about 28 percent below its 1929 detrended level. By 1939, consumption remained about 25 percent below this level.

These unique and large changes in economic activity during the Depression also changed the *composition* of output—the shares of output devoted to consumption, investment, government purchases, and exports and imports. These data are presented in Table 3. The share of output consumed rose considerably during the early 1930s, while the share of output invested, including consumer durables, declined sharply, falling from 25 percent in 1929 to just 8 percent in 1932. During the 1934–39 recovery, the share of output devoted to investment averaged about 15 percent, compared to its postwar average of 20 percent. This low rate of investment led to a decline in the *capital stock*—the gross stock of fixed reproducible private capital declined more than 6 percent between 1929 and 1939, representing a decline of more than 25 percent relative to trend. Foreign trade comprised a small share of economic activity in the United States during the 1929–39 period. Both exports and imports accounted for about 4 percent of output during the decade. The increase in government purchases, combined with the decrease in output, increased the government's share of output from 13 percent to about 20 percent by 1939.

These data raise the possibility that the recovery was a weak one. To shed some light on this possibility, in Table 4, we show the recovery from a typical postwar recession. The data in Table 4 are average detrended levels relative to peak measured quarterly from the trough. A comparison of Tables 2 and 4 shows that the recovery from a typical postwar re-

Table 4 Detrended Levels of Output and its Components in a Typical Postwar Recovery: Measured Quarterly From Trough, Peak = 100

Quarters from Trough	Output	Consumption	Investment	Government purchases
0	95.3	96.8	84.5	98.0
1	96.2	98.1	85.2	97.9
2	98.3	99.5	97.3	98.0
3	100.2	100.8	104.5	99.0
4	102.1	102.7	112.1	99.2

Source: U.S. Department of Commerc. Bureau of Economic Analysis

cession differs considerably from the 1934–39 recovery during the Depression. First, output rapidly recovers to trend following a typical postwar recession. Second, consumption grows smoothly following a typical postwar recession. This contrasts sharply to the flat time path of consumption during the 1934–39 recovery. Third, investment recovers very rapidly following a typical postwar recession. Despite falling much more than output during a recession, investment recovers to a level comparable to the output recovery level within three quarters after the trough. During the Depression, however, the recovery in investment was much slower, remaining well below the recovery in output.

Tables 2 and 4 indicate that the 1934–39 recovery was much weaker than the recovery from a typical recession. One interpretation of the weak 1934–39 recovery is that the economy was not returning to its pre-1929 steady-state growth path, but was settling on a considerably lower steady-state growth path.

The possibility that the economy was converging to a lower steady-state growth path is consistent with the fact that consumption fell about 25 percent below trend by 1933 and remained near that level for the rest of the decade. (See Figure 1.) Consumption is a good barometer of a possible change in the economy's steady state because household dynamic optimization implies that all future expectations of income should be factored into current consumption decisions.[7]

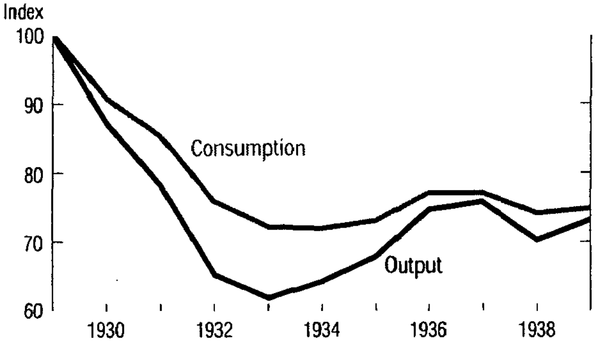


Figure 1 Convergence to a new growth path? Detrended levels of consumption and output in 1929–39. Source: U.S. Department of Commerce, Bureau of Economic Analysis.

Labor Input

Data on labor input are presented in Table 5. We use Kendrick's (1961) data on labor input, capital input, productivity, and output.[8] We present five measures of labor input, each divided by the working-age population. We don't detrend these ratios because theory implies that they will be constant along the steady-state growth path.[9] Here, again, data are expressed relative to their 1929 values.

The three aggregate measures of labor input declined sharply from 1929 to 1933. *Total employment,* which consists of private and government workers, declined about 24 percent between 1929 and 1933 and remained 18 percent below its 1929 level in 1939. *Total hours,* which reflect changes in employment and changes in hours per worker, declined more sharply than total employment, and the trough didn't occur until 1934. Total hours remained 21 percent below their 1929 level in 1939. *Private hours,* which don't include the hours of government workers, declined more sharply than total hours, reflecting the fact that government employment did not fall during the 1930s. Private hours fell more than 25 percent between 1929 and 1939.

These large declines in aggregate labor input reflect different changes across sectors of the economy. *Farm hours* and *manufacturing hours* are shown in the last two columns of Table 5. In addition to being divided by the working-age population, the farm hours measure is adjusted for an annual secular decline in farm employment of about 1.8 percent per year. In contrast to the other measures of labor input, farm hours remained near trend during much of the decade. Farm hours were virtually unchanged between 1929 and 1933, a period in which hours worked in other sectors fell sharply. Farm hours did fall about 10 percent in 1934 and were about 7 percent below their 1929 level by 1939. A very different picture emerges for manufacturing hours, which plummeted more than 40 percent between 1929 and 1933 and remained 22 percent below their detrended 1929 level at the end of the decade.

These data indicate important differences between the farm and manufacturing sectors during the Depression. Why didn't farm hours decline more during the Depression? Why did manufacturing hours decline so much?

Table 5 Five Measures of Labor Input in 1929–39* (Index, 1929 = 100)

	Aggregate Measures			Sectoral Measures	
Year	Total employment	Total hours	Private hours	Farm hours†	Manufacturing hours
1930	93.2	91.9	91.5	99.0	84.6
1931	85.7	83.5	82.8	101.7	68.7
1932	77.5	73.4	72.4	98.7	54.7
1933	76.2	72.6	70.8	99.0	58.4
1934	79.9	71.7	68.7	89.3	61.2
1935	81.4	74.7	71.4	93.3	68.6
1936	83.9	80.6	75.8	91.1	79.2
1937	86.4	83.0	79.5	99.1	85.3
1938	80.4	76.3	71.7	92.7	67.6
1939	82.1	78.7	74.4	93.6	78.0

* Data are divided by this working-age (16 years and older) population. † Farm hours are adjusted for a secular decline in farm employment of about 1.8 percent per year. Source: Kendrick 1961

Finally, note that the changes in nonfarm labor input are similar to changes in consumption during the 1930s. In particular, after falling sharply between 1929 and 1933, measures of labor input remained well below 1929 levels in 1939. Thus, aggregate labor input data also suggest that the economy was settling on a growth path lower than the path the economy was on in 1929.

Productivity

In Table 6, we present two measures of productivity: *labor productivity* (output per hour) and *total factor productivity.* Both measures are detrended and expressed relative to 1929 measures. These two series show similar changes during the 1930s. Labor productivity and total factor productivity both declined sharply in 1932 and 1933, falling about 12 percent and 14 percent, respectively, below their 1929 detrended levels. After 1933, however, both measures rose quickly relative to trend and, in fact, returned to trend by 1936. When we compare 1939 data to 1929 data, we see that the 1930s were a decade of normal productivity growth. Labor productivity grew more than 22 percent between 1929 and 1939, and total factor productivity grew more than 20 percent in the same period. This normal growth in productivity raises an important question about the lack of a recovery in hours worked, consumption, and investment. In the absence of a large negative shift in the long-run path of productivity, why would the economy be on a lower steady-state growth path in 1939?

AN INTERNATIONAL COMPARISON

Many countries suffered economic declines during the 1930s; however, there are two important distinctions between economic activity in the United States and other countries during the 1930s. The decline in the United States was much more severe, and the recovery from the decline was weaker. To see this, we examine average real per capita output relative to its 1929 level for Belgium, Britain, France, Germany, Italy, Japan, and Sweden. The data are from Maddison 1991 and are normalized for each country so that per capita output

Table 6 Detrended Measures of Productivity (Index, 1979 = 100)

Year	Labor productivity*	Total factor productivity
1930	95.9	94.8
1931	95.4	93.5
1932	90.7	87.8
1933	87.9	85.9
1934	96.7	92.6
1935	98.4	96.6
1936	101.6	99.9
1937	100.7	100.5
1938	102.4	100.3
1939	104.6	103.1

* *Labor productivity* is defined as output per hour.

Sources: Kendrick 1961: U.S. Department of Commerce. Bureau of Economic Analysis

Table 7 U.S. vs. International Decline and Recovery: Annual Real per Capita Output in the 1930s (Index, 1929 = 100)

Year	United States	International average*
1932	69.0	91.3
1933	66.7	94.5
1935	76.3	101.0
1938	83.6	112.4

* International average includes Belgium, Britain, France, Germany, Italy, Japan, and Sweden. Source: Maddison 1991

is equal to 100 in 1929. Since there is some debate over the long-run growth rate in some of these countries, we have not detrended the data.

Table 7 shows the U.S. data and the mean of the normalized data for other countries. The total drop in output is relatively small in other countries: an 8.7 percent drop compared to a 33.3 percent drop in the United States. The international economies recovered quickly: output in most countries returned to 1929 levels by 1935 and was above those levels by 1938. Employment also generally recovered to its 1929 level by 1938.[10]

While accounting for other countries' economic declines is beyond the scope of this analysis, we can draw two conclusions from this comparison. First, the larger decline in the United States is consistent with the view that the shocks that caused the decline in the United States were larger than the shocks that caused the decline in the other countries. Second, the weak recovery in the United States is consistent with the view that the shocks that impeded the U.S. recovery did not affect most other countries. Instead, the post-1933 shock seems to be largely specific to the United States.

The data we've examined so far suggest that inputs and output in the United States fell considerably during the 1930s and did not recover much relative to the increase in productivity. Moreover, the data show that the decline was much more severe and the recovery weaker in the United States than in other countries. To account for the decade-long Depression in the United States, we conclude that we should focus on domestic, rather than international, factors. We turn to this task in the next section.

CAN REAL SHOCKS ACCOUNT FOR THE DEPRESSION?

Neoclassical theory and the data have implications for the plausibility of different sources of real shocks in accounting for the Depression. Since the decline in output was so large and persistent, we will look for large and persistent negative shocks. We analyze three classes of real shocks considered important in typical business cycle fluctuations: technology shocks, fiscal policy shocks, and trade shocks.

Technology Shocks? Perhaps Initially

First we consider *technology shocks,* defined as any exogenous factor that changes the efficiency with which business enterprises transform inputs into output. Under this broad definition of technology shocks, changes in productivity reflect not only true changes in technology, but also such other factors as changes in work rules and practices or government regulations that affect the efficiency of production but are exogenous from the perspective of business enterprises. How do technology shocks affect economic activity? The key element that leads to a decline in economic activity in models with technology shocks is a negative shock that reduces the marginal products of capital and labor. Shocks that reduce the efficiency of transforming inputs into output lead households to substitute out of market activities into nonmarket activities and result in lower output. Recent research has identified these shocks as important factors in postwar business cycle fluctuations. Prescott (1986), for example, shows that a standard one-sector neoclassical model with a plausibly parameterized stochastic process for technology shocks can account for 70 percent of postwar business cycle fluctuations. Can technology shocks account for the Depression?

If these shocks were responsible, we should see a large and persistent drop in *technology*—the efficiency of transforming inputs into output—during the 1930s. To see if such a drop occurred, we first need a measure of technology for this period. Under the neoclassical assumptions of constant returns to scale in production and perfectly competitive markets, theory implies that changes in total factor productivity are measures of changes in technology. The data do show a drop in total factor productivity—a 14 percent (detrended) drop between 1929 and 1933 followed by a rapid recovery. What is the quantitative importance of these changes in accounting for the Depression?

To address this question, we present the prediction for output for 1930–39 from a real business cycle model. (See Hansen 1985, Prescott 1986, or King, Plosser, and Rebelo 1988 for a discussion of this model.) Our model consists of equations (A1)–(A5) and (A9) in the Appendix, along with the following preference specification:

$$u(c_t, l_t) = \log(c_t) + A \log(l_t). \tag{1}$$

We use the following Cobb-Douglas production function specification:

$$z_t f(k_t, n_t) = z_t k_t^{\theta}(x_t n_t)^{1-\theta}. \tag{2}$$

The household has one unit of time available each period:

$$1 = l_t + n_t. \tag{3}$$

And we use the following specification of the stochastic process for the technology shock:

$$z_t = (1 - \rho) + \rho z_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2). \tag{4}$$

With values for the parameters of the model, we can use numerical methods to compute an approximate solution to the equilibrium of this economy.[11] We set $\theta = 0.33$ to conform to the observation that capital income is about one-third of output. We set $\sigma = 1.7$ percent and $\rho = 0.9$ to conform to the observed standard deviation and serial correlation of total factor productivity. We choose the value for the parameter A so that households spend about one-third of their discretionary time working in the deterministic steady state. Labor-augmenting technological change (x_t) grows at a rate of 1.9 percent per year. The population (n_t) grows at a rate of 1 percent per year. We set the depreciation rate at 10 percent per year.

We conduct the analysis by assuming that the capital stock in 1929 is equal to its steady-state value, and then we feed in the sequence of observed levels of total factor productivity as measures of the technology shock. Given the initial condition and the time path of technology, the model predicts labor input, output, consumption, and investment for each year during the 1930s. We summarize the results of the analysis in Figure 2, where we plot the detrended predicted level of output from the model between 1929 and 1939. For

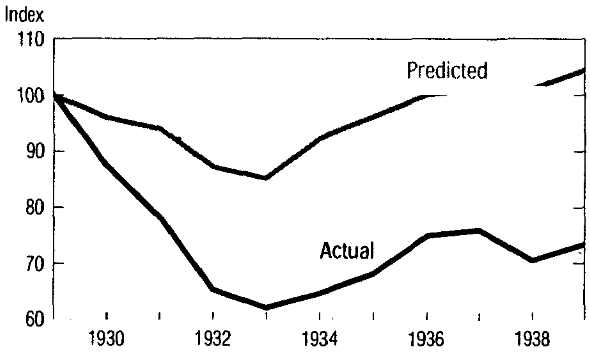


Figure 2 Predicted and actual output in 1929–39: Detrended levels, with initial capital stock in the model equal to the actual capital stock in 1929. Source of basic data: U.S. Department of Commerce, Bureau of Economic Analysis.

comparison, we also plot the actual detrended level of output. Note that the model predicts a significant decline in output between 1929 and 1933, although the decline is not as large as the observed decline in the data: a 15 percent predicted decline compared to a 38 percent actual decline. Further, note that as a consequence of rapid growth in total factor productivity after 1934, the model predicts a rapid recovery: output should have returned to trend by 1936. In contrast, actual output remained about 25 percent below trend during the recovery.

One factor that may be contributing to the rapid recovery in the model is the fact that the capital stock in the model falls less than in the data. Consequently, output predicted by the model may be relatively high because the capital stock is high. To correct for this difference, we conduct another analysis in which we also feed in the sequence of total factor productivity measures between 1934 and 1939, but we use the actual capital stock in 1934 (20 percent below trend) as the initial condition for 1934. Figure 3 shows that this change reduces output predicted by the model by about 3 percent at the beginning of the recovery. But because the initial capital stock in this analysis is lower, the marginal product of capi-

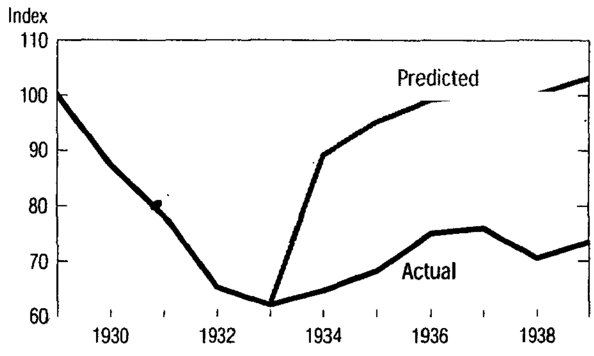


Figure 3 Predicted and actual recovery of output in 1934–39: Detrended levels, with initial capital stock in the model equal to the actual capital stock in 1934. Source of basic data: U.S. Department of Commerce, Bureau of Economic Analysis.

tal is higher, and the predicted rate of output growth in the recovery is faster than in the first analysis. This recovery brings output back to its trend level by 1937. The predicted output level is about 27 percent above the actual data level in 1939.[12] Thus, the predicted recovery is stronger than the actual recovery because predicted labor input is much higher than actual labor input.

Based on measured total factor productivity during the Depression, our analysis suggests a mixed assessment of the technology shock view. On the negative side, the actual slow recovery after 1933 is at variance with the rapid recovery predicted by the theory. Thus, it appears that some shock other than to the efficiency of production is important for understanding the weak recovery between 1934 and 1939. On the positive side, however, the theory predicts that the measured drop in total factor productivity can account for about 40 percent of the decline in output between 1929 and 1933.

Note, however, one caveat in using total factor productivity as a measure of technology shocks during periods of sharp changes in output, such as the 1929–33 decline: An imperfect measurement of capital input can affect measured aggregate total factor productivity. Because total factor productivity *change* is defined as the percentage change in output minus the percentage change in inputs, overstating the inputs will understate productivity, while understanding the inputs will overstate productivity. During the 1929–33 decline, some capital was left idle. The standard measure of capital input is the capital stock. Because this standard measure includes idle capital, it is possible that capital input was overstated during the decline and, consequently, that productivity growth was understated.[13] Although there are no widely accepted measures of capital input adjusted for changes in utilization, this caveat raises the possibility that the decline in aggregate total factor productivity in the early 1930s partially reflects mismeasurement of capital input.[14] Without better data on capital input or an explicit theoretical framework we can use to adjust observed measured total factor productivity fluctuations for capital utilization, we can't easily measure how large technology shocks were in the early 1930s and, consequently, how much of a drop in output technology shocks can account for.

It is important to note here that these results give us an important gauge not only for the technology shock view, but also for any other shock which ceased to be operative after 1933. The predicted rapid recovery in the second experiment implies that any shock which ceased to be operative after 1933 can't easily account for the weak recovery.

Fiscal Policy Shocks? A Little

Next we consider *fiscal policy shocks*—changes in government purchases or tax rates. Christiano and Eichenbaum (1992) argue that government purchase shocks are important in understanding postwar business cycle fluctuations, and Braun (1994) and McGrattan (1994) argue that shocks to distorting taxes have had significant effects on postwar cyclical activity.

To understand how government purchases affect economic activity, consider an unexpected decrease in government purchases. This decrease will tend to increase private consumption and, consequently, lower the marginal rate of substitution between consumption and leisure. Theory predicts that this will lead households to work less and take more leisure. Conversely, consider an increase in government purchases. This increase will tend to decrease private consumption and reduce the marginal rate of substitution between consumption and leisure. In this case, theory predicts that this will lead households to work more and take less leisure.

Historically, changes in government purchases have had large effects on economic activity. Ohanian (1997) shows that the increase in government purchases during World War II can account for much of the 60 percent increase in output during the 1940s. Can changes in government purchases also account for the decrease in output in the 1930s?

If government purchase shocks were a key factor in the decline in employment and output in the 1930s, government purchases should have declined considerably during the period. This did not occur. Government purchases declined modestly between 1929 and 1933 and then rose sharply during the rest of the decade, rising about 12 percent above trend by 1939. These data are inconsistent with the view that government purchase shocks were responsible for the downturn.[15]

Although changes in government purchases are not important in accounting for the Depression, the way they were financed may be. Government purchases are largely financed by *distorting taxes*—taxes that affect the marginal conditions of households or firms. Most government revenue is raised by taxing factor incomes. Changes in factor income taxes change the net rental price of the factor. Increases in labor and capital income taxes reduce the returns to these factors and, thus, can lead households to substitute out of taxed activities by working and saving less.

If changes in factor income taxes were a key factor in the 1930s economy, these rates should have increased considerably in the 1930s. Tax rates on both labor and capital changed very little during the 1929–33 decline, but rose during the rest of the decade. Joines (1981) calculates that between 1929 and 1939, the average marginal tax rate on labor income increased from 3.5 percent to 8.3 percent and the average marginal tax rate on capital income increased from 29.5 percent to 42.5 percent. How much should these increases have depressed economic activity? To answer this question, we consider a deterministic version of the model we used earlier to analyze the importance of technology shocks. We augment this model to allow for distortionary taxes on labor and capital income. The values of the other parameters are the same. We then compare the deterministic steady state of the model with 1939 tax rates to the deterministic steady state of the model with 1929 tax rates. With these differences in tax rates, we find that steady-state labor input falls by 4 percent. This suggests that fiscal policy shocks account for only about 20 percent of the weak 1934–39 recovery.

Trade Shocks? No

Finally, we consider *trade shocks*. In the late 1920s and early 1930s, *tariffs*—domestic taxes on foreign goods—rose in the United States and in other countries. Tariffs raise the domestic price of foreign goods and, consequently, benefit domestic producers of goods that are substitutes with the taxed foreign goods. Theory predicts that increases in tariffs lead to a decline in world trade. International trade did, indeed, fall considerably during the 1930s: the League of Nations (1933) reports that world trade fell about 65 percent between 1929 and 1932. Were these tariff increases responsible for the 1929–33 decline?

To address this question, we first study how a contraction of international trade can lead to a decline in output. In the United States, trade is a small fraction of output and is roughly balanced between exports and imports. Lucas (1994) argues that a country with a small trade share will not be affected much by changes in trade. Based on the small share of trade at the time, Lucas (1994, p. 13) argues that the quantitative effects of the world trade contraction during the 1930s are likely to have been "trivial."[16]

Can trade have an important effect even if the trade share is small? Crucini and Kahn

(1996) argue that a significant fraction of imports during the 1930s were intermediate inputs. If imported intermediate inputs are imperfect substitutes with domestic intermediate inputs, production can fall as a result of a reduction in imported inputs. Quantitatively, the magnitude of the fall is determined by the elasticity of substitution between the inputs. If the goods are poor substitutes, then a reduction in trade can have sizable effects. Little information is available regarding the substitution elasticity between these goods during the Depression. The preferred estimates of this elasticity in the postwar United States are between one and two. (See Stern, Francis, and Schumacher 1976.) Crucini and Kahn (1996) assume an elasticity of two-thirds and report that output would have dropped about 2 percent during the early 1930s as a result of higher tariffs.

This small decline implies that extremely low substitution elasticities are required if the trade disruption is to account for more than a small fraction of the decline in output. How plausible are very low elasticities? The fact that tariffs were widely used points to high, rather than low, elasticities between inputs. To see this, note that with high elasticities, domestic and foreign goods are very good substitutes, and, consequently, tariffs should benefit domestic producers who compete with foreign producers. With very low elasticities, however, domestic goods and foreign goods are poor substitutes. In this case, tariffs provide little benefit to domestic producers and, in fact, can even hurt domestic producers if there are sufficient complementarities between inputs. This suggests that tariffs would not be used much if substitution elasticities were very low.

But even if substitution elasticities were low, it is unlikely that this factor was responsible for the Depression, because the rise in the prices of tariffed goods would ultimately have led domestic producers to begin producing the imported inputs. Once these inputs became available domestically, the decline in output created by the tariff would have been reversed. It is hard to see how the disruption of trade could have affected output significantly for more than the presumably short period it would have taken domestic producers to change their production.

Our analysis thus far suggests that none of the real shocks usually considered important in understanding business cycle fluctuations can account for macroeconomic performance during the 1930s. Lacking an understanding of the Depression based on real shocks, we next examine the effects of monetary shocks from the neoclassical perspective.

CAN MONETARY SHOCKS ACCOUNT FOR THE DEPRESSION?

Monetary shocks—unexpected changes in the stock of money—are considered an alternative to real shocks for understanding business cycles, and many economists think monetary shocks were a key factor in the 1929–33 decline. Much of the attraction to monetary shocks as a source of business cycles comes from the influential narrative monetary history of the United States by Friedman and Schwartz (1963). They present evidence that declines in the money supply tend to precede declines in output over nearly a century in the United States. They also show that the money supply fell sharply during the 1929–33 decline. Friedman and Schwartz (1963, pp. 300–301) conclude from these data that the decline in the money supply during the 1930s was an important cause of the 1929–33 decline (contraction):

> The contraction is in fact a tragic testimonial to the importance of monetary forces. . . . Prevention or moderation of the decline in the stock of money, let alone the substitution of monetary expansion, would have reduced the contraction's severity and almost as certainly its duration.

Maybe for the Decline . . .

We begin our discussion of the monetary shock view of the decline by presenting data on some nominal and real variables. We present the data Friedman and Schwartz (1963) focus on: money, prices, and output. We also present data on interest rates.

In Table 8, we present the *nominal data:* the monetary base, which is the monetary aggregate controlled by the Federal Reserve; M1, which is currency plus checkable deposits; the GNP deflator, or price level; and two interest rates: the rate on three-month U.S. Treasury bills and the rate on commercial paper. The money supply data are expressed in per capita terms by dividing by the working-age population. The money data are also expressed relative to their 1929 values. The interest rates are the annual average percentage rates. These nominal data do, indeed, show the large decline in M1 in the early 1930s that led Friedman and Schwartz (1963) to conclude that the drop in the money supply was an important cause of the 1929–33 decline.[17]

In Table 9, we present the *real data:* the real money supply, which is the two nominal series divided by the GNP deflator; real output; and the ex post real rate of interest, which is the commercial paper rate minus the realized inflation rate. Note that the real money stock fell considerably less than the nominal stock during the early 1930s and then rose between 1933 and 1939. In fact, the variation in the real money stock during the decline is quite similar to the variation in real output.

To understand the empirical relationship between money and output reported by Friedman and Schwartz (1963), economists have developed theoretical models of monetary business cycles. In these models, money is *nonneutral*—changes in the money supply lead to changes in allocations and relative prices. For money to have important nonneutralities, there must be some mechanism that prevents nominal prices from adjusting fully to a change in the money supply. The challenge of monetary business cycle theory is to generate important nonneutralities not by assumption, but as an equilibrium outcome.

The first monetary business cycle model along these lines was developed by Lucas and Rapping (1969). This model was later extended into a fully articulated general equilibrium model by Lucas (1972). Two elements in the Lucas-Rapping model generate

Table 8 Nominal Money, Prices, and Interest Rates in 1929–39

				Annual % interest rate	
Year	Monetary base*	M†*	Price level	3-month U.S. T-bill	Commercial paper
1929	100.0	100.0	100.0	4.4%	6.1%
1930	95.9	94.4	97.0	2.2	4.3
1931	98.7	85.6	88.1	1.2	2.6
1932	104.3	74.5	78.4	.8	2.7
1933	108.9	69.9	76.7	.3	1.7
1934	119.8	78.0	83.2	.3	2.0
1935	139.2	91.0	84.8	.2	.8
1936	157.2	102.1	85.2	.1	.8
1937	168.5	102.9	89.4	.5	.9
1938	181.5	102.2	87.2	.1	.8
1939	215.5	113.7	86.6	.0	.6

* Money measures are divided by the working-age (16 years and older) population. Source: Board of Governors of the Federal Reserve System

Table 9 Real Money, Output, and Interest Rates in 1929–39

Year	Monetary base*	M†*	Output	Interest rate†
1929	100.0	100.0	100.0	6.0%
1930	98.8	97.3	87.3	7.3
1931	112.0	97.1	78.0	11.8
1932	133.1	95.1	65.1	13.8
1933	142.1	91.2	61.7	3.9
1934	144.0	93.8	64.4	–6.5
1935	164.1	107.3	67.9	–1.1
1936	184.4	119.8	74.7	.3
1937	188.6	115.2	75.7	–3.9
1938	208.1	117.2	70.2	3.2
1939	248.7	131.2	73.2	1.3

* Money measures are divided by the working-age (16 years and older) population.
† This is the interest rate on commercial paper minus the realized inflation rate.
Sources: Board of Governors of the Federal Reserve System U.S. Department of Commerce. Bureau of Economic Analysis

cyclical fluctuations: intertemporal substitution of leisure and unexpected changes in wages. The basic idea in the Lucas-Rapping model is that agents' decisions are based on the realization of the real wage relative to its normal, or expected, level. Suppose that the wage turns out to be temporarily high today relative to its expected level. Since the wage is high, the opportunity cost of not working—*leisure*—is also high. If preferences are such that leisure today is substitutable with leisure in the future, households will respond by intertemporally substituting leisure today for future leisure and, thus, will work more today to take advantage of the temporarily high wage. Similarly, if the wage today is temporarily low relative to the normal wage, households will tend to take more leisure today and less leisure in the future when wages return to normal.

How does the money supply in the 1929–33 decline figure into this model? Lucas and Rapping (1969) model households' expectation of the real wage as a weighted average of the real wage's past values. Based on this construction of the weighted average, the rapid decline in the money supply resulted in the real wage falling below its expected level, beginning in 1930. According to the model, the decline in the real wage relative to the expected wage led households to work less, which reduced output.

. . . But Not for the Recovery

Quantitatively, Lucas and Rapping (1969) find that the decline in the real wage relative to the expected wage was important in the 1929–33 decline. The Lucas-Rapping model predicts a large decline in labor input through 1933. The problem for the Lucas-Rapping model is what happened after 1933. The real wage returned to its expected level in 1934, and for the rest of the decade, the wage was either equal to or above its expected level. According to the model, this should have resulted in a recovery that quickly returned output to its 1929 (detrended) level. This did not happen. (See Lucas and Rapping 1972.) The Lucas-Rapping (1969) model can't account for the weak recovery.

Another model that connects changes in money to changes in output is Fisher's (1933) debt-deflation model. In this model, deflation shifts wealth from debtors to creditors by increasing the real value of nominal liabilities. In addition to making this wealth transfer, the increase in the real value of liabilities reduces net worth and, according to Fisher, leads to lower lending and a higher rate of business failures. Qualitatively, Fisher's view matches up with the 1929–32 period, in which both nominal prices and output were falling. The quantitative importance of the debt-deflation mechanism for this period, however, is an open question. Of course, Fisher's model would tend to predict a rapid recovery in economic activity once nominal prices stopped falling in 1933. Thus, Fisher's model can't account for the weak recovery either.[18]

ALTERNATIVE FACTORS

Factors other than those considered important in postwar business cycles have been cited as important contributors to the 1929–33 decline. Do any provide a satisfactory accounting for the Depression from the perspective of neoclassical theory? We examine two widely cited factors: financial intermediation shocks and inflexible nominal wages.

Were Financial Intermediation Shocks Important?

Bank Failures? Maybe, But Only Briefly

Several economists have argued that the large number of bank failures that occurred in the early 1930s disrupted financial intermediation and that this disruption was a key factor in the decline. Bernanke's (1983) work provides empirical support for this argument. He constructs a statistical model, based on Lucas and Rapping's (1969) model, in which unexpected changes in the money stock lead to changes in output. Bernanke estimates the parameters of his model using least squares, and he shows that adding the dollar value of deposits and liabilities of failing banks as explanatory variables significantly increases the fraction of output variation the model can account for.

What economic mechanism might have led bank failures to deepen the 1929–33 decline? One view is that these failures represented a decline in information capital associated with specific relationships between borrowers and intermediaries. Consequently, when a bank failed, this relationship-specific capital was lost, and the efficiency of intermediation declined.

It is difficult to assess the quantitative importance of bank failures as a factor in deepening the 1929–33 decline because the output of the banking sector, like broader measures of economic activity, is an endogenous, not an exogenous, variable. Although bank failures may have exacerbated the decline, as suggested by Bernanke's (1983) empirical work, some of the decline in the inputs and output of the banking sector may also have been an endogenous response to the overall decline in economic activity.[19] Moreover, bank failures were common in the United States during the 1920s, and most of those bank failures did not seem to have important aggregate consequences. Wicker (1980) and White (1984) argue that at least some of the failures during the early 1930s were similar to those during the 1920s.

However, we can assess the potential contribution of intermediation shocks to the 1929–33 decline with the following growth accounting exercise. We can easily show that under the assumption of perfect competition, at least locally, the percentage change in aggregate output, $\hat{Y}$, can be written as a linear function of the percentage change in the sector i outputs, $\hat{y}_i$, for each sector $i = 1, \ldots, n$ and the shares γ_i for each sector as follows:

$$\hat{Y} = \sum_{i=1}^{n} \gamma_i \hat{y}_i. \tag{5}$$

The share of the entire finance, insurance, and real estate (FIRE) sector went from 13 percent in 1929 to 11 percent in 1933. This suggests that the appropriate cost share was 12 percent. The real output of the FIRE sector dropped 39 percent between 1929 and 1933. If we interpret this fall as exogenous, we see that the drop in the entire FIRE sector reduces output by 4.7 percent. Thus, in the absence of large aggregate externalities that would amplify this effect, the contribution of the FIRE sector was small.[20]

To better understand the importance of bank failures, especially for the recovery, we next examine data on financial intermediation during the Depression to determine how the capacity of the banking sector changed as a result of exiting institutions; how the quantity of one productive input into the banking sector, deposits, changed; and how the portfolios of banks changed.

In Table 10, we present data on deposits in operating banks, deposits in suspended banks, the stock of total commercial loans, and federal government securities held by banks. All data are measured relative to nominal output. To measure the flow change in loans, we also present the percentage change in the ratio of loans to output. We note four interesting features of these data.

The decline in deposits during the 1929–33 decline was small relative to the decline in output. The ratio of deposits of operating banks to output rose from 0.57 in 1929 to 0.77 in 1932.

Deposits of suspended institutions were less than 2 percent of deposits of operating banks in every year of the decline except 1933, when the president declared a national bank holiday. Moreover, failures disappeared after 1933, reflecting the introduction of federal deposit insurance.

Loans as a fraction of output did not begin to drop much until 1933, but dropped sharply during the 1934–39 recovery.

Table 10 Bank Assets and Liabilities Relative to Nominal Output in 1929–39

	Deposits				
Year	Operating banks	Suspended banks	Loans	% Change in loans	Federal securities
1929	.57	.00	.41	–6%	.05
1930	.64	.01	.42	3	.06
1931	.62	.02	.48	13	.09
1932	.77	.01	.45	6	.12
1933	.69	.06	.40	–13	.15
1934	.74	.00	.31	–23	.17
1935	.73	.00	.28	–11	.20
1936	.71	.00	.26	–9	.21
1937	.65	.00	.24	–6	.19
1938	.71	.00	.25	3	.20
1939	.73	.00	.25	–2	.21

Source: Board of Governors of the Federal Reserve System

The fraction of federal government securities held by banks as a fraction of output increased steadily during the Depression, rising from 0.05 in 1929 to 0.20 by 1935.

The data in the first two rows of Table 10 suggest that funds available for loans were relatively high during the Depression and that the overall capacity of the banking sector, measured in terms of deposits lost in exiting institutions, did not change much. Why, then, did banks not make more loans during the Depression? Was it because a loss of information capital associated with exiting banks caused a reduction in the efficiency of intermediation? Unfortunately, we can't measure this information capital directly. We can, however, assess this possibility with a very simple model of intermediation, in which loans made at bank i, l_i, and intermediated government debt held by bank i, b_i, are produced from a constant returns to scale technology using deposits, d_i, and exogenous information capital, x_i, such that $l_i + b_i = f(d_i, x_i)$. The total stock of information capital is the sum of information capital across all banks, and the information capital of any bank that exits is destroyed. With competition, the ratio of productive inputs, d_i/x_i, will be identical across banks. This implies that the fraction of information capital in banking lost due to exiting banks is equal to the fraction of deposits lost in exiting banks. Theory thus suggests that, except during 1933, the loss of information capital as a direct result of exiting banks was low during the Depression.[21]

There are other channels, however, through which bank failures could have had important aggregate affects. For example, failures caused by bank runs may have led solvent banks to fear runs and, therefore, shift their portfolios from illiquid loans to liquid government bonds. However, this shift doesn't explain the low level of loans relative to output that persisted during the 1934–39 recovery. Moreover, during the recovery, federal deposit insurance eliminated bank runs. Why would banks still fear runs years later?

This analysis raises some questions about the view that bank runs had very large effects during the 1929–33 decline. It also shows that there is little evidence to support the view that the intermediation shock associated with these bank runs had persistent effects which slowed the recovery after 1933. We next turn to the other intermediation shock that some researchers argue is important for understanding the weak recovery.

Reserve Requirements? Not Much

In August 1936, the Federal Reserve increased the required fraction of net deposits that member banks must hold as reserves from 10 percent to 15 percent. This fraction rose to 17.5 percent in March 1937 and then rose to 20 percent in May 1937. Many economists, for example, Friedman and Schwartz, attribute some of the weak macroeconomic performance during 1937 and 1938 to these policy changes.

These economists argue that these policy changes increased bank reserves, which reduced lending and, consequently, reduced output. If this were true, we would expect to see output fall shortly after these changes. This did not happen. Between August 1936, when the first increase took place, and August 1937, industrial production rose about 12 percent. It is worth noting that industrial production did fall considerably between late 1937 and 1938, but the downturn did not begin until October 1937, which is 14 months after the first and largest increase in reserve requirements. (Industrial production data are from the October 1943 *Federal Reserve Index of Industrial Production* of the Board of Governors of the Federal Reserve System.)

Another potential shortcoming of the reserve requirement view is that interest rates did not rise after these policy changes. Commercial loan rates fell from 2.74 percent in January 1936 to 2.65 percent in August 1936. These rates then fell to 2.57 percent in March 1937 and rose slightly to 2.64 percent in May 1937, the date of the last increase in reserve requirements. Lending rates then ranged between 2.48 percent and 2.60 percent over the rest of 1937 and through 1938. Interest rates on other securities showed similar patterns: rates on Aaa-, Aa-, and A-rated corporate debt were roughly unchanged between 1936 and 1938.[22] (Interest rate data are from *Banking and Monetary Statistics, 1914–1941* of the Board of Governors of the Federal Reserve System.) These data raise questions about the view that higher reserve requirements had important macroeconomic effects in the late 1930s and instead suggest that some other factor was responsible for the weak 1934–39 recovery.

Were Inflexible Nominal Wages Important? Hard to Know

The other alternative factor cited as contributing to the Depression is inflexible nominal wages. This view dates back to Keynes 1935 and more recently to Bernanke and Carey 1996 and Bordo, Erceg, and Evans 1996. The basic idea behind this view is that nominal wages are inflexible—a decline in the money supply lowers the price level but does not lower the nominal wage. This inflexibility suggests that a decline in the price level raises the real wage and, consequently, reduces labor input. Were inflexible nominal wages a key factor in the Depression?

To address this question, in Table 11, we present data on real wages in manufacturing, nonmanufacturing, and the total economy. The data for the manufacturing sector, from Hanes 1996, are divided by the GNP deflator, adjusted for long-run real wage growth of 1.9 percent per year, and measured relative to 1929. The wage rate for the total economy is constructed as real total compensation of employees divided by total hours worked. The total economy rate is also adjusted for long-run real wage growth and measured relative to 1929.

We use the data for the manufacturing wage, the constructed total economy wage, and the employment shares for manufacturing and nonmanufacturing to construct the wage rate for the nonmanufacturing sector. The percentage change in the total wage ($\%\Delta w^{tot}$) between dates t and t–1 is equal to the sum of the percentage change in the manufacturing

Table 11 Detrended Real Wage Rates in 1929–39 (Index, 1929 = 100)

Year	Manufacturing	Total economy	Nonmanufacturing
1930	101.6	99.1	97.6
1931	105.7	98.6	94.5
1932	105.0	97.0	92.6
1933	102.3	91.0	85.2
1934	108.5	95.5	88.1
1935	108.0	94.8	86.9
1936	106.9	97.3	91.4
1937	112.6	97.6	87.9
1938	117.0	98.9	86.9
1939	116.1	99.9	90.2

Source of basic data: Hanes 1996, U.S. Department of Commerce. Bureau of Economic Analysis wage ($\%\Delta w^{mfg}$) weighted by its share of employment (*shm*) at date t–1 and the percentage change in the nonmanufacturing sector weighted by its share of employment at date t–1. Thus, the percentage change in the nonmanufacturing wage ($\%\Delta w^{nonmfg}$) is given by

$$\%\Delta w^{nonmfg} = [\%\Delta w^{tot} - shm_{t-1}(\%\Delta w^{mfg})]/(1 - shm_{t-1}) \quad (6)$$

The economywide real wage was roughly unchanged during 1930 and 1931, and fell 9 percent by 1933. This aggregate measure, however, masks striking differences between the manufacturing and nonmanufacturing sectors. The nonmanufacturing wage fell almost 15 percent between 1929 and 1933 and remained almost 10 percent below trend in 1939. This decline was not unusual: postwar data indicate that real wages are moderately procyclical, which suggests that the large drop in output during the 1929–33 decline would likely have been accompanied by a considerable drop in the real wage.[23]

In contrast, real wages in manufacturing rose above trend during the 1929–33 decline and continued to rise during the rest of the decade. By 1939, manufacturing wages were 16 percent above trend. These data raise questions about the manufacturing sector during the Depression. Why did real wages in manufacturing rise so much during a decade of poor economic performance? Why was the increase only in manufacturing? It seems unlikely that the standard reasons for nominal wage inflexibility—money illusion and explicit nominal contracts—were responsible for the decade-long increase in the manufacturing real wage.[24]

We conclude that neither alternative factor, intermediation shocks or inflexible nominal wages, sheds much light on the weak 1934–39 recovery.[25]

A POSSIBLE SOLUTION

Neoclassical theory indicates that the Depression—particularly the recovery between 1934 and 1939—is a puzzle. The conventional shocks considered important in postwar business cycles do not account for the decade-long drop in employment and output. The conventional shocks are too small. Moreover, the effects of monetary shocks are too transient. Nor does expanding our analysis to consider alternative factors account for the Depression. The effects of alternative factors either are too transient or lack a sufficient theoretical framework.

Where do we go from here? To make progress in understanding the Depression, we identify the observations that are puzzling from the neoclassical perspective and then determine which direction these puzzles point us in. Our analysis identifies three puzzles in particular: Why did labor input, consumption, and investment remain so low during a period of rapid productivity growth? Why did agricultural employment and output remain near trend levels during the early 1930s, while nonagricultural employment and output plummeted? Why did the manufacturing real wage increase so much during the 1930s? With competitive markets, theory suggests that the real wage should have decreased, rather than increased.

These puzzles suggest that some other shocks were preventing a normal recovery. We uncover three clues that may aid in future hunts for the shocks that account for the weak 1934–39 recovery. First, it seems that we can rule out shocks that hit all sectors of the economy proportionately. During the 1929–33 decline, for example, agricultural employment and output fell very little, while manufacturing output and employment fell substantially. Second, our view that the economy was settling on a new, much lower growth path during the 1930s indicates that the shocks responsible for the decline were perceived by households and businesses to be permanent, rather than temporary. Third, some of the puzzles may be related—the fact that investment remained so low may reflect the fact that the capital stock was adjusting to a new, lower steady-state growth path.

To account for the weak recovery, these clues suggest that we look for shocks with specific characteristics, for example, a large shock which hits just some sectors of the economy, in particular, manufacturing, and which causes wages to rise and employment and investment to fall in those sectors. We conjecture that government policies toward monopoly and the distribution of income are a good candidate for this type of shock.

Government policies toward monopoly changed considerably in the 1930s. In particular, the NIRA of 1933 allowed much of the U.S. economy to cartelize. For over 500 sectors, including manufacturing, antitrust law was suspended and incumbent business leaders, in conjunction with government and labor representatives in each sector, drew up codes of fair competition. Many of these codes provided for minimum prices, output quotas, and open price systems in which all firms had to report current prices to the code authority and any price cut had to be filed in advance with the authority, who then notified other producers. Firms that attempted to cut prices were pressured by other industry members and publicly berated by the head of the NIRA as "cut-throat chiselers." In return for government-sanctioned collusion, firms gave incumbent workers large pay increases.

How might this policy change have affected the economy? By permitting monopoly and raising wages, the NIRA would be expected to have depressed employment, output, and investment in the sectors the act covered, including manufacturing. In contrast, economic activity in the sectors not covered by the act, such as agriculture, would probably not have declined as much. Qualitatively, this intuition suggests that this government policy shock has the right characteristics. The key issue, however, is the quantitative impact of the NIRA on the macroeconomy: How much did it change employment, investment, consumption, output, and wages? How did the impact differ across sectors of the economy? Addressing these questions is the focus of our current research.

ACKNOWLEDGMENTS

The authors acknowledge the tremendous contribution Edward Prescott made to this project in the many hours he spent talking with them about the Depression and in the input and guidance he generously provided. The authors also thank Andy Atkeson, Russell Cooper, Ed Green, Chris Hanes, Patrick Kehoe, Narayana Kocherlakota, Art Rolnick, and Jim Schmitz for comments. The authors also thank Jesús Fernández-Villaverde for research assistance and Jenni Schoppers for editorial assistance; both contributed well beyond the call of duty.

NOTES

1. For other studies of the Depression and many additional references, see Brunner 1981; Temin 1989, 1993; Eichengreen 1992; Calomiris 1993; Margo 1993; Romer 1993; Bernanke 1995; Bordo, Erceg, and Evans 1996; and Crucini and Kahn 1996.
2. The National Bureau of Economic Research (NBER) defines a *cyclical decline,* or *recession,* as a period of decline in output across many sectors of the economy which typically lasts at least

six months. Since the NBER uses a monthly frequency, we convert to a quarterly frequency for our comparison by considering a peak (trough) quarter to be the quarter with the highest (lowest) level of output within one quarter of the quarter that contains the month of the NBER peak (trough). We define the *recovery* as the time it takes for output to return to its previous peak.

3. Note that in the closed economy framework of the neoclassical growth model, savings equals investment.
4. We end our analysis in 1939 to avoid the effects of World War II.
5. We make the trend adjustment by dividing each variable by its long-run trend growth rate relative to the reference date. For example, we divide GNP in 1930 by 1.019. This number is 1 plus the average growth rate of 1.9 percent over the 1947–97 period and over the 1919–29 period. For 1931, we divide the variable by 1.019^2, and so forth.
6. To obtain this measure, we divide per capita output in 1939 by per capita output in 1929 (0.89) and divide the result by 1.019^{10}.
7. This point is first stressed in Hall 1978.
8. Kendrick's (1961) data for output are very similar to those in the NIPA.
9. Hours will be constant along the steady-state growth path if preferences and technology satisfy certain properties. See King, Plosser, and Rebelo 1988.
10. The average ratio of employment in 1939 to employment in 1929 was one in these countries, indicating that employment had recovered.
11. Cooley 1995 contains detailed discussions of computing the solution to the stochastic growth model.
12. Some researchers argue that there are many other forms of capital, such as organizational capital and human capital, and that the compensation of labor also includes the implicit compensation of these other types of capital. These researchers argue, therefore, that the true capital share is much higher, around two-thirds, and note that with this higher capital share, convergence in the neoclassical model is much slower. To see what a higher capital share would imply for the 1934–39 recovery, we conducted our recovery exercise assuming a capital share of two-thirds rather than one-third. While slower, the recovery was still much faster than in the data. This exercise predicted output at 90 percent of trend by 1936 and at 95 percent of trend by 1939.
13. Bernanke and Parkinson (1991) estimate returns to scale for some manufacturing industries during the Depression and also find evidence that productivity fell during this period. They attribute at least some of the decline to mismeasurement of capital input or increasing returns.
14. An extreme approach to evaluating the effects of idle capital on total factor productivity measurement is to assume that output is produced from a Leontief technology using capital and labor. Under this Leontief assumption, the percentage decline in capital services is equal to the percentage decline in labor services. Total hours drop 27.4 percent between 1929 and 1933. Under the Leontief assumption, total factor productivity in 1933 is about 7 percent below trend, compared to the 14 percent decline under the opposite extreme view that all capital is utilized. This adjustment from a 14 percent decline to a 7 percent decline is almost surely too large not only because it is based on a Leontief technology, but also because it does not take into account the possibility that the capital left idle during the decline was of lower quality than the capital kept in operation.
15. One reason that private investment may have fallen in the 1930s is because government investment was substituting for private investment; however, this seems unlikely. Government investment that might be a close substitute for private investment did not rise in the 1930s: government expenditures on durable goods and structures were 3 percent of output in 1929 and fluctuated between 3 percent and 4 percent of output during the 1930s.
16. To understand why a trade disruption would have such a small effect on output in a country with a small trade share, consider the following example. Assume that final goods are produced with both domestic (Z) and foreign (M) intermediate goods and that the prices of all goods are nor-

malized to one. Assuming an elasticity of substitution between home and foreign goods of one implies that the production for final goods, Y, is Cobb-Douglas, or

$$Y = Z^{\alpha} M^{1-\alpha}$$

where α is the share parameter for intermediate inputs. This assumption implies that with the level of domestic intermediate goods held fixed,

$$\%\Delta Y = (1 - \alpha)\%\Delta M.$$

That fact that U.S. imports were 4 percent of total output and U.S. exports 5 percent in 1929 suggests that the highest the cost share of inputs in production could have been is $0.04/0.95 \approx 0.04$. Hence, an extreme disruption in trade that led to an 80 percent drop in imports would lead to only a 3.2 percent drop in output. (See Crucini and Kahn 1996 for more on this issue.)

17. Note that the monetary base, which is the components of M1 controlled by the Federal Reserve, grew between 1929 and 1933.
18. In addition to Lucas and Rapping's (1969) findings and Fisher's (1933) debt-deflation view, we have other reasons to question the monetary shock view of the Depression. During the mid- and late-1930s, business investment remained more than 50 percent below its 1929 level despite short-term real interest rates (commercial paper) near zero and long-term real interest rates (Baa corporate bonds) at or below long-run averages. These observations suggest that some other factor was impeding the recovery.
19. Bernanke (1983) acknowledges the possibility of an endogenous response but argues that it was probably not important, since problems in financial intermediation tended to precede the decline in overall activity and because some of the bank failures seem to have been due to contagion or events unrelated to the overall downturn.
Recent work by Calomiris and Mason (1997) raises questions about the view that bank runs reflected contagion and raises the possibility that productive, as well as unproductive, banks could be run. Calomiris and Mason analyze the bank panic in Chicago in June 1932 and find that most of the failures were among insolvent, or near-insolvent, banks.
20. To see how we derive the linear expression for $\hat{Y}$, note that if $Y = F(y_i, \ldots, y_n)$, then

$$dY = \sum_{i=1}^{n} F_i d_{yi}.$$

Note also that if goods are produced competitively, then the price of each factor i is given by its marginal product F_i. Hence, $\gamma_i = F_i y_i / Y$, and the result follows.
Note that the fact that the cost shares didn't change very much is inconsistent with the notion that there was extremely low elasticity of substitution for this input and that the fall in this input was an important cause of the fall in output. For example, a Leontief production function in which $F(y_1, \ldots, y_n) = \min_i y_i$ implies that the cost share of input y_i would go to one if that input was the input in short supply.
21. Cooper and Corbae (1997) develop an explicit model of a financial collapse with a high output equilibrium associated with high levels of intermediation services and a low output equilibrium associated with low levels of intermediation services and a sharp reduction in the size of the banking sector. Their model also implies that the ratio of total deposits to output is a measure of the available level of intermediation services.
22. Interest rates on Baa debt, which is considered by investment bankers to have higher default risk than these other debts, did begin to rise in late 1937 and 1938.
23. While Kendrick's (1961) data on aggregate hours are frequently used in macroeconomic analyses of the pre–World War II economy, we point out that the Bureau of Labor Statistics (BLS) did not estimate broad coverage of hours until the 1940s. Thus, Kendrick's data are most likely of lower quality than the more recent BLS data.
24. Decade-long money illusion is hard to reconcile with maximizing behavior. Regarding nominal contracts, we are unaware of any evidence that explicit long-term nominal wage contracts

were prevalent in the 1930s. This prevalence would seem unlikely, since only about 11 percent of the workforce was unionized in the early 1930s.

25. Alternative views in the literature combine a variety of shocks. Romer (1990, 1992) suggests that the 1929 stock market crash increased uncertainty, which led to a sharp decline in consumption. She argues that this shock, combined with monetary factors, is a key to understanding the 1930s. To assess Romer's view, which is based in part on the large drop in stock prices, we need a well-established theory of asset pricing. Existing theories of asset pricing, however, do not conform closely to the data. (See Grossman and Shiller 1981 or Mehra and Prescott 1985.) Given existing theory, a neoclassical evaluation of Romer's view is difficult.

APPENDIX A: THE NEOCLASSICAL GROWTH MODEL

Here we describe the neoclassical growth model, which provides the theoretical framework in the preceding paper.

The neoclassical growth model has become the workhorse of macroeconomics, public finance, and international economics. The widespread use of this model in aggregate economics reflects its simplicity and the fact that its long-run predictions for output, consumption, investment, and shares of income paid to capital and labor conform closely to the long-run experience of the United States and other developed countries.

The model includes two constructs. One is a production function with constant returns to scale and smooth substitution possibilities between capital and labor inputs. Output is either consumed or saved to augment the capital stock. The other construct is a representative household which chooses a sequence of consumption, savings, and leisure to maximize the present discounted value of utility.*

The basic version of the model can be written as maximizing the lifetime utility of a representative household which is endowed initially with k_0 units of capital and one unit of time at each date. Time can be used for work to produce goods (n_t) or for leisure (l_t). The objective function is maximized subject to a sequence of constraints that require sufficient output [$f(k_t, n_t)$] to finance the sum of consumption (c_t) and investment (i_t) at each date. Each unit of date t output that is invested augments the date $t + 1$ capital stock by one unit. The capital stock depreciates geometrically at rate δ, and β is the household's discount factor. Formally, the maximization problem is

$$max_{\{ct,lt\}} \sum_{t=0}^{\infty} \beta^t u(c_t, l_t) \tag{A1}$$

subject to the following conditions:

$$f(k_t, n_t) \geq c_t + i_t \tag{A2}$$

$$i_t = k_{t+1} - (1 - \delta)k_t \tag{A3}$$

$$1 = n_t + l_t \tag{A4}$$

$$c_t \geq 0, n_t \geq 0, k_{t+1} \geq 0. \tag{A5}$$

* Solow's (1956) original version of this model features a representative agent who inelastically supplies one unit of labor and who consumes and saves a fixed fraction of output. Cass (1965) and Koopmans (1965) replace the fixed savings formulation of Solow with an optimizing representative consumer.

Under standard conditions, an interior optimum exists for this problem. (See Stokey, Lucas, and Prescott 1989). The optimal quantities satisfy the following two first-order conditions at each date:

$$u_{lt} = u_{ct} f_2(k_t, n_t) \tag{A6}$$

$$u_{ct} = \beta u_{ct+1}[f_1(k_{t+1}, n_{t+1}) + (1 - \delta)]. \tag{A7}$$

Equation (A6) characterizes the trade-off between taking leisure and working by equating the marginal utility of leisure, u_{lt}, to the marginal benefit of working, which is working one additional unit and consuming the proceeds: $u_{ct} f_2(k_t, n_t)$. Equation (A7) characterizes the trade-off between consuming one additional unit today and investing that unit and consuming the proceeds tomorrow. This trade-off involves equating the marginal utility of consumption today, u_{ct}, to the discounted marginal utility of consumption tomorrow and multiplying by the marginal product of capital tomorrow. This version of the model has a steady state in which all variables converge to constants. To introduce steady-state growth into this model, the production technology is modified to include labor-augmenting technological change, x_t:

$$x_{t+1} = (1 + \gamma)x_t \tag{A8}$$

where the variable x_t represents the efficiency of labor input, which is assumed to grow at the constant rate γ over time. The production function is modified to be $f(k_t, x_t n_t)$. King, Plosser, and Rebelo (1988) show that relative to trend growth, this version of the model has a steady state and has the same characteristics as the model without growth.

This very simple framework, featuring intertemporal optimization, capital accumulation, and an aggregate production function, is the foundation of many modern business cycle models. For example, models with technology shocks start with this framework and add a stochastic disturbance to the production technology. In this case, the resource constraint becomes

$$z_t f(k_t, n_t) \geq c_t + i_t \tag{A9}$$

where z_t is a random variable that shifts the production function. Fluctuations in the technology shock affect the marginal products of capital and labor and, consequently, lead to fluctuations in allocations and relative prices. (See Prescott 1986 for details.)

Models with government spending shocks start with the basic framework and add stochastic government purchases. In this case, the resource constraint is modified as follows:

$$f(k_t, n_t) \geq c_t + i_t + g_t \tag{A10}$$

where g_t is stochastic government purchases. An increase in government purchases reduces output available for private use. This reduction in private resources makes households poorer and leads them to work more. (See Christiano and Eichenbaum 1992 and Baxter and King 1993 for details.)

Because these economies do not have distortions, such as distorting taxes or money, the allocations obtained as the solution to the maximization problem are also competitive equilibrium allocations. (See Stokey, Lucas, and Prescott 1989). The solution to the optimization problem can be interpreted as the competitive equilibrium of an economy with a large number of identical consumers, all of whom start with k_0 units of capital, and a large number of firms, all of whom have access to the technology $f(k, n)$ for transforming inputs into output. The equilibrium consists of rental prices for capital $r_t = f_1(k_t, n_t)$ and labor w_t

$= f_2(k_t, n_t)$ and the quantities of consumption, labor, and investment at each date $t = 0, \ldots, \infty$. In this economy, the representative consumer's budget constraint is given by

$$r_t k_t + w_t n_t \geq c_t + i_t. \tag{A11}$$

The consumer's objective is to maximize the value of discounted utility subject to the consumer's budget constraint and the transition rule for capital (A3). The firm's objective is to maximize the value of profits at each date. Profits are given by

$$f(k_t, n_t) - r_t k_t - w_t n_t. \tag{A12}$$

The effects of monetary disturbances can also be studied in the neoclassical growth framework by introducing money into the model. The introduction of money, however, represents a distortion; consequently, the competitive equilibrium will not generally coincide with the solution to the optimization problem. (See Stokey, Lucas, and Prescott 1989.) In this case, the equations for the competitive equilibrium, rather than the optimization problem, are used in the analysis.

One widely used approach to adding money to the equilibrium model is to introduce a cash-in-advance constraint, which requires that consumption be purchased with cash:

$$m_t \geq p_t c_t \tag{A13}$$

where m_t is the money supply and p_t is the price (in dollars) of the physical good. In this model, changes in the money stock affect expected inflation, which, in turn, changes households' incentives to work and thus leads to fluctuations in labor input. (See Cooley and Hansen 1989 for details.) More-complex monetary models, including models with imperfectly flexible prices or wages or imperfect information about the stock of money, also use the basic model as a foundation.

APPENDIX B THE RECESSION OF 1921: THE RECOVERY PUZZLE DEEPENS

Many economists, including Friedman and Schwartz (1963), view the 1921 economic downturn as a classic monetary recession. Under this view, the 1921 recession and subsequent recovery support our view in the accompanying article that the weak 1934–39 recovery is puzzling.

In 1921, the monetary base fell 9 percent, reflecting Federal Reserve policy which was intended to reduce the price level from its World War I peak. This decline is the largest one-year drop in the monetary base in the history of the United States. The price level did fall considerably, declining 18.5 percent in 1921. Real per capita output also fell in 1921, declining 3.4 percent relative to trend.

Since many economists assume that monetary factors were important in both the 1929–33 decline and the 1921 recession, we compare these two downturns and their recov price level normalized to 100 in the year before the downturn and normalized detrended real per capita output.

There are two key differences between these periods. One is that the decrease in output relative to the decrease in the price level during the 1920s is small compared to the decrease in output relative to the decrease in the price level that occurred during the 1930s. The 18.5 percent decrease in the price level in 1921 is more than five times as large as the 3.4 percent decrease in output in 1921. In contrast, the decrease in the price level is only about 62 percent of the average decrease in output between 1929 and 1933. The other difference is that the 1921 recession was followed by a fast recovery. Even before the price level ceased falling, the economy began to recover. Once the price level stabilized, the economy grew rapidly. Real per capita output was about 8 percent above trend by 1923, and private investment was nearly 70 percent above its 1921 level in 1923. This pattern is qualitatively consistent with the predictions of monetary business cycle theory: a drop in output in response to the price level decline, followed immediately by a significant recovery.

In contrast, the end of the deflation after 1933 did not bring about a fast recovery after the 1929–33 decline. This comparison between these two declines and subsequent recoveries supports our view that weak post-1933 macroeconomic performance is difficult to understand. The recovery from the 1921 recession offers evidence that factors other than monetary shocks prevented a normal recovery from the 1929–33 decline.

A Strong vs. a Weak Recovery: Price Levels and Detrended Real Output in the Early 1920s . . . (index, 1920 = 100)

Year	Price Level	Real Output
1921	81.5	96.6
1922	75.6	99.0
1923	78.6	108.2

Sources: Kendrick 1961; Romer 1989

. . . And in the 1930s (Index, 1929 = 100)

Year	Price Level	Real Output
1930	97.0	87.3
1931	88.1	78.0
1932	78.4	65.1
1933	76.7	61.7
1934	83.2	64.4
1935	84.8	67.9
1936	85.2	74.7

Source: U.S. Department of Commerce, Bureau of Economic Analysis

REFERENCES

Baxter, M., and King, R. G. 1993. Fiscal policy in general equilibrium. *American Economic Review* 83 (June): 315–34.

Bernanke, B. S. 1983. Nonmonetary effects of the financial crisis in propagation of the Great Depression. *American Economic Review* 73 (June): 257–76.

Bernanke, B. S. 1995. The macroeconomics of the Great Depression: A comparative approach. *Journal of Money, Credit, and Banking* 27 (February): 1–28.

Bernanke, B. S., and Carey, K. 1996. Nominal wage stickiness and aggregate supply in the Great Depression. *Quarterly Journal of Economics* 111 (August): 853–83.

Bernanke, B. S., and Parkinson, M. L. 1991. Procyclical labor productivity and competing theories of the business cycle: Some evidence from interwar U.S. manufacturing industries. *Journal of Political Economy* 99 (June): 439–59.

Bordo, M., Erceg, C. and Evans, C. 1996. Money, sticky wages, and the Great Depression. Discussion paper, Rutgers University.

Braun, R. A. 1994. Tax disturbances and real economic activity in the postwar United States. *Journal of Monetary Economics* 33 (June): 441–62.

Brunner, K., ed. 1981. *The Great Depression revisited.* Rochester Studies in Economics and Policy Issues, Vol. 2. Boston: Martinus Nijhoff Publishing.

Calomiris, C. W. 1993. Financial factors in the Great Depression. *Journal of Economic Perspectives* 7 (Spring): 61–85.

Calomiris, C. W., and Mason, J. R. 1997. Contagion and bank failures during the Great Depression: The June 1932 Chicago banking panic. *American Economic Review* 87 (December): 863–83.

Cass, D. 1965. Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies* 32 (July): 233–40.

Christiano, L. J., and Eichenbaum, M. 1992. Current real-business-cycle theories and aggregate labor-market fluctuations. *American Economic Review* 82 (June): 430–50.

Cooley, T. F., ed. 1995. *Frontiers of business cycle research.* Princeton, N.J.: Princeton University Press.

Cooley, T. F., and Hansen, G. D. 1989. The inflation tax in a real business cycle model. *American Economic Review* 79 (September): 733–48.

Cooper, R. and Corbae, D. 1997. Financial fragility and the Great Depression. Working Paper 6094. National Bureau of Economic Research.

Crucini, M. J., and Kahn, J. 1996. Tariffs and aggregate economic activity: Lessons from the Great Depression. *Journal of Monetary Economics* 38 (December): 427–67.

Eichengreen, B. 1992. *Golden fetters: The gold standard and the Great Depression, 1919–1939.* NBER Series on Long-Term Factors in Economic Development. Oxford: Oxford University Press.

Fisher, I. 1933. The debt-deflation theory of great depressions. *Econometrica* 1 (October): 337–57.

Friedman, M. and Schwartz, A. J. 1963. *A monetary history of the United States, 1867–1960.* Princeton, N.J.: Princeton University Press (for NBER).

Grossman, S. J., and Shiller, R. J. 1981. The determinants of the variability of stock market prices. *American Economic Review* 71 (May): 222–27.

Hall, R. E. 1978. Stochastic implications of the life cycle–permanent income hypothesis: Theory and evidence. *Journal of Political Economy* 86 (December): 971–87.

Hanes, C. 1996. Changes in the cyclical behavior of real wage rates, 1870–1990. *Journal of Economic History* 56 (December): 837–61.

Hansen, G. D. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics* 16 (November): 309–27.

Joines, D. H. 1981. Estimates of effective marginal tax rates on factor incomes. *Journal of Business* 54 (April): 191–226.

Kendrick, J. W. 1961. *Productivity trends in the United States.* Princeton, N.J.: Princeton University Press (for NBER).

Keynes, J. M. 1935. *The general theory of employment, interest, and money.* New York: Harcourt, Brace.

King, R. G., Plosser, C. I., and Rebelo, S. T. 1988. Production, growth and business cycles: I. The basic neoclassical model. *Journal of Monetary Economics* 21 (March/May): 195–232.

Koopmans, T. C. 1965. On the concept of economic growth. In *The econometric approach to development planning.* Chicago: Rand-McNally.

Kydland, F. E., and Prescott, E. C. 1982. Time to build and aggregate fluctuations. *Econometrica* 50 (November): 134–70.

League of Nations. 1933. *World economic survey, 1932–33,* p. 211. Geneva: League of Nations.
Lucas, R. E., Jr. 1972. Expectations and the neutrality of money. *Journal of Economic Theory* 4 (April): 103–24.
Lucas, R. E., Jr. 1994. Review of Milton Friedman and Anna J. Schwartz's *A monetary history of the United States, 1867–1960. Journal of Monetary Economics* 34 (August): 5–16.
Lucas, R. E., Jr., and Rapping, L. A. 1969. Real wages, employment, and inflation. *Journal of Political Economy* 77 (September/October): 721–54.
Lucas, R. E., Jr. and Rapping, L. A. 1972. Unemployment in the Great Depression: Is there a full explanation? *Journal of Political Economy* 80 (January/February): 186–91.
Maddison, A. 1991. Dynamic forces in capitalist development: A long-run comparative view. Oxford: Oxford University Press.
Margo, R. A. 1993. Employment and unemployment in the 1930s. *Journal of Economic Perspectives* 7 (Spring): 41–59.
McGrattan, E. R. 1994. The macroeconomic effects of distortionary taxation. *Journal of Monetary Economics* 33 (June): 573–601.
Mehra, R., and Prescott, E. C. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15 (March): 145–61.
Ohanian, L. E. 1997. The macroeconomic effects of war finance in the United States: World War II and the Korean War. *American Economic Review* 87 (March): 23–40.
Prescott, E. C. 1986. Theory ahead of business cycle measurement. *Federal Reserve Bank of Minneapolis Quarterly Review* 10 (Fall): 9–22.
Romer, C. D. 1989. The prewar business cycle reconsidered: New estimates of gross national product, 1869–1908. *Journal of Political Economy* 97 (February): 1–37.
Romer, C. D. 1990. The Great Crash and the onset of the Great Depression. *Quarterly Journal of Economics* 105 (August): 597–624.
Romer, C. D. 1992. What ended the Great Depression? *Journal of Economic History* 52 (December): 757–84.
Romer, C. D. 1993. The nation in depression. *Journal of Economic Perspectives* 7 (Spring): 19–39.
Solow, R. M. 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70 (February): 65–94.
Stern, R., Francis, J., and Schumacher, B. 1976. *Price elasticities in international trade.* London: Macmillan.
Stokey, N. L., Lucas, R. E., Jr., and Prescott, E. C. 1989. *Recursive methods in economic dynamics.* Cambridge, Mass.: Harvard University Press.
Temin, P. 1989. *Lessons from the Great Depression.* Lionel Robbins Lectures. Cambridge, Mass.: MIT Press.
Temin, P. 1993. Transmission of the Great Depression. *Journal of Economic Perspectives* 7 (Spring): 87–102.
White, E. N. 1984. A reinterpretation of the banking crisis of 1930. *Journal of Economic History* 44 (March): 119–38.
Wicker, E. R. 1980. A reconsideration of the causes of the banking panic of 1930. *Journal of Economic History* 40 (September): 571–83.

10

Does the Yield Spread Predict Real Economic Activity? A Multicountry Analysis

Catherine Bonser-Neal
Associate Professor, Indiana University, Indianapolis, Indiana

Timothy R. Morley
Research Associate, Federal Reserve Bank of Kansas City, Kansas City, Missouri

Forecasts of real economic activity are a critical component of many decisions. Businesses rely on such forecasts in forming their production plans. Policymakers rely on such forecasts when choosing the path of monetary policy or when forming the national budget. The appropriateness of these choices depends, in large part, on the quality of the forecast.

Despite their importance, forecasts of real economic activity can be unreliable. Forecasts based on macroeconomic models are often hindered by the lack of timely and accurate data and the complexity of the forecasting model. These difficulties have led to a growing interest in using financial variables to supplement traditional model-based forecasts of real economic activity. The advantages of forecasts based on financial variables are that such forecasts are simple to implement, and the data are readily available and less prone to measurement error.

One financial variable that has been particularly successful in forecasting U.S. real economic growth is the difference between long-term and short-term interest rates, or the yield spread. In general, a positive yield spread—that is, higher long-term interest rates than short rates—is associated with future economic expansion, while a negative yield spread is associated with future economic contraction. In addition, the magnitude of the spread is related to the level of real economic growth: the larger the spread, the faster the rate of real economic growth in the future.

While evidence on the usefulness of the yield spread as a predictor of real economic activity for the United States is now well-established, evidence outside the United States is limited. Few studies have examined the forecast power of the yield spread in other countries, and those that have are limited in either the sample of countries, the measure of real economic activity, or the length of the forecast horizon. Such evidence on the predictive

Reprinted from: *Federal Reserve Bank of Kansas City Economic Review,* vol. 82, no. 3 (Third Quarter 1997) 37–53.

ability of the yield spread, however, would be useful to businesses and policymakers in the United States as well as abroad. For example, U.S. businesses and policymakers would benefit from better forecasts of foreign real economic activity because projections for U.S. exports depend on forecasts of foreign economic growth. In addition, foreign businesses and policymakers would benefit from knowing which variables are useful in forecasting real economic activity in their country.

To obtain such evidence, this article evaluates the ability of the yield spread to forecast real economic activity in 11 industrial countries. The first section of this article defines the yield spread and explains why the spread may be a useful predictor of real economic activity. The second section describes the data and criteria used to evaluate the predictive power of the yield spread. The third section examines whether yield spreads have reliably forecast real economic activity in the 11 countries, using several measures of real economic activity and alternative forecast horizons. The empirical results indicate the yield spread is a statistically and economically significant predictor of real economic activity in several industrial countries besides the United States. In addition, the yield spread forecasting model generally outperforms two alternative forecasting models in predicting future real GDP growth.

THE LINK BETWEEN THE YIELD SPREAD AND REAL ECONOMIC ACTIVITY

Understanding the relationship between the yield spread and the economy involves understanding the yield curve and what movements in it may reflect.[1] This section defines the yield curve and the yield spread and discusses explanations for why the yield spread could reliably forecast real economic activity.

What is the yield curve? A yield *curve* plots the yields on debt securities with similar risk, liquidity, and tax considerations relative to the securities' time to maturity. For example, suppose today the yield on a 3-month Treasury bill is 6 percent and the yield on a 10-year Treasury bond is 8 percent. The yield curve for these two securities appears as the solid line in Figure 1. If the 10-year bond rate is 7 percent, then the yield curve is given by the lower dotted line. The yield curve in both cases is linear since the yields of only two securities are plotted. With more than two securities of different maturities, the yield curve need not be linear.

The yield *spread* is the difference at a point in time between the yields on two securities with different maturities. In the above example, the yield spread between the 10-year Treasury bond with a yield of 8 percent and the 3-month Treasury bill with a yield of 6 percent is two percentage points. If the 10-year bond yield falls to 7 percent, with no change in the short rate, the yield spread falls to 1 percent.

The yield spread also measures the steepness—or the slope—of the yield curve. The larger the spread, the steeper the yield curve. Figure 1 illustrates the relation between the yield spread and the slope of the yield curve. The yield curve with a spread of 2 percent is steeper than the curve with a spread of 1 percent.

What does the yield spread reflect? Recent research has shown that the yield spread between long-term and short-term bonds helps predict future real activity in the United States and in some other OECD countries. A positive spread between long-term and short-term interest rates (a positively sloped yield curve) is associated with an increase in real economic activity, while a negative spread (a negatively sloped yield curve) is associated

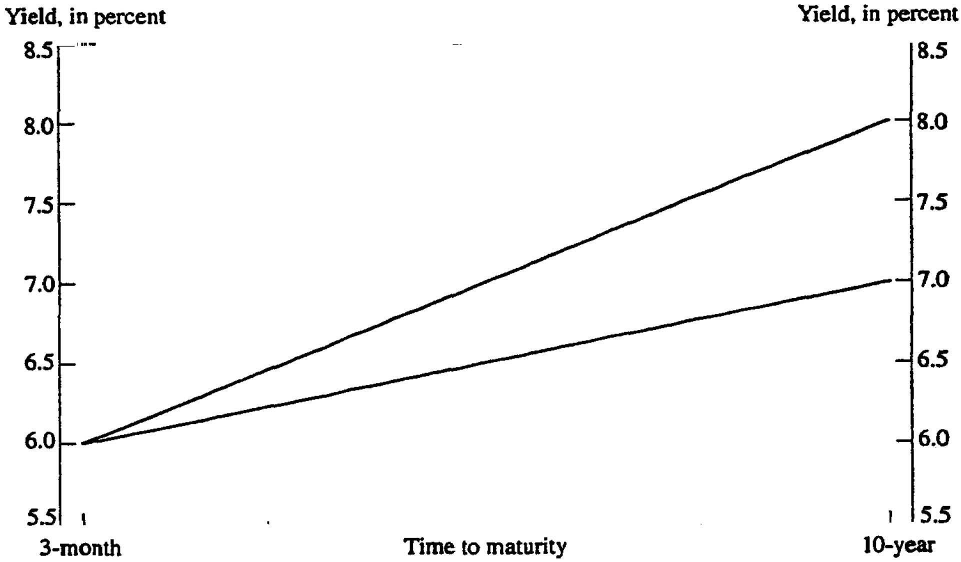


Figure 1 Sample yield curves.

with a decline in real activity. In addition, the larger the yield spread, the higher the level of future real economic activity.[2]

Researchers have offered two reasons for this empirical relationship. First, the yield spread may reflect the stance of monetary policy. When monetary policy is tightened, short-term interest rates rise; long-term rates also typically rise but usually by less. As a result, the yield spread narrows or even turns negative. In time, higher interest rates reduce spending in interest-rate sensitive sectors of the economy, causing economic growth to slow. Consequently, a small, or negative, yield spread will be associated with slower real economic growth in the future.

An alternative explanation for the link between the yield spread and future growth is that the yield spread reflects market expectations of future economic growth. For example, suppose market participants expect real income to rise in the future. An increase in expected future real income implies an increase in profitable investment opportunities today. In order to take advantage of these investment opportunities, businesses increase their borrowing and issue more bonds. Since these investments are typically longer term, the bonds issued will also be longer term. An increase in the supply of longer term bonds reduces their price and increases their yield. Long-term rates will therefore rise relative to short-term rates, and the yield curve will steepen. As long as these expectations for economic growth are at least partially realized, a steepening of the yield curve will be associated with a future increase in real economic activity.[3]

Which theory is the most likely explanation for the ability of the yield spread to forecast real economic growth? Recent evidence suggests both theories may have merit. Estrella and Hardouvelis (1991) and Estrella and Mishkin (1995), for example, show that proxies for current monetary policy do help forecast future real GDP growth; however, the inclusion of these proxies does not eliminate the significance of the yield curve. These results suggest the yield curve reflects more than just the effects of current monetary policy actions.

EVALUATING THE PREDICTIVE POWER OF THE YIELD SPREAD: DATA AND CRITERIA

Before evaluating the forecasting power of the yield spread across countries, several decisions must be made regarding the countries and the variables to be considered. This section describes the data and criteria used to evaluate the predictive power of the yield spread.

Which Countries?

A proper evaluation of the predictive power of the yield spread requires a lengthy time series of market-determined interest rates and accurate measures of real economic activity. Two criteria guided the selection of countries to be included in the analysis. First, only industrial countries with well-developed financial markets were included. This criterion ensured that the interest rates used in the forecasts were determined in a liquid and transparent market, and hence reflect market expectations rather than government controls. Second, data on interest rates and real economic activity had to be available for at least the last 20 years. This criterion ensured a sample size large enough that the forecast power of the yield spread could be reliably assessed. The resulting sample consists of 11 industrial countries: Australia, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, Switzerland, the United Kingdom, and the United States.[4]

Which Yield Spread?

In testing the predictive power of the yield spread, it is important to choose the yields of debt securities which are actively traded and which reflect market expectations. Because several debt securities can satisfy this criterion, there are several alternative measures of the yield spread. In the United States, for example, some of the important interest rates monitored by market participants include the federal funds rate; the 3-month Treasury bill rate; the 1-year, 5-year, 10-year Treasury note rates; and the 30-year Treasury bond rate. A similar range of assets exists in many foreign countries.

In choosing the yield spread, this study seeks to balance comparability with previous studies along with data availability. Previous research on the predictive power of the spread in the United States has focused on the spread between the 10-year Treasury bond rate and the 3-month Treasury bill rate. Consequently, when possible, the yield spread examined for each country is the spread between the rate on the 10-year government bond and the 3-month government bill rate. In countries where these assets are not actively traded, an alternative asset must be used. For example, in countries where a 10-year government bond is not actively traded, the long-term government bond which is actively traded is used. Similarly, in countries where a 3-month government bill rate is not available, an alternative 3-month rate such as the rate on interbank deposits is used.[5]

Which Measure of Real Economic Activity?

How well the yield spread predicts economic activity also depends on the measure of real economic activity examined. This article uses real GDP growth as the primary measure of real economic activity. Real GDP is the broadest measure of economic activity, and it is closely monitored by market participants. Real GDP, however, is available only on a quarterly basis. Alternative measures of real economic activity, such as the index of industrial production and the unemployment rate, are available monthly. Hence, to estimate the predictive power of the spread on a monthly basis, this article also measures real economic activity with the index of industrial production and the unemployment rate.

Which Forecast?

Once the yield spread and economic activity measures are selected, the next step is to choose forecasting techniques and forecasting horizons. Two types of forecasting techniques can be employed to evaluate the forecast power of the yield spread: in-sample and out-of-sample forecasts. An in-sample forecast estimates the average relationship between the yield spread and real economic activity over the entire period for which data are available. Since it measures an average relationship over the full period, an in-sample forecast is calculated using information which was not available at the time market participants formed their forecast. For example, given quarterly data from 1971:Q1 to 1996:Q1, an in-sample forecast for real economic growth in 1985:Q1 would be calculated based on the yield spread in prior quarters, using the relationship of the yield spread to economic activity estimated over the entire sample through 1996:Q1. An out-of-sample forecast, in contrast, only uses information available to market participants at the time of the forecast. For example, an out-of-sample forecast for real economic growth in 1985:Q1 would be calculated based on the yield spread in prior quarters, using the estimated relationship of the yield spread to economic activity only through 1984:Q4. Because both types of forecasts provide insight into the relationship between the yield spread and real economic activity, this article presents the results of both in-sample and out-of-sample forecasts.

The decision on the type of forecast must also consider the length of the forecast horizon. For example, does the yield curve predict real economic activity one quarter or five years into the future? To keep the presentation of the results manageable, this article estimates the ability of the yield curve to forecast one, two, and three years into the future. Previous research by Estrella and Mishkin (1995), and by Estrella and Hardouvelis (1991) shows that most of the forecast power occurs after one year.

Given the above decisions with respect to the sample of countries, yield spread, measure of real economic activity, and type of forecast, the next section estimates the ability of the yield spread to forecast real economic activity in 11 industrial countries.

INTERNATIONAL EVIDENCE ON THE YIELD SPREAD AS A PREDICTOR OF REAL ECONOMIC ACTIVITY

While several studies have examined the forecasting power of the yield spread in other countries, they are limited in either the choice of countries, the measure of real economic activity, or in the forecast horizon examined. This section extends previous work by estimating the predictive power of the yield spread in a wide range of countries across different measures of real economic activity and over different forecast horizons.[6]

To estimate the forecast power of the yield spread, the following equation is estimated for each country:

$$(\textit{change in real economic activity})_{t,t+k} = \alpha + \beta \cdot \textit{spread}_t + \textit{error}, \qquad (1)$$

where the change in real economic activity is defined as the average annualized percentage change in real GDP or industrial production or as the cumulative change in the unemployment rate from today to k periods in the future. The subscript *k* represents the forecast horizon in quarters or months, and the variable *spread* is defined as the difference between the long-term and the short-term interest rates.[7] After the equation is estimated, it is used to generate in-sample and out-of-sample forecasts.[8]

In-Sample Forecasting Results

Equation 1 is estimated for each country over the 1971–96 time period.[9] Data over this time period exist for most countries, thereby ensuring that differences across countries are not due to differences in the sample period. In addition, the 1971–96 time period corresponds to the post-Bretton Woods period of floating exchange rates.

In order to judge the overall performance of the forecasting equation, Figure 2 plots the R-square values from estimating the forecasting equation 1 using real GDP growth as the measure of the change in real economic activity.[10] The R^2 from the regression equation measures the proportion of the variation in real GDP growth that is explained by the yield spread. The R^2 from the estimation of equation 1 range from 1 to 50 percent, depending on the country and the forecasting horizon. In the top panel, for example, the R^2 for Japan is close to zero, indicating that today's yield spread explains virtually none of the changes in real GDP growth over the following year.[11] In Australia, Netherlands, Sweden, Switzerland, and the United Kingdom, the yield spread explains less than 20 percent of the variation in the following year's real GDP growth. In contrast, the yield spread in Canada explains roughly 50 percent of the variation in Canada's real GDP growth over the following year. In France, Germany, Italy, and the United States, the yield curve explains between 25 and 40 percent of the variation in the following year's real GDP growth.

The middle and bottom panels of Figure 2 plot the R-squares for the two-year and three-year forecasts, respectively. In general, the explanatory power of the yield spread falls with the lengthening of the forecast horizon. In the United States, for example, the proportion of variation in future real GDP explained by the yield spread is 40 percent at the one-year horizon, but only 20 percent at the three-year horizon. The explanatory power of the yield spread also falls with the lengthening of the forecast horizon in Canada, France, Italy, Sweden, and the United Kingdom, although in Canada the R^2 falls only slightly from 50 percent at the one-year horizon to 42 percent at the three-year horizon. For some countries (Australia, Germany, the Netherlands, and Switzerland), however, the predictive power of the yield spread is slightly stronger at the two-year horizon, compared with the one-year horizon. In the Netherlands, the R^2 of the yield spread forecasting equation at the two-year and three-year horizons is twice that of the one-year horizon.

While the R^2 statistic provides an indication of the explanatory power of the spread for real GDP growth, the coefficient β from equation 1 measures how much real GDP growth changes following a one-percentage-point change in the yield spread. A positive β would imply a positive relationship between the current yield curve and future economic growth. That is, the larger the spread is between long-term and short-term interest rates, the stronger real growth will be in the future.

Figure 3 provides estimates of β for the one-year-ahead, two-year-ahead, and three-year-ahead forecasts for each country. The coefficient β is positive in all three panels, and the statistical significance of β is indicated by a solid bar. For the one-year forecast, the solid bars also show that the yield spread is a significant predictor of real GDP growth in all countries except Japan. The middle and bottom panels show that the number of countries for which the yield spread is a statistically significant predictor of future real GDP

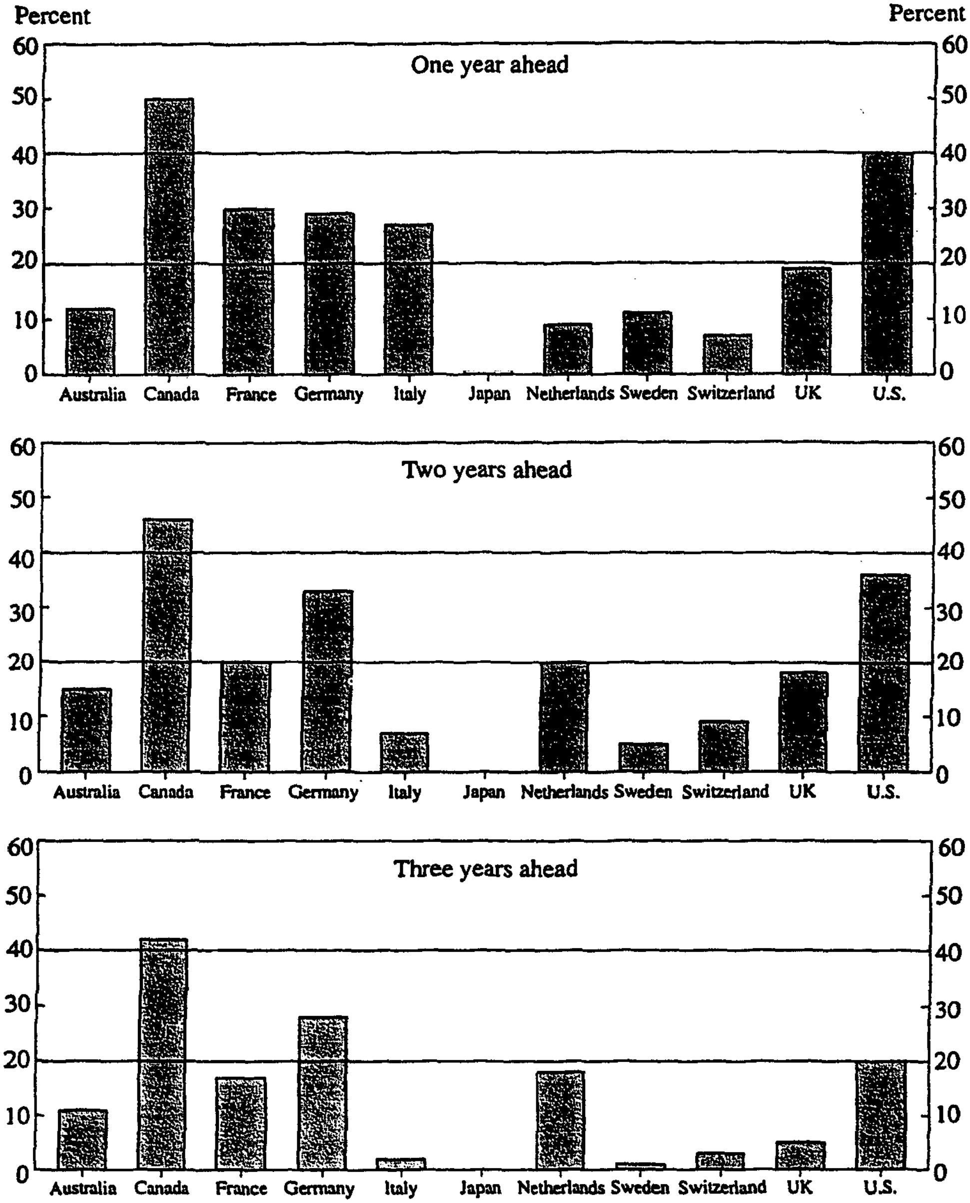


Figure 2 Explanatory power of the yield spread for real GDP. Note: Charts give the adjusted R^2 from the regression of future real GDP growth on the yield spread. Source: See appendix and authors' calculations.

growth declines with the forecast horizon. The yield spread remains a statistically significant predictor of real GDP growth over a two-year horizon in seven countries (Canada, France, Germany, Netherlands, Switzerland, the United Kingdom, and the United States), and is a statistically significant predictor of real GDP growth over a three-year horizon in only five countries (Canada, France, Germany, Netherlands, and the United States).

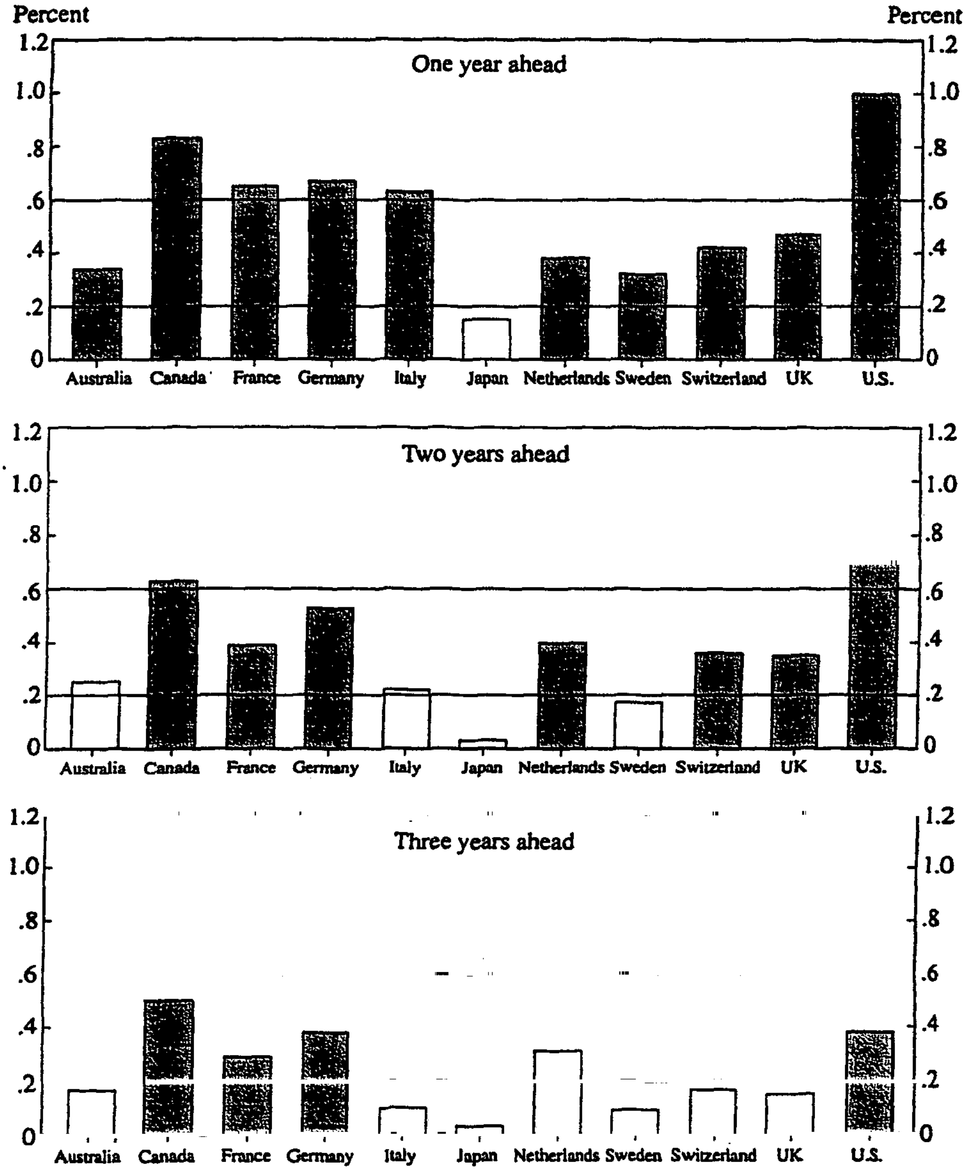


Figure 3 Change in future real GDP growth following a one-percentage-point change in the yield spread. Note: Each bar represents the beta coefficient from the regression of future real GDP growth on the yield spread. Statistical significance is indicated by a shaded bar. Source: See appendix and authors' calculations.

Estimates of the β's themselves from equation 1 also provide an indication of the *economic* significance of the yield curve as a predictor of future real economic growth. In particular, the coefficient β measures the change in real GDP growth for a given one-percentage-point change in the yield spread. In the United States, for example, the three panels show that a one-percentage-point increase in the yield spread today is associated with a one-percentage-point increase in real GDP growth next year, an annualized 0.7-percentage-point increase in growth over the next two years, and an annualized 0.4-percentage-point increase in real GDP growth over the next three years. Hence, all else constant, if real GDP growth in the United States was 2 percent, a widening of the yield spread by one percentage point would imply an increase in real GDP growth to 3 percent ($2 + 1 \times 1$) over the next year, to 2.7 percent ($2 + 0.7 \times 1$) on average over the next two years, and to 2.4 percent ($2 + 0.4 \times 1$) on average over the next three years.[12]

For the remaining countries, a one-percentage-point increase in the yield spread is associated with a less-than-one-percentage-point increase in real GDP growth at all forecast horizons. For example, the top panel shows that in Canada, France, Germany, and Italy a one-percentage-point increase in the yield spread is associated with a 0.6- to 0.8-percentage-point increase in real GDP growth one year from now. In the Netherlands, Sweden, Switzerland, and the United Kingdom, a one-percentage-point increase in the yield spread is associated with a 0.3- to 0.5-percentage-point increase in real GDP growth. As with the United States, the β coefficients fall as the forecast horizon is lengthened.

Together the results indicate that while the yield spread does help explain future real GDP growth in many countries, the strength of the predictive power varies by country. The explanatory power of the yield spread is highest in Canada and the United States, and lowest in Japan. However, real GDP growth is not the only measure of real economic activity of concern to market participants: changes in industrial production and the unemployment rate provide alternative measures of real economic activity.

Because different measures of real economic activity may produce different forecasts of real economic growth, the ability of the yield spread to predict real economic activity may be sensitive to the measure of real activity employed in the forecasting equation. The forecasting equation 1 is therefore reestimated using either industrial production or the unemployment rate as the measure of real economic activity. The R-squares for the one-year forecasts, displayed in Figure 4, indicate that yield spread forecasts are sensitive to the measure of real economic activity employed in the forecasting equation. In Germany and Sweden, for example, the share of variation in real activity explained by the yield spread is greatest when real activity is measured by the change in the unemployment rate. In Germany, the yield spread explains 40 percent of the variation in the unemployment rate over the following year compared with 29 percent of real GDP growth. In Japan, the use of an alternative measure of real economic activity increases the ability of the yield spread to predict future real activity. When the percent change in industrial production is used as the measure of real economic activity, the yield spread explains 16 percent of the variation in real activity compared with none of the variation using real GDP growth. In the United States, the predictive power of the yield spread is similar across the three measures of real activity, but is slightly higher with unemployment as a measure of real activity.

Out-of-Sample Forecasting Results

As noted earlier, one disadvantage with in-sample forecasts is that they allow the forecasts to depend on data which were not available at the time of the forecast. As a result, the empirical results of the previous section may provide a misleading indication of the true ability of the yield curve to forecast real activity. To provide a check on the usefulness of the yield spread as a predictor of activity, out-of-sample forecasts of real GDP growth are estimated. Specifically, forecasts for each period are based on an estimate of equation 1 using only data up to the previous period. For example, the forecast for 1980:Q1 is

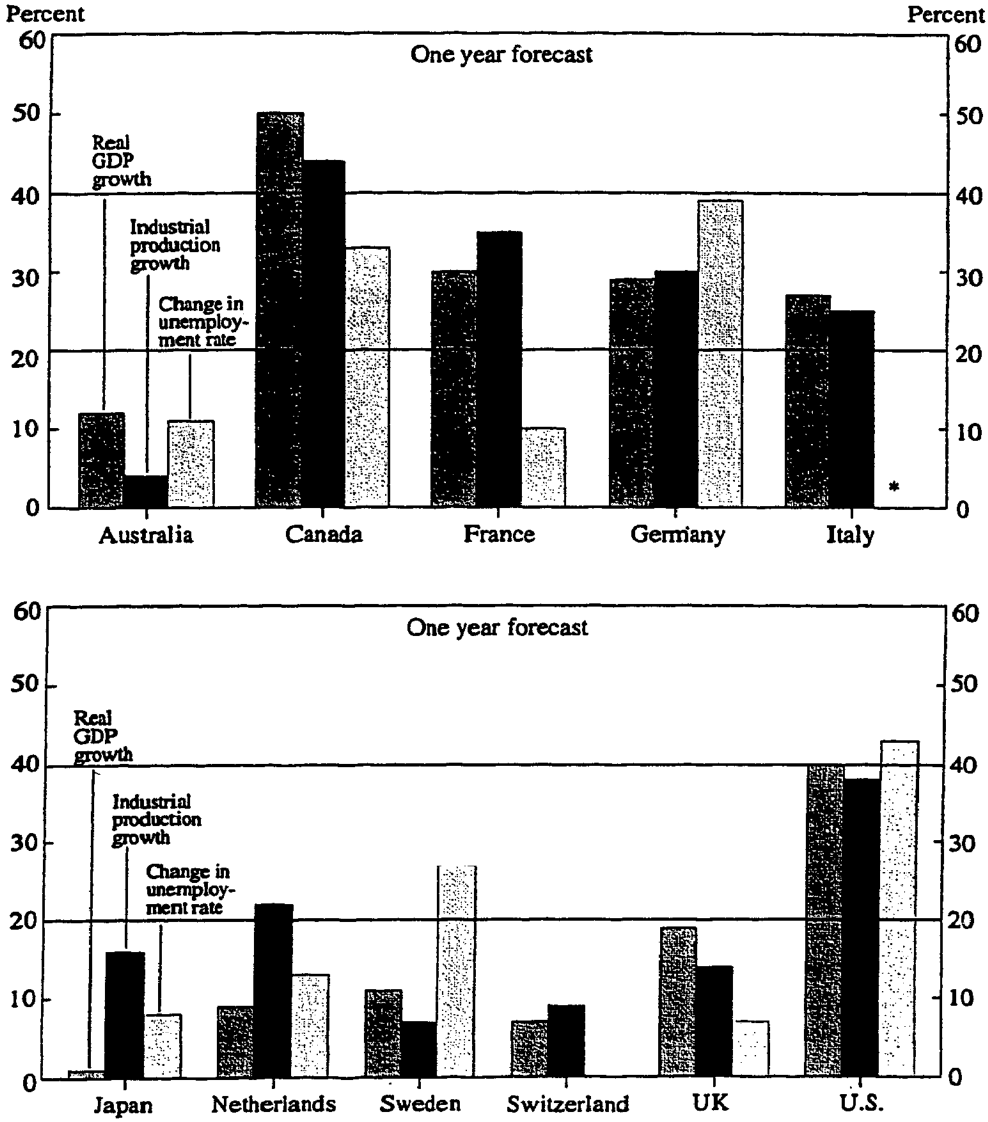


Figure 4 Explanatory power of the yield spread for different measures of economic activity. Note: For each country, the bars indicated the adjusted R^2 from the regression of real GDP growth, industrial production growth, or the change in the unemployment rate, on the yield spread. There is no estimate of the unemployment rate in Switzerland. *The R^2 for Italy for the change in the unemployment rate is zero. Source: See appendix and authors' calculations.

estimated using coefficients from the regression estimated over the 1971:Q1 to 1979:Q4 period, while the forecast for 1980:Q2 is estimated using coefficients estimated over the 1971:Q1 to 1980:Q1 period.

The quality of the out-of-sample forecast is evaluated using the root mean squared error (RMSE) statistic. The RMSE provides an estimate of the out-of-sample forecast error, and hence measures the accuracy of the forecast. The lower the RMSE, the better the forecast. In evaluating the out-of-sample forecast power of the yield spread, the RMSE from the yield spread forecast is compared with the RMSEs of alternative forecasts of real economic activity. Indeed, one advantage of the RMSE measure is that, for a given country, it can be compared across different forecasting models.

In this study, the out-of-sample predictive power of the yield spread model is compared with that of two alternative forecasting models over a one-year horizon.[13] In the first alternative model, past changes in real economic activity are used to predict future changes. Specifically, the forecast for next year's real GDP growth is estimated using the following equation:

$$(\%\ change\ real\ GDP)_{t,t+k} = \alpha + \gamma \cdot (\%\ change\ real\ GDP)_{t-k,t} + error. \tag{2}$$

The second alternative model combines both forecasting variables—the yield spread and current real GDP growth—into one forecasting equation:

$$\begin{aligned}(\%\ change\ in\ real\ GDP)_{t,t+k} &= \alpha + \gamma \\ &\times (\%\ change\ in\ real\ GDP)_{t-k,t} + \beta \cdot (spread)_t + error.\end{aligned} \tag{3}$$

To determine the relative forecast performance of the three forecasting models, the yield spread model (equation 1), the GDP growth model (equation 2), and the combined yield spread plus GDP growth model (equation 3) were estimated across three forecast horizons and their out-of-sample RMSE's were compared. Table 1 shows the results of these model comparisons. The table indicates that the yield spread model generally has the lowest RMSE and hence the best out-of-sample forecast performance. The GDP growth model outperforms the yield spread model in only 4 out of 33 cases. In addition, adding lagged real GDP growth to the yield spread as in the combined model (equation 3) generally does not improve the out-of-sample forecast performance: the combined model has the lowest RMSE in only 8 out of 33 cases. These results indicate the ability of the yield spread forecasting model to predict real GDP growth generally exceeds that of both the lagged real GDP model and the combined forecasting model.

Table 1 Model with Lowest Root Mean Squared Error

	Forecast horizon		
Country	1 year	2 years	3 years
Australia	Spread	Spread + GDP	Spread + GDP
Canada	Spread	Spread	Spread
France	Spread	Spread	Spread
Germany	Spread + GDP	Spread	Spread + GDP
Italy	Spread	Spread	Spread
Japan	GDP	Spread	Spread
Netherlands	Spread	Spread + GDP	Spread + GDP
Sweden	Spread	Spread	Spread + GDP
Switzerland	GDP	GDP	GDP
UK	Spread + GDP	Spread	Spread
U.S.	Spread	Spread	Spread

Note: This table shows for each country which of the following models had the lowest root mean squared error:
Spread: $(\%\ change\ real\ GDP)_{t,t+k} = \alpha + \beta^*\text{spread}_t + error$
GDP: $(\%\ change\ real\ GDP)_{t,t+k} = \alpha + \gamma^*(\%\ change\ real\ GDP)_{t-k,t} + error$
Spread + GDP: $(\%\ change\ real\ GDP)_{t,t+k} = \alpha + \beta^*\text{spread}_t + \gamma^*(\%\ change\ real\ GDP)_{t-k,t} + error$

CONCLUSION

This article has provided evidence on the ability of the yield spread to predict future real economic activity in 11 industrial countries. The results are broadly consistent with the results of previous studies, but are also more comprehensive in that they evaluate the predictive power of the yield spread across multiple countries, measures of real economic activity, and types of forecasts.

The results indicate the yield spread is a statistically and economically significant predictor of economic activity in several countries besides the United States. Increases in the yield spread are followed by increases in real economic growth, while decreases in the spread are followed by decreases in growth. The size of the spread is also related to the level of real economic growth: the larger the spread between long-term and short-term interest rates, the higher the future level of real economic growth. In addition, out-of-sample forecasts of real GDP growth based on the yield spread generally beat forecasts based on past real GDP growth.

The empirical results of this study also show that the strength of the relationship between the yield spread and future real economic growth varies across the 11 countries. The predictive power of the spread is strongest in Canada, Germany, and the United States. In these three countries the yield spread consistently explains roughly 30 to 50 percent of the variation in future real economic activity. Individuals interested in forecasting real economic activity in these countries would benefit by supplementing their existing model forecasts with forecasts based on the yield spread. The ability of the yield spread to forecast real economic activity is weakest in Japan and Switzerland, where the yield spread, on average, explains less than 10 percent of variations in future real economic activity. Thus, in these countries, the yield spread is not a useful indicator of future growth. In the remaining countries, the results are mixed. For example, the yield spread in France explains roughly 30 percent of the following year's real GDP growth but only 10 percent of the change in the unemployment rate. In these countries the benefit to supplementing existing

Appendix A Data Sources

	GDP	Unemployment rate	Industrial production	Short-term interest rate	Long-term interest rate
Australia	BIS	BIS	IFS	IFS	BIS
Canada	IFS	BOG	BIS	IFS	BIS
France	BOG	BOG	BOG	BIS	BIS
Germany	BOG	BOG	BOG	BOG	OECD, BOG
Italy	BIS	BIS	BIS	IFS	BIS, IFS
Japan	BIS	BIS	BIS	BIS	BIS
Netherlands	BIS	BIS	IFS	BIS	IFS
Sweden	IFS	BIS	BOG	IFS	IFS
Switzerland	IFS	N.A.	BIS	BIS	BIS
UK	IFS	BOG	BOG	BOG	OECD
U.S.	BEA	BLS	BOG	BOG	BOG

Note: BEA is the U.S. Dept. of Commerce. Bureau of Economic Analysis. BLS is the U.S. Dept. of Labor, Bureau of Labor Statistics. BOG is the Federal Reserve Board of Governors. BIS is the Bank for International Settlements. IFS is International Financial Statistics. OECD is the OECD Main Economic Indicators. N.A. indicates that data were not available. Two series separated by a comma indicate that a series was formed by combining two sets of data.

Appendix B Coefficient Estimates for a Regression of Real GDP Growth on the Yield Spread

	Forecast horizon								
	1 year			2 years			3 years		
Country	Constant (α)	Slope (β)	$\bar{R}^2$	Constant (α)	Slope (β)	$\bar{R}^2$	Constant (α)	Slope (β)	$\bar{R}^2$
Australia	3.08	.34	.12	3.08	.25	.15	3.02	.17	.11
	(9.22)	(2.62)		(10.56)	(1.69)		(12.43)	(1.37)	
Canada	2.82	.83	.50	2.88	.63	.46	2.88	.50	.42
	(9.31)	(6.55)		(8.65)	(4.95)		(8.04)	(4.35)	
France	1.64	.65	.30	1.89	.39	.20	1.94	.29	.17
	(5.82)	(4.52)		(7.35)	(2.87)		(7.17)	(2.80)	
Germany	1.61	.67	.29	1.71	.53	.33	1.81	.38	.28
	(4.21)	(5.48)		(5.13)	(6.42)		(6.98)	(4.24)	
Italy	3.53	.63	.27	2.89	.22	.07	2.60	.10	.02
	(8.13)	(4.39)		(6.59)	(1.49)		(7.24)	(1.24)	
Japan	3.54	.15	.01	3.46	.03	.00	3.37	.03	.00
	(9.14)	(1.26)		(8.64)	(.44)		(8.84)	(.47)	
Netherlands	2.02	.38	.09	2.04	.40	.20	2.06	.31	.18
	(5.89)	(2.35)		(6.04)	(2.56)		(5.78)	(1.95)	
Sweden	1.59	.32	.11	1.55	.17	.05	1.48	.09	.01
	(4.60)	(2.34)		(3.93)	(1.28)		(3.64)	(.71)	
Switzerland	.97	.42	.07	1.00	.36	.09	1.07	.17	.03
	(2.23)	(3.16)		(2.25)	(3.48)		(2.49)	(1.30)	
UK	1.82	.47	.19	1.86	.35	.18	1.86	.15	.05
	(4.40)	(3.33)		(3.96)	(2.73)		(4.07)	(1.09)	
U.S.	1.46	1.00	.40	1.78	.70	.36	2.12	.38	.20
	(4.36)	(6.14)		(6.00)	(6.21)		(6.00)	(3.47)	

Note: T-statistics are in parentheses. The time period is 1972:1 to 1996:4 for all countries except: Sweden (1972:1 to 1995:3), Switzerland (1972:1 to 1995:4), and the Netherlands (1977:1 to 1996:4). The following model was estimated: (*% change real GDP*)$_{t,\,t+k}$ = α + β**spread*$_t$ + *error*. The results for industrial production growth and unemployment rate change, as the dependent variable, are available upon request from the authors

model forecasts with forecasts based on the yield spread will depend on the measure of real economic activity and the forecast horizon.

NOTES

1. Kessel (1965) was the first to note the relation between the yield curve and future real economic activity.
2. Estrella and Hardouvelis (1991), Harvey (1989), and Haubrick and Dombrosky (1996) find the yield spread predicts real GDP growth in the United States, while Estrella and Hardouvelis (1991) and Estrella and Mishkin (1996) and Dueker (1997) find that the U.S. yield spread forecasts the probability of a U.S. recession. Studies which examine the predictive power of the yield spread in non-U.S. countries include Caporale (1994), Estrella and Mishkin (1995), Hu (1993), and Plosser and Rouwenhorst (1994).
3. Harvey (1988, 1989) presents a related explanation for the relation between the slope of the yield curve and future economic growth. Suppose, as above, bond market participants expect real income to rise in the future. The expectation of increased future income will reduce today's

demand for long-term bonds which pay off in the future. A decrease in the demand for the long-term bonds will cause the price of the bonds to fall, or their yield to rise. Thus, the yield curve steepens as long-term interest rates rise. If the expectations for economic growth are realized, a steepening of the yield curve will be associated with a future increase in real economic activity.

4. Because of the difficulties in interpreting data on East Germany prior to the unification of East and West Germany, the analysis focuses on West German economic growth only.
5. When possible, the interest rate data used averages of daily data over the quarter or month, depending upon whether quarterly or monthly forecasts are evaluated. Averaged data are used because it is more likely that measures of real economic activity are related to average yield spreads, rather than to a single end-of-month value.
6. In particular, Caporale (1994) examines the in-sample and out-of-sample forecast power of the yield spread in 13 countries, but restricts her forecast horizon to one year and her measure of real economic activity to real GDP. Estrella and Mishkin (1995) examine the predictive power of the yield spread in five countries, France, Germany, Italy, the United Kingdom, and the United States, over the 1973 to early 1995 period, and also examine the sensitivity of the empirical results to real GDP, industrial production, and unemployment measures of real activity, as well as to forecast horizons up to five years. They do not, however, examine the out-of-sample forecast power of the yield spread. Hu (1993) examines the in-sample and out-of-sample forecast power of the yield spread from the earliest data point possible to 1991 for the G-7 countries, but restricts his forecast horizon to one year. Finally, Plosser and Rouwenhorst (1994) examine in-sample yield spread forecasts for three countries, Germany, the United Kingdom, and the United States, between 1973 and 1988. They consider both real GDP and industrial production measures of real economic activity and forecast horizons of up to five years, but they restrict their forecasts to in-sample forecasts.
7. When the percent change in real GDP is the dependent variable, the long-term and short-term interest rates used to calculate the spread are the quarterly averages of monthly rates in that quarter. When industrial production or the unemployment rate is the dependent variable, the spread is calculated from daily averages of the long-term and short-term interest rates over the month.
8. Estimating this forecasting equation for $k = 1, 2$, or 3 years with quarterly or monthly data causes the error term to be serially correlated. Consequently, the standard errors from the estimation are corrected following Hansen (1982) and Newey and West (1987).
9. Due to data availability, the sample period for the Netherlands is 1977:1 to 1996:4.
10. The R-squares plotted are actually the R-bar squares from the regression, the R-square values adjusted for the degrees of freedom.
11. Prior to 1980, Japanese financial markets were heavily regulated and may not have reflected market expectations.
12. These average growth rates imply real GDP growth in year 2 will be 2.4 percent ($2.7 \times 2 - 3.0$), while real GDP growth in year 3 will be 1.8 percent ($2.4 \times 3 - 3.0 - 2.4$).
13. The results for the two-year and three-year forecasts are very similar.

REFERENCES

Caporale, G. M. 1994. The Term Structure as a Predictor of Real Economic Activity: Some Empirical Evidence. London Business School, Discussion Paper no. 4–94, February.

Dueker; M. 1997. Strengthening the Case for the Yield Curve as a Predictor of U.S. Recession. Federal Reserve Bank of St. Louis, Review, March/April, pp. 41–50.

Estrella, A., and Mishkin, F. 1996. Predicting U.S. Recessions: Financial Variables as Leading Indicators. Federal Reserve Bank of New York, working paper no. 9609, May.

Estrella, A., and Mishkin, F. 1995. The Term Structure of Interest Rates and Its Role in Monetary Policy for the European Central Bank. National Bureau of Economic Research, working paper no. 5279, September.

Estrella, A., and Mishkin, F. and Hardouvelis, G. A. 1991. The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance,* no. 46, June, pp. 555–76.

Hansen, L. P. 1982. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica,* vol. 50, pp. 1029–54.

Harvey, C. R. 1989. Forecasts of Economic Growth from the Bond and Stock Markets. *Financial Analysis Journal,* September/October, pp. 38–45.

Harvey, C. R. 1988. The Real Term Structure and Consumption Growth. *Journal of Financial Economics,* vol. 22, pp. 305–53.

Haubrich, J. G., and Dombrosky, A. M. 1996. Predicting Real Growth Using the Yield Curve. Federal Reserve Bank of Cleveland, *Economic Review,* First Quarter, pp. 26–35.

Hu, Z. 1993. The Yield Curve and Real Activity. International Monetary Fund, Staff Paper no. 40, December, pp. 781–806.

Kessel, R. A. 1965. The Cyclical Behavior of the Term Structure of Interest Rates. *National Bureau of Economic Research,* Occasional Paper 91.

Newey, W., and West, K. 1987. A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica,* vol. 55, May, pp. 703–78.

Plosser, C. I., and Rouwenhorst, K. G. 1994. International Term Structures and Real Economic Growth. *Journal of Monetary Economics,* vol. 33, pp. 133–55.

This Page Intentionally Left Blank

11

Interest Rate Spreads as Indicators for Monetary Policy

Chan Guk Huh
Economist, Federal Reserve Bank of San Francisco, San Francisco, California

Traditionally, economists have focused on aggregate money stock measures such as M1 and M2 as indicators of future economic activity. However, the relationship between these aggregates and real GDP has deteriorated in recent years. Thus there is a growing interest in alternative indicators, some of which are conceptually quite new compared to the conventional financial market aggregates. For example, Kashyap, Stein, and Wilcox (1993) examine the ratio of bank loans to the sum of both bank loans and funds raised through issuing commercial paper by firms.

This *Weekly Letter* examines related indicators, namely, the spread between the 6-month commercial paper and the 6-month Treasury bill rates, as well as the spread between yields on long-term and short-term Treasury securities. Historically, both spreads have been useful leading indicators of economic activity. However, they are not infallible, and they failed to predict the most current recession. Furthermore, continuing financial market innovations and the changing market environment might further undermine the usefulness of some of these indicators as more assets become available and as portfolio choices and financing sources become more diverse. Thus, policymakers will have to rely on a broad range of indicators, including these new indicators.

INTEREST RATE SPREADS

There are several reasons to consider interest rates as indicators of monetary policy and future economic growth. First, the Federal Reserve has used an interest rate as one of its policy instruments. Second, macroeconomic theory suggests it is through interest rates that monetary policy actions are transmitted to the economy. For example, when the Fed increases the money supply, short-term rates drop, which stimulates activity in interest-sensitive sectors. Third, studies of the determinants of output movements conducted since the early 1980s found that when interest rates are considered, the monetary aggregates lose most of their explanatory power, suggesting that interest rates contain important information about future output (Sims 1980).

Reprinted from: *FRBSF Weekly Letter,* No. 93-12 (March 26, 1993).

Finally, the relationship between output and monetary aggregates has deteriorated in recent years (Judd and Trehan 1992). Econometric studies have revealed a loosening of the long-term relationship between money and income when the data for the 1980s are included (Friedman and Kuttner, 1992). Deregulation and innovation in financial markets are perceived to have contributed to this deterioration. The same changes also prompted policymakers to shift their focus from a narrowly defined monetary aggregate, M1, which consists of fully checkable deposits and currency in the hands of the public, to a more broadly defined measure, M2, in an effort to find a measure that retained a stable relationship with output and prices. Even with M2, however, studies on money demand have found instability in the relationship in the late 1980s.

Prompted by this experience, economists have looked at alternative indicators, such as interest rate spreads. The two spreads examined here are the difference between rates on the 6-month commercial paper and 6-month Treasury bills (the paper-bill spread), and the difference between the yield on 10-year Treasury notes and the yield on the 3-month Treasury bills (the yield curve). Figures 1 and 2 plot these spreads over the past thirty years; the shaded areas denote recessions as designated by the National Bureau of Economic Research.

Before 1990, there is a comovement over time between the indicators and detrended output. Thus, a distinct increase in the paper-bill spread was followed by a recession (Figure 1). In the case of the yield curve, it turned negative immediately prior to each of the recessions in that period (Figure 2). Generally, most large movements in the two rate spreads were associated with slowdowns in detrended output. Such relationships between the spreads and output have been confirmed by formal statistical studies (Esteller and Hardouvelis 1991; Friedman and Kuttner 1992).

WHY ARE INTEREST RATE SPREADS USEFUL?

Interest rate spreads may be helpful for predicting future movements in output for a number of reasons. First, the paper-bill spread is affected by the overall level of risk in the

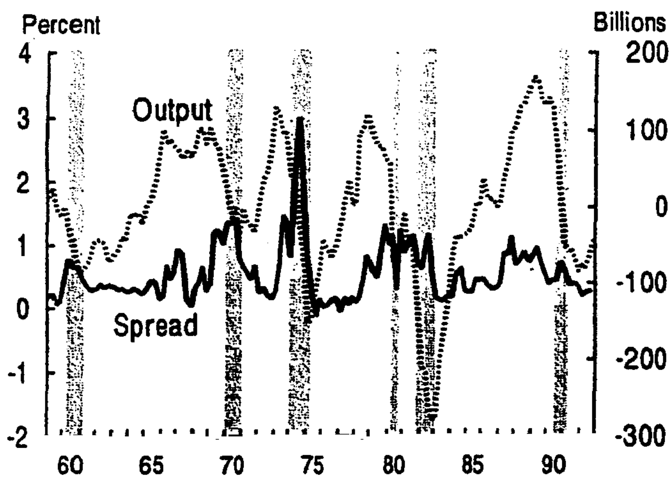


Figure 1 Paper bill spread and output.

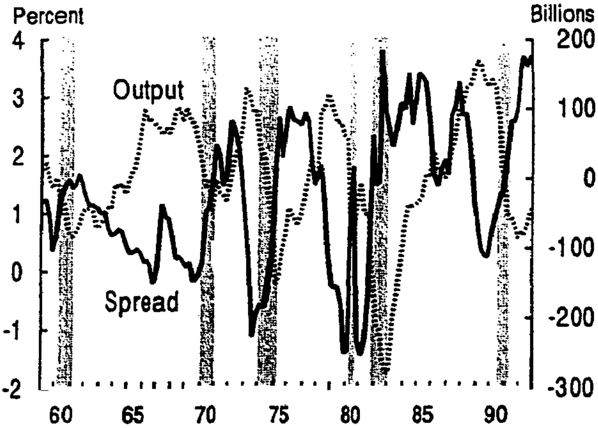


Figure 2 Yield curve spread and output.

economy, which rises and falls with the contractions and expansions in real economic activity. The default risk of commercial paper tends to increase when a downturn in the economy is imminent, driving its rate up; but, since the default risk of the government-backed Treasury bill does not rise, its rate does not go up. Consequently, the difference between the two rates tends to widen before the onset of a recession.

Second, the paper-bill spread may serve as an indicator of the stance of monetary policy. When there is a monetary policy tightening, bank lending contracts in response, and some firms issue more commercial paper to raise funds. The increase in the demand for credit in the commercial paper market will raise the commercial paper rate. This increase raises the paper-bill spread if the T-bill rate does not rise proportionately. The T-bill rate could rise, for example, if commercial banks and other investors sell T-bills from their portfolios and substitute for them commercial paper to take advantage of the higher rates of return. However, Treasury bills and commercial paper are not perfect substitutes in the portfolios of investors and banks, because the two types of securities differ substantially in terms of tax treatment, liquidity, and regulatory considerations. Thus, it is likely that a contraction in economic activity caused by a tightening of monetary policy would be accompanied by a rise in the paper-bill spread.

Movements in the paper-bill spread as well as the bank loan ratio mentioned earlier are related to the so-called "credit channel" view of how monetary policy tightening affects output. The option of borrowing in the private open market, which can mitigate a cutback in bank lending, is not fully available to all firms. Small firms, in particular, have limited access to open financial markets since, unlike large firms, they lack an established name. Thus, when monetary policy is tightened, and some larger borrowers switch to commercial paper, some small firms are denied credit and must curtail their business activities. These declines in spending then contribute to a slowdown in the pace of overall economic activity.

The third reason why the spreads may be useful is related to the yield curve, which depicts the relationship between the yields on securities of comparable risk and their terms to maturity. Most authors have attempted to capture yield curve effects by using the yield spread between long-term and short-term Treasury securities. The "expectations" theory of the term structure of interest rates argues that the expected returns from holding a long-term security until maturity should equal the returns realized from investing in a series of short-term securities for the same period of time. Thus, the difference between, say, the yields on 3-month Treasury bills and 10-year Treasury notes reflects the path of expected yields for the short-term instrument in the future.

For example, if the 10-year rate is lower than the short-term rate, it suggests that investors expect the short-term rate in the future to be lower than it is today. One reason that investors might expect short-term interest rates to fall in the future is that they expect an economic downturn. Thus, an "inversion" of the yield curve often represents a forecast of an economic slowdown.

SPREADS DURING THE RECENT DOWNTURN

As shown in the figures, prior to the 1990 business cycle peak, the interest rate spread variables did a good job of predicting recessions. Prior to each of the last five recessions the paper-bill spread shot up and the yield curve consistently turned negative. However, these spreads did *not* anticipate the 1990–1991 recession: The paper-bill spread did not show a clear increase prior to the 1990 recession; the yield curve was not "inverted." Also, there was an unusually long lead time between the dip in the spread and the onset of the downturn.

These observations are confirmed by the recent performance of sophisticated econometric models in which similar interest rate spreads were used as key information variables. One example is the National Bureau of Economic Research Experimental Recession Probability Index. The index includes both the paper-bill spread and a measure of the yield curve. The index failed to anticipate the 1990 downturn (Huh 1991).

IMPLICATIONS FOR MONETARY POLICY

This Chapter briefly examined two newly proposed indicators that are potentially useful for the conduct of monetary policy. Although these alternative indicators seem to contain information about the future condition of the real economy when looking at earlier recessions, their performance in predicting the 1990 downturn was disappointing.

Perhaps it is unrealistic to expect to find an indicator that would remain consistently useful in forecasting future movements in output for an extended period of time. First, fluctuations in output are caused by myriad factors, such as real shocks like the oil price shock of the early 1970s, and monetary shocks, like the Volcker deflation of the early 1980s. Each of those factors may affect aggregate demand and supply conditions and hence can influence financial market quantity and price variables differently. Thus, it is remarkable that interest rate spreads have been as consistently informative as they have in the past.

Second, not only can the key factors behind business cycles vary over time, but so can the overall thrust of monetary policy, which influences general financial market conditions. For example, monetary policy since the early 1980s has placed greater emphasis on controlling inflation compared to the 1970s. Thus, the information content of some long term rates might have shifted in the recent period due to changes in the expected inflation rate that makes up a part of long rates.

Third, since the 1970s, financial markets have been evolving rapidly, and the trend continues. The introduction of more sophisticated financial instruments is broadening the spectrum of available asset choices, as well as financing sources, and hence makes substitutions between assets more feasible and desirable and also makes the prices and quantities of these assets adjust more rapidly. This changing environment can make the interest spreads less informative over time.

The discussion illustrates the difficulty that monetary policymakers face in the current environment. In the absence of consistently reliable indicators to gauge future changes in economic conditions, it becomes necessary to monitor and interpret a wide set of potentially useful indicators with changing information content.

REFERENCES

Esteller, A., and Hardouvelis, G. A. 1991. The Term Structure as a Predictor of Real Economic Activity. *Journal of Finance* 46, pp. 555–576.

Friedman, B., and Kuttner, K. 1992. Money, Income, Price and Interest Rates. *American Economic Review* 82, pp. 472–492.

Huh, C. 1991. Recession Probability Index: A Survey. Federal Reserve Bank of San Francisco *Economic Review* (Fall), pp. 31–40.

Judd, J., and Trehan, B. 1992. Money, Credit, and M2. Federal Reserve Bank of San Francisco *Weekly Letter* 92-30 (September 4).

Kashyap, A., Stein, J., and Wilcox, D. 1993. Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance. *American Economic Review* 83, pp. 78–98.

Sims, C. 1980. Comparison of Interwar and Post-War Business Cycles: Monetarism Reconsidered. *American Economic Review* 70, pp. 250–257.

This Page Intentionally Left Blank

12
A Dynamic Multivariate Model for Use in Formulating Policy

Tao Zha
Economist, Federal Reserve Bank of Atlanta, Atlanta, Georgia

On March 25, 1997, the Federal Open Market Committee (FOMC) raised its key short-term interest rate target—the federal funds rate—by 25 basis points. The *Wall Street Journal* called the move chairman Alan Greenspan's "preemptive strike against inflation" (Wessel 1997). According to Greenspan, the FOMC "believes it is crucial to keep inflation contained in the near term and ultimately to move toward price stability" (1997a, 1). The FOMC described this increase "as a prudent step that affords greater assurance of prolonging the current economic expansion by sustaining the existing low inflation environment through the rest of this year and next" (Wessel 1997).

The notion of "preemptive strike" or "prudent step" connotes the most important part of policy making: the process of looking forward. Because the Federal Reserve's monetary policy has effects on the overall economy only through long and variable delays, policymakers must look forward to forecast, to the best of their abilities, how today's policy actions will affect economic conditions such as inflation in the future. This process of anticipating the future is indispensable in formulating sound monetary policy (see, for example, Cecchetti 1995, King 1997, and Blinder 1997).

The Humprey-Hawkins Act has set out multiple objectives for the Federal Reserve, including balanced growth and stable prices (Board of Governors 1994). A policy action by the Fed consists of using any one of various instruments, such as the federal funds rate and different measures of money, to pursue its multiple objectives. However, to provide clearer analysis this article characterizes monetary policy actions more narrowly as changes in the federal funds rate and the discussion concentrates on only one of the Federal Reserve's objectives—keeping inflation, as measured by the consumer price index (CPI), low and stable. In such a framework, one aspect of effectively advising policymakers is to provide a forecast of how inflation outlook changes if the Federal Reserve adopts different paths of the federal funds rate over the next two or three years. By consulting a menu of such projected outcomes, called policy projections, policymakers can decide which particular policy actions are most likely to keep inflation around the level commensurate with their objective.

Reprinted from: *Federal Reserve Bank of Atlanta Economic Review* (First Quarter 1998) 16–29.

Policy projections are essential in helping policymakers decide on policy actions. Unfortunately, obtaining an accurate estimation of such projections is a daunting task. Because the projections are based on various forecasts under different scenarios—here, alternative federal funds rate paths—the first and critical step is to develop good forecasting models (Sims 1980). It is therefore the purpose of this article to present a forecasting model that seems to overcome conceptual and empirical difficulties encountered in other models and promises to provide policymakers with a more useful tool for anticipating effects of policy.

The model, one of a class of models called dynamic multivariate models, introduces new techniques that offer two distinctive advantages. One is the ability to forecast the values of key macroeconomic variables such as inflation and output beyond a period over which these values are known, on the assumption that the trends followed within the period continue beyond it. These extrapolated forecasts are known in technical jargon as out-of-sample forecasts. The model's other advantage is its explicit structure that allows empirically coherent ways to assess the uncertainty of forecasts through error bands. These error bands are constructed so that there is a two-thirds probability that actual outcome is contained within the band.

The article first discusses dynamic multivariate modeling in general and reviews other approaches to forecasting. The discussion then turns to the model itself. After describing the specifics of the model, the article presents the model's point forecasts through the 1980s and 1990s. These forecasts represent the scenarios most likely to develop. Finally, the article shows how to use probability distributions to gauge forecast errors.

DYNAMIC MULTIVARIATE MODELING

The term *dynamic* means that economic variables influence one another through variable lags over time. For example, today's change in the federal funds rate will have consequences on the path of inflation in a year or two. The term *multivariate* implies that a set of multiple variables are examined together, not separately, in one framework. By *dynamic multivariate models* this article means a class of models that are designed to capture, in a single framework, joint movements and dynamic patterns in an array of multiple key macroeconomic variables over a particular period of time. (Technical details are discussed in Box 1 in relation to the specific model presented here.)

Other Approaches

Before explaining the key aspects of dynamic multivariate modeling, it is perhaps useful to review briefly two other approaches to forecasting and policy analysis. One approach is to use rules of thumb. Rules of thumb are often used in actual policy discussions because they may be based on theoretical work and thus can provide compelling stories to policymakers. Unfortunately, they are generally insufficient for characterizing the actual economy, and therefore forecasts derived from these rules are likely to be quantitatively unreliable. For example, one rule of thumb often referred to in the popular press is the Phillips curve relationship, which implies that whenever the unemployment rate is low (high), inflation will soon rise (fall).[1] Figure 1 displays annual inflation and the annual unemployment rate from 1960 to 1996. As the chart shows, there were times when inflation and unemployment tended to move in the same direction, not in opposite directions.

Box 1 Details of the Model

This box, heavily drawn from Sims and Zha (1998), describes the important features of the model that is used to produce the results presented in Charts 6–10. The dynamic multivariate model takes the following simultaneous equations form:

$$y(t)A(L) = \varepsilon(t), t = 1, \ldots, T, \qquad (1)$$

where $A(L)$ is an $m \times m$ matrix polynomial of parameters in lag operator L, $y(t)$ is a $1 \times m$ vector of observations of m variables at time t, and $\varepsilon(t)$ is a $1 \times m$ vector of independently, identically distributed (i.i.d.) structural shocks so that

$$E\varepsilon(t) = 0, E\varepsilon(t)' \varepsilon(t) = \underset{m \times m}{I}. \qquad (2)$$

Note that T is the sample size. To estimate system (1), the likelihood function is multiplied by a probability density function. This probability density, formally a Bayesian prior distribution, aims at eliminating the undesirable problems associated with the estimation. These problems are discussed in detail below.

The number of parameters in $A(L)$ grows with the square of the number of variables in system (1). Given the short period of macroeconomic data after World War II, traditional, ordinary least squares (OLS) estimation of a large model (for example, the eighteen-variable model studied by Leeper, Sims, and Zha 1996) becomes imprecise because of relatively low degrees of freedom and a large number of parameters. Thus, models used in macroeconomics are often of small size (say, six variables). For small models like the six variable model presented in this article, error bands on the OLS estimates of parameters are usually tight, and thus quantitative analysis from these models can be informative. Nonetheless, when a model is used for out-of-sample forecasting, one can no longer take comfort in "good" in-sample properties of the OLS estimates. Three major problems prevent reasonable out-of-sample forecasting especially over long horizons (such as two or three years out).

The first problem is a familiar one: overfitting. Because of a large number of parameters, the model tends to fit the sample unrealistically well but fails badly for out-of-sample forecasting.[1] To see how unbelievable the overfitting problem could become, Figure A displays actual values and in-sample (not out-of-sample) forecasts of the stock of M1 from January 1960 to March 1996. These in-sample forecasts, drawn directly from Sims and Zha (1998), are made as of 1959:12 from the estimated model (using the data from 1959:7 to 1996:3) without any prior distribution (that is, with OLS estimates). As shown in Figure A, one could, in 1959, predict with almost perfect precision the level of M1 stock in 1996—an incredible outcome.

Another aspect of overfitting, which has not been addressed in the textbooks, is an unreasonable extraction of business cycles into deterministic components (see Sims and Zha 1998 for technical details). This undesirable feature may have contributed to findings about substantial differences in OLS estimates across different subsamples. It may distort long-run relationships among variables in the model as well. To deal with these overfitting problems, the model used here, following Sims and Zha (1998), uses priors that favor unit roots and cointegration.[2] At the same time, the model avoids imposing exact, but likely spurious, unit roots and cointegrated relationships with a probability of one.

The third problem relutes to low degrees of freedom in most macroeconomic models. Typically, OLS estimates tend to produce large coefficients on distant lags and erratic sampling errors. One of the prior distributions used in the model here is to downweight the influences of distant lags or the unreasonable degree of explosiveness. This prior distribution is essential for ensuring reasonable small-sample properties of the model, especially when degrees of freedom are relatively low.

The prior distribution used here do not intend to encompass all briefs that are likely to improve out-of-sample forecasts. Rather, they reflect some widely held briefs that are likely to be uncontroversial. In this sense, the prior distributions are of a reference nature, and such an approach closely follows the likelihood principle.

1. Dynamic multivariate models are not the only types that produce overfitting. This problem is common across many empirical models (are Dirboid 1996b).
2. From a different perspective. Christofferson and Dirboid (1997) discuss why cointegrated relationships are important for short-term forecasts.

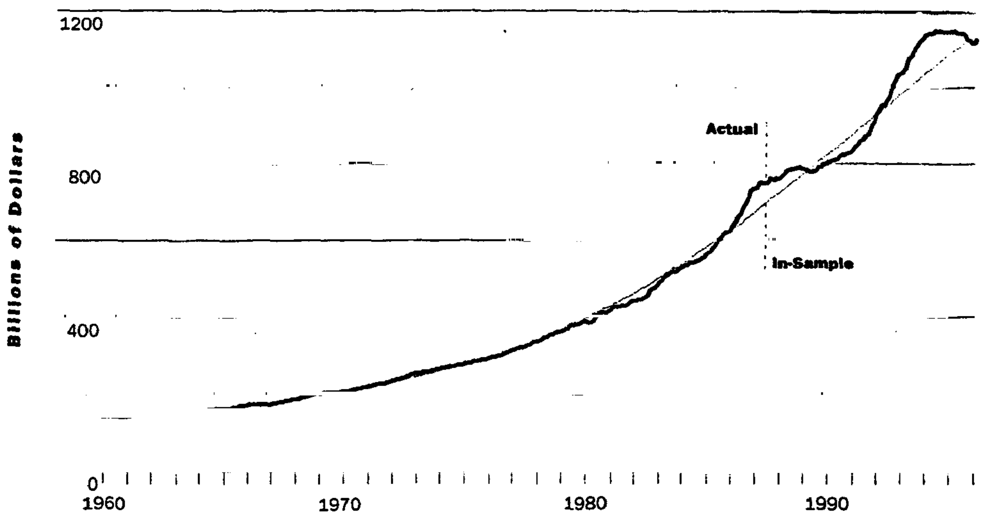


Figure A Actual and forecast M1 monthly series (1960:1–1996:3). Source: Sims and Zha 1996.

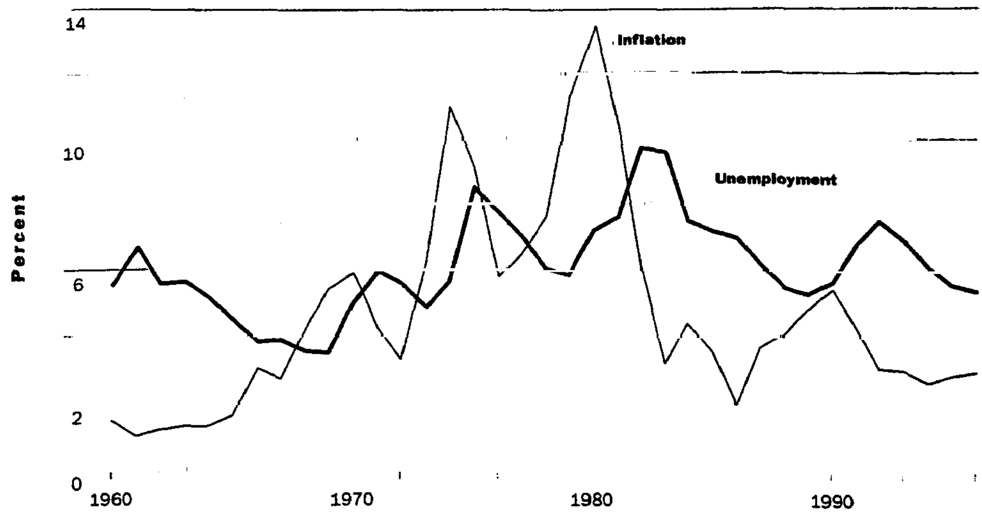


Figure 1 Annual inflation and unemployment rates, 1960–1996. Source: See Box 2.

For instance, from the early to mid-1970s, rising unemployment was coupled with rising inflation; from 1982 to 1986 both inflation and the unemployment rate fell. During other times inflation and unemployment moved in quite different fashions. Consider 1992–96, for example. During this period, the unemployment rate fell steadily but inflation stayed almost flat. If one used the negative relationship between inflation and unemployment in the 1987–91 period to predict inflation, the result would be to overpredict inflation for 1992–96.[2]

Another example of rules of thumb is the bivariate relationship between inflation and the growth rate of money. A number of economists (for example, Friedman 1992) have argued that the M2 growth rate in particular appears to have a stable relationship to infla-

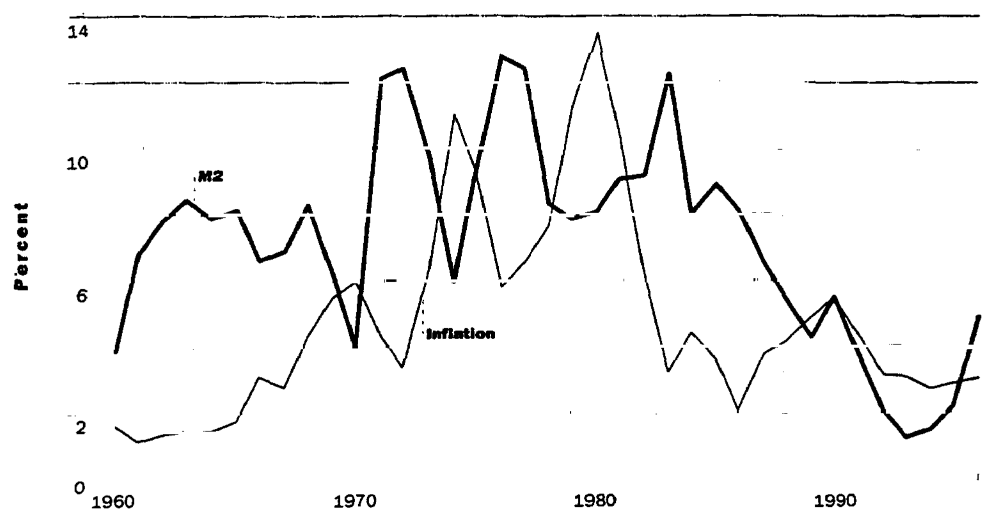


Figure 2 Annual inflation and M2 growth rates, 1960–1996. Source: See Box 2.

tion. Figure 2 displays time-series patterns of inflation and the M2 growth rate from 1960 to 1996. The M2 growth rate reached a peak three times—in 1972, 1976, and 1983. But the path of inflation after each peak was quite different. Clearly, past M2 growth rates predict future inflation through variable lags, and there are no regular patterns.

Another approach to forecasting is to link forecasts of macroeconomic variables to a large array of other variables through econometric techniques. This approach usually involves many strong assumptions or judgmental adjustments. Large-scale structural econometric models are examples of this approach. The goal of these models is to not only provide forecasts of key macroeconomic variables but also examine in detail many different sectors of the economy (Diebold 1998a). Because of their detailed, intricate nature, however, these models are often difficult to produce and evaluate independently. Furthermore, strong assumptions contained in these kinds of models, such as the Phillips curve relationship, may be at odds with the data. Judgmental adjustments consequently play roles in the model's outcomes from period to period. Such periodical adjustments make it difficult to gauge the quality of the model itself.

Distinctive Aspects of Dynamic Multivariate Modeling

Dynamic multivariate modeling offers a different approach. It is not designed to study every detail of the economy. Rather, it is designed to capture only essential elements so that the model can be readily understood and reproduced. It is closely connected to modern economic theory and usually involves only six to eight variables.[3] After the model—the array of variables, the lag length, and other assumptions—is set up, forecasts from the model will not be altered from period to period on the basis of judgments or assumptions outside the model itself. Thus, the model can be evaluated objectively.

At the same time, dynamic multivariate modeling has complex structures in the sense that it allows both contemporaneous and dynamic interactions among the macroeconomic variables. In relation to rules of thumb, dynamic multivariate models capture the relationships implied by these rules if such relationships exist in the data. In contrast to large-scale models, dynamic multivariate modeling avoids imposing strong assumptions that may be at odds with the data. Consequently, both the Federal Reserve's complex behavior and the public's expectations about future policy actions are implicitly embedded in dynamic multivariate models.

More important, dynamic multivariate modeling provides empirically coherent ways to assess the uncertainty about forecasts (Sims and Zha 1998). All forecasts have errors. The errors usually come from two sources—uncertainty about model parameters and uncertainty emanating from exogenous shocks (that is, those that cannot be predicted by the model). Dynamic multivariate modeling lays out a probabilistic structure that takes both types of uncertainties into account explicitly. When probability distributions or error bands are attached to point forecasts, policymakers will be well informed of the likelihood of future inflation.

THE MODEL

The dynamic multivariate model used in this article employs monthly data with the six key macroeconomic variables often used in the literature: the federal funds rate, the stock of M2, the consumer price index, real (inflation-adjusted) gross domestic product, the unem-

Box 2 Data Description

The model uses monthly data from 1959:1 to 1997:9 for six macroeconomic variables:

CP1 Consumer price index for urban consumers (CP1)-U, seasonally adjusted. Source: Bureau of Economic Analysis, the Department of Commerce (BEA).

Commodity Prices International Monetary Fund's index of world commodity prices. Source: International Financial Statistics.

Federal Funds Rate Effective rate, monthly average. Source: Board of Governors of the Federal Reserve System.

GDP Real GDP, seasonally adjusted, billions of chain 1992 dollars. Monthly real GDP is interpolated using the procedure described in Leeper, Sims, and Zha (1996). Source: BEA.

M2 M2 money stock, seasonally adjusted, billions of dollars. Source: Bureau of Labor Statistics (BLS).

Unemployment Civillian unemployment rate (ages sixteen and over), seasonally adjusted. Source: BLS.

ployment rate, and an index of commodity prices (see Box 2 for a precise description of the data set). The data begin at 1959:1 and end at the time when the forecast is made. The model allows these variables to interact with one another both simultaneously and through lags.[4] The lag length is thirteen months, meaning that variables in the past thirteen months are allowed to affect those in the current month.

Because the model does not allow for judgmental adjustments periodically, it aims at strong performance of out-of-sample forecasting by the model itself (see Box 1 for details). When decision making is guided by forecasts extrapolated from the model, actual data for the future period are of course not available to policymakers. Therefore, out-of-sample forecasts, with probability distributions or error bands attached, can be invaluable. The error bands of forecasts give policymakers an indication of the range of the future data. Before the discussion turns to greater detail about the use of probability distributions of forecasts, the next three sections discuss out-of-sample point forecasts produced from the specific dynamic multivariate model presented here.

OUT-OF-SAMPLE POINT FORECASTS

The 1980s

In the late 1970s inflation was accelerating to rates unprecedented in the period since 1960. Then in the 1980s inflation slowed down more quickly than the public thought possible.

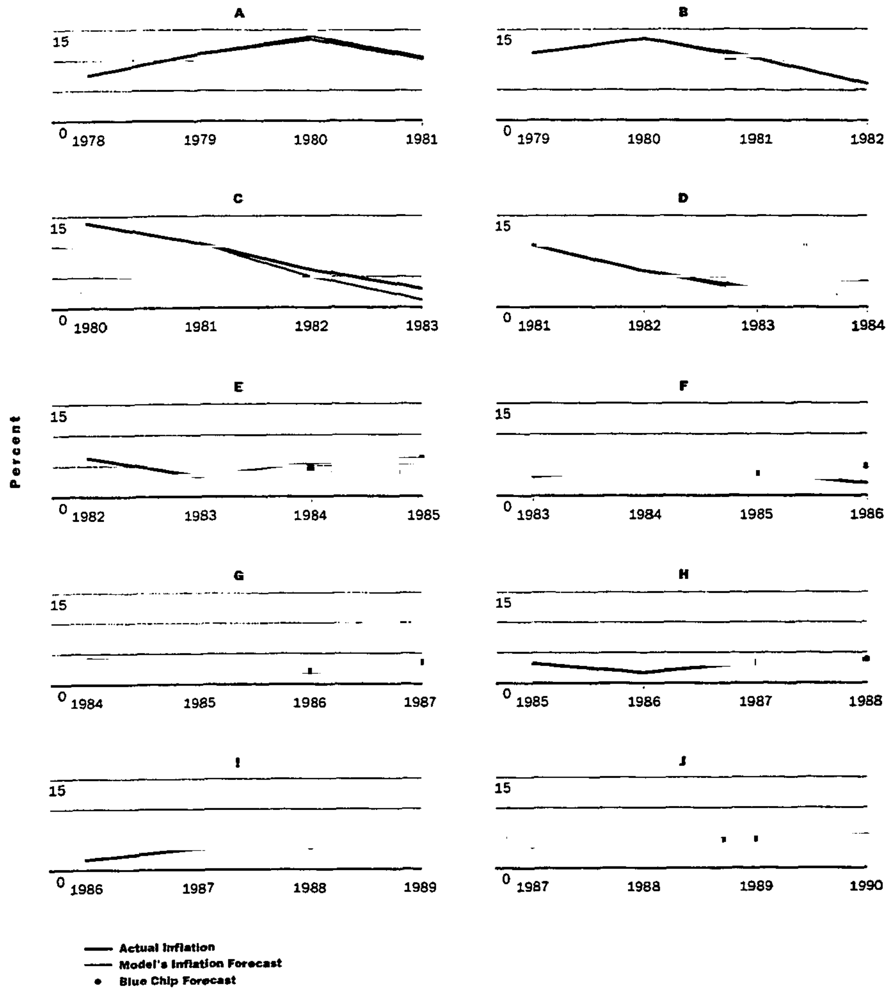


Figure 3 Point forecasts of annual inflation rate, 1980s. Source: See Box 2; *Blue Chip Forecasts.*

Thus, 1980s inflation is difficult to forecast. Figure 3 displays the model's forecasts of annual inflation through the 1980s. In each panel of figure 3, the thick line represents actual outcomes of inflation, the thin line represents the model's forecasts for the next two years, and the dots are the Blue Chip forecasts for the next two years.[5] Note that the Blue Chip forecasts at the beginnings of 1980, 1981, 1982, and 1983 are not displayed here because the new methodology introduced to compute the CPI has significantly changed figures for actual inflation before 1984. New definitions or revisions of the data always affect the accuracy of evaluating the forecasts that were made using old data at the time. Inflation figures after 1983, however, have not been altered much by subsequent data revisions. In Panel E, for instance, the Blue Chip forecasts were made at the beginning of 1984. To be comparable, the model's forecasts are also made at the beginning of 1984. In addition, Panel E displays the actual data in the two years (1982 and 1983) prior to the forecast year. Similarly, in all other panels, the forecasts for the next two years are displayed along with the actual data in the two years prior to the forecast year. For example, in Panel F, inflation forecasts in 1985 and 1986 (the thin line and dots) are made at the beginning of 1985 along with actual inflation in 1983 and 1984.

As Figure 3 shows, without periodic judgmental adjustments the dynamic multivariate model here produces quite reasonable results that are as least as accurate as the Blue Chip forecasts. In particular, the model forecasts the slowdown of inflation in the 1980s fairly well. Because the model is dynamic, it adjusts its forecasts accordingly by systematically incorporating the most recent data. For example, at the beginning of 1981 the model tends to predict that the trend of inflation will be higher than that of actual outcome (Panel B): by the time 1981 is over, the model is able to predict the downturn of future inflation (Panel C).

How do the new data in 1981 help ameliorate the forecasting performance? Remember that the model is not only dynamic but also multivariate. The new data include prices as well as the model's other macroeconomic variables, such as output, the interest rate, and the unemployment rate. The model systematically explores the dynamic relationships among these other variables and the CPI, complex though they might be. It is therefore unsurprising that inflation forecasting can be further improved by the model's ability to capture multivariate relationships in new data.

The 1990s

1990 was a turning point for inflation. Since then, inflation has declined steadily, from 5.4 percent in 1990 to 2.9 percent in 1996. Such a favorable environment has, to a large extent, surprised the public and professional forecasters as well. Indeed, many forecasting firms have overpredicted inflation for this period. The 1990s is thus considered another very difficult inflation period to forecast. Nonetheless, the model's forecasts for this period, as shown in Figure 4, look reasonable in capturing the steadily declining pattern of inflation.

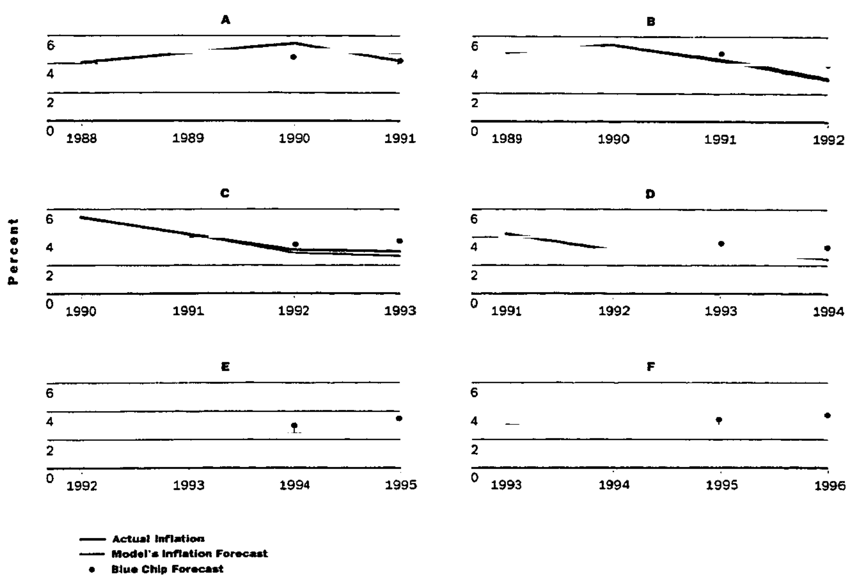


Figure 4 Point forecasts of annual inflation rate, 1990s. Source: See Box 2; *Blue Chip Forecasts.*

From Figure 4 one can see that since 1991, Blue Chip forecasts have been consistently higher than actual outcomes. The overprediction of inflation in the 1990s is consistent with simple rules of thumb such as the Phillips curve trade-off, given the declining unemployment rate after 1992. In contrast, the model's dynamic forecasts are more optimistic about the downward trend in inflation and closer to actual outcomes.

Regime Shifts

There is a common view that monetary policy follows simple rules and that these rules change from time to time in an exogenous fashion. For example, the 1979–82 period is often regarded as one in which the policy "rule" was completely changed because the Federal Reserve adopted new operating procedures to target nonborrowed reserves rather than the federal funds rate. After 1982 the Federal Reserve returned to targeting the federal funds rate. By this argument, the period after 1982 has been under a different regime than the 1979–82 period, and some empirical modelers use a sample period that begins only after 1982 as if the data before 1983 were irrelevant.

To examine this idea, the model here is reestimated using the data starting in 1983. Figure 5 reports inflation forecasts out of sample (indicated by the dots). Evidently, throwing away the data before 1983 does not improve out-of-sample forecasting in general and worsens it considerably in some cases (Panels D, E, and F).[6] One interpretation of these findings is that the Federal Reserve's behavior is complicated and cannot be characterized by discontinuous or abrupt changes in simple rules. Even among economists there is no agreement on whether the Federal Reserve's behavior during the 1979–82 period was actually different (Cook 1989). For example, Goodfriend (1993, 4) argues that "it is more accurate to refer to the period from October 1979 to October 1982 as one of aggressive federal funds rate targeting than one of nonborrowed reserve targeting." From a

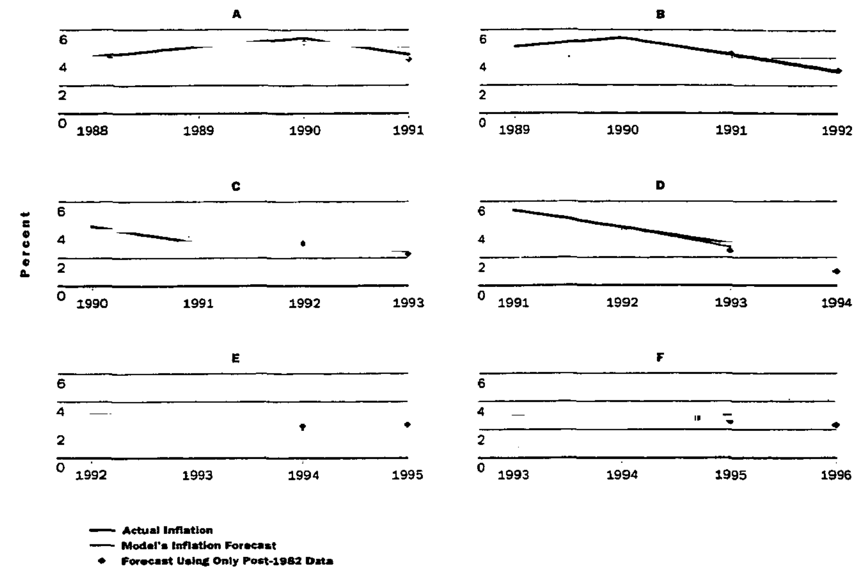


Figure 5 Point forecasts of annual inflation rate, 1990s (using post-1982 data). Source: See Box 2.

forecasting point of view, Figures 3 and 4 show that including data in this period helps forecast inflation in the 1980s and 1990s; Figure 5 suggests that in dismissing the data simply by a priori reasoning valuable information may be lost.

In a nutshell, the dynamic multivariate model that generates results in Figures 3–5 aims at accounting for both short-run dynamics and long-run relationships among the six key macroeconomic variables. Such a modeling strategy may explain the model's reasonable performance in forecasting inflation. Model-based forecasts provide benchmarks by which policymakers can decide on the best policy action given all current information. Furthermore, explicit modeling makes it easy to document the model's forecasting performance (as in Figures 3–5) and to continue improving the model or replace it by a better model when available.

THE DISTRIBUTIONS OF FORECASTS

All models at best only approximate the actual economy. No model can forecast economic conditions with perfect accuracy. Thus, policymakers must use point forecasts cautiously and carefully. When a model is used to advise policymakers, it is desirable that an explicit measure of uncertainty about the model's forecasts be provided. One effective way to measure uncertainty is to provide probability distributions of particular forecasts. With such a distribution, one is able to construct an error band on the forecast or to infer how likely the forecast is to be above or below a certain number. Error bands provide a sense of the uncertainty of economic conditions in the future and where the distribution of, say, inflation lies. Producing realistic error bands on forecasts has been a difficult technical problem. In a recent paper Sims and Zha (1998) provide ways to compute probability distributions of forecasts from dynamic multivariate models (see also Box 1).

Given probability distributions of forecasts, error bands can be constructed for any desired probability. The purpose of constructing such a band is to demarcate reasonably high and low probability regions usable for policy deliberations. The error bands used in this article are constructed so that there is a two-thirds probability that the realized value is contained within the band. With this demarcation, events outside the band are given low probability and thus should be given less weight in decision making. One should bear in mind that low probability events do occur at times but less frequently.

As an example. Figure 6 presents the same forecasts as in Panel B of Figure 3 but with error bands attached. Whereas actual inflation for 1981 falls within the error band, actual 1982 inflation lies outside the error band. The error band at the two-year forecast horizon (that is, 1982) suggests that it is unlikely that 1982 inflation would return to the 1980 level, which indeed did not occur. At the same time, the model gives low probability to values far below 7.9 percent (the lower bound of the 1982 error band). But actual inflation in 1982 did occur at the level of 6.2 percent.

Most of the time, however, actual outcomes of inflation fall within error bands. This evidence is clear from Figures 3 and 4, in which point forecasts are often close to actual values of inflation. In addition to assessing quantitatively the uncertainty of forecasts, error bands provide ways of evaluating forecasts from other sources.[7] To show an example, Figure 7 displays the model's forecasts for the real GDP growth rate in 1995 and 1996 with error bands and Blue Chip predictions.[8] Actual GDP growth is inside the error bands, but the Blue Chip 1995 forecast of GDP growth at about 3.2 percent is far outside the error band. The model suggests that such a high growth rate is unlikely for 1995.

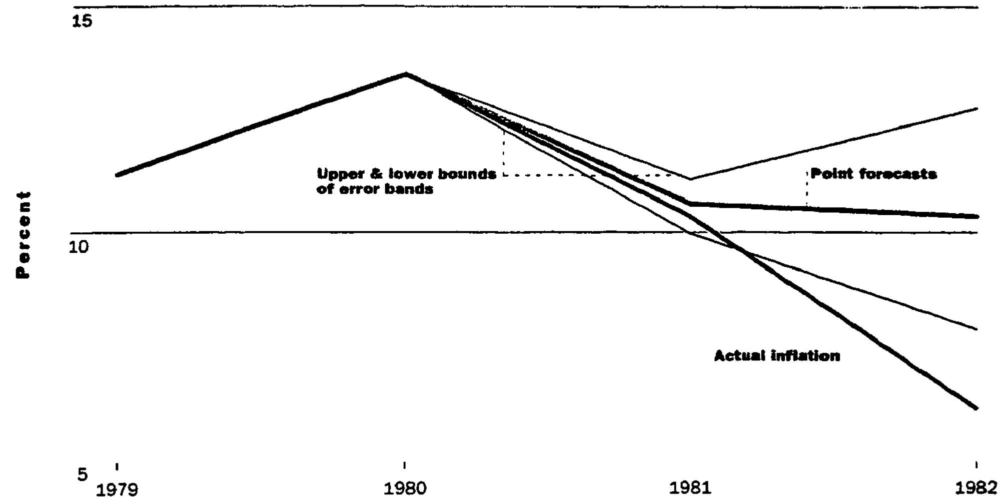


Figure 6 Inflation forecasts with error bands for 1981 and 1982.

Although the error bands considered here are sufficient for most purposes, it is sometimes useful to know the entire distribution or likelihood that a particular forecast is going to be realized. Figures 8 and 9 provide two examples. Corresponding to Figure 6. Figure 8 presents the distribution of the inflation forecast for 1982. The two dashed vertical lines mark the band that contains two-thirds probability, and the solid vertical line marks the actual outcome of inflation in 1982. The dispersed distribution in Figure 8 reflects a great uncertainty about inflation shortly after the high volatility of inflation during the late 1970s and early 1980s. Note that although actual inflation is outside the band, it is close to the lower bound of the band (that is, far away from the tail of the distribution).

Figure 9, corresponding to Figure 7, displays the distribution of the forecast of the real GDP growth rate in 1995. Again, the two dashed vertical lines mark the two-thirds

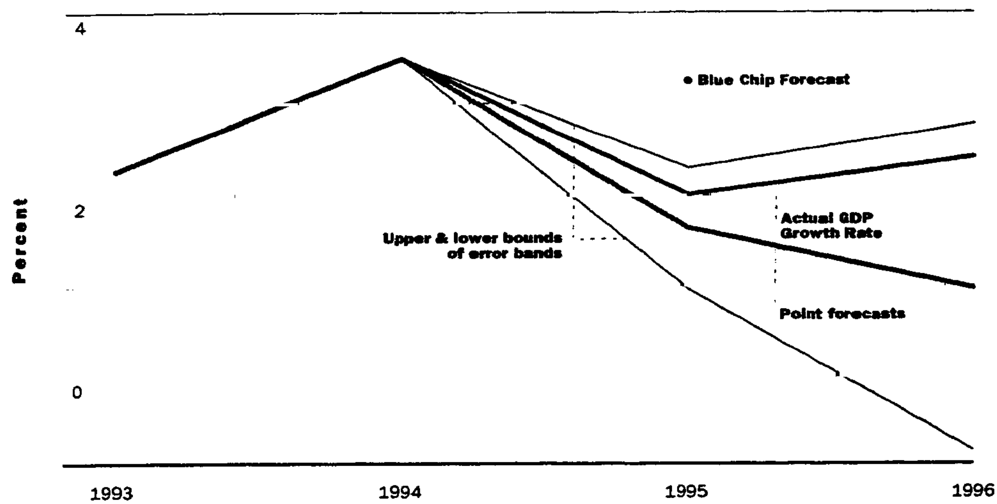


Figure 7 Real GDP forecasts with error bands for 1995 and 1996.

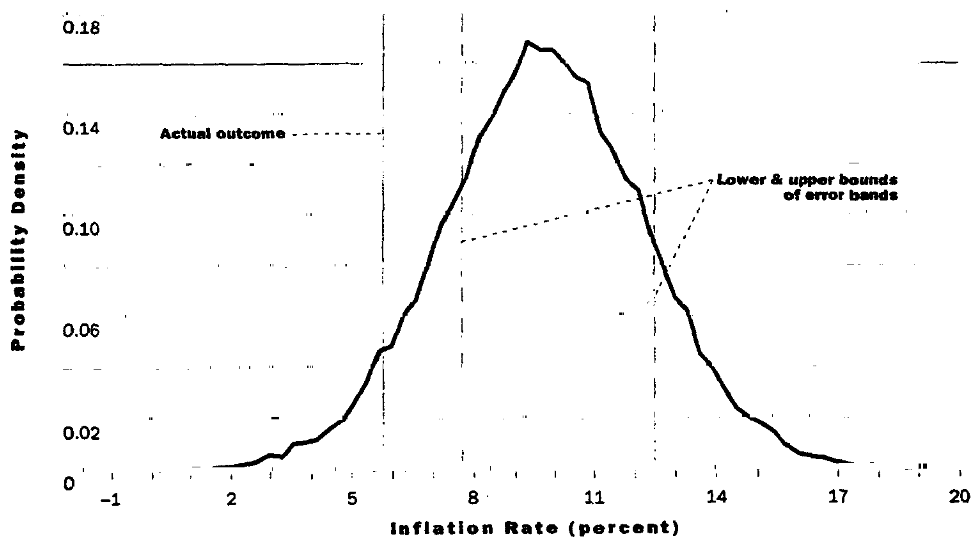


Figure 8 Distribution of inflation rate forecast for 1982.

probability band, the solid vertical line at 2 marks actual output growth in 1995, and the outer vertical line indicates the Blue Chip forecast. As can be seen in Figure 9, the Blue Chip forecast is near the tail of the distribution, implying that by the model's criterion such a forecast is very unlikely to be realized.

The discussion so far has been concerned exclusively with probability distributions or error bands around individual forecasts. While this focus is sufficient and effective for most policy analyses, it is important to bear in mind that individual forecasts are not inde-

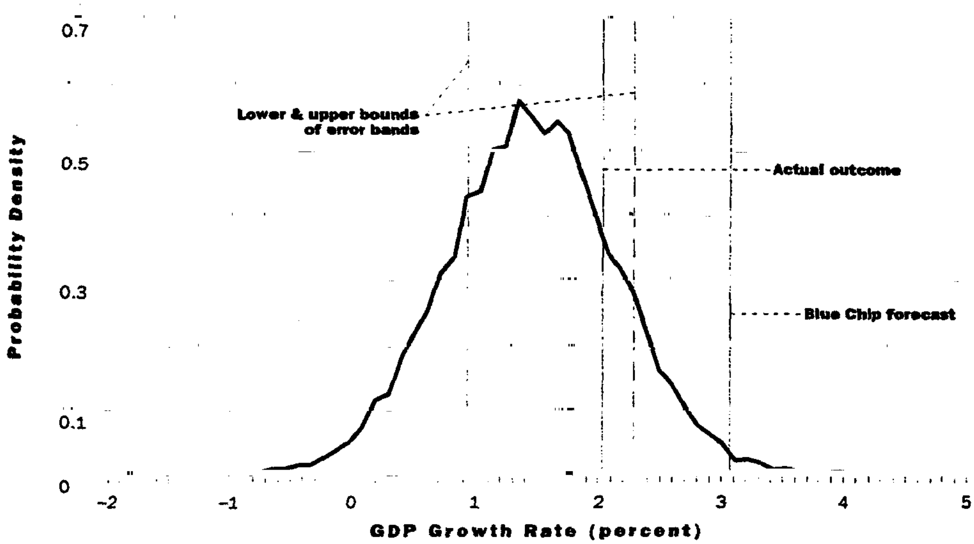


Figure 9 Forecast distribution of GDP growth rate for 1995.

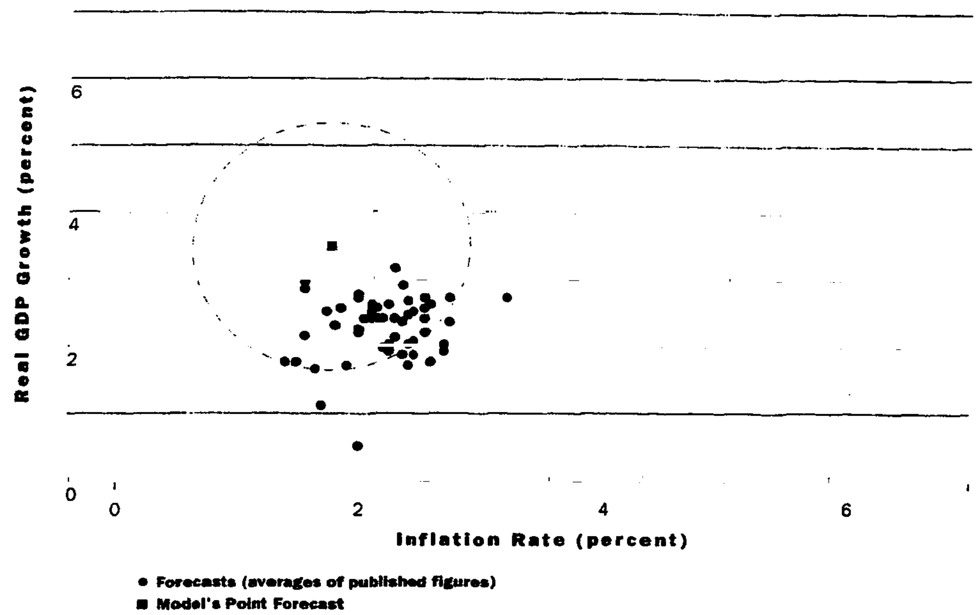


Figure 10 Error region for forecasts of real GDP growth and inflation rates for 1998. Source: *Wall Street Journal,* January 2, 1998.

pendent of one another. Indeed, because of the multivariate nature of the model, forecasts of a set of variables of interest have a joint distribution. Such a distribution can be used to construct an error region that describes how likely forecasts of, say, both high output growth and low inflation are. Figure 10, for example, displays the error region that contains both real GDP growth and inflation for 1998 with a two-thirds probability.[9] The square represents the model's point forecast. The scattered circles are forecasts of real GDP growth and inflation for 1998 from fifty-five different firms, published by the *Wall Street Journal* on January 2, 1998. Because these forecasts were submitted by December 18, 1997, the model's 1998 forecasts and error region in Figure 10 were made as of December 1997 to be as compatible with the *Wall Street Journal* forecasts as possible.[10] According to the error region, the model gives as much probability to the scenario of high GDP growth (3.5–5.5 percent) and low inflation (around 2 percent) as to that of medium GDP growth (2–3.5 percent) and low inflation (around 2 percent). But the model gives low probability to the scenario of low GDP growth (under 2 percent). The *Wall Street Journal* forecasts are unequally dispersed. At least one-fifth of the firms produced forecasts outside the model's error region. None of the firms produced forecasts that fall within the top half of the error region implied by the model.

CONCLUSION

The real world of monetary policy is complex. Because of long and variable lags in the effects of policy actions, the Federal Reserve faces a difficult task in trying to achieve its multiple objectives. The foregoing discussion concentrates on only one of these objectives—to keep the path of inflation low and stable. Given this objective, policy projections under different paths of a policy instrument (for example, the federal funds rate) are an integrated part of forward-looking policy formation. And reliable forecasts of the path of inflation are the first step in this process (Bernanke and Mishkin 1997).

The dynamic multivariate model discussed here is transparent enough to be reproduced and improved. At the same time, it is sufficiently complex to capture dynamic interplay between policymakers and the private sector. Consequently, it shows reasonable performance in forecasting as compared with other forecasts. More important, this approach provides empirically coherent ways to assess the uncertainty inherent in forecasts. Error bands or distributions of forecasts are essential for gauging this uncertainty in at least two aspects. First, they offer an assessment of how likely or realistic other forecasts are. Second, error bands inform policymakers of the uncertainty they face, reminding them of the "need to be flexible in revising forecasts and the policy stance in response to new information contradicting their previous predictions" (Kohn 1995, 233).

As Chairman Greenspan has observed, "Operating on uncertain forecasts, of course, is not unusual. . . . [I]n conducting monetary policy the Federal Reserve needs constantly to look down the road to gauge the future risks to the economy and act accordingly" (1997b, 17). The dynamic multivariate model presented in this article provides a useful tool for gauging future uncertainty and an empirically consistent way to update forecasts. It is hoped that future research will apply such a model to tasks of policy projections.

ACKNOWLEDGMENTS

The author is indebted to Eric Leeper for instrumental suggestions. He is also grateful to Rob Bliss, Jerry Dwyer, Bob Eisenbeis, Frank King, Will Roberds, Mary Rosenbaum, Ellis Tallman, and Dan Waggoner, whose detailed comments have significantly improved the article's argument and exposition. Bryan Acree and Jeff Johnson provided able research assistance.

NOTES

1. A.W. Phillips first noted such a relationship in 1958. His original study examined a temporary trade-off between changes in nominal wages and the unemployment rate in the United Kingdom over a period from 1861 to 1957.
2. The literature presents several versions of the birariate relationship between unemployment and inflation. For critical discussions consult, for example, Chang (1997), Espinosa and Russell (1997), and Staiger, Stock, and Watson (1997).
3. See, for example, Diebold (1998a) and Sims and Zha (1996) for detailed discussions.
4. The mathematical structure is similar to Sims and Zha (1998). See Box 1 for details.
5. Blue Chip Forecasts is a monthly publication based on a survey of a number of forecasters from different industries. The Blue Chip forecasts displayed in this article are the consensus forecasts.
6. Technically, these two sets of forecasts may not be statistically different when error bands are considered. Small samples such as the data after 1982 tend to give unreliable results due to erratic sampling errors, as found in, say, Cecchetti (1995). The fact that the model with only the post-1982 data delivers reasonable results may be due to recent developments in Bayesian

methods that deal with problems associated with low degrees of freedom (see Sims and Zha 1998 and also Box 1). This feature is still largely unexplored and deserves further research.

7. These sources can be various commercial firms, particular economic theories, institutional knowledge, or even ad hoc views.
8. All forecasts are made at the beginning of 1995. Although this article concentrates on inflation for simplicity of the analysis, forecasts of other macroeconomic variables such as output and unemployment are equally important for monetary policy. In particular: a number of economists believe that there is a short-term trade-off between inflation and output, especially when unexpected large shocks hit the economy (King 1997).
9. Similar to error bands of individual forecasts, error regions of joint forecasts can be constructed for any desired probability. Again, the discussion here focuses on two-thirds probability.
10. The forecasts displayed in Chart 10 are the 1998 averages of published figures in the *Wall Street Journal.*

REFERENCES

Bernanke, B. S., and Mishkin, F. S. 1997. Inflation Targeting: A New Framework for Monetary Policy? *Journal of Economic Perspectives* 11 (Spring): 97–116.

Blinder, A. S. 1997. What Central Bankers Could Learn from Academics—and Vice Versa. *Journal of Economic Perspectives* 11 (Spring): 3–19.

Board of Governors of the Federal Reserve System. 1994. *The Federal Reserve System: Purposes and Functions.* Washington, D.C.

Cecchetti, S. G. 1995. Inflation Indicators and Inflation Policy. *NBER Macro Annual 1995,* 189–219.

Chang, R. 1997. Is Low Unemployment Inflationary? Federal Reserve Bank of Atlanta *Economic Review* 82 (First Quarter): 4–13.

Christofferson, P. F., and Diebold, F. X. 1998. Cointegration and Long-Horizon Forecasting. *Journal of Business and Economic Statistics,* Vol. 16, No. 4, pp. 450–458.

Cook, T. 1989 Determinants of the Federal Funds Rate: 1979–1982. Federal Reserve Bank of Richmond *Economic Review* 75 (January/February): 3–19.

Diebold, F. X. 1998a. The Past, Present, and Future of Macroeconomic Forecasting. *Journal of Economic Perspectives,* Vol. 12, No. 2, pp. 175–192.

Diebold, F. X. 1998b. *Elements of Forecasting.* Cincinnati, Ohio: South-Western College Publishing.

Espinosa, M. A., and Russell, S. 1997. History and Theory of the NAIRU: A Critical Review. Federal Reserve Bank of Atlanta *Economic Review* 82 (Second Quarter): 4–25.

Friedman, M. 1992. *Money Mischief: Episodes in Monetary History.* New York: Harcourt Brace Jovanovich.

Goodfriend, M. 1993. Interest Rate Policy and the Inflation Scare Problem: 1979–1992. Federal Reserve Bank of Richmond *Economic Quarterly* 79 (Winter): 1–24.

Greenspan, A. 1997a. *Statement before the Senate Committee on Banking, Housing, and Urban Affairs.* February 26.

Greenspan, A. 1997b. *Statement before the House Subcommittee on Domestic and International Monetary Policy, Committee on Banking and Financial Services.* July 22.

King, M. 1997. The Inflation Target Five Years On. *Bank of England Quarterly Bulletin* (November): 434–42.

Kohn, D. L. 1995. Comment on "Inflation Indicators and Inflation Policy" by Cecchetti. *NBER Macro Annual 1995,* 227–35.

Leeper, E. M., Sims, C. A. and Zha, T. 1996. What Does Monetary Policy Do? *Brookings Papers on Economic Activity* 2:1–63.

Sims, C. A. 1980. Macroeconomics and Reality. *Econometrica* 48 (January): 1–48.

Sims, C. A., and Zha, T. 1996. Does Monetary Policy Generate Recessions? Yale University and Federal Reserve Bank of Atlanta, manuscript.

Sims, C. A., and Zha, T. 1998. Bayesian Methods for Dynamic Multivariate Models. *International Economic Review,* forthcoming.

Staiger, D., Stock, J. and Watson, M. 1997. The NAIRU, Unemployment, and Monetary Policy. *Journal of Economic Perspectives* 11 (Winter): 33–50.

Wessel, D. 1997. Fed Lifts Key Short-Term Interest Rate Amid Speculation More Rises Are Likely. *Wall Street Journal,* March 26.

This Page Intentionally Left Blank

13

Recent U.S. Intervention
Is Less More?

Owen F. Humpage
Economic Adviser, Federal Reserve Bank of Cleveland, Cleveland, Ohio

INTRODUCTION

In the past seven years, U.S. interventions in the foreign exchange market have become increasingly rare.[1] This paper offers an explanation for the reluctance to intervene. The apparent frequency with which recent U.S. interventions have stabilized key dollar exchange rates seems attributable primarily to the random-walk nature of movements in these rates. Official transactions by U.S. monetary authorities generally do not appear to improve the efficiency with which the foreign exchange market obtains information.

As discussed in the next section, U.S. interventions do not seem to affect fundamental determinants of exchange rates; rather, they change the way the market perceives and interprets information about those fundamentals. Sections II and III offer a definition of a successful intervention and ask if exchange rate movements consistent with this definition occur more frequently when the United States intervenes. Although the success criterion used is somewhat arbitrary, it encompasses outcomes that most economists would consider desirable. The empirical tests follow a methodology proposed by Merton (1981) and Henriksson and Merton (1981) and applied by Leahy (1995) in a study of U.S. profits from intervention. The results are given in section IV, and section V concludes with a brief discussion of some shortcomings that limit the interpretation of the results.

INTERVENTION AND THE CHANNELS OF INFLUENCE

Economists' doubts about the effectiveness of U.S. intervention originate with the Federal Reserve's practice of preventing official exchange-market transactions from interfering with monetary policy.[2] When, for example, the United States sells German marks in an attempt to prevent a dollar depreciation, the Federal Reserve receives payment in dollars by debiting the reserve accounts of the appropriate commercial banks. Other things being

Reprinted from: *Federal Reserve Bank of Cleveland Economic Review,* v. 33, n. 3 (Quarter 3, 1997) 2–10.

equal, this action shrinks bank reserves, the monetary base, and ultimately the U.S. money stock. The German money stock will tend to rise. Although dollar exchange rates should respond favorably, the mechanism can interfere with the inflation objectives of monetary policy when the initial underlying cause of the dollar's depreciation is anything other than a domestic monetary impulse. Moreover, if the Federal Reserve tolerated such interference, the U.S. Treasury, which has primacy regarding intervention in this country, could influence monetary policy and violate the Fed's independence (see Humpage [1994]).

To avoid possible conflicts between exchange rate and domestic price objectives, the Federal Reserve routinely offsets the monetary-base effects of U.S. intervention through open-market transactions in Treasury securities. To continue with the example of the mark begun earlier, the Fed purchases Treasuries and credits banks' reserve accounts. (The Bundesbank tends to do likewise.) Although this eliminates the most obvious, direct influence on exchange rates—relative changes in the U.S. and German money stocks—the process alters the currency composition of publicly held government debt. After the offset, the public holds fewer dollar-denominated securities and more mark-denominated securities. According to the portfolio-balance approach to determining exchange rates, if Ricardian equivalence does not hold and if investors regard these bonds as imperfect substitutes, changes in the currency composition of outstanding debt will cause the dollar to depreciate, independent of our monetary policy stance. Unfortunately, empirical studies find virtually no evidence that intervention alters exchange rates through this channel (see Edison [1993]).

Even if intervention does not alter market fundamentals, it could still influence exchange rates by affecting either the market's perception of current fundamentals or expectations about how they might change. Foreign exchange dealers face strong incentives to acquire all possible information about current and anticipated economic developments that could influence exchange rates. If these dealers are successful, current quotations will incorporate all available information, and only new information that revises traders' expectations will affect exchange rates. To the extent that traders formulate their expectations without systematic errors, revision will be random, and exchange rate changes will approximate a random walk.

Although economists generally regard foreign exchange markets as highly efficient processors of information, markets do not always respond to news instantaneously or completely. Information is costly, and some time must elapse—whether minutes, hours, or days—between the receipt of new information and its full incorporation into exchange rates. Traders' expectations can be dissimilar or highly uncertain. Consequently, monetary authorities could sometimes possess better information than other market players and could use intervention to convey it to the market. For example, a central bank could have superior knowledge about an impending change in monetary policy. Nevertheless, the notion that it *routinely* has better information than the market—even about monetary policy—remains debatable.

SUCCESS CRITERION

If U.S. monetary authorities can routinely affect the information flow within the foreign exchange market, then one would regularly observe an adjustment in the spot exchange rate when intervention occurs. Furthermore, if intervention can promote an exchange rate policy, one would expect these adjustments to conform to that policy's objective.

The stated aim of U.S. intervention policy is to counter disorderly market conditions, a goal that eludes a simple, precise, or even impartial definition. Sometimes, reported interpretations of this objective, such as reintroducing a sense of two-way risk, also elude a verifiable description in terms of exchange rate movements. At other times—especially over the period considered in this study—U.S. actions seem to signal support for the initiative of other central banks, rather than ardent conviction about the dollar. Nevertheless, official descriptions of efforts since May 1, 1990, suggest that U.S. monetary officials usually determine the success or failure of their interventions with reference to movements in spot dollar exchange rates.[3] Although the success criterion offered below is somewhat arbitrary, it is nevertheless consistent with the objective of countering disorderly markets and is readily verifiable. One could, of course, propose and test other criteria.

A General Success Criterion

Since one is never precisely certain whether the intended goal of intervention on any given day is to dampen exchange rate movements, to reverse their direction, or to encourage them along their present path, I adopt a broad success criterion—jointly expressed by (1a) and (1b) below—that subsumes all of these purposes. For U.S. sales of foreign exchange,

$$ws_t = \begin{cases} 1 & \text{if } I_t > 0 \text{ and } \Delta S_t > 0 \text{ or} \\ & \quad \Delta S_t > \Delta SAM_t, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{1a}$$

For U.S. purchases of foreign exchange.

$$wb_t = \begin{cases} 1 & \text{if } I_t < 0 \text{ and } \Delta S_t < 0 \text{ or} \\ & \quad \Delta S_t < \Delta SAM_t, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{1b}$$

The variables are defined as follows:
ws_t and wb_t are dichotomous success variables.

- I_t is official U.S. intervention, with positive values indicating sales of foreign exchange and negative values designating purchases,
- ΔS_t measures the change in the exchange rate between the morning opening of the New York market (9:00 a.m.) and the afternoon closing (4:00 p.m.), and
- ΔSAM_t measures the change in the exchange rate from the morning opening on day $t - 1$ to the morning opening on day t.

The respective parts of the dichotomous success criterion (ws_t and wb_t) take a value of one if U.S. intervention sales or purchases of foreign exchange are successful. An intervention sale of foreign exchange ($I_t > 0$) is successful if it is associated with a dollar appreciation ($\Delta S_t > 0$) or a smaller depreciation ($\Delta S_t > \Delta SAM_t$) when both ΔS_t and ΔSAM_t are negative. An intervention purchase of foreign exchange ($I_t < 0$) is successful if it is associated with a dollar depreciation ($\Delta S_t < 0$) or a smaller appreciation ($\Delta S_t < \Delta SAM_t$) when both ΔS_t and ΔSAM_t are positive.

In this paper, I_t refers to official data on actions against German marks or Japanese yen, the only foreign currencies that are subject to U.S. intervention. I assume that all such events occur in the New York market between its morning opening and afternoon closing (see Goodhart and Hesse [1993]). All exchange rates are bid quotes in German marks per dollar or Japanese yen per dollar.

The criterion pertains to movements in the exchange rate during the current day or compares current changes with movements over the previous 24 hours. In a highly efficient market, dealers' quotations will quickly incorporate useful information arising from intervention. In considering U.S. actions alone, the following tests assume that the market fully processes the relevant news about intervention on the day of the official transaction.

Because the United States and foreign monetary authorities closely coordinated their interventions during the sample period, I modified the success criterion slightly to lengthen the timing convention and to capture possible effects of foreign intervention. This was done by substituting ΔSPM_t for ΔS_t and ΔSPM_{t-1} for ΔSAM_t in expressions (1a) and (1b), where SPM is the afternoon closing exchange-rate quotation for the New York market. These substitutions measure success by comparing changes in today's closing quotation with yesterday's and by comparing movements in today's and yesterday's exchange rate. Foreign interventions, undertaken before the U.S. market opens and possibly with the acquiescence of U.S. officials, could affect the opening quotation in New York before American authorities act. Subsequent U.S. intervention may not supply any further information to the market or have any effect on the exchange rate, but one might wish to consider the overall intervention (domestic and foreign) a success.[4]

Sample Period: May 1, 1990–March 19, 1997

I applied the success criterion described in expressions (1a) and (1b) to U.S. interventions between May 1, 1990 and March 19, 1997. During this period, the nation demonstrated a growing reluctance to intervene. Initially, this hesitation appears to have resulted from a series of dissents on Federal Open Market Committee votes related to U.S. intervention policies in late 1989 and early 1990. These dissents touched on various aspects of official policy, but generally expressed skepticism about the efficacy of intervention and concern about adverse spillovers onto monetary policy (see Humpage [1994]). At this writing, the United States has not intervened in the foreign exchange market since August 15, 1995—the longest period of abstinence since the dollar began to float.[5]

Between May 1, 1990 and March 19, 1997, the United States intervened on 45 occasions against the mark and on 21 occasions against the yen (see Table 1). The vast majority of these events involved official sales of marks or yen. The number of actions during this period was far smaller than in the previous one, which had been influenced by the Louvre Accord of February 1987. In addition, instances of intervention in the sample period usually did not persist as they did in the 1987–90 period immediately following the Accord (see Figure 1). Often, they lasted no more than a single day.

Although the frequency was lower, the average amount of intervention *sales* of marks or yen was substantially greater during the sample period than in 1987–90. The average amount of intervention *purchases* of foreign exchange was smaller during the sample period, but the United States undertook very few of these.

I also break the sample period into subperiods: May 1 to July 31, 1990 and August 1, 1990 to March 19, 1997. During the first subperiod, the only U.S. intervention involved selling marks on 17 occasions. The Federal Reserve undertook these sales as agent for the Treasury's ESF. The operations were intended to adjust ESF balances and to facilitate a reversal of outstanding warehousing operations with the Federal Reserve System. A warehousing operation is a swap transaction between the ESF and the System, whereby the System acquires foreign exchange (German marks in this case) and the ESF receives U.S. dollars. The warehousing operation unwinds at a set future date (see Humpage [1994]).

Table 1 U.S. Intervention Amounts and Frequencies

	Count	Mean[a]	Standard Deviation[a]	Min[a]	Max[a]
Louvre Period: February 23, 1987–April 30, 1990					
Against German marks					
Absolute value	147	146.6	114.2	15.0	695.0
Sales of marks	110	155.1	116.6	25.0	695.0
Purchases of marks	37	121.2	104.1	15.0	395.0
Against Japanese yen					
Absolute value	147	148.6	121.8	3.0	720.2
Sales of yen	83	156.7	110.7	6.0	555.0
Purchases of yen	64	138.2	135.0	3.0	720.2
Number of observations = 805					
Sample Period: May 1, 1990–March 19, 1997					
Against German marks					
Absolute value	45	220.6	226.5	20.0	850.0
Sales of marks	39	241.2	236.8	20.0	850.0
Purchases of marks	6	86.7	21.6	21.6	100.0
Against Japanese yen					
Absolute value	21	331.6	215.5	30.0	800.0
Sales of yen	17	396.1	186.2	165.0	800.0
Purchases of yen	4	57.5	29.9	30.0	100.0
Number of observations = 1.733					

[a] In millions of dollars.
Source: Author's calculations.

Although these transactions were not designed to affect the mark-dollar exchange rate, they remain interesting because even interventions without any intended effect on exchange rates should frequently appear successful when rates follow a random walk.[6]

Success and the Random Walk

In a highly efficient market, exchange rate changes will approximate a random walk (see Baillie and McMahon [1989]). Consequently, even completely ineffectual interventions frequently seem successful.

Figure 2 illustrates this point. Imagine that at the beginning of day $t - 1$. 1.85 German marks trade for one dollar, but that over the day, the dollar depreciates 5 percent against the mark. At the start of day t, therefore, 1.76 marks trade for one dollar. Under the random-walk hypothesis, the best guess for the mark-dollar exchange rate on day $t + 1$ is 1.76, but an appreciation or a depreciation away from 1.76 is equally probable. Consequently, the chance of observing a dollar appreciation (depreciation) following the sale (purchase) of foreign exchange—even when the effort is completely ineffectual—will approach 50 percent. (One must also allow for the chance of no change.) Indeed, during the sample period, the dollar depreciated against the mark 48 percent of the time, appreciated against the mark 48 percent of the time, and otherwise remained unchanged. The results are similar for the yen–dollar exchange rate, and dropping observations that include intervention does not substantially alter the proportions.

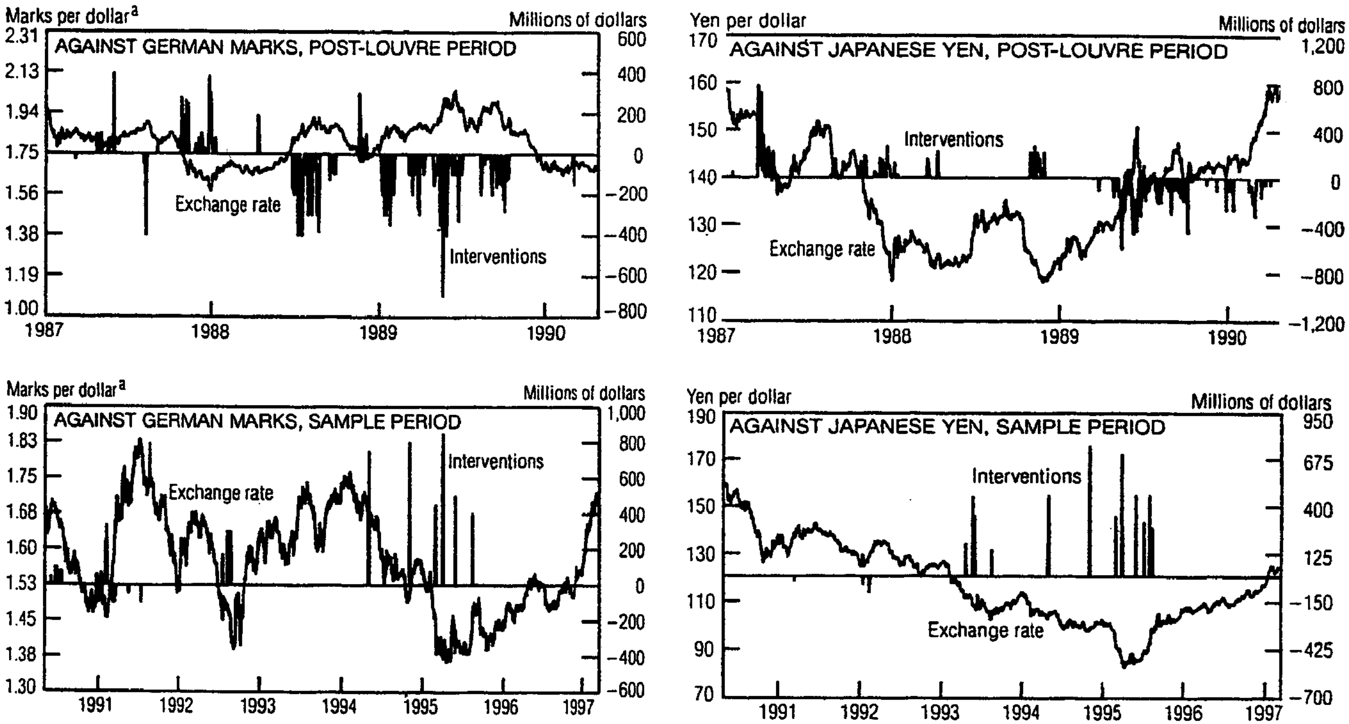


Figure 1 U.S. interventions. [a]Rounding to two decimal places causes the appearance of variation among intervals. SOURCES: Board of Governors of the Federal Reserve System; and U.S. Department of Commerce, Bureau of Economic Analysis.

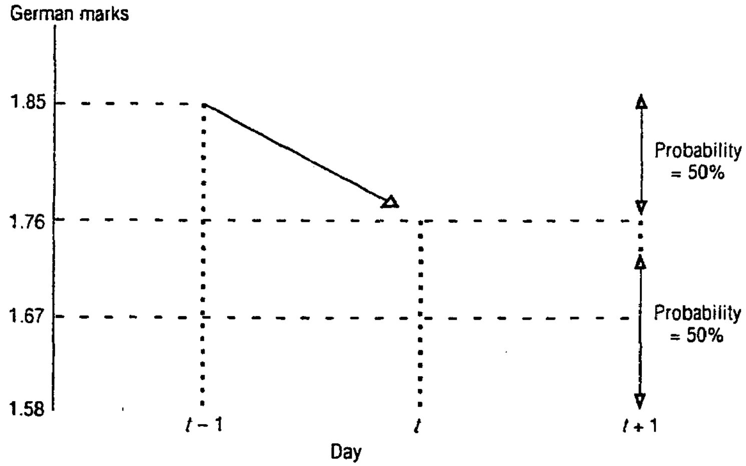


Figure 2 Random walk and the probability of success.

If the 5 percent depreciation of the dollar on day t–1 continued throughout day t, the exchange rate would be 1.67 at the start of day $t + 1$. As figure 2 demonstrates, the probability of seeing an appreciation (or a smaller depreciation) on day $t + 1$ must be greater than 50 percent. Hence, we expect that the probability of observing a success according to the general criterion—expressions (1a) and (1b)—will exceed 50 percent. The frequency of observing exchange rate movements consistent with these definitions is approximately 63 percent for the entire sample (that is, with or without interventions). This probability does not change when one drops intervention days from the sample (see Humpage [1996]).

THE PROBABILITY OF SUCCESS

If exchange rate changes followed a random walk without any drift, one could view each change as an independent event and analyze the frequency of success using standard statistical distributions (see Humpage [1996]). Exchange rate changes, however, are generally not strict random walks. Even when they exhibit such behavior over an entire, lengthy sample, they may deviate from a random walk around times when intervention occurs.[7]

Merton (1981) and Henriksson and Merton (1981) develop a nonparametric test to evaluate investment managers' ability to predict the relative performance of stocks and bonds, which have statistical properties similar to those of exchange rates. To apply the test, I treat expression (1) as an official forecast of near-term exchange rate movements that U.S. monetary authorities reveal by intervening. When the Federal Reserve sells foreign exchange, for example, it forecasts a near-term appreciation of the dollar or a smaller depreciation than recently observed. A purchase of foreign exchange has a corresponding interpretation. Evidence of exceptional forecasting skills would suggest that U.S. monetary authorities act with better information than the market and successfully convey that information to it.

The chief advantage of this procedure is that it does not require specific assumptions about either the distribution of exchange rate changes or the probabilities of individual events. A disadvantage is that it investigates only the number of times intervention is successful, not the magnitude of any effect.

To illustrate the test, consider U.S. intervention sales of foreign exchange. Following Merton, I define the conditional probabilities as

$$p1 = \text{prob } [I > 0 \mid \Delta S > 0 \quad \text{or} \quad \Delta S > \Delta SAM], \text{ and} \tag{2a}$$

$$1 - p1 = \text{prob } [I \leq 0 \mid \Delta S > 0 \quad \text{or} \quad \Delta S > \Delta SAM]. \tag{2b}$$

$$p2 = \text{prob } [I \leq 0 \mid \Delta S \leq 0 \quad \text{or} \quad \Delta S \leq \Delta SAM], \text{ and} \tag{2c}$$

$$1 - p2 = \text{prob } [I > 0 \mid \Delta S \leq 0 \quad \text{or} \quad \Delta S \leq \Delta SAM]. \tag{2d}$$

Expression (2a) is the probability that the exchange rate behaves in a manner consistent with the criterion for success—expression (1a)—and the United States intervenes. Expression (2c) is the probability that the exchange rate does *not* conform with the success criterion and the United States does *not* sell foreign exchange. The conditional probabilities defined in (2b) and (2d) are for events complementary to those considered in (2a) and (2c).

U.S. intervention sales would have no value as a forecast of the success criterion if the probability of observing an official sale of foreign exchange given a dollar appreciation or smaller depreciation ($p1$) was no greater than the probability of observing an official sale of foreign exchange given exchange rate behavior inconsistent with the success criterion ($1 - p2$). In a test of the forecast value of intervention, the null hypothesis—that U.S. intervention has no predictive value—becomes

$$H_0: p1 = 1 - p2 => p1 + p2 = 1. \tag{3}$$

In this case, traders would not modify their prior estimates of the distribution of exchange rate movements as a result of intervention.[8] Intervention has positive forecast value if $p1 + p2 > 1$. If, for example, intervention conveyed perfect information to the market, then $p1 = 1, p2 = 1$, and $p1 + p2 = 2$. Similarly, intervention would have negative forecast value if $p1 + p2 < 1$.[9]

I obtain estimates of conditional probabilities from the sample data (see Table 2). In the case of U.S. sales of German marks, for example, $n1$ is the number of successful mark sales (23): $n2$ represents unsuccessful mark sales ($16 = 39 - 23$): $N1$ is the number of *virtual successes,* that is, the number of days on which the dollar–mark exchange rate appreciates or dampens a depreciation (1,101): and $N2$ is the remaining number of observations ($632 = 1.733 - 1.101$). It follows that $E(n1/N1) = p1$ and $E(n2/N2) = 1 - p2$. Hence, $\hat{p}1 + \hat{p}2 = 0.996$.

Henriksson and Merton (1981) show that under the null hypothesis ($p1 + p2 = 1$), the number of correct interventions will have a hypergeometric distribution. This provides a direct test of the null hypothesis that depends neither on the underlying exchange-rate process nor on an underlying guess of the probability of an individual success (see Humpage [1996]). Assuming that $n1$ is a hypergeometric random variable, I reject the null hypothesis that $p1 + p2 = 1$ in favor of $p1 + p2 > 1$, if the probability of observing an equal or greater number of successes—that is, one minus the cumulative density function (1–*CDF*)—is very small. I reject the null hypothesis in favor of $p1 + p2 < 1$ if the probability of observing an equal or greater number of successes (1–*CDF*) is very large.

EMPIRICAL RESULTS

Table 2 reports the results of the experiment for the entire period (May 1, 1990–March 19, 1997), and table 3 breaks out two subperiods. As the first column of each table indicates. I

Table 2 Analysis of U.S. Interventions: May 1, 1990–March 19, 1997

	Count	Successes	Intervention percentage	Successes	Virtual percentage	$p1$	$p2$	$p1+p2$	$1-CDF$
Against German marks									
Opening to closing									
Purchases	6	4	66.7	1.147	66.2	0.003	0.997	1.000	0.341
Sales	39	23	59.0	1.101	63.5	0.021	0.975	0.996	0.670
Closing to closing									
Purchases	6	3	50.0	1.058	61.1	0.003	0.996	0.998	0.566
Sales	39	24	61.5	1.048	60.5	0.023	0.978	1.001	0.385
Number of observations = 1.733									
Against Japanese yen									
Opening to closing									
Purchases	4	3	75.0	1.131	65.3	0.003	0.998	1.001	0.181
Sales	17	12	70.6	1.143	66.0	0.010	0.992	1.002	0.260
Closing to closing									
Purchases	4	2	50.0	1.048	60.5	0.002	0.997	0.999	0.483
Sales	17	13	76.5	1.066	61.5	0.012	0.994	1.006	0.059
Number of observations = 1.733									

Source: Author's calculations.

test both purchases and sales of German marks and Japanese yen against the success criterion defined with opening-to-closing and with closing-to-closing changes in the exchange rate. As noted above, the longer time frame accommodates cooperation between U.S. and foreign monetary authorities, which occurred frequently over the sample period. Columns 2 and 3 indicate the number of interventions and successful interventions, respectively, for each category listed in column 1. Approximately 64 percent of the interventions succeeded according to criteria (1a) and (1b). In column 4, this statistic ranged from a low of 50 percent to a high of 76.5 percent. Over the sample period, U.S. interventions generally seem more successful against yen than against marks. Column 5 counts the virtual successes, that is, the number of days over which exchange rate movements conformed with the general success criteria, irrespective of U.S. intervention. When I measure exchange rate changes from opening to closing, the frequency of a virtual success is approximately 65 percent. When I measure exchange rate changes from closing to closing, the frequency is somewhat lower (approximately 61 percent). In general, therefore, the frequency of a successful intervention is not substantially different from the frequency of a virtual success. Random interventions would seem to have done as well.

Estimates of the relevant conditional probabilities follow in columns 7 and 8. It is unsettling that the value of $p1$ is very low in cases where the number of interventions is small, but nearly all are successful, as in the case of U.S. purchases of Japanese yen. Nevertheless, if all interventions were successful, $p2$ alone would equal one, and the statistical test would always reject the null hypothesis.

Column 9 records the test statistic for the null hypothesis of no forecast value, which I assume to have a hypergeometric distribution. As the final column indicates, I can reject the null hypothesis in only one case—that of U.S. sales of Japanese yen when exchange rates are measured from closing to closing. Here, one can reject the null with 94 percent confidence in favor of the positive forecast value. The inability to reject the null hypothesis for U.S. sales of Japanese yen when the tests include opening-to-closing exchange rate

Table 3 Analysis of U.S. Intervention: Two Subperiod Samples

	Count	Successes	Intervention percentage	Successes	Virtual percentage	$p1$	$p2$	$p1+p2$	$1-CDF$
				May 1, 1990–July 31, 1990					
Against German marks									
Opening to closing									
Purchases	0	0	—	40	62.5	—	—	—	—
Sales	17	11	64.7	39	60.9	0.282	0.996	1.279	0.256
Closing to closing									
Purchases	0	0	—	40	62.5	—	—	—	—
Sales	17	10	58.8	38	59.4	0.263	0.996	1.259	0.411
Number of observations = 64									
				August 1, 1990–March 19, 1997					
Against German marks[a]									
Opening to closing									
Purchases	6	4	66.7	1.107	66.3	0.004	0.997	1.000	0.344
Sales	22	12	54.5	1.062	63.6	0.011	0.985	0.996	0.751
Closing to closing									
Purchases	6	3	50.0	1.018	61.0	0.003	0.996	0.999	0.565
Sales	22	14	63.6	1.010	60.5	0.014	0.989	1.003	0.306
Number of observations = 1.669									
Against Japanese yen									
Opening to closing									
Purchases	4	3	75.0	1.090	65.3	0.003	0.998	1.001	0.182
Sales	17	12	70.6	1.104	66.1	0.011	0.992	1.003	0.265
Closing to closing									
Purchases	4	2	50.0	1.008	60.4	0.002	0.997	0.999	0.482
Sales	17	13	76.5	1.027	61.5	0.013	0.994	1.007	0.059
Number of observations = 1.669									

[a] Excludes interventions associated with warehousing.
Source: Author's calculations.

movements suggests that foreign, not U.S. intervention may have provided the forecast value, but I did not test this proposition directly.

The findings do not change when I remove the 17 sales of German marks that were undertaken to adjust the ESF's portfolio and unwind its warehousing operation. The results shown in Table 3 parallel those in Table 2. implying that, with the exception already noted, recent U.S. intervention did not systematically affect the mark–dollar or yen–dollar exchange rates.

CONCLUSION

This paper investigates the forecast value of U.S. intervention policy, using a methodology that Merton (1981) and Henriksson and Merton (1981) proposed and that Leahy (1995) applied to an analysis of intervention profits. Evidence of superior forecasting skill would imply that U.S. monetary authorities typically act with better information than the market and that intervention could alter foreign exchange traders' expectations about rates. My analysis, however, indicates that for recent U.S. interventions (May 1, 1990–March 19, 1997), this was not the case. The random-walk nature of exchange rate movements—rather than superior information—seems capable of explaining the frequency of success.

This chapter has some shortcomings that limit its interpretation. For one thing, although broad and readily verifiable, the success criterion used is necessarily arbitrary. Under some alternative criteria, intervention could appear successful and have positive forecast value. In addition, the time frame for analysis is short. A narrow period—opening to closing or closing to closing—is consistent with the notion that exchange markets are highly efficient processors of information. A broader time frame, however, might produce different results. A third shortcoming is my treatment of success as a dichotomous variable. I do not consider the possibility that the magnitude of exchange rate movements during the limited instances of successful interventions may be substantially different than at other times. Moreover, this study does not condition the probability of success on whether the United States coordinated its interventions or on the size of its transactions. Humpage (1996) found that coordination—and, to a lesser extent, large dollar amounts—increased the probability of an intervention's success. Despite these shortcomings, the results offer a plausible reason for not expecting more from less intervention.

ACKNOWLEDGMENTS

The author thanks Michael Leahy for a helpful discussion of the Merton and Henriksson tests and William Osterberg and an anonymous referee for their comments on the initial draft.

NOTES

1. Under the Gold Reserve Act of 1934, the U.S. Treasury, through its Exchange Stabilization Fund (ESF), maintains primary responsibility for the nation's interventions. The Federal Reserve intervenes both as the ESF's agent and on its own behalf, typically splitting any transactions equally between the two accounts.
2. Almekinders (1995), Dominguez and Frankel (1993), and Edison (1993) provide useful surveys of exchange-market intervention.
3. This statement is based on a survey of "Treasury and Federal Reserve Foreign Exchange Operations." which appeared quarterly in the *Federal Reserve Bulletin* between October 1990 and June 1997.
4. The tests utilize only official U.S. intervention data because foreign data are unavailable.
5. I base this statement on official published summaries of "Treasury and Federal Reserve Foreign Exchange Operations" and news accounts of currency markets. Official data used in this paper terminate in December 1995.
6. The "Foreign Exchange" column of the *Wall Street Journal* made no mention of these interventions on the days they took place.
7. The author thanks an anonymous referee for comments about the random-walk hypothesis.
8. Merton (1981; proposition III. p. 384) shows this to be a necessary and sufficient condition for the forecast to have no value.
9. Ironically, an intervention that is consistently wrong also conveys useful information to the market. The market can profit by belting against the intervention: Buy when the Federal Reserve sells foreign exchange.

ACKNOWLEDGMENTS

Owen F. Humpage is an economic adviser at the Federal Reserve Bank of Cleveland. The author thanks Michael Leahy for a helpful discussion of the Merton and Henriksson tests and William Osterberg and an anonymous referee for their comments on the initial draft.

NOTES

1. Under the Gold Reserve Act of 1934, the U.S. Treasury, through its Exchange Stabilization Fund (ESF), maintains primary responsibility for the nation's interventions. The Federal Reserve intervenes both as the ESF's agent and on its own behalf, typically splitting any transactions equally between the two accounts.
2. Almekinders (1995). Dominguez and Frankel (1993) and Edison (1993) provide useful surveys of exchange-market intervention.

REFERENCES

Almekinders, G.J. *Foreign Exchange Intervention: Theory and Evidence.* Brookfield. Vt.: Elgar, 1995.

Baillie, R.T., and McMahon, P.C. *The Foreign Exchange Market: Theory and Econometric Evidence.* New York: Cambridge University Press, 1989.

Dominguez, K.M., and Frankel, J.A. *Does Foreign Exchange Intervention Work?* Washington, D.C.: Institute for International Economics, 1993.

Edison, H.J. The Effectiveness of Central Bank Intervention: A Survey of the Literature after 1982. Princeton University. Special Papers in International Economics, No. 18, July 1993.

Goodhart, C.A.E., and Hesse, T. Central Bank Forex Intervention Assessed in Continuous Time. *Journal of International Money and Finance.* vol. 12, no. 4 (August 1993), pp. 368–89.

Henriksson, R.D., and Merton, R.C. On Market Timing and Investment Performance. Part II. Statistical Procedures for Evaluating Forecasting Skills. *Journal of Business.* vol. 54, no. 4 (October 1981), pp. 513–33.

Humpage, O.F. Intervention and the Dollar's Decline. Federal Reserve Bank of Cleveland. *Economic Review.* vol. 24, no. 2 (Quarter 2 1988), pp. 2–17.

Humpage, O.F. Institutional Aspects of U.S. Intervention. Federal Reserve Bank of Cleveland, *Economic Review.* vol. 30, no. 1 (Quarter 1 1994), pp. 2–19.

Humpage, O.F. U.S. Intervention: Assessing the Probability of Success. Federal Reserve Bank of Cleveland, Working Paper No. 9608. September 1996.

Leahy, M.P. The Profitability of US Intervention in the Foreign Exchange Markets. *Journal of International Money and Finance.* vol. 14, no. 6 (December 1995), pp. 823–41.

Merton, R.C. On Market Timing and Investment Performance. Part I: An Equilibrium Theory of Value for Market Forecasts, *Journal of Business.* vol. 54, no. 3 (July 1981), pp. 363–406.

14

Forecasts, Indicators, and Monetary Policy

Keith Sill

Senior Economist, Federal Reserve Bank of Philadelphia, Philadelphia, Pennsylvania

Conducting monetary policy is a difficult business. It's easy enough to set a policy goal such as price stability or low and stable inflation, but because monetary policy affects the economy with a lag, achieving those goals requires an ability to peer into the future. A change in the money supply or interest rates today won't affect inflation for months, or even years, down the road. Consequently, policymakers use economic models and forecasts to help them make decisions.

Historically, economists and policymakers have used two major approaches to help predict future outcomes. The first approach relies on large-scale statistical models of the economy that capture historical relationships among hundreds of economic variables. The second approach is simpler: focus on small models containing just a few variables, such as money growth and interest rates, that seem to provide information about future output growth, employment, and inflation.

Unfortunately, such models can fail to give reliable predictions, especially when factors affecting the economy change in a major way. Thus, after the oil-price shocks of the 1970s, forecasting models had difficulty predicting the simultaneous occurrence of high unemployment and high inflation. In the 1980s, major changes were made to regulations governing financial institutions, and many models were affected by the breakdown of formerly trusted links between money, inflation, and output growth. When forecasting models become unreliable, policymakers and economists find it more difficult to predict how today's actions will affect the future.

Nature abhors a vacuum in economics no less than in physics. As the old ways of divining the future came under fire, new proposals quickly emerged. Many of these proposals suggested monetary policy be guided by, or even target, certain variables that are sensitive to the market's expectation of inflation, such as commodity prices and interest rate spreads. Perhaps these indicators could give advanced warning if inflation was about to accelerate, allowing policymakers to take steps to ward it off.

Focusing on, or targeting, a small number of expectations-sensitive indicators seems to make life a lot simpler for policymakers. Why bother with complicated economic

Reprinted from: *Federal Reserve Bank of Philadelphia Business Review,* (May/June 1999) 3–14.

models or unreliable measures of the money supply if you can act on a few indicators that send good signals about future inflation or output growth? After all, building economic models is difficult, as is trying to understand the causes of inflation and growth.

Unfortunately, there is little reason to believe a monetary policy based primarily on expectations-sensitive forecasting indicators would be successful in the long run. The link between these indicators and economic outcomes is sensitive to many factors in the economy, including the way in which monetary policy is conducted. If policymakers change their method of implementing monetary policy in a substantive way, previously reliable forecasting relationships could easily break down. In addition, market expectations can change for reasons unrelated to future inflation or output growth. A policy focused too closely on expectations-driven indicators could easily be led astray. Another concern is the possibility of bad interaction between market expectations, forecasting indicators, and monetary policy. The result might be self-fulfilling changes in expectations and greater economic instability, precisely the outcome these new approaches seek to avoid. Monetary policy is more likely to succeed if it is guided primarily by variables tied to the underlying causes of inflation and economic growth and not by variables tied to *expectations* of inflation and growth.

MONETARY POLICY: A FORWARD-LOOKING VENTURE

Many economists and policymakers are convinced monetary policy can best contribute to maximum sustainable economic growth by delivering a stable price level or low and stable inflation. (See Box 1.) Because monetary policy affects inflation with a lag, policymakers who want to keep inflation low and stable must rely, to some extent, on forecasts or indicators of future inflation. In the 1960s and early 1970s, one standard way of predicting inflation relied on the Phillips curve. The Phillips curve posits a link between the unemployment rate and inflation: when the unemployment rate is low, inflation is high, and when the unemployment rate is high, inflation is low. Before the 1970s, a stable Phillips curve seemed evident in the data. But the relationship broke down following the oil-price shocks and high inflation of the 1970s.[1]

The breakdown of the Phillips curve returned attention to the relationship between money growth and inflation. A key concept for understanding the link between money and inflation is money demand, which links money to prices, interest rates, and output. Individuals and businesses need money to carry on their daily activities. The amount they wish to hold (demand) depends on the level of interest rates, prices, overall activity in the economy, and the technologies used to make payments. How does the amount of money demanded increase? The more activity there is in the economy, the more transactions are occurring, so people need more money to make purchases. When prices are higher, people need more money to buy the same goods and services. When interest rates are low, people hold more money because the cost of doing so is low.[2] Evolving payments technologies that allow people to buy and sell without holding cash or checking account balances mean less money is held. The increasingly widespread use of credit cards is one example.

These relationships are summarized by the concept of the velocity of money, a shorthand measure of how many times a year the average dollar changes hands. Velocity plays a key role in the link between money supply growth and inflation:

$$\text{inflation} = \text{money supply growth} + \text{growth of velocity} - \text{output growth.}$$

Box 1 The Benefits of Low and Stable Inflation

High and variable inflation hurts economic performance in several ways.[a] Variable inflation raises the uncertainty faced by debtors and creditors. Loan agreements typically specify an interest rate-based on the expected rate of inflation over the life of the loan. If inflation differs from what was expected. Since people generally don't like uncertainty, variable inflation can reduce the flow of credit in the economy. U.S. taxes on interest income and capital gains are not indexed to inflation, so high inflation reduces the after-tax return on saving. As a result, peoples' decisions on how much to save and spend are distorted. And then there's what economists call the "shoe-leather" cost of inflation: the higher the inflation rate, the more time and resources people spend minimizing their holdings of currency and demand deposits; that leaves less effort and fewer resources devoted to productive activity. Price stability avoids these distortions.

Monetary policy in many countries reflects an emerging consensus on these benefits. In the United States, central bankers have emphasized a commitment to price stability as a means to achieve maximum sustainable growth. The governments or central banks of some countries, including the United Kingdom, New Zealand, Sweden, and Canada, have adopted numerical targets for very low to zero inflation. Such goals do not mean policymakers are unconcerned about short-run economic developments. But policymakers recognize that monetary policy affects the economy with a lag, making it impossible to smooth out all short-run fluctuations. Moreover, they realize that sometimes a policy can have beneficial short-term results but unacceptable long-run consequences.[b]

[a]While some empirical studies, such as the one by Robert Sarro, find moderate rates of inflation are inversely related to economic growth, others find no significant correlation between the two (see the paper by McCandless and Weber). However, very high rates of inflation are associated with lower economic growth.

[b]This conclusion is based on the debate over rules vs. discretion and the time consistency of monetary policies. See chapter 12 in Bennett McCallum's book for a good discussion of discretionary vs. rule-based policymaking. The article by Herb Taylor and the one by Charl, Kehoe, and Prescott discuss time consistency and economic policy.

In principle, policymakers can use their knowledge of money velocity to control long-run inflation. If velocity is constant and money grows faster than output, the result is more money chasing the same amount of goods and services in the economy, so prices rise. Too rapid growth in the money supply eventually leads to higher prices and inflation.[3]

However, if money demand moves unpredictably, policymakers can't predict velocity and so are uncertain about the link between growth in the money supply and future inflation. Unfortunately, this appears to be what happened beginning in the 1980s (see Box 2). Economists Benjamin Friedman and Kenneth Kuttner, focusing on the period since

Box 2 Unstable Money Demand and Velocity

The increasing instability of money demand can be easily seen in plots of short-term interest rates and the velocity of money (Figures 1 and 2). Velocity is defined as nominal GDP divided by the money supply. It is a measure of how frequently money changes hands in the economy. When velocity is high, money is turning over rapidly because a relatively small quantity of money suffices to do a year's transactions in goods and services in the economy. Alternatively, low velocity means a relatively large quantity of money is held to buy goods and services, so money must not be changing hands as frequently.

Velocity is positively related to nominal interest rates. When interest rates are high, people tend to reduce money holdings because money generally earns less interest than other assets. With less money held for economywide spending, velocity tends to be high. As long as velocity can be reliably related to interest rates (or other variables), all is well. But if the relationship changes, the Fed will have trouble figuring out how much money to supply to the economy.

For M1, the trends of velocity and interest rates were similar up until the early 1980s: interest rates were generally rising as was velocity. In addition, velocity as was not very sensitive to swings in interest rates. The combination made it easy to predict M1 velocity. After that, velocity started to show much wilder swings than before, and even though interest rates came down quite a bit in the 1990s, velocity didn't. The relationship between M2 velocity and interest rates also broke down, though the timing is different. The link between velocity and intrest rates held up fairly well until the early 1990s: M2 velocity moved closely with interest rates. After that, M2 velocity increased dramatically at the same time interest rates declined.

The breakdown of the link between velocity and interest rates is a symptom of unstable money demand. Policymakers found the money supply was becoming unreliable in that it did not do well in predicting future economic outcomes.

1975, have shown that the money supply was a useful predictor of the future economy from 1975 until the mid-1980s. But from about the mid-1980s on, they found that the money supply (M1 and M2) had little or no predictive power for future inflation and output. A study by economists Arturo Estrella and Frederic Mishkin, which used the monetary base and M2 as measures of the money supply, came to a similar conclusion.[4] Estrella and Mishkin looked at data since 1979 and found that the money supply didn't predict future inflation or output well, and it's not a very good indicator of the stance of monetary policy either. In short, something happened in the 1980s that caused a serious breakdown in the link between money and economic activity.

Why did the breakdown occur? Likely candidates include technological innovation, changing financial regulations, and the financial innovation that both responded to and

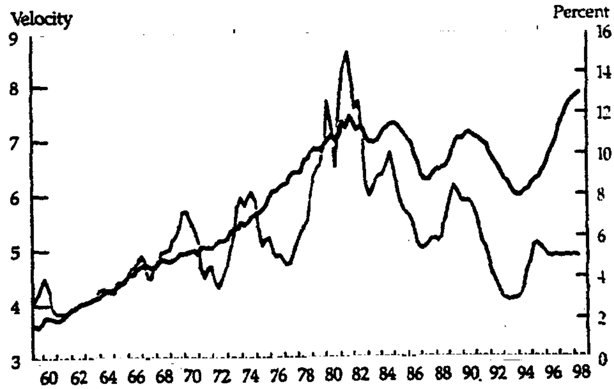


Figure 1 M1 velocity and 3-month T-bill (two-month moving average) interest rate.

drove regulatory change. For example, before 1980, banks faced regulatory limits on the interest rates they could pay on certain types of deposits counted in the money supply. In the late 1970s, short-term interest rates were very high, so consumers were reluctant to hold funds in deposit accounts that didn't pay competitive rates. Banks responded by finding a way to, in effect, pay interest on checking deposits: they created negotiable order of withdrawal (NOW) accounts. Many other financial innovations have occured over the last 20 years or so, including money market mutual funds, expanded use of credit cards, and widespread use of ATM machines. Financial innovations such as these can have a great impact on money demand and make that demand unstable.[5]

Before the 1980s, simple models that used the growth rate of the money supply to predict future inflation worked fairly well, once they accounted for the oil-price shocks of the 1970s. The instability of money demand in the 1980s meant the models started to have larger-than-usual forecasting errors. In short, the breakdown of the money-demand relationship meant money growth was no longer a reliable variable on which to anchor monetary policy.

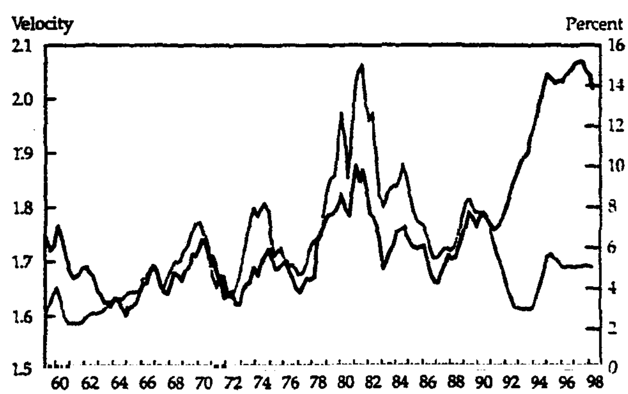


Figure 2 M2 velocity and 3-month T-bill (two-month moving average) interest rate.

IF NOT MONEY GROWTH OR THE PHILLIPS CURVE, WHAT?

Once it became clear that the simple Phillips curve and growth of the money supply had become increasingly unreliable for setting monetary policy, a search for alternative indicators commenced. In the mid-1980s, inflation averaged about 4 percent, but people retained vivid memories of inflation that exceeded 10 percent in 1979 and 1980. Concerns about inflation were reflected in early suggestions that monetary policymakers try to stabilize commodity prices, an action that, it was argued, would lead to general price stability in the economy.[6]

The notion that monetary policy should stabilize commodity prices or commodity-price indexes met with a good deal of skepticism. Many commodity prices are very volatile and seem to move in response to factors other than movements in the general price level. If policymakers tried stabilizing commodity prices, they would have to respond to sources of volatility that have little to do with inflation. Though the proposal to stabilize commodity prices was flawed, some economists and policymakers argued that commodity prices could still play a useful role in formulating monetary policy. If commodity prices were good *predictors* of inflation, they might be useful indicators of inflationary pressures in the economy, even if they were not good targets for policy.[7] After all, a desirable monetary policy is one that responds to inflationary pressures well before actual inflation begins to rise. Perhaps by focusing on other predictors of inflation, rather than variables such as money growth and unemployment, policymakers could better achieve goals of low inflation and long-run price stability.

EXPECTATIONS, INDICATORS, AND MONETARY POLICY: CAVEATS

Unfortunately, even if a proposed indicator has been a good predictor of inflation, it may not be a good guide for monetary-policy decisions. If the indicator is driven by market expectations of inflation, it may respond to changes in underlying causes of inflation. But it may also respond to other factors, ones unrelated to future inflation. In the latter case, policymakers face greater uncertainty about the signal an indicator gives: false signals are likely.

Suppose policymakers rely on variables, such as the price of gold, that are driven by market expectations about inflation. If expectations about future inflation change, it doesn't matter whether those expectations are well founded—the price of gold today will change. If many people start to believe inflation will be higher in the future, that belief will be reflected in higher gold prices, even if the fundamentals of the economy are unchanged. Other potential indicators that are sensitive to market expectations are spreads between long-term and short-term interest rates and survey-based measures of inflation expectations. Of course, these variables respond to other economic factors besides changes in inflation expectations. Gold prices respond to discoveries of new gold deposits and extraction techniques and to wars. Interest rate spreads may change in response to financial uncertainties, such as those brought about by the recent Asian crisis. In addition, changes in the way expectations are formed can lead to large swings in expectations-driven indicators.

Pitfalls When Using Expectations-Driven Indicators for Policymaking

Michael Woodford's article discusses several pitfalls for policymakers who focus too narrowly on expectations-driven indicators that have forecasted inflation in the past.

First, as people recognize that monetary policy has shifted its focus to a new set of forecasting indicators, those indicators can become unreliable. The basis for this idea is the Lucas critique, named after Nobel Prize-winning economist Robert E. Lucas, who pointed out that the beliefs of households and firms about the conduct of monetary and fiscal policy are important determinants of economic behavior. (See Box 3.) In making economic decisions, households and firms take into account their beliefs about the course of economic policy and how it will affect them. When policy changes, new beliefs are formed, and economic decisions can then change.

What does this mean for monetary policy? Suppose policymakers announced that henceforth the price of gold will be targeted because gold prices are sensitive to expecta-

Box 3 The Lucas Critique

In 1976, Robert Lucas published a now-famous article arguing that the large-scale macroeconomic models then in vogue could, in principle, provide no usful information about the actual consequences of alternative monetary and fiscal policies, even though the models might be very good at short-term forecasting.[a] The argument became known as the Lucas critique and has been very influential in macroeconomics.

The core of Lucas's argument is the recognition that expectations about policy are an important ingredient in the behavior of households and firms. Implicit in the estimated large-scale models were a set of beliefs by households and firms about the future course of monetary and fiscal policy. If policies changed substantially, new beliefs would be formed, and households and firms would change their behavior. Thus, it made little sense to use models estimated under one set of beliefs to evaluate the consquences of economic policies that would generate new beliefs and different behavior. Models estimated under the old beliefs would likely give incorrect answers.

One way to think about Lucas's point is in terms of a game: if the rules of the game change, the players will adapt to the new environment by changing their behavior. We can develop an analogy in terms of a football game.[b] Football fans will have noticed that, during the 1998 season, the Philadelphia Eagles almost always punted when confronted with a fourth down in their own territory. It didn't matter whom the Eagles were playing or where; it was a safe bet they would punt on fourth down when in their own end of the field. On the basis of this history, we would do well to predict that the Eagles would continue to punt on fourth down, and we wouldn't need any understanding of football to be right most of the time. But suppose we want to analyze a change in the rules. The new rules state that teams get six attempts to make a first down. For anyone who understands football, it is clear that our old model, which says the Eagles punt on fourth down, will not predict well in the new environment: the Eagles will change their behavior in response to the new policy.

[a]See the article "Econometric Policy Evaluation: A Critique."
[b]The football analogy was originally developed in a 1986 article by Thomas Sargent.

tions about inflation. Once policy is switched, the old link between gold prices and inflation would almost certainly change, since individuals would buy and sell gold with the knowledge that policymakers are responding to its price. Policymakers would then have to figure out how to interpret the new relationship between gold prices and inflation expectations. The same potential problem would be faced with any variables driven primarily by expectations about inflation or output growth. If policymakers begin to target these variables, it may well happen that their usefulness as forecasters of future inflation and growth will change dramatically.

A classic example of this phenomenon is what happened to the Phillips curve in the early 1970s. As discussed earlier, the Phillips curve suggests an inverse link between the unemployment rate and inflation. Before the 1970s, this relationship was apparent in the data. But once policymakers accepted higher inflation to avoid higher unemployment in the 1970s, they found the relationship changed. Households and firms revised their expectations about future inflation as inflation remained high. Expecting higher inflation, workers adjusted their views about the real payoff to working additional hours and about their wage demands; the result was that higher unemployment no longer meant lower inflation. The old link between inflation and unemployment deteriorated when workers changed their beliefs about future inflation. Nowadays, most economists reject the notion of a stable long-run tradeoff between inflation and unemployment.

A second pitfall is that just because a variable doesn't forecast inflation or output growth doesn't mean it should be ignored. An example from outside the world of economics might be helpful. If you step on the gas pedal of a running car, the car will accelerate. But suppose you are in a dual-control car, and whenever you step on the gas, someone sitting next to you steps on the brake. Someone who only sees you pushing on the gas pedal might falsely conclude there isn't much of a relationship between that action and the car's acceleration.

So, too, with economic variables. Suppose accelerating wage growth indicates higher future inflation. If the central bank used tighter monetary policy to offset accelerating wage growth, it would appear as if there were little or no relationship between wage growth and inflation—even though wage growth was an important variable for policymaking. So, a variable may lack forecasting power because it is already being used in making decisions. In that case, it looks as if the variable is not helpful in making forecasts, even though it is tied to subsequent outcomes. In such a case, we would need a detailed statistical model of the underlying relationships in order to uncover the usefulness of the variable.

There is a third pitfall to beware of when using expectations-driven variables to guide monetary policy: the possibility for bad interactions between policy actions and forecasting variables. Expectations could become self-fulfilling, and the economy could become more volatile. Suppose the central bank uses the following policy rule: whenever three-month interest rates rise more than one percentage point above 10-year interest rates, the central bank will increase the growth rate of the money supply, and whenever 10-year interest rates rise more than one percentage point above three-month interest rates, the central bank will decrease the growth rate of the money supply. This policy might seem sensible, since, in the past, when three-month rates rose well above 10-year rates, the economy often ended up in a recession, and when 10-year rates rose much higher than three-month rates, it often signaled accelerating inflation.

Imagine that for some reason, perhaps an unexpected jump in oil prices, people think there will be a temporary rise in inflation, one that will taper off in the next year or so. The rise in expected inflation causes three-month interest rates to jump up right away; once inflation tapers off, rates are expected to fall back to normal levels. As a result of this expected pattern, three-month interest rates rise above 10-year interest rates. If policymakers respond too aggressively and pump a lot of money into the economy, the belief that there will be higher inflation could be ratified by substantially higher actual inflation. Thus, the mere expectation of inflation could become self-fulfilling, and the economy could destabilize under this policy rule. The key factor is that interest rate spreads are heavily influenced by expectations. Policymakers may find they are reacting to these expectations in a way that affirms changes in beliefs that aren't tied to economic fundamentals.

BETTER ECONOMIC MODELS TO THE RESCUE? NOT YET

The many problems with using expectations-influenced forecasting variables to guide monetary policy suggest we shouldn't be too quick to turn away from economic models when trying to understand the behavior of the economy and the implementation of monetary policy. Many economists argue that forecasting models broke down in the 1970s and 1980s because they were not built up from first principles of household and business behavior: they were not sufficiently explicit about the underlying factors governing household choice and business investment. Ideally, the way to assess how forecasting relationships change when policy changes is to have a model that is precise about what causes what in the economy. Such models can trace the effects of policy changes on the economy because they are explicit about the links between variables that cause inflation and growth; the decisions of households, firms, and policymakers; and final outcomes.[8] Since the models are explicit about the fundamental determinants of economic behavior and the formation of expectations, alternative monetary policies can be fed into the models, and the effects on variables such as inflation, employment, and output growth can be examined. In addition, we could, in principle, trace how expectations-driven variables are affected by alternative policies.

This is a great idea in theory, but where do we get such models? Macroeconomists have been hard at work on models of the economy for quite some time, but there is, as yet, no consensus on models or modeling strategies.[9] Some economists believe business-cycle fluctuations are driven primarily by demand-side factors, such as the money supply, investment, and consumption. Others attribute business-cycle fluctuations primarily to supply-side factors, such as technological progress. As a result, we have a menu of models from which to choose when investigating the consequences of alternative monetary and fiscal policies. This variety of models may not reveal truth, but it does help us analyze which differences in models are important for generating different predictions about policy responses and outcomes. Different models give us alternative ways to sort out and interpret economic data and allow us to frame questions and investigate implications more clearly.

WHAT CAN POLICYMAKERS DO?

Absent a reliable link between the money supply and inflation or a completely trustworthy model of the economy, how can a successful policy be implemented? Since the early 1990s, policymakers have found themselves in the position of not having a completely trustworthy indicator for use in setting monetary policy. Despite this, the performance of the economy has been quite good: inflation is low and the economy has been expanding since April 1991.

Economists have investigated an approach that potentially avoids the pitfalls of narrowly focusing on expectations-driven indicators to guide monetary policy: design explicit rules that tell policymakers how to adjust variables they control directly (the federal funds rate or the monetary base) in response to observed deviations of target variables (inflation or nominal GDP) from values policymakers consider desirable.[10] Such rules are less prone to the problems that arise when people change the way they form expectations. However, these rules can still be adversely affected by a change in the economy that alters the links between the variables policymakers control and the variables they ultimately care about.

Frederic Mishkin, in a 1997 article, argues that many central bankers have been following what he calls the "just do it" strategy, in which preemptive moves are made against inflation and recession. This strategy differs significantly from one that focuses narrowly on a small set of expectations-driven indicators. Under the "just do it" strategy, policymakers pay attention to *many* economic variables and use a *variety* of models in an effort to cut off inflationary pressure well before inflation actually increases. They also act to forestall a recession before it begins. There is no clearly articulated, explicit strategy for making monetary policy, but the strategy is coherent nonetheless. The policy is not very transparent, since markets don't know for certain how policymakers are reading the economy at any given moment. But the policy has been pretty successful so far: U.S. inflation is below 2 percent, the unemployment rate is well below 5 percent, and real output has been growing for eight years.

But even under such a strategy, policymakers have to look at economic indicators to judge the stance of monetary policy and assess inflationary pressures. Thus, the warnings we raised about expectations-driven indicators are still important. That is not to say that such indicators are not useful. Policymakers must get a sense of how financial markets expect the future to unfold. However, focusing too narrowly on these variables can quickly lead to the problems enumerated above.

Mishkin points out some drawbacks to the "just do it" approach that could lead to trouble in the future. For one, the lack of transparency—the fact that markets can't look at a specific set of indicators and infer how policymakers are reading the economy—could result in greater financial and economic uncertainty. As a result, the economy may not operate as efficiently as it otherwise might. Another drawback is that the success of the "just do it" approach depends on the individuals who make policy decisions. Individual policymakers can differ in their abilities and in their ranking of various policy objectives. Under a "just do it" approach, a change in the persons making policy can more easily lead to a change in economic outcomes.

CONCLUSION

The fact that monetary policy affects the economy with a lag means that central bankers have to make decisions today based on how they expect policy to affect output and inflation in the future. In assessing the likely consequences of policy actions, policymakers must pay attention to variables tied to the ultimate causes of inflation. In the past, the money supply filled that role, but it became increasingly unreliable as a tool for the conduct of policy. The alternative of focusing on a narrow set of expectations-driven indicators raises its own set of problems. Forecasting indicators sensitive to expectations can easily lead policy astray and become unreliable when they become the focus of policy.

The best way to understand the interrelationship of monetary policy and forecasting indicators is to develop models of the economy that are explicit about the fundamental determinants of households' and firms' choices about spending and investment. By focusing on variables tied to the fundamental causes of inflation and not solely on expectations-driven variables, monetary policymakers are more likely to achieve the goal of low and stable inflation. Building better models of the economy is an ongoing project for macroeconomists.

Unfortunately, policymakers do not yet have the luxury of models that closely match the U.S. economy and are always usable for policy analysis. Consequently, the current policy regime is eclectic: policymakers look at many variables and different models to gauge the appropriate stance of policy and the degree of inflationary pressure.

NOTES

1. For more on the Phillips curve, see the article by Robert King and Mark Watson. The 1999 article by Thomas Sargent contains a discussion of the breakdown of the Phillips curve in the context of monetary policy.
2. Currency and demand deposits, a significant portion of the money stock, pay no interest, but nonmonetary assets, such as Treasury bills and bonds, do. The higher interest rates are, the more interest income people forgo by holding currency and demand deposits, so they hold less money. If interest rates decline, people will hold more money.
3. The story is more complicated than we have made it seem. When the Fed increases the money supply, it does so by purchasing bonds in an open-market operation, which increases bank reserves. Banks then increase lending, which results in more money circulating in the economy. In addition, changes in the money supply are associated with changes in real activity, such as employment and output, at least in the short run.
4. The monetary base consists of the currency component of the money supply plus bank reserves.
5. There is an extensive literature on the instability of money demand. See the review article by Stephen Goldfeld and Daniel Sichel.
6. Michael Woodford's article has references to proposals that monetary policy be guided by commodity prices.
7. Some advocates for using commodity prices to gauge inflationary pressures came from within the Federal Reserve System. See the speeches by Federal Reserve Board governors Wayne Angell, Robert Heller, and Manuel Johnson. In general, the empirical evidence suggests that commodity prices are not very good predictors of inflation. See the article by Michael Woodford for a brief review of the literature and references.
8. Robert E. Lucas's 1977 article contains a nontechnical discussion of many of these points.
9. Much work has been done over the last 20 years on building better foundational models for policy analysis, for example, the paper by Eric Leeper and Christopher Sims, and the book edited by Thomas Cooley. These sources offer examples of what economists call stochastic, dynamic, general equilibrium models. One key difference between these newer models and older, large-scale statistical models is that the newer models have restrictions across equations that account for how people respond to perceived changes in monetary and fiscal policies. In addition, the new models are based explicitly on the utility and profit-maximizing behavior of households and firms.
10. We have in mind Taylor's rule and McCallum's rule. See the article by John Taylor and the one by Bennett McCallum for details.

REFERENCES

Angell, W. A Commodity Price Guide to Monetary Aggregate Targeting. Paper presented to the Lehrman Institute, 1987.

Barro, R. J. Inflation and Economic Growth. Bank of England *Quarterly Bulletin,* 35 (May 1995), pp. 166–76.

Chari, V. V., Kehoe, P. J. and Prescott, E. C. Time Consistency and Policy. In Barro, R. ed., *Modern Business Cycle Theory.* Cambridge: Harvard University Press, 1989, pp. 265–305.

Cooley, T. F., ed. *Frontiers of Business Cycle Research.* Princeton: Princeton University Press, 1995.

Estrella, A., and Mishkin, F. S. Is There a Role for Monetary Aggregates in the Conduct of Monetary Policy? National Bureau of Economic Research Working Paper 5845 (November 1996).

Friedman, B., and Kuttner, K. N. A Price Target for U.S. Monetary Policy? Lessons from the Experience with Money Growth Targets. *Brookings Papers on Economic Activity,* 1 (1996), pp. 77–125.

Goldfeld, S., and Sichel, D. E. The Demand for Money. In B.J. Friedman and F.H. Hahn, eds., *Handbook of Monetary Economics.* Vol 2. Amsterdam: North-Holland, 1990.

Heller, R. Anchoring the International Monetary System. Paper presented to the Heritage Foundation, Washington, D.C. (March 1987).

Johnson, M. Current Perspectives on Monetary Policy. *Cato Journal,* 8 (1988), pp. 253–60.

King, R. G., and Watson, M.W. The Post-War U.S. Phillips Curve: A Revisionist Econometric History. *Carnegie Rochester Conference on Public Policy,* 41 (1994), pp. 157–219.

Leeper, E. M., and Sims, C. A. Toward a Modern Macroeconomics Model Usable for Policy Analysis. *NBER Macroeconomics Annual* (1994), pp. 81–118.

Lucas, R. E. Econometric Policy Evaluation: A Critique. *Carnegie Rochester Conference on Public Policy,* 1 (1976), pp. 19–46.

Lucas, R. E. Understanding Business Cycles. *Carnegie Rochester Conference on Public Policy,* 5 (1977), pp. 7–29.

McCallum, B. T. *Monetary Economics: Theory and Policy.* New York: Macmillan Publishing Co., 1989.

McCallum, B. T. The Case for Rules in the Conduct of Monetary Policy: A Concrete Example. Federal Reserve Bank of Richmond *Economic Review,* 73 (September/October 1987).

McCandless, G. T., and Weber, W.E. Some Monetary Facts. Federal Reserve Bank of Minneapolis *Quarterly Review* (Summer 1995).

Mishkin, F. S. Strategies for Controlling Inflation. National Bureau of Economic Research Working Paper 6122 (August 1997).

Sargent, T. J. Rational Expectations and the Reconstruction of Macroeconomics. In *Rational Expectations and Inflation.* New York: Harper and Row, 1986, pp. 1–18.

Sargent, T. J. The Temporary (?) Conquest of American Inflation. Web Site Working Paper, 1999.

Taylor, H. E. Time-Inconsistency: A Potential Problem for Policymakers. Federal Reserve Bank of Philadelphia *Business Review* (March/April 1985).

Taylor, J. B. Estimation and Control of a Macroeconomic Model with Rational Expectations. *Econometrica* 47 (September 1979), pp. 1267–86.

Woodford, M. Nonstandard Indicators for Monetary Policy: Can Their Usefulness Be Judged from Forecasting Regressions? In N.G. Mankiw, ed., *Monetary Policy.* Chicago: University of Chicago Press, 1994.

15

Inflation Targets
The Next Step for Monetary Policy

Mark S. Sniderman
Vice President and Director of Research, Federal Reserve Bank of Cleveland, Cleveland, Ohio

In September 1995, Senator Connie Mack (R–Fla.) introduced legislation that would require the Federal Reserve to provide a numerical definition of price stability, to set a timetable for achieving it, and to subordinate other monetary policy objectives to it. The merits of these provisions have been debated at length by economists and policymakers alike. Some believe that a monetary policy based on price stability must compromise economic growth, while others note that the differences between a *price-level* objective and an *inflation* target go way beyond semantics. I believe that the nation's economic prosperity will be enhanced by a price-stability approach to monetary policy regardless of how the objective is defined.

I would like to preface my remarks on this subject with a few comments about national economic conditions. The expansion is entering its sixth year, making it one of the longest upturns in the last half of this century. By almost all accounts, this situation will continue for another year or two.

The expansion has been unusual in a few ways that I regard as being positive for the long term. Business investment has been very strong, enhancing the prospect that productivity growth will improve in the years ahead. Inflation has not accelerated over the course of the expansion, further encouraging productive capital accumulation. Indeed, this impressive inflation performance has been achieved despite an unemployment rate of 5.5 percent—lower than current estimates of the rate considered to be consistent with maintaining stable prices (the nonaccelerating inflation rate of unemployment, or NAIRU).[1] Is this expansion an aberration, or evidence of structural economic change? It's too soon to know, of course, but it is promising to see a long expansion without accelerating inflation in a low-unemployment environment.

I believe that the Federal Reserve has materially contributed to the longevity and strength of the expansion by pursuing a monetary policy focused on achieving price stabil-

---

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (August 1, 1996). The views stated herein are those of the author and not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

ity. If, as a nation, we could find a means for institutionalizing this type of policy, I think we would be pleased with the effects on our living standards over time.

Material increases in a nation's standard of living stem from increases in the quality and supply of labor, the quantity of capital that it works with, and changes in technology. The primary wealth driver in our market economy is the market structure of the economy itself. Poor economic policies (of *all* kinds) lead to resource misallocations and can inhibit economic growth. Poor monetary policy can make resource-allocation decisions that appear sensible in the short run very costly to undo when accelerating inflation is eventually stopped and unwound.

Sound monetary policy, by providing a consistent value for dollar-based transactions, encourages long-range planning and leads to what economists call *intertemporal utility maximization:* the highest sustainable position of economic well-being over time. An environment of low or no inflation can enhance the functioning of our economy just as surely as would a consumption-based tax system with a broad base and low marginal rates.

TEXTBOOK ECONOMICS

We do not, however, usually experience the economic policies prescribed by our textbooks. Why don't we have optimal monetary, fiscal, and regulatory policies? One reason, of course, is that many policymakers and the general public do not have a grasp of basic economic principles. But this is probably less important than two other explanations. First, textbook economic assumptions are rarely, if ever, satisfied in the real world. Second, economists disagree about how our economy actually works. What is optimal economic policy to one may be misguided opinion in the eyes of another. Although abstract economic theorizing and professional disagreements are fine in academic debates, policymakers have to make decisions in real time. Despite being relatively insulated from political influence, central bankers still confront a less-than-ideal economy that no one fully understands.

CENTRAL BANK GOALS

All of this leads us back to the topic of price-level targeting and the legislation introduced by Senator Mack. A few obvious considerations confront those who must take sides on this sort of issue. Should a central bank be asked to subordinate other objectives to price stability? Should price stability be expressed as a *price-level* target or as an *inflation-rate* target? What price index should be used, and what should be done about biases in its measurement?

These important questions deserve answers, but I want to approach the subject from a less obvious direction. I think that some historical reflection can provide guidance: How did we get to this legislative juncture, and why should we believe that a legal mandate to achieve price stability will yield desirable outcomes?

It is useful to recall that inflation is a *persistent* increase in the general level of prices, not an occasional rise in the Consumer Price Index (CPI) induced by transitory factors. The average price level could be kept constant by regulating the supply of money to always equal the public's demand for it when the economy is operating at that price level. If we lived in a world where "money" meant only "central bank liabilities" and its demand were perfectly forecastable, price-level stability would be easy to accomplish, because the central bank has nearly perfect control over its own liabilities. In our actual economy, where money includes the liabilities of both the central bank and private financial institutions, we need stability in the money multiplier (the relationship between central bank money and circulating, privately issued liabilities) and in money demand (the amount of money the public wants to hold under current economic conditions). Alas, we have stability in neither.

Over the last several decades, it has become apparent that whatever stability we thought existed in these relationships was the result of the regulations and technology that determined the financial instruments available to the public. Every time innovation and deregulation swept through the financial services industry, new products became available, the relative prices of those products changed, and the public altered its holdings of financial assets. As a result, what came to be regarded as money changed. Using data from the 1950s and 1960s, some economists (known as "monetarists") established the feasibility of basing monetary policy on the observed movement of certain monetary aggregates, only to see the stability of the underlying relationships collapse repeatedly during the 1980s and 1990s.

It is important to recognize that these relationships have gone awry before. Breakdowns in historical relationships have sometimes led to changes in the definitions of particular monetary aggregates—changes that were specifically designed to restore stability in these relationships! The limitations of this framework became apparent only with the passage of time and poor results.

KEYNESIAN ECONOMICS

Monetarism's failure to provide a reliable policy framework was unfortunate, if only because the reigning Keynesian paradigm was also deficient. In the orthodox Keynesian view, inflation is caused not by "too much money chasing too few goods," but rather by "too many jobs chasing too few people." Inflation control required aggregate-demand management. Vintage Keynesian policy of the 1960s and 1970s hinged on a known potential output path and a stable relationship between unemployment and inflation (known as the Phillips curve).[2] The government decided which unemployment/inflation combination it wanted, and used both fiscal and monetary policies to control the demand for aggregate output. Monetary policy did its part by tightening and easing credit conditions.

We learned two lessons about this framework from the sixties and seventies. First, the Phillips curve is not stable in the normal sense of the term. Output and inflation do tend to move inversely over short periods, but not predictably enough over the course of an entire business cycle, and particularly not from cycle to cycle. This means that the Phillips curve is simply not exploitable by policymakers.

The second lesson is that estimates of potential output, and therefore the "output gap," are subject to significant errors.[3] Moreover, when the actual rate of unemployment is considered high in absolute terms, the output gap tends to be regarded as large. Judgments of this sort seem especially prevalent during election years.

The Keynesian legacy lives on, but in a more sophisticated form. Expectations now play an important role, as well they should. In the long run, the inflation rate is independent of the unemployment rate, and inflation is, once again, regarded as a monetary phenomenon.[4] Today, policymakers are no longer thought capable of selecting whatever combination of unemployment and inflation they want. Any inflation rate can be sustained indefinitely, as long as an equilibrium condition is met: People must expect that rate to continue. If policymakers want to lower the inflation rate from wherever it happens to be in equilibrium, the contemporary Keynesian perspective stipulates that they will still need to create a gap between actual and potential output and drive the actual unemployment rate above the NAIRU value for some period.

Conventional estimates suggest that each percentage-point reduction in inflation requires a 2 or 2.5 percentage-point increase in unemployment for a year. For example, when inflation and inflation expectations coincide at 5 percent, the output cost of moving to 3 percent would typically be gauged at 4 to 5 percentage points of extra unemployment annually—enough to question the efficacy of initiating such a policy. The benefits of lower inflation are commonly regarded as trivial. And yet, in the United States we have reduced the equilibrium inflation rate along this path during the last decade without, I believe, incurring exorbitant output costs.

The output-cost argument against attempts to continue on a disinflationary path begs the following question. If the NAIRU model and these cost estimates are correct, why would policymakers not want to *raise* the equilibrium inflation rate from 5 percent to 7 percent? After all, the neo-Keynesian framework implies that the public would reap the benefits of an output surge and bear only a trivial cost. And why stop at 7 percent? This logic brings us right back to the failure of the original Phillips-curve framework.

The missing link is the failure to stress the important benefits of price stability. The Phillips-curve framework itself offers nothing to tie down the inflation trend or anchor inflation expectations.[5] It brings to mind an old song: "If you can't be with the one (inflation rate) you love, . . . love the one you're with." I believe there are significant benefits to be realized from lowering inflation and making it more predictable. Granted, these benefits have proven difficult to evaluate, because few periods in the post-gold standard era have been marked by inflation sustained in equilibrium and at different rates. However, logic and historical experience convince me that benefits do exist. Between 1953 and 1965, for example, real economic growth averaged slightly more than 3 percent per year, while inflation averaged less than 2 percent. Remember, it took a while for physicists to find those quarks, but they turned out to be there, too.

INFLATION TARGETS

It's ironic that America's success in reducing inflation from 5 percent to 3 percent may inhibit us from doing better, while the failure of some other nations to produce low inflation has induced them to take stronger steps to tie their monetary authorities to the price stability mast. The countries that have adopted inflation targets—Australia, Canada, Finland, Israel, Mexico, New Zealand, Spain, Sweden, and the United Kingdom—generally did so because they had failed to achieve low or even moderate inflation through business as usual (Australia's target is self-imposed by the central bank, whereas the others stem from legislation).[6]

The basic idea of inflation targeting is that the monetary authority is precommitted to achieve a specified inflation objective. The central bank bases its policy actions on predicted deviations between actual inflation and this prespecified target. The public, knowing the objective (and in some instances receiving the central bank's inflation forecasts), can anticipate policy actions.

For some of these nations, inflation targeting came as part of a package bestowing greater independence on the central bank from the rest of government. Direct inflation targeting was regarded as a statement that neither intermediate monetary-aggregate targeting, exchange-rate targeting, nor aggregate-demand management policies provided a sufficient framework for monetary policy. But it meant more than that. It demonstrated that these nations desired a credible nominal anchor for the purchasing power of their currencies and were willing to stake their prestige on explicit public inflation targets to get it. The goals are highly ambitious: With the exception of Israel, the targets center around 2 percent per year (see Table 1).

Does inflation targeting represent the grasp of the obvious or the clutch of the desperate? It is far too early to tell. New Zealand has had the most experience, and that amounts to just over five years. I can well imagine that these nations recognize the fragility of their situations. After all, they are aware of the high hopes that previous central bankers had for the various frameworks that preceded this one. Some of these approaches, like the use of intermediate monetary targets, died a quiet death; others, like the European Community's exchange-rate mechanism, died violently.

Anticipating potential problems, policymakers generally build certain flexibilities into inflation-targeting regimes. Many nations adopt a price index that not only excludes such volatile components as energy and food prices, but also eliminates indirect government taxes and mortgage interest payments. In New Zealand, attempts are made to

Table 1 Inflation Objectives in Selected Countries

Country	Price index	Quantitative objective[a]	Time-specific?	Exemptions and caveats
Australia	CPI	2%–3% average	No: medium-term	Mortgage interest payments, government-controlled prices, and energy prices
Canada	CPI	1%–3% between 1995 and 1998	Yes	Indirect taxes, food and energy prices (operational exemption)
Finland	CPI	About 2% from 1995	No	Housing capital costs, indirect taxes, and government subsidies
Israel	CPI	8%–11% for 1995	Yes: updated annually	None
New Zealand	CPI	0%–2%	Yes: updated annually	Commodity prices, government-controlled prices, interest and credit charges
Spain	CPI	Below 3% by 1997	Yes	Mortgage interest payments
Sweden	CPI	2% ± 1% from 1995	No	None
United Kingdom	RPIX[b]	Lower half of 1%–4% range by spring 1997; $2^1/_2$% or less thereafter	No	Mortgage interest payments

[a] For annual inflation.
[b] Retail Price Index.
Source: Andrew G. Haldane, ed., Introduction to *Targeting Inflation*, London: Bank of England, 1995, p. 8. Reprinted with permission.

eliminate price-level movements due to supply shocks. The rationale is that supply shocks can shift the price level without altering the underlying inflation rate, and the central bank is not to be held responsible for such supply shocks. One also supposes that supply shocks are expected to be both positive and negative over time, so that the price level itself will not drift away from its intended long-run path as a result of such shocks alone.

It is apparent from examining these various inflation-targeting systems that nations are wrestling with a perceived trade-off between flexibility and credibility. Price-level targets with a narrow tolerance limit and no exceptions for special factors would be the most constraining and arguably would produce the greatest credibility. At the same time, the constraints could lead the central bank to force the real economy to adjust so abruptly to an unanticipated event that the public would disapprove. A very flexible system based on inflation rates and allowing for various disturbances, by contrast, would permit greater output smoothing at the cost of less certainty about the price level itself five to 10 years hence.

It seems that by choosing inflation-rate targets, these nations fear the consequences of returning to a prespecified price level after being forced away by unexpected events. However, their annual inflation-rate targets, generally 2 percent or smaller, are so low that price-level drift should be minimal over time, particularly after taking price-index measurement biases into account.[7] It seems clear, however, that these nations wisely do not regard their inflation-targeting systems as a panacea. They have witnessed other frameworks crack and collapse under stress. What these countries seem to desire above all else is a policy process based on a goal that their monetary authorities can actually deliver over time—price stability—conducted with as much openness as possible regarding how the central bank will respond to evolving economic conditions.

CONCLUSION

It may be another 10 years or more before we have enough data to begin scientifically evaluating the performance of these systems. By what standards shall we judge the success of this experiment? Do nations that adopt inflation targets get any interest-rate benefit in the form of a lower inflation or inflation uncertainty premium? Does the greater transparency with which monetary policy is conducted enable the real economy to adjust more efficiently to unexpected macroeconomic events and signals about inflationary pressures? What kinds of shocks will the public allow its central bank to deflect away from the price level, and which others can be absorbed into it? Answers to these questions may differ from country to country.

I don't know whether the United States ever will be included in studies of inflation targeting. Our most recent monetary policy experiences have been sufficiently promising to prompt many observers to say "if it ain't broke, don't fix it."

This prescription has some appeal, but it may be a sign of complacency. As economists, we know that our current policy framework is inadequate because we have no way of anchoring the purchasing power of the dollar. Inflation-targeting systems are worth our time and attention precisely for that reason. Providing an anchor is a more modest policy objective for a central bank than keeping the economy at full employment, but experience reminds us that it is the more realistic goal.

FOOTNOTES

1. See Douglas Staiger, James Stock, and Mark Watson, "How Precise Are Estimates of the Natural Rate of Unemployment?" National Bureau of Economic Research, Working Paper No. 5477, March 1996. The author estimates NAIRU's current value at 6.1 percent. They also acknowledge that due to the difficulty of obtaining precise measurements, the actual figure may lie anywhere between 4.6 and 6.9 percent. NAIRU is also known as the natural rate of unemployment.
2. The Phillips curve postulates a trade-off between unemployment and inflation: Lower rates of unemployment are possible only with higher rates of inflation.
3. The *output gap* refers to the difference between actual and potential output. The Keynesian approach contends that a negative gap (where actual output exceeds potential output) would have inflationary consequences.
4. In technical language, the long-run Phillips curve is vertical at NAIRU.
5. Remember that under this framework, any inflation rate is sustainable in equilibrium as long as the public expects that rate to continue indefinitely.
6. For a very readable and useful survey of current practices among these countries, see Andrew G. Haldane, ed., *Targeting Inflation,* London: Bank of England, 1995.
7. For a more detailed discussion of the upward biases that may occur in measuring price changes (especially in the CPI), see Michael F. Bryan and Jagadeesh Gokhale, "The Consumer Price Index and National Saving," Federal Reserve Bank of Cleveland, *Economic Commentary,* October 15, 1995.

This Page Intentionally Left Blank

16

Central Bank Independence and Inflation

Robert T. Parry
President and Chief Executive Officer, Federal Reserve Bank of San Francisco, San Francisco, California

My comments focus on the link between central bank independence and the pursuit of low-inflation policies. This is a timely issue, because a number of countries—in Western Europe, Central Europe, and even China—have been considering changes to their central banks' structures.

I think central bank independence is an especially interesting issue because a lot of the research on it revolves around an age-old human problem—temptation. The temptation in this case isn't like the one Eve faced in the Garden of Eden, when the snake offered her fruit from the Tree of Knowledge. The temptation instead is to allow a high-inflation environment to develop in the economy.

You might think the comparison of Eve's temptation and the temptation to inflate is ridiculous. After all, her punishment was mortality and eternal banishment from Eden. But according to Dante's *Inferno,* the punishment for people who debase the currency could be pretty rough, too—the tenth pit of the eighth circle of Hell, to be exact.

As a conservative central banker, I can easily think of a number of reasons to justify Dante's harsh treatment of inflators. To begin with, a higher level of inflation is generally associated with higher uncertainty about inflation. And higher uncertainty about inflation means more difficult planning and contracting by business and labor. It also means higher risk premia in long-term interest rates, which depress capital formation and productivity.

In fact, 600 years after Dante, Lenin said the best way to destroy the capitalist system was to debauch the currency. And according to Keynes, Lenin was right.

TEMPTATION AND CENTRAL BANK INDEPENDENCE

I'd like to start by describing how temptation plays a role in justifying central bank independence. Basically it has to do with the difference between monetary policy's effects in the short run and the long run. In the short run—say, six to nine months—an "easy" monetary policy can stimulate the economy to grow faster and to lower the unemployment rate.

Reprinted from: *FRBSF Weekly Letter,* No. 95–16 (Apr. 21, 1995).

But *persistently* stimulating the economy to get gains in output and employment is an exercise in futility. The reason is that, in the long run, output and unemployment depend mainly on *real* factors—such as population growth and improvements in technology and productivity. So eventually the monetary effects on the real economy dissipate, and rather than long-run gains in output and unemployment, all we end up with is persistently high and volatile inflation. And that in itself takes a toll on economic growth.

Furthermore, as we learned from the early 1980s, once inflation gets going, it's costly to bring it under control. Back in 1980, CPI inflation was running at over 12 percent a year, and 30-year fixed-rate mortgages were as high as 16 percent. By 1982, the annual inflation rate was down to 3.8 percent. But to get there we went through a double-dip recession with over 10 million in the ranks of the unemployed.

Temptation enters the picture because it's natural to let near-term concerns obscure considerations about long-term consequences. In the standard story involving monetary policy, it's elected officials who could be tempted to be short-sighted in their efforts to expand the economy. With elections on the near horizon, they'd be tempted to try to ensure their return to office by pushing the central bank to follow an "easy" policy, figuring they can deal with the resulting inflation when it shows up later on.

The more political pressure on the central bank, the more likely this short-sighted strategy will be followed. As I said, in the long run, this strategy doesn't lead to any gains in employment or output, it just leads to higher, more volatile inflation.

But that's not all. It also leads people to expect higher inflation. Because the public believes that if politicians can pressure the central bank into an "easy" policy once, they can do it again.

Under these circumstances, a central bank would have a hard time convincing anybody that it was serious about maintaining low inflation. In other words, the public wouldn't think that the central bank's commitment to a low-inflation policy was credible. For a central banker, this is the worst of all possible worlds. Without a credible low-inflation policy, the only way the central bank can slow inflation down is to step on the economic brakes *hard.* And the result can be a deep recession with many jobs lost.

THE FEDERAL RESERVE'S STRUCTURAL INDEPENDENCE

One way to insulate monetary policy from day-to-day political pressures is to give the central bank a more independent structure. In this respect, the Fed ranks toward the top among the world's central banks. So I'll draw from the Fed to describe some of the elements of an independent structure.

First, the Governors on the Board are appointed to 14-year terms—not as long as Supreme Court Justices, but much longer than congressional or presidential terms. Although they're political appointees—in that the U.S. President nominates them and the Senate confirms them—their terms aren't tied to an election cycle. Second, the Reserve Bank presidents—like me—aren't appointed by any politician. Instead, we're selected by our Reserve Bank's Board of Directors and approved by the Board of Governors. This provides further insulation from partisan politics. Third, the Fed covers its own operating expenses, so it isn't dependent on Congressional appropriations.

I want to stress that independence does *not* mean that the Fed is free of *accountability* to the government. For example, the Fed—like virtually all central banks—has its goals set by the government. In the case of the U.S., the goals are laid out in the Humphrey-Hawkins Act. In addition, the Chairman, the Governors, and the Reserve Bank Presidents give testimony before Congress on a number of issues. Furthermore, the Fed meets regularly with senior Administration officials to discuss one another's respective economic programs. In fact, the phrase usually used to describe the Fed is "independent *within* government."

INDEPENDENCE, GOALS, AND CREDIBILITY

There *is* a way beyond such structural means to make monetary policy even more independent of political pressure to inflate. And that way is for the Congress to give us a clear statement that low inflation is our goal. The Humphrey-Hawkins Act actually does this, by calling for an inflation rate in the range of 0 to 3 percent.

But it also calls for unemployment in the range of 3 to 4 percent. Just about any economist you talk to will tell you that, in today's economy, you can't achieve both goals at once—and furthermore, it's virtually impossible to achieve the unemployment goal at all! This combination of goals sends the Fed mixed signals, where almost any policy approach—from "easy" to "tight"—is technically consistent with the instructions from Congress.

In this session of Congress, Senator Connie Mack (R) of Florida actually has proposed legislating low inflation as the Fed's only goal. Such a change would not only reduce the Fed's exposure to pressures to inflate, but some economists also have argued that it may add credibility to the Fed's low-inflation policy.

In fact, on just those grounds a number of other countries recently gave their central banks specific low-inflation goals. So now let me take a look at how their experiences with independence, goals, and credibility have worked out.

Since 1990, the governments of New Zealand, Canada, and the United Kingdom have worked with their respective central banks to adopt specific low-inflation targets. And, so far, all three countries have achieved their inflation objectives.

But the evidence also suggests that the low-inflation regimes in these countries aren't fully credible (Ammer and Freeman 1994). For example, surveys show that the public's expectation of inflation remains somewhat above actual inflation, although expectations have come down since the introduction of the targets. That probably means that it takes more than inflation targets and a few years of success to establish full credibility.

The example of Japan drives home the point. The Bank of Japan has neither political independence nor formal inflation mandates. Yet because inflation in Japan has averaged only around 2 percent for the last 10 years, the Bank of Japan is among the world's most credible central banks.

But that's not to say that an explicit low-inflation mandate or target from Congress wouldn't help—and help significantly—in the U.S. I personally support proposals that make low inflation the central bank's primary goal. In fact, when former Representative Stephen Neal (D) of North Carolina held hearings on a similar plan a few years ago, I testified in favor of it. I think such a goal would help in a number of ways: It would help us avoid the pressures for higher inflation, it would add to the Fed's accountability to the Congress and the public, and it would help us solidify and extend the significant gains against inflation that the Fed has made in the last decade and a half.

CONCLUSION

In closing, let me restate what I believe are the most important issues in discussions of central bank independence: First, the temptation to focus on the short run rather that the long run can thwart the best efforts to maintain low and stable inflation. Second, some approaches—namely, structural independence and low-inflation goals—can help fend off those temptations. Finally, although the Fed has made considerable progress against inflation over the last 15 years, we still have our work cut out for us, and we can't afford to relax our vigilance.

To return to Dante, I urge continued vigilance *not* because I think the alternative is spending eternity in the eighth circle of Hell. I urge it because achieving and maintaining low and stable inflation is the best way the Fed—or any other central bank—can contribute to long-run, sustainable growth in the economy.

REFERENCE

Ammer, J., and Freeman, R.T. 1994. Inflation Targeting in the 1990s: The Experiences of New Zealand, Canada, and the United Kingdom. *International Finance Discussion Paper* Number 473. Federal Reserve Board (June).

17
Monetary Policy in the Cold War Era

Mark S. Sniderman
Senior Vice President and Director of Research, Federal Reserve Bank of Cleveland, Cleveland

The Soviet Union officially disbanded on December 26, 1991, one day after the resignation of Mikhail Gorbachev. Ever since, the countries that made up the former USSR have been struggling both to govern themselves and to find their places in the world. The cold war against communism was over.

The palpable threat of nuclear attack by the Soviet Union brought a high degree of cohesion to U.S. foreign and defense policies. Now, the vacuum created by the "evil empire's" collapse is prompting questions that remain largely unanswered. Do we still have adversaries, and, if so, what harms can they inflict? How can we best achieve our objectives, and to what lengths are we willing to go to fulfill them? How much will these efforts cost? In a dangerous world, how much risk should we bear? Answering these questions requires making choices, and each choice comes with its own price tag. Like the former Soviet republics, we too are struggling to define our relationships with the rest of the world.

Complicating the reconstruction of a new foreign-policy framework is the fact that, seven years after the fall of the Berlin Wall, the United States faces fewer evident threats to its national security than at any time since World War II. Some argue that our defense establishment is paranoid when it seeks public support for more resources. It's not that Americans no longer care about national security. Rather, the public expects the Defense and State Departments to justify their policy stances in terms of a new world order—one that no one yet fully grasps.

I suggest that there is an analogy between the search for this new world order in foreign policy and recent attitudes about monetary policy. The cold war against communism is indisputably over; however, can the same be said about the war against inflation? The U.S. economy has been expanding almost continuously for 15 years, the unemployment rate lies near 5 percent, and inflation pressures appear scant. Yet, to take a hard line against inflation today is, like being opposed to communism, passé. Are people who advocate a price-stability objective for monetary policy indeed fighting the last war?

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (June 1997).

THE WAR AGAINST INFLATION

In 1979, in the midst of the cold war, the United States initiated a "hot war" against another seemingly implacable foe—inflation. President Carter appointed Paul Volcker to head the Federal Reserve, giving him a mandate to eliminate double-digit inflation. In conducting that war, the Fed relied on demonstrably tight monetary policy and the public's willingness to suffer temporarily higher unemployment rates if warranted. Inflation was so intolerable that having a numerical goal was unimportant; all that mattered was bringing it down. With the support of President Reagan, the Volcker-led Fed continued its use of heavy artillery to end the inflationary spiral, reducing the core inflation rate from 11 percent to 5 percent by 1983.

Under the leadership of Alan Greenspan, who took the helm in 1987, the Federal Reserve continued its battle against inflation, which it described as a campaign to achieve price stability. With inflation now a lower-level threat to economic progress, the Fed could squeeze it down more gradually. Initially, the Greenspan Fed followed a course of limited aggression and persistently combative rhetoric. This strategy finally paid off in 1991. As Boris Yeltsin faced down a tank in Red Square, the U.S. inflation trend collapsed from 5 percent to 3 percent, capitulating to a seven-year siege. The Federal Reserve reduced inflation to levels not seen since Sputnik. The monetary policy hot war was over, and the United States could feel proud of its victory.

THE MONETARY COLD WAR ERA

The surprisingly swift transition to lower and more stable inflation rates caused some to declare that inflation was dead. The economy's pace faltered after the Gulf War, and the nation's attention was focused on expansion and employment, not inflation.

The inflation rate has not varied much during the last six years, despite predictions that it would advance when unemployment dropped below 6 percent in mid-1994. In early 1995, it was not uncommon to hear forecasters state that a 7 percent or greater federal funds rate would be required to repel the coming inflationary invasion. The Federal Reserve never raised the funds rate to these heights, but even as the rate crested at 6 percent and monetary policymakers spoke about their commitment to stable prices, critics said the Fed was fighting the last war. The public, it seemed, was tired of combat.

Yet, Federal Reserve officials still talk publicly about the importance of achieving price stability, a condition that some have described as inflation so low that it doesn't affect people's economic decisions. However, for inflation not to enter into economic decisions, the Federal Reserve must succeed at informing the public about the value of price stability in a market economy, and at convincing them that its policies will be set to achieve that goal. This is a tall order at a time when many Americans are relatively satisfied with the inflation rate and worry that efforts to contain or reduce it may entail slower economic growth. In their view, "close enough" is "good enough."

Others are pushing for even greater accountability. Price stability as a goal does not lend itself as readily to accountability and oversight as a numerical inflation objective. Some observers decry this imprecision as a shortcoming of the current monetary policy regime, and, believing that it lessens the Fed's credibility, have proposed revising the legislative framework within which policy decisions are made. Advocates of stricter accountability attribute a fair portion of the nation's favorable economic performance over the last decade to monetary policy. Hence, they are looking for ways to institutionalize the goal of price stability in the policy-setting process.

These contrasting views about the nature and desirability of an inflation objective illustrate an often underappreciated aspect of policymaking, namely, that policies must be understood and supported by the public. Americans eventually accepted the Federal Reserve's hot war against inflation, but only after they became convinced that an accelerating price level would not be accompanied by more output and employment growth. I think it is reasonable to characterize the post–1991 policy regime as a monetary cold war—a strategy designed to attain policy objectives through less forceful means than strenuously and persistently tightening money and credit conditions. Public acceptance of this war has been easier to achieve and maintain, I believe, because inflation has continued to drift down throughout the course of a lengthy economic expansion. But the conflict will not be complete until inflation psychology itself is undermined, so that the public sees no reason to legitimize it or embrace its cause.

THE IMPORTANCE OF PRICE STABILITY

Most economists agree that once inflation is fully anticipated, employers, employees, savers, and borrowers simply adjust the prices at which they are willing to transact with one another to reflect their expectations about the currency's declining purchasing power. If this is true, inflation imposes no real effects on economic activity.

But the premise is not true, for several important reasons. When a monetary authority debases the purchasing power of its currency, it drives a wedge between what people will realize from a monetary transaction and what that transaction is actually worth to the economy. For example, the U.S. tax code contains an indexing provision for labor income (personal exemptions, income brackets, and so on), but levies tax obligations for capital income in nominal dollars. As a result, inflation—even if fully anticipated—increases the effective tax rate on capital income, which discourages capital formation and long-term economic growth. The potential impact is huge.

Another distortion to economic decisions comes in the form of an inflation-uncertainty premium. Even though two parties may have the same expectation regarding inflation's average rate over time, they may have different degrees of confidence about their estimate, or different tolerances for being wrong. Periods of high inflation tend to be periods in which the price changes of individual goods and services vary considerably. As inflation accelerates, one party in a transaction may demand a premium from his counterparty for bearing the risks of error. People devote time and real resources to avoiding the costs of uncertain inflation, and these costs—like a rising flood plain—can accumulate and become large.

Accelerating inflation is like the game of musical chairs: Everyone knows that when the music stops, someone will come up short. For an individual, it is rational not to want the music to stop, but collectively, society is wasting its resources. Once inflation reaches high levels, its distortions are so substantial that everyone is dizzy and wants the game to end.

Ending inflation can be costly, however, because doing so disrupts plans and decisions that have already been made. An excellent example can be found in the housing markets of the 1970s, when many people thought that home ownership would be an effective hedge against inflation. These buyers sought houses not because they wanted the shelter or amenities that a home offers, but because they assumed that the property could be readily sold at a profit. The boom brought land, labor, and financing into housing markets from other uses merely to satisfy the demand for an inflation hedge. When the boom ended, many people suffered a sudden reversal of fortune, including those who entered at the tail end and never benefited at all.

But when the musical chairs game is played at a slow pace, few seem to mind. And, to be honest, economists have had difficulty quantifying large social losses in low-inflation circumstances. The tax and uncertainty distortions I've mentioned are proportional to the amount of inflation. So, what's wrong with a little inflation?

I will cite two reasons for opposing this attitude. The first has to do with unbounded expectations. What is a "low" inflation rate? If 3 percent inflation is thought to cause little harm, then neither will 4 percent; after a while, 5 percent becomes only a small differential from 4 percent, and so on. Regarding our current 3 percent inflation rate as just the happenstance of where we are economically imparts an ephemeral quality to it. Although very low inflation, per se, may cause few distortions, this "here today, gone tomorrow" mindset would likely inject an inflation risk premium into interest rates and economic decisions. Zero inflation need not to be the only acceptable rate: The criterion should be rates so low that they do not alter economic decisions.

The second reason for resisting inflation tolerance has to do with the false notion that inflation can be traded off permanently for something of value, such as faster economic growth or lower unemployment rates. It is really likely that debasing the purchasing power of money will lead to more wealth creation?

Yes, easy money can temporarily stimulate economic activity, just as tight money can temporarily retard growth. And a sequence of stop/go monetary actions can be very destabilizing to economic activity. But over time, wealth creation depends on the availability of skilled labor, productive capital, and a legal infrastructure that facilitates economic exchange. Stable expectations about money's purchasing power—especially over long horizons—enable people to make decisions that better reflect the value of the resources called into play.

Recall the previous example of housing markets in the 1970s. With hindsight, it should be obvious that our country would have been better off had more savings been channeled into the creation of productive business capital, instead of being poured into housing markets as an inflation hedge. Unfortunately, rationally formed expectations about future inflation meant that the proper incentives were not in place.

When the inflation rate hit double digits and the pace was accelerating, the Federal Reserve simply aimed to get the rate down, and down fast. No one asked where inflation would settle out, and no one bothered to set a target. At the time, policymakers realized that a gradual approach would not work, since that strategy had been tried unsuccessfully during the 1970s. The failures of that era stemmed not from the absence of a strategy, but rather from a two-pronged strategy of first, thinking that more economic growth could be purchased with a little more inflation, and second, demonstrating an unwillingness to risk disrupting the pace of economic activity in order to reduce inflation.

The irony is that economic activity was being disrupted in a very serious way, but not an obvious one. The disruption came in the form of escalating prices for homes, art objects, and farmland. The psychology of the times was to become a debtor and to use someone else's savings to acquire hard assets. Business plans were premised on rising prices. The game was to raise your prices faster than your suppliers could raise theirs. Accelerating inflation also transferred resources from the private sector to the government through the unindexed tax code. Output and real incomes were lower than they otherwise would have been because resources were diverted from wealth-creating activities to wealth-protecting ones. By the end of the 1970s, it had become painfully clear that our political leaders' unwillingness to risk any slowdown in output was shortsighted. Moreover, the public's change of heart illustrates that what is regarded as politically expedient at one moment may become political poison the next.

A JUST PEACE

I am certainly not dismissing the prospect that inflation might accelerate again, perhaps even imminently. I am trying to point out why a relaxed view about inflation is misguided. In my opinion, the Federal Reserve is not engaged in a conventional war against inflation, but rather in a cold war. One difference is the seriousness of the threat we're facing. The Fed's conventional war was launched only after inflation spiraled seemingly out of control, while today's cold-war policy is directed against a lower-grade enemy. A related difference can be seen in public attitudes: The Federal Reserve's 1979–90 anti-inflationary policy enjoyed strong popular support, whereas today's climate is not so universally accepting.

How does an honorable monetary authority achieve a responsible peace with inflation? A workable compromise requires that the public and its central bank understand one another's aspirations and limitations. After all, nations create independent central banks to prevent the popular wish for easy money from running amok. An unduly restrictive monetary policy will eventually lose popular support, but so will policies of appeasement, as the choices of the 1970s illustrate. Although there is more than a little room for misunderstanding and mischief in the goal-setting process, an honorable monetary authority attempts to be as transparent as possible about both its intentions and its operations.

Transparency, unfortunately, does not always equal precision. The Federal Reserve has not declared a numerical objective for inflation or a timetable for reaching its goal of price stability. Although I believe that a more clearly articulated framework would enhance its actions, the Fed's monetary policy is *realpolitik*—rooted in what can work rather than relying on ideals alone.

I might say that the Fed has achieved détente with inflation, but I would also remind you that détente does not equal peace. In my opinion, peace will be attained when the public supports the principle that monetary policy best contributes to national prosperity by eliminating inflation and the expectation that it will reemerge.

Perhaps the current expansion, whose features include low and stable inflation, a capital spending boom, and strong employment growth, will instill confidence in the merits of such an approach. And perhaps it will not be long before we can welcome, to paraphrase the title of a famous novel, the Fed that Came in from the Cold:

> It was *man* who ended the Cold War in case you didn't notice. It wasn't weaponry, or technology, or armies or campaigns. It was just *man*. Not even Western man either, as it happened, but our sworn enemy in the East, who went into the streets, faced the bullets and the batons and said: we've had enough. It was *their* emperor, not ours, who had the nerve to mount the rostrum and declare he had no clothes. And the ideologies trailed after these impossible events like condemned prisoners, as ideologies do when they've had their day.[1]

FOOTNOTE

1. John le Carré. *The Secret Pilgrim.* New York: Knopf, 1991, p. 321.

This Page Intentionally Left Blank

18

Federal Reserve Independence and the Accord of 1951

Carl E. Walsh
Professor of Economics, University of California, Santa Cruz, California

Since the establishment of the Federal Reserve System in 1913, the Fed's relationship with both Congress and the President has gone through many phases, and proposals to change this relationship, or the way the Fed conducts monetary policy, have appeared frequently in Congress. For example, Congress amended the Federal Reserve Act in 1977 to require the Fed to establish money supply targets and to report those targets to Congress every six months. In the early 1980s, legislation was introduced, but not passed, that would have required the Fed to establish targets for real interest rates. Other proposals have focused less on the actual implementations and more on the ultimate objectives of monetary policy. For instance, four years ago, Representative Neal (D) held hearings on a bill that would have established zero inflation as the official, and sole, policy objective for the conduct of monetary policy.

In contrast to these earlier legislative attempts to affect either the Fed's objectives or its implementation of policy, Senator Sarbanes (D) of Maryland and Representative Gonzalez (D) of Texas have proposed changing the structure of monetary policy decisionmaking. The intent of their proposals is to provide greater political control over the conduct of monetary policy by increasing the role of the President in determining who makes the decisions about monetary policy.

During other episodes in the Federal Reserve's history, Congress has supported moves designed to increase the Fed's independence from Executive Branch influence. One of the most important of these moves occurred in 1951. In March of that year, the Federal Reserve System and the U.S. Treasury reached an agreement, known as the Accord, that recognized the independence of the Federal Reserve to conduct monetary policy. In the Fed's negotiations with the Treasury, the Fed was bolstered by Congressional support for an independent monetary policy. The modern conduct of discretionary monetary policy in the United States can be dated from the Accord.

THE PRE-ACCORD PERIOD

During the decade before the 1951 Accord, Federal Reserve actions were dominated by considerations arising from the government's World War II financing needs. The Treasury,

Reprinted from: *FRBSF Weekly Letter,* No. 93–21 (May 28, 1993).

faced with the need to raise funds far in excess of tax receipts in order to finance the war effort, wanted to keep interest rates on government securities at low levels. The Treasury view was supported by the Federal Reserve, and the Fed adopted an explicit policy of supporting the government bond market. Particularly during 1942 to 1945 when the government was engaging in massive borrowing, this was clearly an important consideration. As expressed by G.L. Bach, "In this period, Federal Reserve and Treasury officials agreed, with perhaps more patriotic fervor than foresight, that there must be no shortage of money to buy the weapons of war . . ." (Bach 1971, p. 78). In April 1942, the Fed announced that it would maintain the rate on 90 day government bills at 3/8 percent. It did so for the next 5 years.

Whenever a central bank adopts a policy of pegging market interest rates, it gives up control over the supply of money. If the pegged rates are set too low, private sector demand for new government debt issues will be too small to take up the entire issue. To prevent bond prices from falling (and yields rising), the Fed must serve as the residual purchaser. In so doing, the Fed automatically increases the reserves of the banking system, allowing an expansion of the money supply. If the pegged rates are set too high, there will be an excess private sector demand for government securities, and the Fed must sell from its own portfolio of government security holdings in order to prevent interest rates from falling. In the process, banking sector reserves are reduced, leading to a fall in the supply of money.

The Fed continued to support bond prices after the war for several reasons. First, the policy facilitated government borrowing. The low interest rates reduced the cost of government borrowing, the Fed commitment ensured that the Treasury could always sell its new bond issues since the Fed served as the residual purchaser, and, by insuring the market against capital losses that would occur if interest rates rose, the bond support program increased the overall demand for government debt. Second, any increase in interest rates on government debt would also raise interest rates faced by private borrowers, thereby resulting in reduced private sector investment and increased unemployment. This concern reflected the fears of a postwar recession. Third, it was argued that a rise in interest rates was an ineffectual means of combating inflation.

In the immediate postwar period, the Federal Reserve was increasingly concerned about its inability to prevent inflation as long as it was required to support the price of government debt. With the Consumer Price Index rising more than 14 percent during 1947 and nearly 8 percent during 1948, the Fed believed it needed to control money and credit growth. The Treasury continued to argue that low interest rates were necessary to maintain confidence in government credit and to hold down the cost of government debt, and that controlling the money supply was not necessarily an effective means of reducing inflation. Tensions rose between the Federal Reserve and the Treasury over the Fed's desire to establish monetary control. Marriner Eccles, who had been appointed Chairman of the Board of Governors in 1934 and who openly argued against the bond support policy, was not reappointed by President Truman in 1948.

As long as the Fed supported the prices of long-term government debt, holders of the debt could view these assets as very liquid. With the bill rate held to $\frac{3}{8}$ percent since 1942 while the ceiling on long-term government securities was a much higher $2\frac{1}{2}$ percent, there was little demand for Treasury bills. Of the $16 billion in bills outstanding in 1947, the Fed held $15.5 billion. In the middle of 1947, the Fed allowed the bill rate to rise. One consequence of the rate increase was a rise in the Fed's interest income on the Treasury bills it held. To ensure Treasury support for the rate increase, the Fed agreed to turn over 90 percent of its revenue to the Treasury.

In June, 1948, the Federal Open Market Committee, the Fed's policymaking committee, and the Treasury announced that the FOMC would direct open market operations" . . . with primary regard to the general business and credit situation" (*Federal Reserve Bulletin,* 35, July 1949, p. 776). Fed Chairman Thomas McCabe considered this announcement to constitute" . . . the removal of the strait jacket in which monetary policy has been operating for nearly a decade. . . ." At the time, however, the concern was with the economic recession that developed in late 1948. Unemployment rose from 3.8 percent in 1948 to 5.9 percent in 1949, and prices actually declined by 1 percent in 1949. Consequently, the FOMC–Treasury agreement was an agreement to lower interest rates in an attempt to stimulate the economy. It was unclear whether the Fed would have the flexibility to raise interest rates if the problem became one of inflation. In Congressional testimony in 1949, the Treasury Secretary made clear that his interpretation was that the $2\frac{1}{2}$ percent rate on new long-term government securities would not rise.

The conflict between the Treasury and the Fed over interest rate policy led, in 1949, to Congressional hearings on the subject headed by Senator Paul Douglas of Illinois. At this time, Congress was generally viewed as supporting the Fed in its conflicts with the Treasury. According to Stein (1969, p. 258), the hearings" . . . made it clear that any attempt to bring the Federal Reserve forcibly to heel would encounter considerable resistance in the Congress, and that the resistance would have leadership and principles to which there would be a popular response." The Douglas report concluded that the benefits of avoiding inflation were great enough to justify giving the Federal Reserve the freedom to raise interest rates, even at the cost of a rise in the cost of federal debt.

THE ACCORD

In 1950, with the recession over, inflation and the need for monetary restraint once more became a policy concern. During January and February 1951, the Treasury attempted to bind the Fed to the maintenance of low interest rates through public announcements. The Secretary of the Treasury, John Snyder, announced that consultations with President Truman and the Chairman of the Federal Reserve Board had led to a decision that new long-term debt issues would continue to be offered at a $2\frac{1}{2}$ percent interest rate, a view apparently not shared by the Fed. When Fed disagreement became known, President Truman called the entire FOMC to a White House meeting to discuss policy. The White House and the Treasury then announced that the Fed would continue to support government bond prices. Eccles, who was still a member of the Board of Governors, then released the Fed's confidential minutes of the White House meeting, minutes that contradicted the White House and Treasury claims of a Fed commitment to keep rates fixed.

As a result of these public disputes, the Fed asked the President to initiate negotiations between the Treasury and the Federal Reserve. While the President established a formal committee to resolve the issues of conflict, the actual "accord" between the two institutions was worked out directly between Federal Reserve and Treasury officials. On March 4, 1951, the Accord was announced to the public: "The Treasury and the Federal Reserve System have reached full accord with respect to debt-management and monetary policies to be pursued in furthering their common purpose to assure the successful financing of the Government's requirements and, at the same time, to minimize monetization of the public debt" (*Federal Reserve Bulletin,* March 1951, p. 267).

Despite the current view that the Accord enhanced the Fed's ability to conduct an independent monetary policy, the language of the Accord did not specifically address the issue of conflict—would the Fed be expected to continue to support bond prices? In fact, at the time many commentators felt the Accord was not the final resolution of the Treasury-Fed disagreements. However, it soon became clear that the Fed had in fact been freed from its obligation to support the price of government bonds.

AFTER THE ACCORD

Interest rates gradually rose during the two years following the Accord, and market interest rates became much more volatile as the Fed was now able to pursue more activist policies. However, Shiller (1980) shows that once short-term interest rates on commercial paper were corrected for inflation, real interest rates actually became much less volatile over the 20 years following the Accord. More importantly for the longer-term conduct of U.S. monetary policy, the Accord separated the determination of debt-management policy from that of monetary policy. This was a necessary separation for controlling the money supply and for providing the Fed the means for controlling inflation.

Did the Accord actually give the Fed independence from the Executive Branch to conduct monetary policy? At the end of March 1951, just three weeks after the Accord was announced, President Truman appointed William McChesney Martin, the Treasury official who had negotiated the Accord for the Treasury,, as Chairman of the Federal Reserve Board, a position he held until 1970. Martin made clear, however, his view that the Federal Reserve was an independent agency of government, responsible to the Congress.

It is important that monetary policy be unconstrained by debt management considerations. Requiring the Fed to maintain interest rates at levels that are too low runs the risk of increased inflation. The conflict between the Fed and the Treasury that led to the 1951 Accord did, according to Stein, serve a useful purpose: "If monetary policy had floated free of the [interest rate] pegs without a direct confrontation, the importance of flexible monetary policy might never have become so clear as it did to large numbers of people, and the Federal Reserve would not have been left with so vivid a reminder of the dangers of compromising its independence" (Stein 1969, p. 278).

REFERENCES

Bach, G.L. 1971. *Making Monetary and Fiscal Policy.* Washington, D.C.: The Brookings Institution.
Shiller, R.J. 1980. Can the Fed Control Real Interest Rates? In *Rational Expectations and Economic Policy,* ed. S. Fischer. Chicago: The University of Chicago Press.
Stein, H. 1969. *The Fiscal Revolution in America.* Chicago: The University of Chicago Press.

19

Is There a Cost to Having an Independent Central Bank?

Carl E. Walsh
Professor of Economics, University of California, Santa Cruz, California

In recent years, many economists have argued that average inflation can be kept low if central banks are insulated from political pressures. Politically independent central banks, it is argued, will be less likely to give in to pressures to adopt expansionary monetary policies for political purposes, and, therefore, will be able to deliver lower average inflation. Indeed, using a variety of measures of independence, a number of researchers have documented empirically the association between greater central bank independence and lower average inflation (for a survey, see Cukierman 1992).

But are there costs associated with keeping inflation down? This *Weekly Letter* examines recent evidence on the relationship between measures of real economic performance, such as output growth, and measures of central bank independence. The evidence suggests that central bank independence, while associated with low inflation, is not associated with either slower of more volatile economic growth. Moreover, given that low inflation helps to reduce arbitrary wealth transfers and economic uncertainty, an independent central bank is likely to enhance economic welfare.

INDEPENDENCE AND THE CONDUCT OF MONETARY POLICY

The usual reason cited to explain the correlation between independent central banks and lower average inflation is that independent central banks are often viewed as more concerned with achieving and maintaining low inflation than politicians are. Thus, independent central banks are similar to other systems of "checks and balances" in a democratic society: An independent central bank would tend to offset a bias towards excessive inflation that is commonly thought to characterize monetary policy that is not guided by a specific mandate or rule on inflation. While society benefits from lower average inflation as a result, it also gets a less activist monetary policy. By placing greater weight on its inflation goals, the central bank might be willing to tolerate greater fluctuations in real economic activity and unemployment.

Reprinted from: *FRBSF Weekly Letter*, No. 94–05 (Feb. 4, 1994).

A recent study of 16 major industrial economies by Alesina and Summers (1993) attempts to determine if central bank independence is associated with either slower real growth or greater economic fluctuations. Perhaps somewhat surprisingly, they find that the degree of central bank independence is not related to average real GDP growth or average unemployment, nor to measures of the volatility of economic growth or unemployment, nor to the average level of real interest rates. These are important findings, since they would seem to indicate that central bank independence yields lower and less volatile inflation at no cost in terms of slower average real growth, higher average unemployment, or increased economic fluctuations.

Before drawing very strong policy conclusions from this work, however, it is important to determine whether these findings hold for a larger set of countries. And since economists have disagreed over how to measure central bank independence, it is also of interest to determine if these results still hold when using different indexes to measure central bank independence.

WHAT IS CENTRAL BANK INDEPENDENCE?

Most research has focused on two dimensions of central bank independence. One, usually called political independence, represents the degree to which a country's central bank has policy objectives that are insulated from political pressures to expand aggregate demand rapidly. Political independence is influenced by institutional structure, such as the process for appointing the bank's policymaking board, and the existence of explicit policy goals, such as price stability. For example, a central bank with government representatives and political appointees with short terms of office on its policymaking committee would be classified as having a low degree of independence.

The second dimension of independence, called economic independence, is the degree to which the central bank is free to use its policy instruments to pursue its objectives. In some countries, the central bank must finance government deficits; such a central bank lacks economic independence, since it cannot control the extent to which government deficits are monetized.

Alesina and Summers employ as their measure of central bank independence an index that is an average of measures of political and economic independence. Among the 16 countries examined, Germany and Switzerland are ranked as having the most independent central banks, with New Zealand having the least (prior to the 1989 reform of the Reserve Bank of New Zealand). The U.S. is ranked just below Germany and Switzerland.

Formal institutional independence may bear no relationship to the actual degree of independence enjoyed by a central bank. For example, Cukierman, Webb, and Neypati (1992) constructed a measure of independence based on the legal description of the bank's structure and a measure based on surveys of experts in each country who were asked to assess the independence the central bank had *in practice*. The analysis covered 24 countries and found no correlation. That is, a central bank ranked as highly independent by one measure was no more likely to be highly ranked than lowly ranked on the other index (see Walsh 1993).

The Bank of Japan (BOJ) provides an interesting case in point. The BOJ is legally subordinate to the Ministry of Finance. For this reason, Cukierman, Webb, and Neypati (CWN) rank it as having little formal independence. In fact, they rank it below even the pre-reform Reserve Bank of New Zealand, the Bank of Italy and the Bank of Spain. In contrast, Alesina and Summers rank the BOJ as having a degree of independence slightly above the average in their sample, because, over the past 20 years, it has operated much like the most independent central banks in keeping inflation at low levels. Part of the explanation may lie in the connection found by Cukierman, Edwards, and Tabellini (1992) between inflation and political instability. Countries with unstable political cultures, with frequent changes in office by competing political parties, are more likely to generate pressures on the central bank to follow inflationary policies. Until the elections of 1993, Japan had been governed by a single party for nearly 40 years. Cargill and Hutchison (1990) argue that, with the LDP party holding a monopoly on power, it was unnecessary to exploit the government's control over the Bank of Japan to engage in expansionary policies. It will be interesting to see if more competition for electoral control also leads to greater political pressure on the Bank of Japan.

INDEPENDENCE AND ECONOMIC ACTIVITY

Converting a description of a central bank's institutional and policymaking structure into an index of independence of necessity involves judgment calls. Therefore it is useful to see if measures of central bank independence other than the one employed by Alesina and Summers display any relationship to average real growth or measures of economic fluctuations. To investigate this, I used the CWN index of legal independence which covers 68 countries, compared to the 16 analyzed by Alesina and Summers. Data on real per capita GDP were obtained from the Penn World Table (Summers and Heston 1991). Focusing on the period from 1971 to 1985 yielded a sample of 60 countries for which both the CWN index and real per capita GDP data were available.

To explore the effects of using an alternative measure of central bank independence the average growth rate of real per capita GDP and its standard deviation (a measure of its volatility) can be compared to the CWN index of independence for the same countries used by Alesina and Summers. Such an analysis shows a slight negative relationship between central bank independence and average GDP growth. However, this is the result of one data point—Japan—which had the highest average growth rate over this period and has the least independent central bank according to CWN. If Japan is dropped from the sample, or if the BOJ is reclassified according to Alesina and Summers' ranking, any relationship between growth and central bank independence disappears.

Using the CWN index of central bank independence in place of the Alesina-Summers measure, therefore, provides some further evidence in support of their finding that, for the industrialized countries, the degree of central bank independence is unrelated to average real economic growth or its fluctuations. This result, however, is based on a small group of countries. Since the CWN index is available for a large number of countries, one can examine the robustness of the finding when the number of countries in the sample is greatly expanded. Figure 1 shows the average growth rate of real per capita GDP and its standard deviation for 60 countries for 1971 to 1985.

Figure 1 shows no apparent relationship between central bank independence and either average income growth or its volatility. More formal statistical analysis is consistent with the absence of any relationship. This holds even if the two outliers with extreme output fluctuations (Nicaragua and Uganda) are excluded.

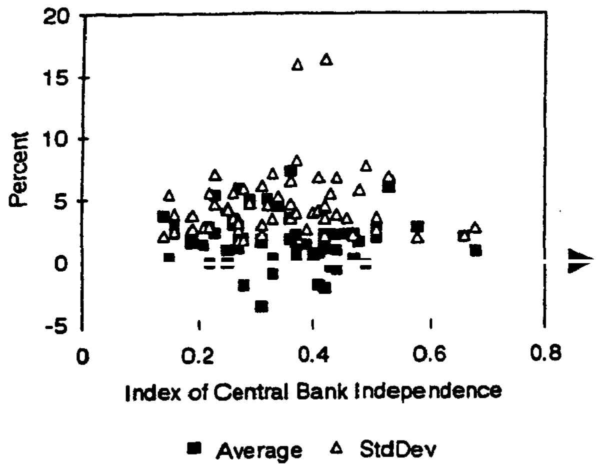


Figure 1 GDP growth and its standard deviation 1971–1985: CWN-based sample.

CONCLUSIONS

Using an alternative measure of central bank independence and examining a larger sample of countries serves to confirm the findings of Alesina and Summers: There appears to be no association between a country's real economic performance, as measured by average growth or its volatility, and the degree of political and economic independence enjoyed by its central bank. Many researchers, however, have documented a negative association between average inflation and central bank independence. While both the degree of independence granted the central bank and the average inflation rate reflect fundamental attitudes of the population, countries that insulate their central banks from direct political pressure do not seem to suffer adverse effects on real incomes.

Inflation imposes significant costs on an economy, particularly by causing arbitrary wealth redistributions and heightened economic uncertainty. These costs clearly lower the level of real income and overall economic welfare, and there is some evidence that inflation may even lower long-term average real growth (see Motley 1993). Consequently, a policy of price stability is seen by many as the most important contribution monetary policy can make to ensuring desirable real economic performance. While there certainly may be important transitional costs to eliminating existing inflation, greater political control over the conduct of monetary policy seems to result in higher average rates of inflation with no apparent gain in average real economic performance.

REFERENCES

Alesina, A., and Summers, L. 1993. Central Bank Independence and Macroeconomic Performance. *Journal of Money, Credit and Banking,* 25 (May) pp. 151–162.

Cargill, T.F., and Hutchison, M. 1990. Monetary Policy and Political Economy: the Federal Reserve and the Bank of Japan. In *The Political Economy of American Monetary Policy,* ed. T. Mayer, pp. 165–180. Cambridge University Press.

Cukierman, A. 1992. *Central Bank Strategy, Credibility and Independence.* Cambridge: MIT Press.

Cukierman, A., Edwards, S. and Tabellini, G. 1993. Seigniorage and Political Instability. *American Economic Review* 82, pp. 537–555.

Cukierman, A., Webb, S., and Neypati, B. 1992. The Measurement of Central Bank Independence and its Effect on Policy Outcomes. *The World Bank Economic Review* 6, pp. 353–398.

Motley, B. 1993. Growth and Inflation: A Cross-Country Study. Federal Reserve Bank of San Francisco Working Paper 93–11.

Summers, R., and Heston, A. 1991. The Penn World Tables (Mark 5): An Expanded Set of International Comparisons, 1950–1988. *Quarterly Journal of Economics* 106 (May) pp. 327–368.

Walsh, C.E. 1993. Central Bank Strategy, Credibility and Independence: A Review Essay. *Journal of Monetary Economics* 32 (October).

This Page Intentionally Left Blank

20

Describing Fed Behavior

John P. Judd

Vice President and Associate Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

Glenn D. Rudebusch

Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

Describing the reasons for the policy actions of the Federal Reserve has long been a popular topic for economists, economic journalists, investors, and others. In particular, there is keen interest in what economists call the Fed's implied "reaction function," which models how the Fed sets monetary policy in response to conditions in the economy. This interest is not surprising given that the reaction function can provide insight into possible future changes in the stance of Fed policy. Also, within the context of a model of the economy, a reaction function provides a basis for evaluating monetary policy (as in Rudebusch and Svensson 1998), as well as for understanding the effects of other policies (for example, fiscal policy) or economic shocks (for example, the 1970s oil embargo) that may induce a monetary policy response. In this chapter, we summarize the results of our research paper—Judd and Rudebusch 1998—which provides estimates of a Fed reaction function.

Large numbers of Fed reaction functions have been estimated by economists. But despite this work, researchers have not been particularly successful in providing a definitive representation of Fed behavior (see Rudebusch 1998). In part, this lack of success stems from the fact that the Fed's specific response to certain economic situations seems to change over time.

One factor that may be associated with changes in the Fed's reaction function over time is changes in the composition of the policymaking body—the Federal Open Market Committee (FOMC). Such compositional changes may bring to the fore policymakers with different preferences and different conceptions of the appropriate operation and the likely transmission of monetary policy. While many people and events influence policy, arguably one of the more important and *identifiable* compositional changes is in the Fed Chairmanship.

In this chapter, we use the Taylor rule as a tool for characterizing Fed policy. In essence, the rule describes a policy regime in which the Fed sets the real (inflation-adjusted) funds rate with an eye toward controlling inflation and stabilizing the business cy-

Reprinted from: *FRBSF Economic Letter*, No. 98-38 (December 25, 1998).

cle. Thus the rule focuses on the variables of primary interest to the Fed. We examine whether the rule is capable of capturing the broad differences in how policy was conducted during the tenures of Fed Chairmen Greenspan, Volcker, and Burns.

TAYLOR RULE

Taylor (1993) suggested a very specific and simple rule for monetary policy. It sets the level of the real funds rate equal to an "equilibrium" real funds rate (a benchmark for neutral policy that is consistent with full employment) plus a weighted average of two gaps: (1) recent inflation less a target rate, and (2) the (percent) deviation of real GDP from an estimate of its potential, or full-employment, level.

Taylor assumed that the equilibrium real interest rate and the inflation target were both equal to 2%, and that the weights the Fed gave to deviations of inflation and output were both equal to 1/2. Thus, for example, if inflation were 1 percentage point above its target and output were at its potential level, the rule would recommend a funds rate of 5 1/2% (3% for inflation plus 2% for the equilibrium real funds rate plus 1/2% for the excess of inflation over its target).

This rule is consistent with a policy regime in which the Fed attempts to control inflation in the long run and to smooth the amplitude of the business cycle in the short run. The arguments in the rule—inflation and the GDP gap—roughly correspond with goals legislated for U.S. monetary policy, namely, stable prices and full employment. In this spirit, Governor Meyer (1998) stresses that stabilizing real GDP around its trend in the short run and controlling inflation in the longer term are important concerns of the Fed. Although U.S. policymakers look at many economic and financial indicators, the two gaps specified in the rule may be stylized measures of their ultimate goals.

Moreover, the GDP gap can be interpreted not only as a measure of business cycle conditions but also as an indicator of future inflation in the context of a Phillips curve model. The productive capacity of the U.S. economy, whether measured by potential GDP, industrial capacity, or the "natural" rate of unemployment, appears to figure prominently in Fed forecasts of future inflation.

In contrast to the Taylor rule, most empirical reaction functions suggest that the Fed responds both to the broad measures of economic performance that are of ultimate interest for policy, such as output and inflation, as well as to so-called intermediate variables, which are not of direct interest to the Fed but may affect or predict the ultimate goal variables. Examples of such intermediate variables include the monetary aggregates, exchange rates, the budget deficit, and commodity prices. However, Fed responses to these intermediate variables are especially likely to change over time because their relationship to the ultimate goal variables may shift.

For example, the monetary aggregates played a more direct role in policy formulation in the 1970s and especially the early 1980s than they do now. Even when the Fed was explicitly targeting the aggregates, it was not ultimately interested in them per se, but instead cared about how the aggregates affected economic performance.

By focusing on the ultimate goals of policy, the Taylor rule may be capable of capturing Fed reactions in a consistent way during periods when the Fed actively targeted money and when it did not. More generally, by eliminating intermediate variables and focusing only on a few basic goal variables, the Taylor rule may be able to avoid some of the instability plaguing previous Fed reaction functions.

FINDINGS

Taylor (1993) showed that his rule does a reasonable job of describing the actual funds rate under Chairman Greenspan. The rule also provides some perspective on policies under Chairmen Burns and Volcker (Judd and Trehan 1995). With regard to the Burns period, the actual funds rate consistently was lower than the rule's recommended rate, which accords with the overall increase in inflation during this period. During the Volcker period, when the Fed significantly reduced inflation, the actual funds rate was consistently higher than what the rule recommended.

But while the original Taylor rule provides a reasonable starting point, it is useful to examine alternatives to Taylor's simple specification by estimating the reaction function weights using the historical record, rather than simply assuming weights as Taylor did. Estimating Taylor-type equations may provide a different or better description of Fed policy. In this chapter, we can only summarize our results. The details of the empirical analysis can be seen in Judd and Rudebusch (1998).

One complication in estimating the Taylor rule is that central banks often appear to adjust interest rates in a gradual fashion—taking small steps toward a desired setting. We allow for such interest rate smoothing by estimating the Taylor rule in the context of a so-called error-correction model. This approach allows for the possibility that the funds rate adjusts *gradually* to achieve the rate recommended by the rule. In fact, such interest rate smoothing is apparent in the regression results for the entire period examined from 1970 to the present.

The estimated Taylor rule for Chairman Greenspan's tenure (1987 to the present) fits the data quite well. The estimated equation explains two-thirds of the quarterly variation in the funds rate during this period. The estimated weight on inflation of 0.54 is very close to what Taylor (1993) assumed (0.5), while the estimated coefficient on the GDP gap of 0.99 is higher than Taylor assumed (0.5). Finally, the data suggest that the equilibrium funds rate and the inflation target both fall in a range of 1 3/4 to 2 3/4%—not far from Taylor's assumption of 2%.

The estimation for the period during which Paul Volcker was Chairman (1979 to 1987) similarly finds evidence that policy was concerned with both the rate of inflation relative to a target and the growth rate of real GDP relative to the growth rate of potential GDP. The coefficient on the inflation gap is again very close to the 0.5 assumed by Taylor, while the response to output growth is 1.5.

However, the equation is estimated for this period with much less precision than for the Greenspan period. For example, the equation explains only slightly less than one-half of the quarterly variation in the funds rate compared with two-thirds for Greenspan. In part, this could be because the problem facing policy in 1979 was so clear. The double-digit inflation prevailing at the time was so far above any reasonable inflation target that policy did not need to be as concerned with the rather refined judgments about funds rate settings provided by a Taylor-style reaction function. Instead, policy could simply keep the real funds rate at a "high" level until inflation began to come down.

A key feature of the reaction function for Chairman Burns's tenure (1970 to 1978) is the clear insignificance of the coefficient on the inflation gap. Instead, the funds rate responded only to the GDP gap. The lack of an implicit or explicit inflation target is consistent with the large increase in inflation during this period.

Of course, other factors may have played a role as well. In particular, there were two large oil shocks that added substantially to the price level.

In addition, the output gap may have been underestimated during this period. The existence of such a mistake has been given an important role during the period by many analysts. Such a consistent string of mistakes would not be too surprising. During this period, productivity and potential output both exhibited a surprising slowdown in growth, a development which is still largely unexplained by economists. At the same time, demographic factors, especially the entrance of the baby boom generation into the labor force, created an increase in the natural rate of unemployment that also was unexpected. Indeed, there was a widespread view that an unemployment rate of 4% was a suitable benchmark rate for policy. In contrast, recent (time-varying) estimates of the natural rate that prevailed during this period are in the 6 to 6 1/2% range (e.g., Gordon 1997). Both of these factors—an underestimate of the GDP gap and the natural rate of unemployment—could have contributed to unduly easy policy, since it may have appeared at the time that inflationary pressures were less severe than they really were.

CONCLUSION

Overall, our analysis finds that the Taylor rule—which describes policy in terms of the Fed's basic goal variables—is a useful framework for summarizing key elements of monetary policy. Estimates of this equation confirm that while inflation was not a key variable guiding policy in the 1970s, policy has focused on controlling inflation and smoothing the business cycle in the 1980s and 1990s.

REFERENCES

Gordon, R. 1997. The Time-Varying NAIRU and Its Implications for Economic Policy. *Journal of Economic Perspectives* 11 (Winter) pp. 11–32.

Judd, J. P., and Trehan, B. 1995. Has the Fed Gotten Tougher on Inflation? *FRBSF Weekly Letter* 95–13 (March 31).

Judd, J. P., and Rudebusch, G. D. 1998. Taylor's Rule and the Fed: 1970–1997. Federal Reserve Bank of San Francisco *Economic Review* 3, pp. 1–16.

Meyer, L. H. 1998. The Strategy of Monetary Policy. The Alan R. Holmes Lecture, Middlebury College, Middlebury, Vermont (March 16).

Rudebusch, G. D. 1998. Do Measures of Monetary Policy Shocks in a VAR Make Sense? *International Economic Review* 39, pp. 907–931.

Rudebusch, G. D. and Svensson, L. E. O. 1998. Policy Rules for Inflation Targeting. NBER Working Paper 6512.

Taylor, J. B. 1993. Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, pp. 195–214.

21
Money

James Madison
U.S. President, 1809–1817

Observations written[1] *posterior to the circular Address of Congress in Sept. 1779, and prior to their Act of March, 1780.*[2]

It has been taken for an axiom in all our reasonings on the subject of finance, that supposing the quantity and demand of things vendible in a country to remain the same, their price will vary according to the variation in the quantity of the circulating medium; in other words, that the value of money will be regulated by its quantity. I shall submit to the judgment of the public some considerations which determine mine to reject the proposition as founded in error. Should they be deemed not absolutely conclusive, they seem at least to shew that it is liable to too many exceptions and restrictions to be taken for granted as a fundamental truth.

If the circulating medium be of universal value as specie, a local increase or decrease of its quantity, will not, whilst a communication subsists with other countries, produce a correspondent rise or fall in its value. The reason is obvious. When a redundancy of universal money prevails in any one country, the holders of it know their interest too well to waste it in extravagant prices, when it would be worth so much more to them elsewhere. When a deficiency happens, those who hold commodities, rather than part with them at an undervalue in one country, would carry them to another. The variation of prices in these cases, cannot therefore exceed the expence and insurance of transportation.

Suppose a country totally unconnected with Europe, or with any other country, to possess specie in the same proportion to circulating property that Europe does; prices there would correspond with those in Europe. Suppose that so much specie were thrown into circulation as to make the quantity exceed the proportion of Europe tenfold, without any change in commodities, or in the demand for them: as soon as such an augmentation had

Reprinted from: *Federal Reserve Bank of Minneapolis Quarterly Review (Fall 1997) 3–17.* This essay originally appeared in two parts, in the December 19 and 22, 1791, issues of Philip Freneau's *National Gazette* of Philadelphia. The version printed here is the edited and annotated version from *The Papers of James Madison,* vol. I, *16 March 1751–16 December 1779,* ed. William T. Hutchinson and William M. E. Rachal (Chicago: University of Chicago Press, 1962), pp. 302–10. The note referenced by a dagger (†) is James Madison's: the numbered notes are those of the University of Chicago Press editors. This version of the essay is reprinted with the permission of the University of Chicago Press.

produced its effect, prices would rise tenfold; or which is the same thing, money would be depreciated tenfold. In this state of things, suppose again, that a free and ready communication were opened between this country and Europe, and that the inhabitants of the former, were made sensible of the value of their money in the latter; would not its value among themselves immediately cease to be regulated by its quantity, and assimilate itself to the foreign value?

Mr. Hume in his discourse on the balance of trade supposes, "that if four fifths of all the money in Britain were annihilated in one night, and the nation reduced to the same condition, in this particular, as in the reigns of the Harrys and Edwards, that the price of all labour and commodities would sink in proportion, and every thing be sold as cheap as in those ages: That, again, if all the money in Britain were multiplied fivefold in one night, a contrary effect would follow." This very ingenious writer seems not to have considered that in the reigns of the Harrys and Edwards, the state of prices in the circumjacent nations corresponded with that of Britain; whereas in both of his suppositions, it would be no less than four fifths different. Imagine that such a difference really existed, and remark the consequence. Trade is at present carried on between Britain and the rest of Europe, at a profit of 15 or 20 per cent. Were that profit raised to 400 per cent, would not their home market, in case of such a fall of prices, be so exhausted by exportation—and in case of such a rise of prices, be so overstocked with foreign commodities, as immediately to restore the general equilibrium? Now, to borrow the language of the same author, "the same causes which would redress the inequality were it to happen, must forever prevent it, without some violent external operation."[3]

The situation of a country connected by commercial intercourse with other countries, may be compared to a single town or province whose intercourse with other towns and provinces results from political connection. Will it be pretended that if the national currency were to be accumulated in a single town or province, so as to exceed its due proportion five or tenfold, a correspondent depreciation would ensue, and every thing be sold five or ten times as dear as in a neighboring town or province?

If the circulating medium be a municipal one, as paper currency, still its value does not depend on its quantity. It depends on the credit of the state issuing it, and on the time of its redemption; and is no otherwise affected by the quantity, than as the quantity may be supposed to *endanger* or *postpone* the redemption.

That it depends in part on the credit of the issuer, no one will deny. If the credit of the issuer, therefore be perfectly unsuspected, the time of redemption alone will regulate its value.

To support what is here advanced, it is sufficient to appeal to the nature of paper money. It consists of bills or notes of obligation payable in specie to the bearer, either on demand or at a future day. Of the first kind is the paper currency of Britain, and hence its equivalence to specie. Of the latter kind is the paper currency of the United States, and hence its inferiority to specie. But if its being redeemable not on demand but at a future day, be the cause of its inferiority, the distance of that day, and not its quantity, ought to be the measure of that inferiority.

It has been shewn that the value of specie does not fluctuate according to local fluctuations in its quantity. Great Britain, in which there is such an immensity of circulating paper, shews that the value of paper depends as little on its quantity as that of specie, when the paper represents specie payable on demand. Let us suppose that the circulating notes of Great Britain, instead of being payable on demand, were to be redeemed at a future day, at the end of one year for example, and that no interest was due on them. If the same assurance prevailed that at the end of the year they would be equivalent to specie, as now prevails that they are every moment equivalent, would any other effect result from such a change, except that the notes would suffer a depreciation equal to one year's interest? They would in that case represent, not the nominal sum expressed on the face of them, but the sum remaining after a deduction of one year's interest. But if when they represent the full nominal sum of specie, their circulation contributes no more to depreciate them, than the circulation of the specie itself would do; does it not follow, that if they represented a sum of specie less than the nominal inscription, their circulation ought to depreciate them no more than so much specie, if substituted, would depreciate itself? We may extend the time from one, to five, or to twenty years; but we shall find no other rule of depreciation than the loss of the intermediate interest.

What has been here supposed with respect to Great Britain has actually taken place in the United States. Being engaged in a necessary war without specie to defray the expence, or to support paper emissions for that purpose redeemable on demand, and being at the same time unable to borrow, no resource was left, but to emit bills of credit to be redeemed in future. The inferiority of these bills to specie was therefore incident to the very nature of them. If they had been exchangeable on demand for specie, they would have been equivalent to it; as they were not exchangeable on demand, they were inferior to it. The degree of their inferiority must consequently be estimated by the time of their becoming exchangeable for specie, that is the time of their redemption.

To make it still more palpable that the value of our currency does not depend on its quantity, let us put the case, that Congress had, during the first year of the war, emitted five millions of dollars to be redeemed at the end of ten years; that, during the second year of the war, they had emitted ten millions more, but with due security that the whole fifteen millions should be redeemed in five years; that, during the two succeeding years, they had augmented the emissions to one hundred millions, but from the discovery of some extraordinary sources of wealth, had been able to engage for the redemption of the whole sum in one year: it is asked, whether the depreciation, under these circumstances, would have increased as the quantity of money increased—or whether on the contrary, the money would not have risen in value, at every accession to its quantity?[4]

It has indeed happened, that a progressive depreciation of our currency has accompanied its growing quantity; and to this is probably owing in a great measure the prevalence of the doctrine here opposed. When the fact however is explained, it will be found to coincide perfectly with what has been said. Every one must have taken notice that, in the emissions of Congress, no precise time has been stipulated for their redemption, nor any specific provision made for that purpose. A general promise entitling the bearer to so many dollars of metal as the paper bills express, has been the only basis of their credit. Every one therefore has been left to his own conjectures as to the time the redemption would be fulfilled; and as every addition made to the quantity in circulation, would naturally be supposed to remove to a proportionally greater distance the redemption of the whole mass, it could not happen otherwise than that every additional emission would be followed by a further depreciation.

In like manner has the effect of a distrust of public credit, the other source of depreciation, been erroneously imputed to the quantity of money. The circumstances under which our early emissions were made, could not but strongly concur, with the futurity of their redemption, to debase their value. The situation of the United States resembled that of an individual engaged in an expensive undertaking, carried on, for want of cash, with bonds and notes secured on an estate to which his title was disputed; and who had besides, a combination of enemies employing every artifice to disparage that security. A train of sinister events during the early stages of the war likewise contributed to increase the distrust of the public ability to fulfill their engagements. Before the depreciation arising from this cause was removed by the success of our arms, and our alliance with France, it had drawn so large a quantity into circulation, that the quantity itself soon after begat a distrust of the *public disposition* to fulfill their engagements; as well as new doubts, in timid minds, concerning the issue of the contest. From that period, this cause of depreciation has been incessantly operating. It has first conduced to swell the amount of necessary emissions, and from that very amount has derived new force and efficacy to itself. Thus, a further discredit of our money has necessarily followed the augmentation of its quantity; but every one must perceive, that it has not been the effect of the quantity, considered in itself, but considered as an omen of public bankruptcy.†[5]

Whether the money of a country, then, be gold and silver, or paper currency, it appears that its value is not regulated by its quantity. If it be the former, its value depends on the general proportion of gold and silver, to circulating property throughout all countries having free inter communication. If the latter, it depend[s] on the credit of the state issuing it, and the time at which it is to become equal to gold and silver.

Every circumstance which has been found to accelerate the depreciation of our currency naturally resolves itself into these general principles. The spirit of monopoly hath affected it in no other way than by creating an artificial scarcity of commodities wanted for public use, the consequence of which has been an increase of their price, and of the necessary emissions. Now it is this increase of emissions which has been shewn to lengthen the supposed period of their redemption, and to foster suspicions of public credit. Monopolies destroy the natural relation between money and commodities; but it is by raising the value of the latter, not by debasing that of the former. Had our money been gold or silver, the same prevalence of monopoly would have had the same effect on prices and expenditures; but these would not have had the same effect on the value of money.

The depreciation of our money has been charged on misconduct in the purchasing departments: but this misconduct must have operated in the same manner as the spirit of monopoly. By unnecessarily raising the price of articles required for public use, it has swelled the amount of necessary emissions, on which has depended the general opinion concerning the time and the probability of their redemption.

The same remark may be applied to the deficiency of imported commodities. The deficiency of these commodities has raised the price of them; the rise of their price has increased the emissions for purchasing them; and with the increase of emissions, have increased suspicions concerning their redemption.

Those who consider the quantity of money as the criterion of its value, compute the intrinsic depreciation of our currency by dividing the whole mass by the supposed necessary medium of circulation. Thus supposing the medium necessary for the United States to be 30,000,000 dollars, and the circulating emissions to be 200,000,000 the intrinsic difference between paper and specie will be nearly as 7 for 1. If its value depends on the time of its redemption, as hath been above maintained, the real difference will be found to be considerably less. Suppose the period necessary for its redemption to be 18 years, as seems to be understood by Congress; 100 dollars of paper 18 years hence will be equal in value to 100 dollars of specie; for at the end of that term, 100 dollars of specie may be demanded for them. They must consequently at this time be equal to as much specie as, with compound interest, will amount, in that number of years, to 100 dollars. If the interest of money be rated at 5 per cent. this present sum of specie will be about 41 1–2 dollars. Admit, however the use of money to be worth 6 per cent. about 35 dollars will then amount in 18 years to 100.35 dollars of specie therefore is at this time equal to 100 of paper; that is, the man who would exchange his specie for paper at this discount, and lock it in his desk for 18 years, would get 6 per cent. for his money. The proportion of 100 to 35 is less than 3 to 1. The intrinsic depreciation of our money therefore, according to this rule of computation, is less than 3 to 1; instead of 7 to 1, according to the rule espoused in the circular address,[6] or 30 or 40 to 1, according to its currency in the market.

I shall conclude with observing, that if the preceding principles and reasoning be just, the plan on which our domestic loans have been obtained, must have operated in a manner directly contrary to what was intended. A loan-office certificate differs in nothing from a common bill of credit, except in its higher denomination, and in the interest allowed on it; and the interest is allowed, merely as a compensation to the lender, for exchanging a number of small bills, which being easily transferable, are most convenient, for a single one so large as not to be transferable in ordinary transactions. As the certificates, however, do circulate in many of the more considerable transactions, it may justly be questioned, even on the supposition that the value of money depended on its quantity, whether the advantage to the public from the exchange, would justify the terms of it. But dismissing this consideration, I ask whether such loans do in any shape, lessen the public debt, and thereby render the discharge of it less suspected or less remote? Do they give any new assurance that a paper dollar will be one day equal to a silver dollar, or do they shorten the distance of that day? Far from it: The certificates constitute a part of the public debt no less than the bills of credit exchanged for them, and have an equal claim to redemption within the general period; nay, are to be paid off long before the expiration of that period, with bills of credit, which will thus return into the general mass, to be redeemed along with it. Were these bills, therefore, not to be taken out of circulation at all, by means of the certificates, not only the expence of offices for exchanging, re-exchanging, and annually paying the interest, would be avoided; but the whole sum of interest would be saved, which must make a formidable addition to the public emissions, protract the period of their redemption, and proportionally increase their depreciation. No expedient could perhaps have been devised more preposterous and unlucky. In order to relieve public credit sinking under the weight of an enormous debt, we invest new expenditures. In order to raise the value of our money, which depends on the time of its redemption, we have recourse to a measure which removes its redemption to a more distant day. Instead of paying off the capital to the public creditors, we give them an enormous interest to change the name of the bit of paper which expresses the sum due to them; and think it a piece of dexterity in finance, by *emitting loan-office certificates,* to elude the necessity of *emitting bills of credit.*

NOTES

1. The original manuscript of the essay is not known to be extract. In the Tracy W. McGregor Library, University of Virginia, is a transcript of about the first one-third of the article, which John C. Payne probably copied from the newspaper version of it.
2. Pledging on September 1, 1779, not to increase its $160 million of outstanding bills of credit by more than 25 percent, and that only in case of a dire emergency, the Continental Congress had John Jay draft a "Circular Address" to the states (adopted September 13) exhorting them to supply enough soldiers, money, and matériel to restore public credit and advance the common cause.

And yet, by March 18, 1780, the gloomy situation obliged Congress to authorize the states to issue new bills of credit and declare that the old continental issues would be redeemed at only one-fortieth of their face value (*Journals of the Continental Congress,* XV, 1052–62: XVL 262–67). Although in the prefatory note Madison declared that he wrote his essay during the six months intervening between these two actions by Congress, he probably could have narrowed the time to the period from late in December 1779 to early in March of the next year. In his brief third-person autobiography, written long afterward. Madison mentioned his election to Congress on December 14, 1779, and then added that "To prepare himself for this service, he employed an unavoidable detention from it in making himself acquainted with the state of Continental affairs and particularly that of the finances which, owing to the depreciation of the paper currency, was truly deplorable. The view he was led to take of the evil, and its causes, was put on paper, now to be found in several periodical publications, particularly Freneau's National Gazette." By "unavoidable detention" he most likely referred to his necessary preparations at Montpelier for his residence in Philadelphia and the heavy snow which delayed his departure for that city until March 6, 1780, or for some days after he had planned to begin the trip. The essay was printed as the fourth in Madison's series of seventeen politically tinged articles appearing in Freneau's newspaper late in President Washington's first term. Even though Madison may have revised his original manuscript before releasing it for publication, it deals with a problem which was much less acuse by 1791 than when he wrote the essay nearly twelve years earlier.

3. Madison accurately reflects the thought, but does not always quote the exact words of David Hume in his *Political Discourses* (Edinburgh, 1752), pp. 82–83.
4. The portion of the essay in the issue of the *National Gazette* for December 19, 1791, ends here. The remainder is in the issue of December 22, 1791.

† As the depreciation of our money has been ascribed to a wrong cause, so, it may be remarked, have effects been ascribed to the depreciation, which result from other causes. Money is the instrument by which men's wants are supplied, and many who possess it will part with it for that purpose, who would not gratify themselves at the expence of their visible property. Many also may acquire it, who have no visible property. By increasing the quantity of money therefore, you both increase the means of spending, and stimulate the desire to spend; and if the objects desired do not increase in proportion, their price must rise from the influence of the greater demand for them. Should the objects in demand happen, at the same juncture, as in the United States, to become scarcer, their prices must rise in a double proportion.

It is by this influence of an augmentation of money on demand, that we ought to account for that proportional level of money, in all countries, which Mr. Hume attributes to its direct influence on prices. When an augmentation of the national coin takes place, it may be supposed either, 1. not to augment demand at all; or, 2. to augment it so gradually that a proportional increase of industry will supply the objects of it; or, 3. to augment it so rapidly that the domestic market may prove inadequate, whilst the taste for distinction natural to wealth, inspires, at the same time, a preference for foreign luxuries. The first case can seldom happen. Were it to happen, no change in prices, nor any efflux of money, would ensue: unless indeed, it should be employed or loaned abroad. The superfluous portion would be either hoarded or turned into plate. The second case can occur only where the augmentation of money advances with a very slow and equable pace: and would be attended neither with a rise of prices, nor with a superfluity of money. The third is the only case, in which the plenty of money would occasion it to overflow into other countries. The insufficiency of the home market to satisfy the demand would be supplied from such countries as might afford the articles in demand: and the money would thus be drained off, till that and the demand excited by it, should fall to a proper level, and a balance be thereby restored between exports and imports. The principle on which Mr. Hume's theory, and that of Montesquieu's before him, is founded, is manifestly erroneous. He considers the money in every country as the representative of the whole circulating property and industry in the country; and thence concludes, that every variation in its quantity must increase or lessen the portion which represents the same portion of property and labor. The error lies in supposing, that because money serves to measure the value of all things, it represents and is equal in value to all things. The circulating property in every country, according to its market rate far exceeds the amount of its money. At Athens oxen, at Rome sheep, were once used as a measure of the value of other things. It will hardly be supposed, they were therefore equal in value to all other things.

5. Madison's entire footnote is in italics in the *National Gazette.* In the last paragraph of the footnote, he refers to Book XXII of Montesquieu's *De l'esprit des lois,* first published in Geneva in 1748 and soon thereafter translated into English. Madison's daring in challenging the correctness of this redoubtable authority is noted by Paul Merrill Spurlin in his *Montesquieu in America, 1760–1801* (Baton Rogue, La. 1940) pp. 175–76.
6. See note 2.

This Page Intentionally Left Blank

22

Monetary Policy and the Great Crash of 1929

A Bursting Bubble or Collapsing Fundamentals?

Timothy Cogley

Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

In recent years, a number of economists have expressed concern that the stock market is overvalued. Some have compared the situation with the 1920s, warning that the market may be headed for a similar collapse. Indeed, some suggest that lax monetary policy contributed to the Great Crash and have argued that current monetary policy is also dangerously lax. For example, an April 1998 *Economist* article stated:

> In the late 1920s, the Fed was also reluctant to raise interest rates in response to soaring share prices, leaving rampant bank lending to push prices higher still. When the Fed did belatedly act, the bubble burst with a vengeance.

To avoid the same mistake, *The Economist* suggested that it would be better for the Fed to take deliberate, preemptive steps to deflate the bubble in share prices. It warned that the bubble could harm the economy if it were to burst suddenly, reducing the value of collateral assets and bringing on a recession. The article went on to say that "the longer that asset prices continue to be pumped up by easy money, the more inflated the bubble will become and the more painful the economic aftereffects when it bursts," and it concluded that "the Fed needs to raise interest rates."

In this *Economic Letter,* I argue that *The Economist* has misinterpreted the lessons of the Great Crash. A closer examination of the events of the late 1920s suggests it is mistaken on at least four points. First, stock prices were not obviously overvalued at the end of 1927. Second, starting in 1928 the Fed shifted toward increasingly tight monetary policy, motivated in large part by a concern about speculation in the stock market. Third, tight monetary policy probably did contribute to a fall in share prices in 1929. And fourth, the depth of the contraction in economic activity probably had less to do with the magnitude of the crash and more to do with the fact that the Fed continued a tight money policy after the

Reprinted from: *FRBSF Economic Letter,* No. 99-10 (Mar. 26, 1999).

crash. Hence, rather than illustrating the dangers of standing on the sidelines, the events of 1928–1930 actually provide a case study of the risks associated with a deliberate attempt to puncture a speculative bubble.

MONETARY POLICY 1927–1930

In 1927, there was a mild recession in the United States. In addition, Britain was threatened by a balance of payments crisis whose proximate cause was a demand by France to convert a large quantity of sterling reserves into gold. Thus, both domestic and international conditions inclined the Fed to shift toward easing. The resulting fall in interest rates helped damp the decline in domestic economic activity and facilitated an outflow of gold toward Britain and France.

Should the Fed have refrained from easing in 1927 because of concerns that the stock market might be overvalued? Measures of conventional valuation suggest the answer is no, for there was no obvious sign of an emerging bubble at that time. For example, Figure 1 illustrates the price-dividend ratio on the value-weighted New York Stock Exchange (NYSE) portfolio. At the end of 1927, the price-dividend ratio was around 23, which is actually a bit below its long-run average of 25. Although share prices had risen rapidly in the 1920s, so had dividends. Given that the price-dividend ratio was slightly below average, the Fed would have had little reason to refrain from easing in a recession year or to decline assistance to a gold standard partner in maintaining balance of payments equilibrium.

In any case, equity prices began to accelerate in January 1928, and they rose by 39% for the year. Dividend payments also grew rapidly that year, and the price-dividend ratio increased by 27%.

Motivated by a concern about speculation in the stock market, the Fed responded aggressively. Between January and July 1928 the Fed raised the discount rate from 3.5% to

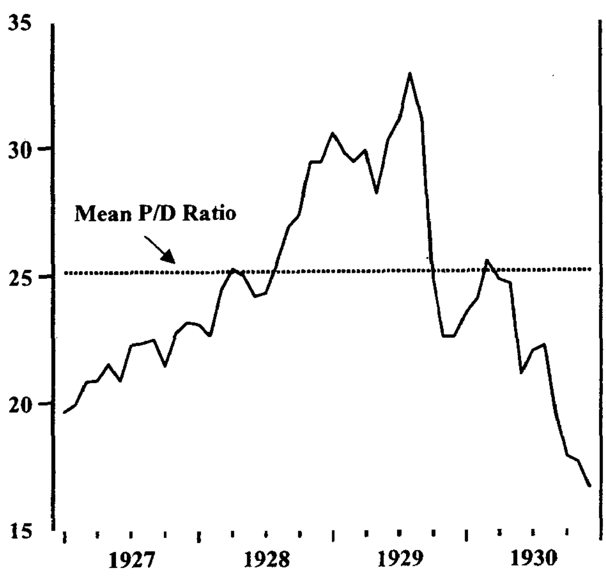


Figure 1 Price-dividend ratio, 1927–1930.

5%. Because nominal prices were falling, the latter translated into a real discount rate of 6%, which is quite high in a year following a recession. At the same time, the Fed engaged in extensive open market operations to drain reserves from the banking system. Hamilton (1987) reports that it sold more than three-quarters of its total stock of government securities: "in terms of the magnitudes consciously controlled by the Federal Reserve, it would have been difficult to design a more contractionary policy."

Furthermore, as Eichengreen (1992) has emphasized, monetary policy was tight not only in the U.S. but also throughout much of the rest of the world. By that time, roughly three dozen countries had returned to the gold standard, and when the Fed tightened, many countries faced a dilemma: Unless their central banks also tightened, lending from the U.S. would be disrupted and their balance of payments would move toward a deficit. In that case, they would either have to devalue or abandon the gold standard altogether. The former option was unattractive for countries with dollar-denominated debts, and the latter was virtually out of the question at the time, especially for countries where restoration of the gold standard had been painful and difficult.

The alternative was to conform with the Fed. By shifting toward more contractionary monetary policies, other gold standard countries could ensure that domestic interest rates would rise in parallel with those in the U.S. and would be able to maintain balance of payments equilibrium. This explains, for example, why the Bank of England shifted toward tighter policy in 1928, three years after Britain had entered a slump. It also explains why countries still rebuilding from WWI would adopt contractionary policies.

The implication is that monetary policy was far more restrictive than a purely domestic perspective might suggest. In 1928 there was a synchronized, global contraction of monetary policy, which occurred primarily because the Fed was concerned about stock prices. These actions had predictable effects on economic activity. By the second quarter of 1929 it was apparent that economic activity was slowing. The U.S. economy peaked in August and fell into a recession in September.

What were the effects on the stock market? At the beginning of 1929, it seemed that the contractionary measures taken in 1928 were working. The NYSE price-dividend ratio reached a local peak in January and then fell gradually through the first half of the year. Thus, it appeared that stock prices had stabilized. Furthermore, shares still were not obviously overvalued. The local peak was reached at 30.5, which is roughly 20% above the long-term average. Dividends had grown rapidly through 1928, and investors projecting similar growth rates forward may have been willing to settle for dividend yields somewhat below the long-run average.

Monetary policy was on hold during the first half of 1929, and some economists have argued that inaction in this period was responsible for the events that followed. But three observations are relevant here. First, as mentioned above, price-dividend ratios had stabilized and were falling gradually. To a contemporary observer, it would have appeared that the actions of 1928 were having the intended effects. Second, it was becoming increasingly apparent that general economic activity was slowing, and many other countries already had entered recessions. And third, while monetary policy was not becoming tighter, it was still quite tight. Short-term real interest rates were still around 6%, and there was no growth in the monetary base.

Price-dividend ratios continued to fall until July 1929, but then prices began to take off. In August, the Fed raised the discount rate by another percentage point to 6%. The stock market peaked in the first week of September. It is worth noting that at its peak the price-dividend ratio was 32.8, which is well below values reached in the 1960s or 1990s.

Share prices declined in a more or less orderly fashion until the end of October, but then the market crashed. From its peak, the price-dividend ratio fell roughly 30%, to a level roughly similar to that prevailing at the beginning of 1928, when the Fed began to tighten.

In the immediate aftermath of the crash, the New York Fed took prompt and decisive action to ease credit conditions. When investors attempted to liquidate their equity holdings, many lenders also called their loans to securities brokers. With the encouragement of officials at the New York Fed, many of these brokers' loans were taken over by New York banks, who were allowed to borrow freely at the discount window for this purpose. The New York Fed also bought government securities on its own account in order to inject reserves into the banking system. In this way, they were able to contain an incipient liquidity crisis and prevent the crash from spreading to money markets.

But this respite from tight money proved to be temporary. After the liquidity crisis had been contained, monetary policy once again resumed a contractionary stance. Throughout 1930, officials at the New York Fed repeatedly proposed that the System buy government securities on the open market, but they were systematically rebuffed. The reasons other members of the Federal Reserve gave for opposing monetary expansion are instructive. Several felt that much of the investment undertaken in the previous expansion was fundamentally unsound and that the economy could not recover until it was scrapped. Others felt that a monetary expansion would only ignite another round of speculative activity, perhaps even in the stock market. In any event, monetary policy remained contractionary; the monetary aggregates fell by 2% to 4%, and long-term real interest rates increased.

By maintaining a contractionary stance throughout 1930, after a recession had already begun, the Fed contributed to a further decline in economic activity and share prices. By the end of the year, the price-dividend ratio had fallen to 16.6, or roughly 34% below the long-run average. By then, there was a consensus that speculative activity had been eliminated!

WERE THE FED'S ACTIONS STABILIZING OR DESTABILIZING?

If one grants that a speculative bubble existed at the beginning of 1928, when the Fed began to tighten, then stocks must have still been overvalued in the aftermath of the crash. After all, price-dividend ratios were about the same in the dark days of November 1929 as at the beginning of 1928, and fundamentals must surely have taken a turn for the worse. If equities were still overvalued, it follows that a further dose of contractionary monetary policy was needed to purge speculative elements from the market. Perhaps this is what motivated the famous advice of Treasury Secretary Andrew Mellon to President Herbert Hoover, to "liquidate labor, liquidate stocks, liquidate the farmers, liquidate real estate." To argue that the actions of 1928–1930 were stabilizing, one must adopt the liquidationist position.

On the other hand, if one interprets the Great Crash as a bursting bubble, so that shares were more or less properly valued in the aftermath, then it follows that they were probably also not far from their fundamental values at the start of 1928, when the Fed began to tighten. Again, prices and price-dividend ratios were about the same after the crash, and fundamentals had surely become less favorable. According to this interpretation, the Fed's initial actions may have been destabilizing, and the actions of 1930 certainly were.

CONCLUSION

In retrospect, it seems that the lesson of the Great Crash is more about the difficulty of identifying speculative bubbles and the risks associated with aggressive actions conditioned on noisy observations. In the critical years 1928 to 1930, the Fed did not stand on the sidelines and allow asset prices to soar unabated. On the contrary, its policy represented a striking example of *The Economist's* recommendation: a deliberate, preemptive strike against an (apparent) bubble. The Fed succeeded in putting a halt to the rapid increase in share prices, but in doing so it may have contributed one of the main impulses for the Great Depression.

REFERENCES

Economist. 1998. America's Bubble Economy (April 18).

Eichengreen, B. 1992. *Golden Fetters: The Gold Standard and the Great Depression, 1919–1939.* Oxford: Oxford University Press.

Hamilton, J. D. 1987. Monetary Factors in the Great Depression. *Journal of Monetary Economics* 19, pp. 145–169.

This Page Intentionally Left Blank

23

A Primer on Monetary Policy Part I: Goals and Instruments

Carl E. Walsh
Professor of Economics, University of California, Santa Cruz, California

Recent interest rate increases have focused the public's attention on the conduct of monetary policy and the role of the Federal Reserve. In February, the Fed's policymaking body, the Federal Open Market Committee (FOMC), raised the federal funds rate a quarter of a percentage point, the first such increase since 1989. The FOMC voted again for two more 25 basis point increases in March and April, and at the May 17th meeting the federal funds rate was pushed up another half a percentage point.

These actions were designed to prevent inflation from rising above its current annual rate of under 3 percent. The interest rate hikes were consistent with the Fed's desire to move the economy gradually towards even lower rates of inflation in order to achieve price stability. Because these actions occurred before any actual increase in inflation, the FOMC's policy has generated wide debate among economists and in the popular press.

In order to grasp the issues at the heart of the recent debates, it is helpful to understand some of the elements involved in setting monetary policy. This chapter provides the first of a two-part primer designed to give an overview of issues in how monetary policy is conducted. Part one reviews the goals of monetary policy and the basic instruments the Fed can use to conduct policy. The next chapter discusses the role of intermediate targets, indicators, rules and forecasts in the implementation of monetary policy.

THE GOALS OF MONETARY POLICY

In the lobby of the Federal Reserve Bank of San Francisco, visitors are exposed to the difficulties of implementing monetary policy through an electronic video game. The object is to time the release of a dart from a moving arm in order to hit the bull's-eye of a moving target. The moving bull's-eye reflects, in a graphic manner, the uncertainty that exists over the appropriate goals of monetary policy, and the changing values of these goals as a result of developments both in the economy and in our understanding of the economy.

Reprinted from: *FRBSF Economic Letter,* No. 94–27 (Aug. 5, 1994).

The "official" goals of monetary policy, enshrined in the Federal Reserve Act (Section 2A(1)), are to "promote . . . maximum employment, stable prices, and moderate long-term interest rates." If maximum employment is interpreted as that level of employment consistent with the numerical goal of 4 percent unemployment established by the Humphrey-Hawkins "Full Employment and Balanced Growth Act of 1978," then few economists believe it is possible to achieve all these goals at once in the U.S. economy.

The belief that 4 percent unemployment and stable prices are inconsistent is shaped by the widely accepted "natural rate hypothesis." It argues that the economy's average equilibrium unemployment rate, often called the natural rate of unemployment, is independent of monetary policy. Most current estimates of the natural rate place it in the range of 6 to $6\frac{1}{2}$ percent. Taken together, these imply that the unemployment rate will tend to average around the 6 to $6\frac{1}{2}$ percent range in the long run, regardless of the conduct of monetary policy.

What if the Federal Reserve tried to achieve 4 percent unemployment in the long run? *Consistent* attempts to expand the economy beyond its potential for production will result in higher inflation while ultimately failing to produce lower average unemployment. In fact, extreme rates of inflation (or deflation) may so disrupt the role of the price system in directing resources in a market economy that the result could be slower average growth *and* higher average unemployment. Although economists continue to debate whether reducing inflation from moderate to low rates would significantly improve the long-run performance of the economy, most believe that there are no long-term gains from consistently pursuing expansionary policies.

While the Fed has little effect on the natural rate of unemployment, the Fed can determine the economy's average rate of inflation. Thus, many recent commentators have emphasized the need to define the goals of monetary policy in terms of low or zero inflation, which *is* achievable. In 1989, Representative Neal (D, NC) actually introduced a bill in Congress that would have amended the Federal Reserve Act to make price stability the sole goal of U.S. monetary policy. The idea that central banks should have goals defined only in terms of price stability or low inflation is not new. John Maynard Keynes wrote in 1924 that ". . . they (Treasury and Bank of England) should adopt the stability of sterling prices as their primary objective" (p. 202).

In practice, the Fed, like most central banks, cares about both inflation and measures of the short-run cyclical performance of the economy. However, pursuing multiple goals can create conflicts for policy; for example, the desire to mitigate short-run downturns raises the issue of whether this goal should take precedence over a low-inflation goal at any particular point in time. Thus, it is important to avoid allowing short-run, temporary successes in preventing employment losses during recessions from leading to longer-run failures in maintaining low inflation. Proponents of more activist policy argue, however, that monetary policy can help to stabilize the economy and should act to offset temporary downturns in economic activity. They argue that responding to temporary, cyclical fluctuations need not be inconsistent with maintaining a firm commitment to low average inflation.

One effect of having multiple, conflicting goals is that it leads to political pressures on the Fed, varying in strength and intensity over time, for lower interest rates, for faster growth, or for lower inflation. Political pressure on monetary policy is usually criticized for its tendency to emphasize short-run considerations over longer-run objectives. For example, politicians who may have time horizons that extend only to the next election will be tempted to push for more expansionary policy, which may produce lower unemployment and faster real growth in the short run, even though in the long run it can only lead to higher inflation. Unless the central bank has sufficient independence from political institutions to resist such pressures, the result is likely to be higher average inflation with no appreciable effect on average unemployment or real growth. Alesina and Summers (1993) find that greater central bank independence is associated with lower average inflation, but they find no association with average real growth among the industrialized economies.

In sum, monetary policy is continually faced with a conflict between what it can temporarily achieve in the short run and what it can permanently achieve in the longer run. That is why even those economists who believe monetary policy has an important role to play in helping to stabilize short-run business cycle fluctuations also have, in recent years, increasingly emphasized the importance of maintaining a commitment to low average rates of inflation.

THE INSTRUMENTS OF MONETARY POLICY

The Fed does not control inflation or unemployment directly; instead, the Fed must decide on the settings for the tools, or instruments, of policy that it does control directly as it attempts to achieve its objectives of low inflation and stable economic growth. It is predominantly through *open market operations*—purchases and sales of government securities—that the Fed attempts to influence the economy and achieve its policy goals. Open market operations influence the level of bank reserves in the economy, which in turn influences the level of interest rates, the provision of money and credit, investment spending, and the pace of economic activity.

Banks are legally required to hold a fraction of certain types of deposit accounts that they issue as reserves. They keep these reserves in the form of vault cash or deposits with the Fed. When banks need additional reserves on a short-term basis, they can borrow them from other banks who happen to have more reserves than they need. The interest rate on the overnight borrowing of reserves is called the federal funds rate. The funds rate adjusts to balance the supply and demand for reserves.

Open market operations affect the supply of reserves in the banking system. In the Fed buys government securities, it pays for them by adding reserves to the banking system; this increases the supply of reserves, which lowers the cost of borrowing reserves—the federal funds rate falls. When the Fed sells government securities, the reverse happens: the supply of reserves falls, and the federal funds rate rises.

If the demand for reserves were perfectly predictable, the Fed could predict exactly the relationship between the quantity of reserves and the funds rate. In this case, it could use either reserves or the funds rate as its policy instrument equally well. But, because reserve demand can fluctuate unpredictably, the choice between the use of a quantity and the use of an interest rate as the chief instrument of policy does make a difference. To see why, suppose the economy grows faster than predicted, putting upward pressure on interest rates as credit demand increases. If the Fed tries to control the quantity of reserves, it will not accommodate the greater demand for credit, and the funds rate will rise. This will tend to push up other interest rates and act as an automatic brake on the economy. If the Fed is using the funds rate as its instrument, this pressure for higher interest rates will automatically produce a rise in the supply of reserves as the Fed acts to prevent the funds rate from rising. In this example, policy that focuses on the quantity of reserves would be less likely to let inflation rise.

If, in contrast, pressures for higher interest rates came from a financial market development, such as tighter regulatory supervision of bank lending practices, a policy that acted to keep the quantity of reserves constant would lead some key lending rates to rise, which would tend to have a contractionary effect on the economy. If policy acted to keep the funds rate constant, then the supply of reserves would automatically increase to offset the contractionary effect of the financial disturbance. During most of the post-war era, the interest rate approach to implementing policy—setting the level of the funds rate—provides a more accurate description of Fed operating procedures.

The link between open market activities and the federal funds rates is fairly straightforward. The other linkages between policy actions and the behavior of the economy are subject to more controversy.

REFERENCES

Alesina, A., and Summers, L. 1993. Central Bank Independence and Macroeconomic Performance. *Journal of Money, Credit and Banking* 25 (May) pp. 151–162.

Keynes, J. M. 1924. *Monetary Reform.* New York: Harcourt Brace and Company.

24

A Primer on Monetary Policy Part II: Targets and Indicators

Carl E. Walsh
Professor of Economics, University of California, Santa Cruz, California

CHANNELS OF MONETARY POLICY

Economists disagree about the exact linkages among monetary policy actions, inflation, and economic activity. Most agree that banks play a critical role in the transmission process, although evidence is inconclusive about whether it is through the liability side of banks' balance sheets (deposits and other components of the money supply) or through the asset side (bank loans). In either case, the general view is that monetary policy works by affecting interest rates. Increases in interest rates raise the cost of borrowing and lead to reductions in business investment spending and household purchases of durable goods such as autos and new homes. These declines in spending reduce the aggregate demand for the economy's output, leading firms to cut back on production and employment. Conversely, interest rate declines stimulate aggregate spending and lead to increases in production and employment.

Since the Fed can control the federal funds rate, it would appear to be a simple matter to link policy actions—changes in the funds rate—to real economic activity. Unfortunately, four critical problems arise in implementing monetary policy. First, while the Fed can affect market interest rates, spending decisions and economic activity depend on *real* interest rates, that is, market rates corrected for expected rates of inflation. Second, economic activity is likely to be related to both short-term and long-term real interest rates, while the Fed most directly controls very short-term market rates. Third, the Fed is interested ultimately in measures of economic performance like inflation, real economic growth and employment, yet data on these variables that might be used to guide policy are not available every day or every week or even every month. And fourth, policy actions taken today will affect the economy only with a significant lag so that policy changes must be made in anticipation of future developments in the economy. Because the first two of these issues have been recently discussed by Trehan (1993) and Cogley (1993), they are touched upon only briefly here.

Reprinted from: *FRBSF Economic Letter,* No. 94-28 (Aug. 19, 1994).

REAL INTEREST RATES

Aggregate spending is related not to market interest rates but to the expected real rate of interest. Since it is difficult to measure expected inflation, it is hard to know the current level of real interest rates. And variations in expected inflation can make a big difference. In 1978, the funds rate averaged 7.93 percent, but the rate of inflation was 9.1 percent; if the inflation was fully anticipated, that 7.93 funds rate was equivalent to an expected real rate of negative 1.17 percent. Today the funds rate is 4.25. If the market expects a continuation of the current 3 percent inflation rate, then today's funds rate translates into a positive 1.25 percent expected real rate. So a funds rate of 4.25 percent today may be more restrictive than the 7.93 funds rate was in 1978. With inflation expectations difficult to measure, economists can disagree about the current level of real rates and therefore the stance of monetary policy.

In addition, the Fed can only influence the level of real interest rates in the short-run; it cannot permanently prevent the real rate from returning to its equilibrium level without risking accelerating inflation or deflation. Persistent attempts to keep real rates too low will initially generate an economic expansion that will lead to more rapid inflation. As individuals come to expect higher inflation, real rates will tend to adjust back to their equilibrium level. Further expansionary policy would be needed to keep the real rate down, leading to further increases in inflation.

Most estimates of expected inflation imply that real short-term interest rates earlier this year were very low, too low to be consistent with steady real growth at a sustainable rate. However, the real rate of interest consistent with sustained growth varies over time in ways that are difficult to measure or predict. Thus, the benchmark against which any estimate of the current real rate should be compared is itself not directly measured. So economists can disagree about whether current real rates are too high or too low.

LONG-TERM INTEREST RATES

Aggregate spending is related both to long-term real interest rates and to the short-term rates the Fed can affect directly. Long-term rates will be equal to the average of the expected future short-term rates plus a risk factor that reflects the premium necessary to induce risk-averse investors to hold long-term bonds. An increase in short-term rates that is viewed to be temporary will have a much smaller impact on long rates than would an increase expected to be relatively persistent.

Long-term interest rates can be expressed as the sum of an expected real return and an adjustment for expected inflation. Long rates have the potential, therefore, to provide information about the market's expectations about inflation. Long rates will tend to rise (fall) if higher (lower) inflation is expected. After the Fed's most recent increase in the funds rate on May 17th, long-term interest rates actually declined. This was interpreted as evidence that financial market participants were confident the Fed had tightened sufficiently to ensure inflation would not increase. Unfortunately, long-term interest rates also vary because of variations in the expected rate of return. Because of the difficulties in predicting these variations, it is not always possible to interpret changes in long-term interest rates as providing information about inflation expectations.

INTERMEDIATE TARGETS

Ideally, the Fed would like to be able to monitor continuously its ultimate goals, like the rate of inflation, in order to adjust its policy instruments. Unfortunately, new data on inflation are available only monthly, while data on GDP growth are available only quarterly. Consequently, the Fed must rely on data available more frequently, such as interest rates, which it can observe continuously, or monetary aggregates, which are available weekly, as *intermediate targets* to help guide policy. An intermediate target is a variable that, while not directly under the control of the Fed (that is, it is not an instrument like the federal funds rate), responds fairly quickly to policy actions, is observable frequently, and bears a predictable relationship to the ultimate goals of policy.

To use an intermediate target, the Fed must first determine the value for the intermediate target consistent with the desired goals. The Fed then adjusts its instruments in order to ensure the intermediate target variable takes on the chosen value. That is, policy is conducted as if the intermediate target value were the goal of policy. If new information suggests the intermediate target variable is diverging from the targeted value, policy instruments are adjusted to return it to target.

During the early 1980s, several different measures of the money supply served as intermediate targets; for example, when M2 was growing above its target range, this signaled a need to tighten policy by contracting the growth of bank reserves. Slow M2 growth was a signal to expand reserve growth. However, the relationship between the monetary aggregates and the ultimate goals of monetary policy became increasingly unpredictable, reducing the value of the aggregates as intermediate targets (see Judd-Trehan 1992).

POLICY INDICATORS

Currently, the Fed has no single reliable intermediate target that could be used to guide policy; consequently, the Fed must rely on many variables for information to guide policy. These variables are indicators of the current state of the economy or of future developments in the economy.

Indicators of future developments are needed because it takes time for a monetary policy action to affect the economy. Policy actions taken in early 1994 are likely to have their greatest effect on the economy in late 1994 and early 1995. This makes it imperative that policy actions be taken not on the basis of current economic conditions, but on the basis of forecasts of future economic conditions. To wait to shift the Fed's policy stance until inflation actually increases, for example, would mean that inflationary momentum will have already developed, making the task of reducing inflation that much harder and more costly in terms of job losses. In the past, the Fed has been criticized for waiting too long before adjusting its policies.

Basing policy on forecasts creates its own difficulties. Because economic developments are difficult to predict, forecasts often turn out to be wrong. And because forecasters often disagree, there will be corresponding differences over the appropriate stance of policy. The current situation is a case in point. The Fed has tightened policy, not in response to any current rise in inflation, but on the basis of its forecast of rising inflation in the future if it maintained its previous policy stance. Critics have claimed that future inflation increases are unlikely. Because the debate is over forecasts of what inflation would have done under the Fed's previous policy, they are difficult to resolve.

In the absence of an agreed upon intermediate target to guide policy, the Fed must evaluate a number of variables that may serve to indicate future economic developments in order to determine if policy changes are appropriate. Among the indicators that have been proposed are nominal income growth, real interest rates, commodity prices, exchange rates and the price of gold. For example, the Fed could use nominal income growth as an indicator by comparing the most recent data on nominal growth to the growth rate consistent with sustained real growth and low inflation. Since most estimates of the economy's long-run sustainable growth rate of real income are in the 2 to $2\frac{1}{2}$ percent range, if the inflation target were 1 percent, nominal income growth should be in the 3 to $3\frac{1}{2}$ percent range.

Because no single indicator variable consistently provides accurate information on the future of the economy or the stance of monetary policy, the Fed must rely on a number of indicators; it "looks at everything." In principle, this is just what the Fed should do. Exchange rates, nominal income growth, real interest rates, money supply growth, commodity prices, and so on all provide some information that is useful for conducting policy.

Unfortunately, each indicator also can provide misleading signals about the economy, and often the signals they give are contradictory. During the last two years, for example, while real interest rates were low indicating expansionary monetary policy, the M2 definition of the money supply was growing very slowly, indicating a more contractionary stance of policy.

As an alternative to using forecasts or relying on a number of indicator variables, many economists have proposed simple rules to guide Fed behavior. The most famous was Milton Friedman's rule of maintaining a constant growth rate of the money supply. More recently, rules for the monetary base (currency plus bank reserves), M2, nominal GDP, and the funds rate have been studied (for example, see Judd and Motley 1991). In general, these alternatives are "feedback rules": The Fed's policy instrument is adjusted on the basis of recent movements in measures of economic activity such as nominal income growth, the unemployment rate, or actual inflation. Such rules can help to reduce the uncertainty associated with monetary policy actions by making policy more predictable.

CONCLUSIONS

The conduct of monetary policy often consists of balancing inconsistent goals using sometimes unreliable indicators to manipulate tools whose effects on the economy are uncertain. Despite these uncertainties, the general conduct of monetary policy in recent years has received surprisingly wide approval. The current controversy over interest rate increases is not about the fundamental need to prevent a resurgence of inflation, but instead has centered on the difficulty of forecasting the future course of inflation.

REFERENCES

Cogley, T. 1993. Monetary Policy and the Long-Term Real Interest Rate. Federal Reserve Bank of San Francisco *Weekly Letter* (December 3).

Judd, P., and Motley, B. 1991. Nominal Feedback Rules for Monetary Policy. Federal Reserve Bank of San Francisco *Economic Review* (3) pp. 3–17.

Judd, P., and Trehan, B. 1992. Money, Credit, and M2. Federal Reserve Bank of San Francisco *Weekly Letter* (September 4).

Trehan, B. 1993. Real Interest Rates. Federal Reserve Bank of San Francisco *Weekly Letter* (November 5).

Walsh, C. E. 1994. A Primer on Monetary Policy Part I: Goals and Instruments. Federal Reserve Bank of San Francisco *Weekly Letter* (August 5).

25

U.S. Monetary Policy
An Introduction

Economic Research Department
Federal Reserve Bank of San Francisco, San Francisco, California

U.S. monetary policy affects all kinds of economic and financial decisions people make in this country—whether to get a loan to buy a new house or car or to start up a company, whether to expand a business by investing in a new plant or equipment, and whether to put savings in a bank, in bonds, or in the stock market, for example. Furthermore, because the United States is the largest economy in the world, its monetary policy also has significant economic and financial effects on other countries.

The object of monetary policy is to influence the performance of the economy, as reflected in such factors as inflation, economic output, and employment. It works by affecting demand across the economy—that is, the population's willingness to spend on goods and services.

While most people are familiar with the fiscal policy tools that affect demand—such as taxes and government spending—many are less familiar with monetary policy and its tools. Monetary policy is conducted by the Federal Reserve System, the nation's central bank, and it influences demand mainly by raising and lowering short-term interest rates.

This *Economic Letter* is an introduction to U.S. monetary policy as it is currently conducted, and it answers a series of questions:

How is the Fed structured to make monetary policy decisions?
What are the Fed's goals?
What tools does it use to implement its policies?
How does monetary policy affect the U.S. economy?
How does the Fed formulate strategies to reach its goals?

HOW IS THE FEDERAL RESERVE STRUCTURED?

The Federal Reserve System (called the Fed, for short) is the nation's central bank. It was established by an Act of Congress in 1913 and consists of the seven members of the Board

Reprinted from: *FRBSF Economic Letter,* No. 99-01 (Jan 1, 1999).

of Governors in Washington, D.C., and twelve Federal Reserve District Banks (for a discussion of the Fed's overall responsibilities, see *The Federal Reserve System: Purposes and Functions*).

The Congress structured the Fed to be independent within the government—that is, although the Fed is accountable to the Congress, it is insulated from day-to-day political pressures. This reflects the conviction that the people who control the country's money supply should be independent of the people who frame the government's spending decisions. Most studies of central bank independence rank the Fed among the most independent in the world.

What Makes the Fed Independent?

Three structural features make the Fed independent: the appointment procedure for governors, the appointment procedure for Reserve Bank Presidents, and funding.

Appointment Procedure for Governors

The seven Governors on the Federal Reserve Board are appointed by the President of the United States and confirmed by the Senate. Independence derives from a couple of factors: First, the appointments are staggered to reduce the chance that a single U.S. President could "load" the Board with appointees; second, their terms of office are 14 years—much longer than elected officials' terms.

Appointment Procedure for Reserve Bank Presidents

Each Reserve Bank President is appointed to a five-year term by that Bank's Board of Directors, subject to final approval by the Board of Governors. This procedure adds to independence because the Directors of each Reserve Bank are not chosen by politicians but are selected to provide a cross-section of interests within the region, including those of depository institutions, nonfinancial businesses, labor, and the public.

Funding

The Fed is structured to be self-sufficient in the sense that it meets its operating expenses primarily from the interest earnings on its portfolio of securities. Therefore, it is independent of Congressional decisions about appropriations.

How Is the Fed "Independent within Government"?

Even though the Fed is independent of Congressional appropriations and administrative control, it is ultimately accountable to Congress and comes under government audit and review. The Chairman, other Governors, and Reserve Bank Presidents report regularly to the Congress on monetary policy, regulatory policy, and a variety of other issues and meet with senior Administration officials to discuss the Federal Reserve's and the federal government's economic programs.

Who Makes Monetary Policy?

The Fed's FOMC (Federal Open Market Committee) has primary responsibility for conducting monetary policy. The FOMC meets in Washington eight times a year and has twelve members: the seven members of the Board of Governors, the President of the Federal Reserve Bank of New York, and four of the other Reserve Bank Presidents, who serve in rotation. The remaining Reserve Bank Presidents attend the meetings and contribute to the Committee's discussions and deliberations.

In addition, the Directors of each Reserve Bank contribute to monetary policy by making recommendations about the appropriate discount rate, which are subject to final approval by the Governors. (See "What Are the Tools of Monetary Policy?")

WHAT ARE THE GOALS OF U.S. MONETARY POLICY?

Monetary policy has two basic goals: to promote "maximum" output and employment and to promote "stable" prices. These goals are prescribed in a 1977 amendment to the Federal Reserve Act.

What Does "Maximum" Output and Employment Mean?

In the long run, the level of output and employment in the economy depends on factors other than monetary policy. These include technology and people's preferences for saving, risk, and work effort. So, "maximum" employment and output means the levels consistent with these factors in the long run.

But the economy goes through business cycles in which output and employment are above or below their long-run levels. Even though monetary policy can't affect either output or employment in the long run, it can affect them in the short run. For example, when demand contracts and there's a recession, the Fed can stimulate the economy—temporarily—and help push it back toward its long-run level of output by lowering interest rates. Therefore, in the short run, the Fed and many other central banks are concerned with stabilizing the economy—that is, smoothing out the peaks and valleys in output and employment around their long-run levels.

If the Fed Can Stimulate the Economy out of a Recession, Why Doesn't It Stimulate All the Time?

Consistent attempts to expand the economy beyond its long-run level will result in capacity constraints that lead to higher and higher inflation, without producing lower unemployment or higher output in the long run. In other words, not only are there no long-term gains from consistently pursuing expansionary policies, but there's also a price—higher inflation.

What's so Bad about Higher Inflation?

High inflation can hinder economic growth. For example, when inflation is high, it also tends to vary a lot, and that makes people uncertain about what inflation will be in the future. That uncertainty can hinder economic growth in a couple of ways—it adds an inflation risk premium to long-term interest rates, and it complicates the planning and contracting by business and labor that are so essential to capital formation.

High inflation also hinders economic growth in other ways. For example, because many aspects of the tax system are not indexed to inflation, high inflation distorts economic decisions by arbitrarily increasing or decreasing after-tax rates of return to different kinds of economic activities. In addition, it leads people to spend time and resources hedging against inflation instead of pursuing more productive activities.

So that's why the Other Goal is "Stable Prices"?

Yes. Although monetary policy cannot expand the economy or reduce unemployment in the long run, it can stabilize prices in the long run. Price "stability" is basically low inflation—that is, inflation that's so low that people don't worry about it when they make decisions about what to buy, whether to borrow or invest, and so on.

If Low Inflation Is the Only Thing the Fed Can Achieve in the Long Run, Why Isn't it the Sole Focus of Monetary Policy?

Because the Fed can determine the economy's average rate of inflation, some commentators—and some members of Congress as well—have emphasized the need to define the goals of monetary policy in terms of price stability, which is achievable. However, volatility in output and employment also is costly to people. In practice, the Fed, like most central banks, cares about both inflation and measures of the short-run performance of the economy.

Are the Two Goals Ever in Conflict?

Yes, sometimes they are. One kind of conflict involves deciding which goal should take precedence at any point in time. For example, suppose there's a recession and the Fed works to prevent employment losses from being too severe; this short-run success could turn into a long-run problem if monetary policy remains expansionary too long, because that could trigger inflationary pressures. So it's important for the Fed to find the balance between its short-run goal of stabilization and its longer-run goal of maintaining low inflation.

Another kind of conflict involves the potential for pressure from the political arena. For example, in the day-to-day course of governing the country and making economic policy, politicians may be tempted to put the emphasis on short-run results rather than on the longer-run health of the economy. The Fed is somewhat insulated from such pressure, however, by its independence, which allows it to achieve a more appropriate balance between short-run and long-run objectives.

Why Don't the Goals Include Helping a Region of the Country that's in Recession?

Often enough, some state or region is going through a recession of its own while the national economy is humming along. But the Fed can't concentrate its efforts to expand the weak region for two reasons. First, monetary policy works through credit markets, and since credit markets are linked nationally, the Fed simply has no way to direct stimulus to any particular part of the country that needs help. Second, if the Fed stimulated whenever any state had economic hard times, it would be stimulating much of the time, and this would mean higher inflation.

But this focus on the well-being of the national economy doesn't mean that the Fed ignores regional economic conditions. Extensive regional data and anecdotal information are used, along with statistics that directly measure developments in the regional economy, to fit together a picture of the national economy's performance. This is one advantage to having regional Federal Reserve Bank Presidents sit on the FOMC: They are in close contact with economic developments in their regions of the country.

WHAT ARE THE TOOLS OF MONETARY POLICY?

The Fed can't control inflation or influence output and employment directly; instead, it affects them indirectly, mainly by raising or lowering short-term interest rates. The Fed affects interest rates mainly through open market operations and the discount rate, and both of these methods work through the market for bank reserves, known as the federal funds market.

What Are Bank Reserves?

Banks and other depository institutions (for convenience, we'll refer to all of these as "banks") are legally required to hold a specific amount of funds in reserve. These funds, which can be used to meet unexpected outflows, are called reserves, and banks keep them as cash in their vaults or as deposits with the Fed. Currently, banks must hold 3–10% of the funds they have in interest-bearing and non-interest-bearing checking accounts as reserves (depending on the dollar amount of such accounts held at each bank).

What Is the Federal Funds Market?

From day to day, the amount of reserves a bank has to hold may change as its deposits change. When a bank needs additional reserves on a short-term basis, it can borrow them from other banks that happen to have more reserves than they need. These loans take place in a private financial market called the federal funds market.

The interest rate on the overnight borrowing of reserves is called the federal funds rate or simply the "funds rate." It adjusts to balance the supply of and demand for reserves. For example, an increase in the amount of reserves supplied to the federal funds market causes the funds rate to fall, while a decrease in the supply of reserves raises that rate.

What Are Open Market Operations?

The major tool the Fed uses to affect the supply of reserves in the banking system is open market operations—that is, the Fed buys and sells government securities on the open market. These operations are conducted by the Federal Reserve Bank of New York.

Suppose the Fed wants the funds rate to fall. To do this, it buys government securities from a bank. The Fed then pays for the securities by increasing that bank's reserves. As a result, the bank now has more reserves than it is required to hold. So the bank can lend these excess reserves to another bank in the federal funds market. Thus, the Fed's open market purchase increases the supply of reserves to the banking system, and the federal funds rate falls.

When the Fed wants the funds rate to rise, it does the reverse, that is, it sells government securities. The Fed receives payment in reserves from banks, which lowers the supply of reserves in the banking system, and the funds rate rises.

What Is the Discount Rate?

Banks also can borrow reserves from the Federal Reserve Banks at their "discount windows," and the interest rate they must pay on this borrowing is called the discount rate. The total quantity of discount window borrowing tends to be small, because the Fed discourages such borrowing except to meet occasional short-term reserve deficiencies (see *The Federal Reserve: Purposes and Functions* for a discussion of other types of discount window borrowing that are unrelated to monetary policy).

The discount rate plays a role in monetary policy because, traditionally, changes in the rate may have "announcement effects"—that is, they sometimes signal to markets a significant change in monetary policy. A higher discount rate can be used to indicate a more restrictive policy, while a lower rate may signal a more expansionary policy. Therefore, discount rate changes are sometimes coordinated with FOMC decisions to change the funds rate.

What about Foreign Currency Operations?

Purchases and sales of foreign currency by the Fed are directed by the FOMC, acting in cooperation with the Treasury, which has overall responsibility for these operations. The Fed does not have targets, or desired levels, for the exchange rate. Instead, the Fed gets involved to counter "disorderly" movements in foreign exchange markets, such as speculative movements that may disrupt the efficient functioning of these markets or of financial markets in general. For example, during some periods of disorderly declines in the dollar, the Fed has purchased dollars (sold foreign currency) to absorb some of the selling pressure.

Intervention operations involving dollars, whether initiated by the Fed, the Treasury, or by a foreign authority, are not allowed to alter the supply of bank reserves or the funds rate. The process of keeping intervention from affecting reserves and the funds rate is called the "sterilization" of exchange market operations. As such, these operations are not used as a tool of monetary policy.

HOW DOES MONETARY POLICY AFFECT THE ECONOMY?

The point of implementing policy through raising or lowering interest rates is to affect people's and firms' demand for goods and services. This section discusses how policy actions affect real interest rates, which in turn affect demand and ultimately output, employment, and inflation.

What Are Real Interest Rates and Why Do They Matter?

For the most part, the demand for goods and services is not related to the market interest rates quoted on the financial pages of newspapers, known as nominal rates. Instead, it is related to real interest rates—that is, nominal interest rates minus the expected rate of inflation.

Variations in expected inflation can make a big difference in interpreting the stance of monetary policy. In 1978, the nominal funds rate averaged 8%, but the rate of inflation was 9%. So, even though nominal interest rates were high, monetary policy actually was stimulating demand with a negative real funds rate of minus 1%.

By contrast, in early 1998, the nominal funds rate was 5% and the inflation rate was running at about 2%. This implied a positive 3% real funds rate. So the nominal funds rate of 8% in 1978 was more stimulative than the 5% nominal funds rate in early 1995.

How Do Real Interest Rates Affect Economic Activity in the Short Run?

Changes in real interest rates affect the public's demand for goods and services mainly by altering borrowing costs, the availability of bank loans, the wealth of households, and foreign exchange rates.

For example, a decrease in real interest rates lowers the cost of borrowing and leads to increases in business investment spending and household purchases of durable goods, such as autos and new homes. In addition, lower real rates and a healthy economy may increase banks' willingness to lend to businesses and households. This may increase spending, especially by smaller borrowers who have few sources of credit other than banks. Lower real rates make common stocks and other such investments more attractive than bonds and other debt instruments; as a result, common stock prices tend to rise. Households with stocks in their portfolios find that the value of their holdings has gone up, and this increase in wealth makes them willing to spend more. Higher stock prices also make it more attractive for businesses to invest in plant and equipment by issuing stock. In the short run, lower real interest rates in the United States also tend to reduce the foreign exchange value of the dollar, which lowers the prices of the exports we sell abroad and raises the prices we pay for foreign-produced goods. This leads to higher aggregate spending on goods and services produced in the United States.

The increase in aggregate demand for the economy's output through these various channels leads firms to raise production and employment, which in turn increases business spending on capital goods even further by making greater demands on existing factory capacity. It also boosts consumption further because of the income gains that result from the higher level of economic output.

How Does Monetary Policy Affect Inflation?

Wages and prices will begin to rise at faster rates if monetary policy stimulates aggregate demand enough to push labor and capital markets beyond their long-run capacities. In fact, a monetary policy that persistently attempts to keep short-term real rates low will lead eventually to higher inflation and higher nominal interest rates, with no permanent increases in output or decreases in unemployment. As noted earlier, in the long run, output and employment cannot be set by monetary policy. In other words, while there is a trade-off between higher inflation and lower unemployment in the short run, the trade-off disappears in the long run.

Policy also can affect inflation directly through people's expectations about future inflation. For example, suppose the Fed eases monetary policy. If consumers and business people expect higher inflation in the future, they'll ask for bigger increases in wages and prices. That in itself will raise inflation without big changes in employment and output.

Doesn't U.S. Inflation Depend on Worldwide Capacity, not Just U.S. Capacity?

In this era of intense global competition, it might seem parochial to focus on U.S. capacity as a determinant of U.S. inflation, rather than on world capacity. For example, some argue that even if unemployment in the U.S. drops to very low levels, U.S. workers wouldn't be able to push for higher wages, because they're competing for jobs with workers abroad, who are willing to accept much lower wages.

This reasoning doesn't hold up too well, however, for a couple of reasons. First, a large proportion of what we consume in the U.S. isn't affected very much by foreign trade. One example is health care, which isn't traded internationally, and which amounts to about 14% of GDP.

Second, even when we consider goods that *are* traded internationally, the effect on U.S. prices is largely offset by flexible foreign exchange rates. Suppose the price of steel, or some other good, is lower in Japan than in the U.S. When U.S. manufacturers buy Japanese steel, they have to pay for it in yen, which they buy on the foreign exchange market. As a result, the value of the yen will climb relative to the dollar, and the cost of Japanese steel to U.S. firms will go up—even though the Japanese have not changed the (yen) price they charge.

How Long Does It Take a Policy Action to Affect the Economy and Inflation?

The Tags in monetary policy are long and variable. The major effects of a change in policy on growth in the overall production of goods and services usually are felt within three months to two years. And the effects on inflation tend to involve even longer lags, perhaps one to three years, or more.

Why Are the Lags so Hard to Predict?

Since monetary policy is aimed at affecting people's demand, it's dealing with human responses, which are changeable and hard to predict.

For example, the effect of a policy action on the economy will depend on what people think the Fed action means for inflation in the future. If people believe that a tightening of policy means the Fed is determined to keep inflation under control, they'll immediately expect low inflation in the future, so they're likely to ask for smaller wage and price increases, and this will help to achieve that end. But if people aren't convinced that the Fed is going to contain inflation, they're likely to ask for bigger wage and price increases, and that means that inflation is likely to rise. In this case, the only way to bring inflation down is to tighten so much and for so long that there are significant losses in employment and output.

HOW DOES THE FED FORMULATE ITS STRATEGIES?

The Fed's job of stabilizing output in the short run and promoting price stability in the long run is made more difficult by two main factors: the long and variable lags in policy, and the uncertain influences of factors other than monetary policy on the economy.

What Problems Do Lags Cause?

The Fed's job would be much easier if monetary policy had swift and sure effects. Policymakers could set policy, see its effects, and then adjust the settings until they eliminated any discrepancy between economic developments and the goals.

But with the long lags and uncertain effects of monetary policy actions, the Fed must be able to anticipate the effects of its policy actions into the distant future. To see why, suppose the Fed waits to shift its policy stance until it actually sees an increase in inflation. That would mean that inflationary momentum already had developed, so the task of reducing inflation would be that much harder and more costly in terms of job losses. Not surprisingly, anticipating policy effects in the future is a difficult task.

What Problems Are Caused by Other Influences on the Economy?

Output, employment, and inflation are influenced not only by monetary policy actions, but also by such factors as our government's taxing and spending policies, the availability and price of key natural resources (such as oil), economic developments abroad, financial conditions at home and abroad, and the introduction of new technologies.

In order to have the desired effect on the economy, the Fed must take into account the influences of these other factors and either offset them or reinforce them as needed. This isn't easy because sometimes these developments occur unexpectedly and because the size and timing of their effects are difficult to estimate.

The recent currency crisis in East Asia is a good example. Over the past year or so, economic activity in those countries has either slowed or declined, and this has reduced their demand for U.S. products. In addition, the foreign exchange value of most of their currencies has depreciated, and this has made Asian-produced goods less expensive for us to buy and U.S.-produced goods more expensive in Asian countries. By themselves, these factors would reduce the demand for U.S. products and therefore lower our output and employment. As a result, this is a factor that the Fed has had to consider in setting monetary policy.

Another example is the spread of new technologies that can enhance productivity. When workers and capital are more productive, the economy can expand more rapidly without creating inflationary pressures. In the last few years, there have been indications that the U.S. economy may have experienced a productivity surge, perhaps brought on by computers and other high-tech developments. The issue for monetary policymakers is how much faster productivity is increasing and whether those increases are temporary or permanent.

With All These Uncertainties, How Does the Fed Know How and When Its Policies Will Affect the Economy?

The Fed looks at a whole range of indicators of the future course of output, employment, and inflation. Among the indicators are measures of the money supply, real interest rates, the unemployment rate, nominal and real GDP growth, commodity prices, exchange rates, various interest rate spreads (including the term structure of interest rates), and inflation expectations surveys. Economic forecasting models help give structure to understanding the interplay of these indicators and policy actions. But these models are far from perfect—so policymakers rely on their own less formal judgments about indicators as well. Indeed, policymakers often disagree about how important one indicator is rather than another—and this isn't surprising, because the indicators can be hard to interpret, and they can even give contradictory signals.

To illustrate the difficulties of interpreting these indicators, consider the problems with three of the most prominent: the money supply measures (M1, M2, and M3), real interest rates, and the unemployment rate.

What Are the Problems of Using the Money Supply as an Indicator of Future Economic Performance?

Before much of the deregulation of the financial markets in the 1980s, measures of the money supply were pretty reliable predictors of aggregate spending; moreover, they could be controlled relatively well by the Fed. So the Fed paid special attention to them and to their annual target ranges during the 1970s and 1980s. In fact, from late 1979 to late 1982 the Fed explicitly targeted money on a short-term basis.

But the predictable relationship between the money supply and aggregate spending began to fall apart once financial markets were deregulated and new financial instruments were introduced. For example, consider M1, the narrowest monetary measure, which includes only currency and (fully) checkable deposits. Before deregulation, banks couldn't pay explicit interest on the deposits in M1, so people tended to keep only as much in them as they needed for their transactions; that made those deposits track spending pretty closely.

Once banks were allowed to pay explicit interest nation-wide on checkable deposits, M1 no longer reflected spending so well because people started to leave money in those deposits over and above what they needed for transactions. Furthermore, once private financial markets started introducing instruments that competed with M1 deposits, some people shifted their funds to those instruments, and that also weakened the relationship between M1 and spending. Ultimately, the same kinds of deterioration occurred with the broader money supply measures, M2 and M3.

The Fed still establishes annual ranges for M2 and M3, as well as for total nonfinancial debt, as required by Congress. However, given the problems with the reliability of the aggregates, they have come to play a less central role in the formulation of monetary policy in the 1990s.

What Are the Problems with Using Real Interest Rates as Indicators of Future Economic Performance?

Real interest rates are natural variables to consider as policy indicators, since they are influenced by the Fed and they are a key link in the transmission mechanism of monetary policy. But real interest rates are problematic as indicators of real GDP for at least two reasons.

First, it is not always obvious when real rates are "high" or "low." The reason is that real rates are figured as the nominal rate minus expected future inflation. The level of expected future inflation may be hard to estimate. Second, it also is not obvious how to determine the equilibrium real interest rate—that is, the rate that would be consistent with the full employment of labor and with real GDP being at its long-run level. This rate is needed as a benchmark to judge whether a given real interest rate is expansionary or contractionary.

The equilibrium real rate varies over time in ways that are difficult to measure or predict, and it depends on many factors, such as the productivity of investment, fiscal policy, tax rates, and preferences for risk and saving. So, unless real interest rates are extremely high or low relative to historical experience, it can be difficult to interpret the implications of observed market interest rates for future economic developments.

Why Is It Hard to Pinpoint the Natural Rate of Unemployment?

The unemployment rate sometimes is used as an indicator of future inflation. In judging the inflationary implications of the unemployment rate, some economists focus on the so-called "natural rate" of unemployment as a benchmark. The natural rate is the umemployment rate that would occur when short-run cyclical factors have played themselves out—

that is, when wages have had time to adjust to balance labor demand and supply. All else equal, if unemployment is below the natural rate, inflation would tend to rise; likewise, if unemployment is above the natural rate, inflation would tend to fall. But it is difficult to know what the natural rate of unemployment is, because it can change if the structure of the labor market changes. For example, the natural rate rose temporarily in the 1970s as more women sought jobs. And in recent years, some economists have argued that the natural rate has fallen because of worker "insecurity" stemming from rapid changes in the job skills needed by firms as computers and other new technologies were introduced.

Is That Why Policymakers Look at so Many Indicators?

Although all of the indicators mentioned above provide some useful information, none is reliable enough to be used mechanically as a sole target or guide to policy.

As a result, each FOMC policymaker must process all the available information according to his or her own best judgment and with the advice of the best research available. They then discuss and debate the policy options at FOMC meetings and try to reach a consensus on the best course of action.

GLOSSARY OF TERMS

Capital market The market in which corporate equity and longer-term debt securities (those maturing in more than one year) are issued and traded.

Central bank Principal monetary authority of a nation, which performs several key functions, including issuing currency and regulating the supply of money and credit in the economy. The Federal Reserve is the central bank of the United States.

Depository institution Financial institution that obtains its funds mainly through deposits from the public; includes commercial banks, savings and loan associations, savings banks, and credit unions.

Discount rate Interest rate at which an eligible depository institution may borrow funds, typically for a short period, directly from a Federal Reserve Bank. The law requires that the board of directors of each Reserve Bank establish the discount rate every fourteen days subject to the approval of the Board of Governors.

Equilibrium real interest rate The level of the real interest rate that is consistent with the level of long-run output and full employment.

Excess reserves Amount of reserves held by an institution in excess of its reserve requirement and required clearing balance.

Federal funds rate The interest rate at which banks borrow surplus reserves and other immediately available funds. The federal funds rate is the shortest short-term interest rate, with maturities on federal funds concentrated in overnight or one-day transactions.

Fiscal policy Federal government policy regarding taxation and spending; set by Congress and the Administration.

Foreign currency operations Purchase or sale of the currencies of other nations by a central bank for the purpose of influencing foreign exchange rates or maintaining orderly foreign exchange markets. Also called foreign-exchange market intervention.

Foreign exchange rate Price of the currency of one nation in terms of the currency of another nation.

Government securities Securities issued by the U.S. Treasury or federal agencies.

Gross Domestic Product (GDP) The total market value of a nation's output of goods and services. GDP may be expressed in terms of product—consumption, investment, government purchases of goods and services, and net exports—or, it may be expressed in terms of income earned—wages, interest, and profits.
Inflation A rate of increase in the general price level of all goods and services. (This should not be confused with increases in the prices of specific goods relative to the prices of other goods.)
Inflationary expectations The rate of increase in the general price level anticipated by the public in the period ahead.
Long-term interest rates Interest rates on loan contracts—or debt instruments, such as Treasury bonds or utility, industrial, or municipal bonds—having maturities greater than one year. Often called capital market rates.
M1 Measure of the U.S. money stock that consists of currency held by the public, travelers checks, demand deposits, and other fully checkable deposits.
M2 Measure of the U.S. money stock that consists of M1, certain overnight repurchase agreements and certain overnight Eurodollars, savings deposits (including money market deposit accounts), time deposits in amounts of less than $100,000, and balances in money market mutual funds (other than those restricted to institutional investors).
M3 Measure of the U.S. money stock that consists of M2, time deposits of $100,000 or more at all depository institutions, term repurchase agreements in amounts of $100,000 or more, certain term Eurodollars, and balances in money market mutual funds restricted to institutional investors.
Market interest rates Rates of interest paid on deposits and other investments, determined by the interaction of the supply of and demand for funds in financial markets.
Monetary policy A central bank's actions to influence short-term interest rates and the supply of money and credit, as a means of helping to promote national economic goals. Tools of U. S. monetary policy include open market operations, discount rate policy, and reserve requirements.
Natural rate of unemployment The rate of unemployment that can be sustained in the long run and that is consistent with constant inflation.
Nominal interest rates Current stated rates of interest paid or earned.
Open market operations Purchases and sales of government securities and certain other securities in the open market, through the Domestic Trading Desk at the Federal Reserve Bank of New York as directed by the Federal Open Market Committee, to influence short-term interest rates and the volume of money and credit in the economy. Purchases inject reserves into the banking system and stimulate growth of money and credit; sales do the opposite.
Productivity The level of output per hour of work.
Real GDP GDP adjusted for inflation. Real GDP provides the value of GDP in constant dollars, which is used as an indicator of the volume of the nation's output.
Real interest rates Interest rates adjusted for the expected erosion of purchasing power resulting from inflation. Technically defined as nominal interest rates minus the expected rate of inflation.
Short-term interest rates Interest rates on loan contracts—or debt instruments such as Treasury bills, bank certificates of deposit, or commercial paper—having maturities less than one year. Often called money market rates.
Total nonfinancial debt Includes outstanding credit market debt of federal, state, and local governments and of private nonfinancial sectors (including mortgages and other kinds

of consumer credit and bank loans, corporate bonds, commercial paper, bankers acceptances, and other debt instruments).

SUGGESTED READING

For an overview of the Federal Reserve System and its functions, see:

The Federal Reserve System: Purposes and Functions, 8th ed. Washington, DC: Board of Governors, Federal Reserve System, December 1994.

The Federal Reserve System in Brief. Federal Reserve Bank of San Francisco.

This Page Intentionally Left Blank

26

Against the Tide Malcolm Bryan and the Introduction of Monetary Aggregate Targets

R. W. Hafer
Professor of Economics, Southern Illinois University, Edwardsville, Illinois

Monetary policy was freed from the straightjacket of pegging U.S. Treasury interest rates following the Treasury–Federal Reserve Accord in 1951. This newfound freedom led to a growing debate inside and outside the Federal Reserve System about the appropriate measures to use as operating guides. As the 1950s progressed, the Federal Reserve focused on controlling market interest rates to achieve its policy goals. This policy, which remains in use today, came under fire from a handful of policymakers and policy watchers. An alternative suggested by some was to place greater emphasis on the behavior of the monetary aggregates in setting policy. This article examines the contributions of Malcolm Bryan, president of the Federal Reserve Bank of Atlanta from 1951 through 1965, to this debate and to the development of monetary policy in the postaccord era.

Bryan parted company with most of his colleagues on the Federal Open Market Committee (FOMC) during the late 1950s and into the 1960s. Bryan tried to steer policy away from focusing on interest rates and money market conditions to placing more weight on the monetary aggregates. A reading of the transcripts of the FOMC meetings during this time reveals that Bryan's alternative policy reflected his desire to prevent the disruptive effects on the economy from short-run fluctuations in money growth and the longer-term effects of expansive Fed actions, namely, inflation.[1] Bryan and a few other committee members considered money market conditions and changes in interest rates to be inadequate indicators of policy actions. Bryan argued that monetary aggregates not only provided better feedback on policy changes than tone and feel but also afforded the FOMC a better gauge for measuring the success of achieving its desired policies.

Bryan was not a lone figure in this debate, but his contributions deserve special mention.[2] For example, in 1959 he became the first member of the FOMC to introduce an explicit quantitative, long-run aggregate target into postwar policy discussions. In 1960 he further separated himself from the majority opinion on the FOMC by developing and in-

Reprinted from: *Federal Reserve Bank of Atlanta Economic Review* (First Quarter 1999) 20–37.

troducing short-run, monetary growth targets—growth cones, as they would become known in the 1970s—into the policy debate.[3] Bryan used these quantitative indicators to help interpret and steer Fed policy during the late 1950s and early 1960s. Reviewing the debate over the usefulness of monetary aggregates as operating guides provides an informative case study of U.S. monetary policy, one that offers valuable lessons for monetary policy even today.

To provide a background for Bryan's introduction of aggregate targeting, the article first describes the state of monetary policy and the economy during the period following the 1953–54 recession and through 1958. Of particular interest is how the FOMC interpreted economic developments as the economy emerged from the recession and the inflationary shadow of the Korean War. Against this backdrop, the article discusses Bryan's 1959 introduction of a long-term aggregate target, one based on the postwar trend in the growth of reserves. The next section focuses on the policy debate in 1960 and Bryan's introduction of a short-term, reserve growth cone as an operating guide. The concluding section offers some final observations.

MONETARY POLICY AND THE ECONOMY: 1956–58

Prior to the Treasury–Federal Reserve Accord, monetary aggregates essentially were ignored as policy operated within the confines of supporting Treasury security prices. Following the accord, the policy role of monetary aggregates expanded, though the shift was more cosmetic than real (Friedman 1982). The Federal Reserve focused on achieving desired market rates and maintaining orderly markets to hit stated policy goals, such as sustained economic growth and low inflation. Federal Reserve policy during the 1950s and into the 1960s often relied on hitting target levels of free reserves—the difference between banks' excess reserve holdings and reserves borrowed through the discount window—in the banking system to bring about changes in financial markets. Goodfriend (1991), for example, argues that free reserves in the 1950s and 1960s, like nonborrowed reserves in the 1980s, provided a distraction to the Fed's primary policy concern of manipulating market interest rates. Thus, even though free reserves are a "monetary" aggregate, their use in policy was conditioned on activity in the financial markets.

Changes in free reserves were brought about as the FOMC and the manager of the Open Market Desk at the Federal Reserve Bank of New York took their cues from events in both the domestic and, later, international financial markets. Monetary policy in the 1950s thus depended heavily on the subjective judgment of the FOMC and the manager of the desk.[4] Their role was to interpret developments in the financial markets—to determine the tone and feel of the markets—and how these would influence and be influenced by policy actions.[5] The combination of free reserves and tone and feel did give the Fed some control over short-term interest rates, essentially the three-month Treasury bill rate. But the policy had its costs. Calomiris and Wheelock (1998) assert that the Fed's reliance on free reserves as an operating guide simply recycled the policies that Benjamin Strong had advocated during the 1920s. Operating under the limited constraints of the gold standard, the Fed manipulated free reserves to achieve desired levels of short-term interest rates with one eye toward the domestic economy and the other toward the growing problems of external imbalances. Such policy choices, Wheelock (1997) suggests, help explain the upward drift in money growth and inflation that lasted for the next two decades.

Dissension among FOMC members arose over the best course for policy following the 1953–54 recession. The economy in 1955 grew quite rapidly after the recession, with real gross national product (GNP) increasing at an annual rate of more than 6 percent. Brisk real growth and the sharp run-up in prices that followed the Korean War made inflation a primary concern at FOMC meetings throughout 1955. Hetzel notes that inflation was "the primary macroeconomic preoccupation of the political system in the 1950s" (1995, 6). By the end of 1956, however, real growth had slowed considerably, increasing at an annual rate of only 1.4 percent for the year. Even so, the members' inflation fears now seemed justified: the price level increased at an annual rate of more than 3.5 percent during 1956, up from a 2.5 percent annual rate of change in 1955.

The FOMC reacted to the potential of higher rates of inflation with a policy of increased restraint during 1956 and into 1957. William McChesney Martin, the chairman of the Board of Governors, voiced the majority opinion that the Fed should not repeat the mistake it made coming out of the last recession, essentially that of not raising market rates fast enough to curb inflation. The increased policy restraint resulted in money growth (M1) falling from a 2.2 percent annual growth rate in 1955 to a 1.1 percent rate in 1956. This constraint persisted into 1957, with M1 decreasing at an annual rate of 0.5 percent in 1957, bank credit flat, and a three-month Treasury bill rate that rose throughout the year.

This episode intensified the committee's internal debate over the choice of policy guides. Bryan's comments at the January 28, 1957, FOMC meeting are representative of the confusion and uncertainty that using free reserves engendered. He argued that the behavior of free reserves was not "particularly useful at the present time" (FOMC 1957, 13) and that some alternative should be discussed.[6] Board economist Woodlief Thomas noted that using tone and feel, the companion operating guide at the time, often led to changes that unfortunately contradicted the policy desires of the committee. Governor J.L. Robertson, among others, also expressed discontent with tone and feel as an effective operating guide. The FOMC's dilemma was that neither approach seemed to provide very reliable signals about policy actions and their effects on the economy.

Chairman Martin took the position often expressed by Fed chairmen: operating guides such as tone and feel may be less than perfect, but they afforded the FOMC and the manager of the Open Market Desk the needed flexibility to respond to unforeseen changes in financial markets or the economy. Robert Roosa recalled that "the Federal Reserve has had to rely primarily on experimental probing. . . . *[U]tilizing its own qualitative concept of pressure,* it has withheld or released new bank reserves . . . the 'feel' arising from participation in securities markets and broader judgments of current economic trends" (1960, 262; emphasis added).[7] Alfred Hayes, president of the New York Fed and vice chairman of the FOMC, gives another perspective on the process of using tone and feel to guide policy: "[T]he tone of the market is a very difficult thing to describe unless you are actually sitting at this trading desk, which is the nerve center of the bank and the nerve center of the System for keeping in touch with credit and banking and money market developments. But I would say that it [tone] is a compound of all kinds of impressions you get from the volume of trading, the speed of trading, what is happening to prices, what the bank's position is, whether the dealers are hard up for financing or have plenty of financing, whether funds are well distributed throughout the country or not well distributed" (Atkinson 1969, 85).[8] Chairman Martin and others extolled the flexibility that tone and feel offered. Martin also dismissed claims that monetary aggregates could serve as credible operating guides, insisting on more than one occasion that monetary policy would not be constrained by the "dead hand of statistics," something that he associated with reliance on the aggregates. Sentiment for a change in operating guides was not overwhelming.

By the end of 1957, a few members of the FOMC advised that the Fed's restrictive policy was having a deleterious effect on the economy and needed to be reversed. Bryan was one of the more vocal critics of current policy. At the November 12, 1957, meeting he warned that the lack of reserve and money growth was creating "a terrific drag" on the economy. Unless the FOMC moved to quickly reverse this policy, Bryan warned, the Federal Reserve would be "party to producing economic convulsions" (FOMC 1957, 695). This view was not new for Bryan (Bryan 1938, 1948). A decade earlier, in a speech before the Alabama Bankers Association, Bryan suggested that "the central bank must *lean against the breeze* in times of boom and inflation and likewise in times of depression and deflation" (1948; emphasis added). At the December 17, 1957, meeting Bryan reiterated his long-held view on the nature of policy effects, further stating that, "I believe [monetary policy] can play its most effective role in a downturn if monetary ease is injected *during the early stages of a downward movement* rather than after the recession is well underway. . . . [I]n the face of a now clearly perceptible economic downturn, our effective policy, whatever our intentions, has been to allow a reserve base providing for no growth whatever in the economy. I believe it is clear that the continuation of such policy must finally be an important causative factor in promoting a serious recession" (FOMC 1957, 801–2; emphasis added).

Most other members of the FOMC, however, pressed for continued restraint since their primary concern was to avoid "sloppy" financial markets. In general, committee members and the chairman believed that reversing the course of monetary policy would have little impact on curtailing any recessionary momentum that might already be under way in the economy.

Bryan's view of monetary policy and its effect on real economic activity was not ordinary for a member of the FOMC at that time. His approach to judging policy, contrary to his colleagues and many others in the economics profession, employed ideas associated with a small number of monetary economists that were beginning to circulate through the profession. For example, compare Bryan's comments cited above with Milton Friedman's testimony before the Joint Economic Committee in March 1958. Friedman asserted that actions taken by the Federal Reserve were a "causative factor" in explaining past recessions. Friedman's analysis of recent Fed policies concluded that "the tight money policy of 1956 and 1957 which coexisted with rising prices . . . [is] with us in the current recession" (Friedman 1958, 250, quoted in Meigs 1976, 445).

Bryan was aware of and closely followed developments in monetary economics, as a correspondence between Bryan and Friedman indicates. In a letter dated April 7, 1959, Bryan wrote to Friedman asking for a copy of his paper "Some Theoretical and Empirical Aspects of the Supply of Money."[9] From Friedman's response, it is clear that Bryan had sent along a copy of his paper "The Sovereign, the Central Bank, and the Monetary Standard" (1959), which he had delivered several times in speeches. Later Friedman wrote to Bryan, sending him a copy of his recently published *A Program for Monetary Stability*. This correspondence, albeit limited, suggests that Bryan followed the developments and debates in the increasingly active area of monetary economics and sought input from one of its leading theorists and proponents for an aggregates-based policy.[10]

Bryan's policy analysis was not solely the product of others' research, however. Bryan's views on the role and effects of monetary policy, as indicated in a series of speeches, was established as early as 1938. By 1957 Bryan already was arguing that significant, short-run changes in money growth were likely to influence real economic activity. This date suggests that Bryan's policy stance predates the monetarist position usually associated with the Federal Reserve Bank of St. Louis. For instance, Homer Jones, a lead-

Box 1 Biography of Malcolm Bryan

Bryan was born in 1902 in Wateska, Illinois, a small town of several thousand a little more than 100 miles south of Chicago. At the age of twenty-two Bryan graduated with a bachelor's degree from the University of Illinois, where he remained for an additional year to earn a master's degree. Presumably the master's degree was in economics since following graduation Bryan took a position in the economics department at the University of Georgia.

During his years at the University of Georgia (1925–36), Bryan served in several positions. In 1929 he was a member of Georgia's Special Tax Commission, and in 1934 he served on the Special Committee on Banking and Taxation under the auspices of the U.S. Treasury. From 1933 to 1937, Bryan served as editor of the *Southern Economic Journal*. Perhaps the most important aspect of this period in Bryan's life is that he spent 1927 and 1928 doing postgraduate work in economics at the University of Chicago, where it is likely that his views on the role of monetary policy and the link between monetary aggregates and the economy were influenced by economists Viner, Knight, Douglas, Mints, and Simon, all faculty members at that time.

Bryan left the University of Georgia in 1936 and began his career in the Federal Reserve System with a two-year stint as an economist at the Board of Governors. Afterward he returned to Georgia as a vice president of the Federal Reserve Bank of Atlanta, a position that he held from 1938 through 1941. In 1941 Bryan was promoted to first vice president of the Atlanta Fed, where he remained until 1946. While Bryan was at the Atlanta Fed, he was elected president of the Southern Economics Association in 1942. He also served on the American Technical Staff, part of the negotiating team at the Bretton Woods Monetary and Financial Conference in 1944.

After the end of World War II, Bryan left the Atlanta Fed to become vice chairman of the Trust Company of Georgia. He remained in this job from 1946 until 1951. During this interval away from the Fed, Bryan served as a member of the Senate Finance Committee's Advisory Committee on Social Security (1947–48) and as part of the Economic and Financial Mission to Peru in 1948. Bryan left the Trust Company of Georgia in 1951, returning to the Atlanta Fed as its president. He remained in this position for the rest of his professional career, retiring in 1965. Bryan died in 1967.

ing proponent of using monetary aggregates, did not begin as research director at St. Louis until 1958. And Meigs's observation that "[t]he *new* element in the St. Louis position in 1960 was a recognition that short run changes in the money stock can have adverse effects on income and employment" (1976, 447; emphasis added) was a conviction Bryan used in his policy analysis. The development suggests that the maverick views so often associated with the St. Louis bank were already operating at the Atlanta bank.[11]

Monetary policy abruptly changed direction in 1958. The growth of M1 increased dramatically: after decreasing at an annual rate of 2.1 percent in the first quarter of the year, the growth rate of M1 jumped to more than 6 percent during the next two quarters. At the same time, real output roared out of the recession, increasing at better than a 10 percent annual rate during the second half of the year. In contrast to the positions of his colleagues, Bryan's statements at FOMC meetings throughout 1958 reflect an evolution in his view about the economic effects of short-run fluctuations in money growth. Eschewing the common practice of measuring policy on a meeting-to-meeting basis—often a period of only weeks—Bryan began to compare the level of reserves at one meeting with that of the previous year. While others focused on measuring reserve growth on a meeting-to-meeting basis, Bryan put current policy analysis and discussion into a longer-term perspective in order to understand the current stance of policy actions. This development is reflected in his introduction of a reserve growth target in 1959.

INTRODUCING A RESERVE GROWTH TARGET: 1959

Bryan's concern about the inflationary effects associated with long-term reserve growth is consistent with his often expressed distress about the dangers of inflation and the Federal Reserve's responsibility to contain it. A popular notion at the time was that a little inflation was good for the economy. Bryan considered this view to be economically naive and morally bankrupt. For example, he publicly argued in a speech that inflation was merely a "transfer of purchasing power from savers in money forms to other classes of society" and a process that erodes the very foundation of a market economy. He believed that "once money is destroyed as a store of value or its function therein seriously impaired . . . the judgment of the consumer, the saver, the businessman and often governments as to their best interests in the presence of inflation as against what their judgment of their best interests would be in the absence of inflation" is negatively affected (1957).

Following the expansionary policies of 1958, the threat of inflation once again pervaded FOMC meetings into 1959. For instance, as early as the January 6, 1959, meeting, the presentation to the FOMC by staff economists characterized the economic situation as a "maturing recovery" with the problems of recession replaced by "problems of sustainable growth" without igniting inflation (FOMC 1959, 5). Money growth continued strong, with M1 increasing at an annual rate of 4 percent during the second half of 1958.

A major problem facing the FOMC was that its desire to reign in potential inflation conflicted with its perceived necessity to accommodate the Treasury's financing needs. In the spring of 1959 a few FOMC members openly expressed dissatisfaction with such policy constraints. D.C. Johns, for example, said at the February 10 meeting that the FOMC should "pay more attention to what was happening to the money supply" in its discussions (FOMC 1959, 92). Governor A.L. Mills suggested at the March 3 meeting that continued use of free reserves would likely repeat the "unhappy experience" of 1958 when Fed actions caused a "supercharged growth in the money supply" (167). Bryan continued in this vein at the April meeting, asserting that recent reserve growth was "sufficient to finance a first-rate inflation and [that] it would easily be possible to get into trouble" (288–89). He repeated this warning in May, noting that policy "had not been particularly restraining" (313).

Bryan also questioned the manner in which policy instructions were communicated to the manager of the desk: the directive. A chronic problem for the committee was that different members often had different interpretations of the directive to which they had agreed. (To provide a perspective on the problem, the directive forwarded to the desk from

Box 2 The Directive of May 5, 1959

Thereupon, upon motion duly made and seconded, the Committee voted unanimously to direct the Federal Reserve Bank of New York, until otherwise directed by the Committee:

(1) To make such purchases, sales, or exchanges (including replacement of maturing securities, and allowing maturities to run off without replacement) for the System Open Market Account in the open market or, in the case of maturing securities, by direct exchange with the Treasury, as may be necessary in the light of current and prospective economic conditions and the general credit situation of the country, with a view (a) to relating the supply of funds in the market to the needs of commerce and business, (b) to fostering conditions in the money market conducive to sustainable economic growth and stability, and (c) to the practical administration of the Account; provided that the aggregate amount of securities held in the System Account (including commitments for the purchase or sale of securities for the Account) at the close of this date, other than special short-term certificates of indebtedness purchased from time to time for the temporary accommodation of the Treasury, shall not be increased or decreased by more than $1 billion;

(2) To purchase direct from the Treasury for the account of the Federal Reserve Bank of New York (with discretion, in cases where it seems desirable, to issue participations to one or more Federal Reserve Banks) such amounts of special short-term certificates of indebtedness as may be necessary from time to time for the temporary accommodation of the Treasury; provided that the total amount of such certificates held at any one time by the Federal Reserve Banks shall not exceed in the aggregate $500 million.

Source: FOMC (1959, 341–42).

the May 5, 1959, meeting is reprinted in Box 2.) For example, the consensus at the May 5 meeting called for an even-keel policy, prompted by upcoming Treasury financing needs. Bryan questioned what this meant: was an even-keel policy "measured by net free reserves, net borrowed reserves, the feel of the market, or the intuition of the Account Manager"? (FOMC 1959, 340). The manager in response, "thought it was a mixture of the things [Bryan] had mentioned" (340). Relying on free reserves to achieve interest rate targets engendered uncertainty over policy directions to the desk.

Bryan became an increasingly outspoken proponent for changing the directive from a qualitative description of policy desires—firm up the markets, achieve some ease in free reserves, and so forth—to one that established numerical targets for policy. In the summer of 1959, Bryan began to base his policy analysis on the short- and long-run growth of a reserve measure that was developed by the Atlanta bank, called total effective reserves.[12] At the August 18, 1959, meeting Bryan introduced something novel in postwar FOMC deliberations: the idea of gauging policy by tracking total effective reserves relative to their postwar trend, an annual average growth rate of 3.6 percent.[13] Bryan argued that when reserves go below their postwar trend, he "would be inclined to resolve any doubts [on the stance of policy] slightly on the side of ease" (FOMC 1959, 558). Bryan pushed for the adoption of his reserve measure and its trend growth as the operating guide for policy to replace financial market behavior and free reserves. Not surprisingly, little was made of his suggestion at this time.

Monetary policy by late 1959 confronted conflicting economic signals. Domestically, money growth had deteriorated sharply, with M1 decreasing at an annual rate of 3.4 percent in the fourth quarter of 1959. The steel strike that began in mid-July and ended in November disrupted the relation between policy actions and the economy. Research out of the St. Louis bank, for example, suggested that the decline in economic activity due to the strike lowered velocity and made the recent slowing in money growth less restrictive than normal. Uncertainty about the domestic economy and the correct policy response was compounded by a worsening balance-of-payments situation. As Hetzel describes it, "monetary policymakers walked a tightrope requiring them to balance internal and external objectives . . . each requiring conflicting policy responses" (1996, 23). By 1959, the currencies of the European Economic Community countries had become fully convertible into U.S. dollars.[14] A persistent balance-of-payments deficit generated significant gold outflows from the U.S. Treasury. This situation not only caused alarm among policymakers but was also politically unacceptable. With domestic interest rates below European rates, the Fed moved to raise the three-month Treasury bill rate in an attempt to quell the gold outflow. Monetary policy thus was conducted with heightened uncertainty (and disagreement) over which objective—internal or external balance—was more important.

Bryan believed that monetary policy did not cause and could do little to solve external imbalances. At the September 22 meeting he asserted that unless reserves increased fast enough to satisfy seasonal needs plus some positive growth, the "System could easily get itself into the position of bringing about greater tightness this fall than it desired" (FOMC 1959, 649). In other words, the current course of policy would, in all likelihood, cause another recession. In contrast, Chairman Martin said at the November 24 meeting that while he "shared some of the apprehension that had been expressed about the money supply and the relationship of credit to growth . . . he did not believe this was the time to correct it" (828). Martin steadfastly relied on the flexibility given by tone and feel to guide policy decisions even though Robert Rouse, manager of the desk, cautioned at the December 15 meeting that "interpretation of the signals given off by the market was by no means easy" (836). Martin simply reiterated his distrust of allowing the behavior of the monetary aggregates to guide policy: "[O]ne should not go overboard on the money supply question unless he was certain that the velocity factor was not playing a part. . . . For this reason he was less wary of restraint" (876). Martin's position, which also was the Committee's consensus view, meant that the FOMC continued to restrain money and credit growth while real economic growth was showing signs of deteriorating.

At this meeting Bryan asserted that the relevant issue had become not whether the Fed should maintain its current policy but how damaging to the economy the Fed's policy of tightening actually would be. Bryan used two analytical tools to drive home his point. One was a set of tables showing the annual growth rates of effective reserves, real GNP, and inflation over the period from 1947 through 1959. These tables are reproduced here as Tables 1–3. The deterioration of reserve growth is evident in Table 1. After increasing at an annual rate of more than 6 percent in 1958, effect reserves were growing only at about a 2 percent rate in 1959. Bryan used the tables to illustrate the connection between fluctuations in reserve growth and economic activity. For example, it is evident from Table 1 that reserve

Table 1 Compounded Annual Growth Rates for Effective Reserves[a]

Base Year	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959[b]
1947	x	1.8	1.0	2.0	2.6	3.2	3.1	3.4	3.3	3.1	2.9	3.3	3.0
1948	x	x	0.2	2.0	2.9	3.5	3.3	3.6	3.5	3.2	3.0	3.4	3.3
1949	x	x	x	3.9	4.3	4.6	4.1	4.3	4.1	3.7	3.4	3.8	3.6
1950	x	x	x	x	4.6	5.0	4.2	4.4	4.1	3.7	3.3	3.8	3.6
1951	x	x	x	x	x	5.4	4.0	4.3	4.0	3.5	3.1	3.7	3.5
1952	x	x	x	x	x	x	2.6	3.8	3.5	3.0	2.6	3.4	3.2
1953	x	x	x	x	x	x	x	5.0	4.0	3.1	2.6	3.6	3.3
1954	x	x	x	x	x	x	x	x	3.0	2.2	1.9	3.3	3.0
1955	x	x	x	x	x	x	x	x	x	1.3	1.3	3.3	3.0
1956	x	x	x	x	x	x	x	x	x	x	2.0	4.3	3.6
1957	x	x	x	x	x	x	x	x	x	x	x	6.6	4.7
1958	x	x	x	x	x	x	x	x	x	x	x	x	2.1
1959	x	x	x	x	x	x	x	x	x	x	x	x	x

[a] Percentage changes, base year to terminal year. The footnote to the original table reads: "Reserve figures exhibited in [this table] and the chart on effective reserves [Chart 1] are total member bank reserves (monthly averages of daily figures) adjusted for changes in reserve requirements and for seasonal influences. No effort was made to remove the expansion potential of total reserves resulting from shifts in deposits among classes of banks and between types of deposits subject to different requirements.

"Method of computation: For May 1958–November 1959, figures used are actual member bank reserves, adjusted for seasonal influences. Monthly values of effective reserves for January 1947 through April 1958 (when reserve requirements were last changed) have been derived by (1) obtaining the ratio of average required reserves to average deposits subject to legal reserves for May 1958–April 1959: (2) multiplying actual reserves by the percentage the above ratio is of the ratio of required reserves to deposits subject to legal reserves for each specified month; and (3) adjusting the values for seasonal influences."

[b] Eleven months

Source: FOMC (1959, 882)

Table 2 Compounded Annual Growth of the U.S. Economy[a]

Base Year	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959[b]
1947	x	3.9	1.8	4.1	4.9	4.6	4.6	3.7	4.2	4.0	3.8	3.2	3.5
1948	x	x	–0.2	4.1	5.2	4.8	4.7	3.6	4.3	4.0	3.7	3.1	3.4
1949	x	x	x	8.7	8.1	6.5	6.0	4.4	5.0	4.6	4.3	3.5	3.8
1950	x	x	x	x	7.5	5.4	5.1	3.4	4.3	3.9	3.6	2.9	3.3
1951	x	x	x	x	x	3.4	3.9	2.0	3.5	3.2	3.0	2.2	2.8
1952	x	x	x	x	x	x	4.4	1.3	3.6	3.2	2.9	2.0	2.7
1953	x	x	x	x	x	x	x	–1.7	3.1	2.8	2.6	1.6	2.4
1954	x	x	x	x	x	x	x	x	8.2	5.1	4.0	2.4	3.2
1955	x	x	x	x	x	x	x	x	x	2.1	2.0	0.5	2.0
1956	x	x	x	x	x	x	x	x	x	x	1.8	–0.2	2.0
1957	x	x	x	x	x	x	x	x	x	x	x	–2.3	2.1
1958	x	x	x	x	x	x	x	x	x	x	x	x	6.7
1959	x	x	x	x	x	x	x	x	x	x	x	x	x

[a] Percentage changes, base year to terminal year, of GNP in 1954 dollars

[b] Three quarters

Source: FOMC (1959, 883)

Table 3 Compounded Annual Growth Rates of Price Inflation[a]

Base Year	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959[b]
1947	x	7.6	3.2	2.5	3.8	3.5	3.1	2.7	2.3	2.2	2.3	2.4	2.2
1948	x	x	−1.0	0.0	2.6	2.5	2.2	2.2	1.6	1.5	1.8	1.9	1.7
1949	x	x	x	1.0	4.4	3.7	3.0	2.4	2.0	1.9	2.1	2.2	2.0
1950	x	x	x	x	8.0	5.1	3.6	2.8	2.2	2.1	2.3	2.3	2.1
1951	x	x	x	x	x	2.3	1.5	1.1	0.8	0.9	1.3	1.5	1.4
1952	x	x	x	x	x	x	0.8	0.6	0.3	0.6	1.2	1.4	1.3
1953	x	x	x	x	x	x	x	0.3	0.0	0.5	1.2	1.5	1.4
1954	x	x	x	x	x	x	x	x	−0.3	0.6	1.5	1.8	1.6
1955	x	x	x	x	x	x	x	x	x	1.5	2.5	2.6	2.1
1956	x	x	x	x	x	x	x	x	x	x	3.4	3.1	2.3
1957	x	x	x	x	x	x	x	x	x	x	x	2.7	1.7
1958	x	x	x	x	x	x	x	x	x	x	x	x	0.7
1959	x	x	x	x	x	x	x	x	x	x	x	x	x

[a] Percentage changes, base year to terminal year, in consumer price index
[b] Ten months
Source: FOMC (1959, 884)

growth in 1956 and 1957—about 1 percent and 2 percent, respectively—was well below the trend rate of growth of about 3 percent. These slow rates of reserve growth precede the downturn in 1958, when real GNP decreased at an annual rate of 2.3 percent (Table 2). Bryan also used the tables to establish the connection between the longer-term movements in reserves and inflation, although the focus at this meeting was on the impending downturn.

The other tool was a chart showing the level of total effective reserves plotted around their postwar trend. Figure 1 reproduces Bryan's chart. Bryan used this chart for two purposes. One was to illustrate the relation between effective reserve growth and real economic activity, hence the appearance of National Bureau of Economic Research (NBER)–designated recessions. Bryan argued, on the basis of the growth triangles and the chart, that "a situation appeared to be approaching in which the matter of the growth factor in reserves should have serious consideration" (FOMC 1959, 871).

Figure 1 also served another purpose for Bryan. He used it to reiterate his misgivings about the qualitative nature of the directive. Bryan proposed that a quantitative tool like Figure 1 gave the FOMC "a means by which instruction can be given in quantitative rather than qualitative terms" (FOMC 1959, 872). By introducing this simple chart Bryan sought to transform the policy debate from money market conditions to the behavior of the monetary aggregates. This chart not only enabled Bryan to provide an explicit, quantitative analysis of past policies but also to illustrate their effects on the economy.

The next section turns to the events of 1960 as Bryan used his new framework to analyze policy and attempted to convince others of its merits. This appraisal details Bryan's introduction of short-term, aggregate growth cones for setting monetary policy as a means of achieving his longer-term targets and the opposition he faced within the FOMC.

INTRODUCING AGGREGATE GROWTH CONES: 1960

The January 12, 1960, FOMC meeting opened with routine reports by staff economists. Guy Noyes, the Board's director of research and statistics, observed that the "customary

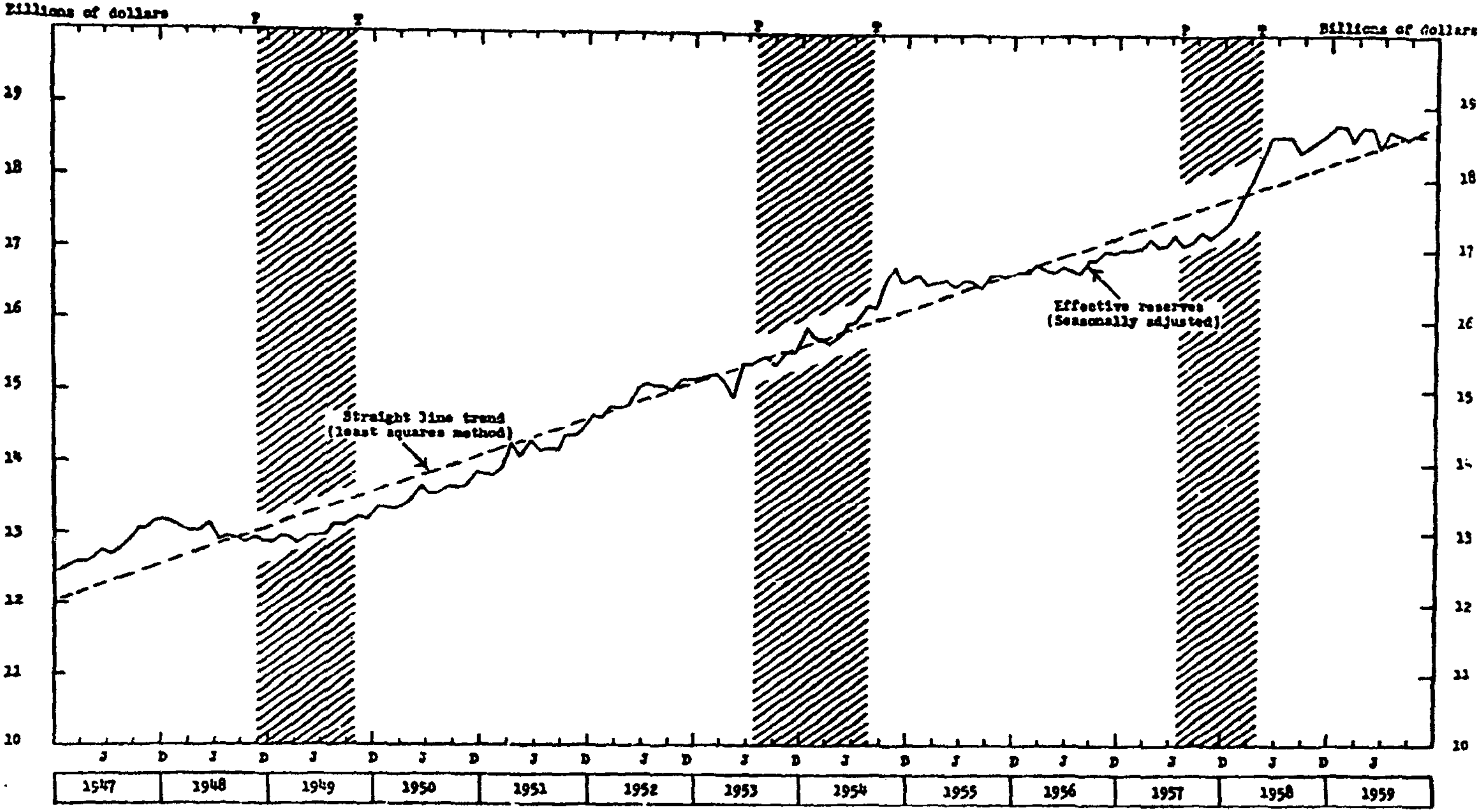


Figure 1 Trend growth of effective reserves. Note: Shaded areas indicate recessions, according to reference dates of National Bureau of Economic Research. P = peak; T = trough. Last month plotted: November 1959. Trend line exhibits an annual growth of 3.6% per year. Source: Reprinted from FOMC (1959, 881).

measures of current economic activity [for example, construction activity, industrial production, and GNP] are all up, and further increases seem as certain for the near term as anything can be" (FOMC 1960, 6). As far as any economist in or out of the Federal Reserve System could predict, this view was correct. (Although not known at the time, real GNP increased at a 7.2 percent rate in the first quarter of 1960). Robust real output growth seemed assured. Still, some FOMC members were warning that past policies, which left the growth rate of the money supply trending down over the second half of 1959, would exert significant downward pressure on economic growth. If the expansion were to continue, it was argued, action to ease needed to be taken immediately. Even Noyes recognized the potential for a slowing in real output growth as he puzzled over the fact that the "high and growing rate of economic activity [stands] in interesting contrast to wholesale prices and the money supply, both of which are substantially unchanged from year-ago levels" (7).

Discussion at this meeting revealed uncertainty among FOMC members over the economic impact of recent policy actions and what direction policy should take. For example, Governor J. L. Robertson thought that "at this particular stage of the business cycle, it [is] incumbent upon the system to maintain a restrictive policy" (FOMC 1960, 30). Johns of St. Louis agreed that continued restraint was needed "to avoid inflationary developments" (16). New York Federal Reserve Bank President Alfred Hayes offered the view that the current level of restraint was appropriate in the face of the apparent economic expansion.

Bryan's evaluation of the situation separates him from his colleagues, even those who also had pressed for quantitative targets. Bryan presented the FOMC with an updated version of his chart of total effective reserves around its trend, reproduced here as Figure 2. Figure 2 clearly is more complex than Figure 1. In Figure 2, for example, Bryan details the level of reserves relative to the trend as "surplus" or "deficit." Figure 2 makes it clear that Bryan's placement of these notes is not random: deficits appear before recessions (marked off in shaded bars using NBER dating), and surpluses follow. This arrangement likely reflects the impact of the recent findings of other monetary economists on Bryan's thinking. Another item of interest in Figure 2 is that the inset box showing the behavior of effective reserves since September 1958 illustrates the restrictiveness of policy over all of 1959. This point was the basis for Bryan's policy position at this meeting. He judged the restrictive policies of 1958–59 as a necessary "mopping-up" operation to get reserves back on trend and dampen any inflationary pressure that might have built up. But enough was enough: the "justifiable mopping-up operation seems to me to be completed" (FOMC 1960, 97). The FOMC, Bryan suggested, must focus on policies to increase total effective reserves or run the risk of inducing another recession.

Meigs (1976) argues that even if the FOMC could have agreed on a target rate of money growth, its members would not have known how to accomplish it. Most members of the FOMC believed that the Fed simply did not possess the tools to effectively influence the behavior of the monetary aggregates. At the January 12 meeting Bryan met this challenge by introducing his most novel contribution, short-run growth targets for effective reserves. This innovation squarely addressed the issue of how the FOMC could assess the desk's success in hitting short-run reserve targets. It also marks the first time aggregate growth cones were used in FOMC policy deliberations.

Bryan argued that his reserve growth cones, reproduced as Figure 3, allowed the FOMC to track the behavior of reserves relative to the path stated in the directive. Three different base periods appear in the original version. The choice of August 1958 and June 1959 was based on the fact that in those months the FOMC changed its directive. The Au-

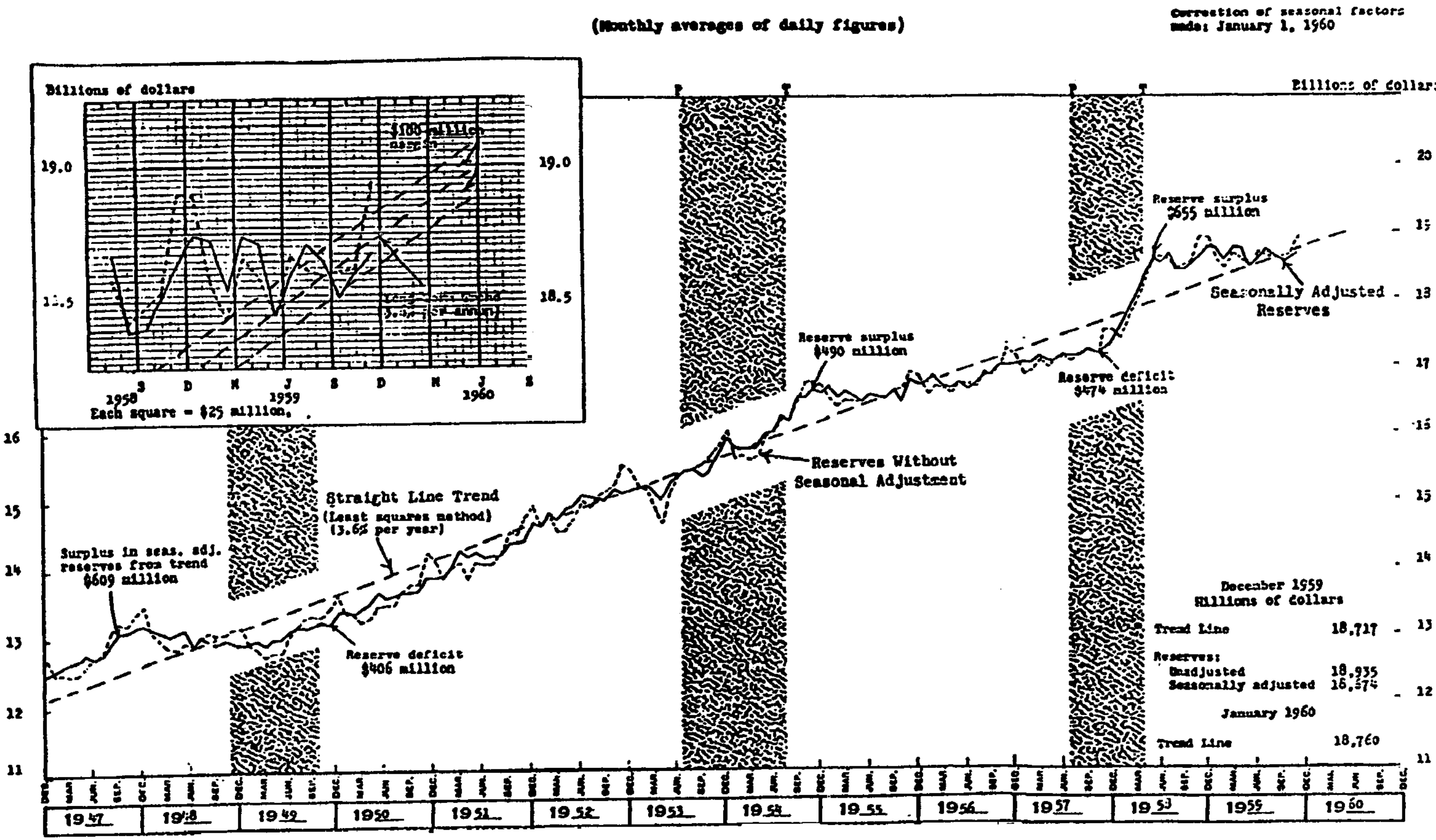


Figure 2 Effective reserves. Note: Shaded areas indicate recessions, according to reference dates of National Bureau of Economic Research. P = peak; T = trough. Last month plotted: December 1959. Trend line exhibits an annual growth of 3.6% per year or $43 million per month. Source: Reprinted from FOMC (1960, 108).

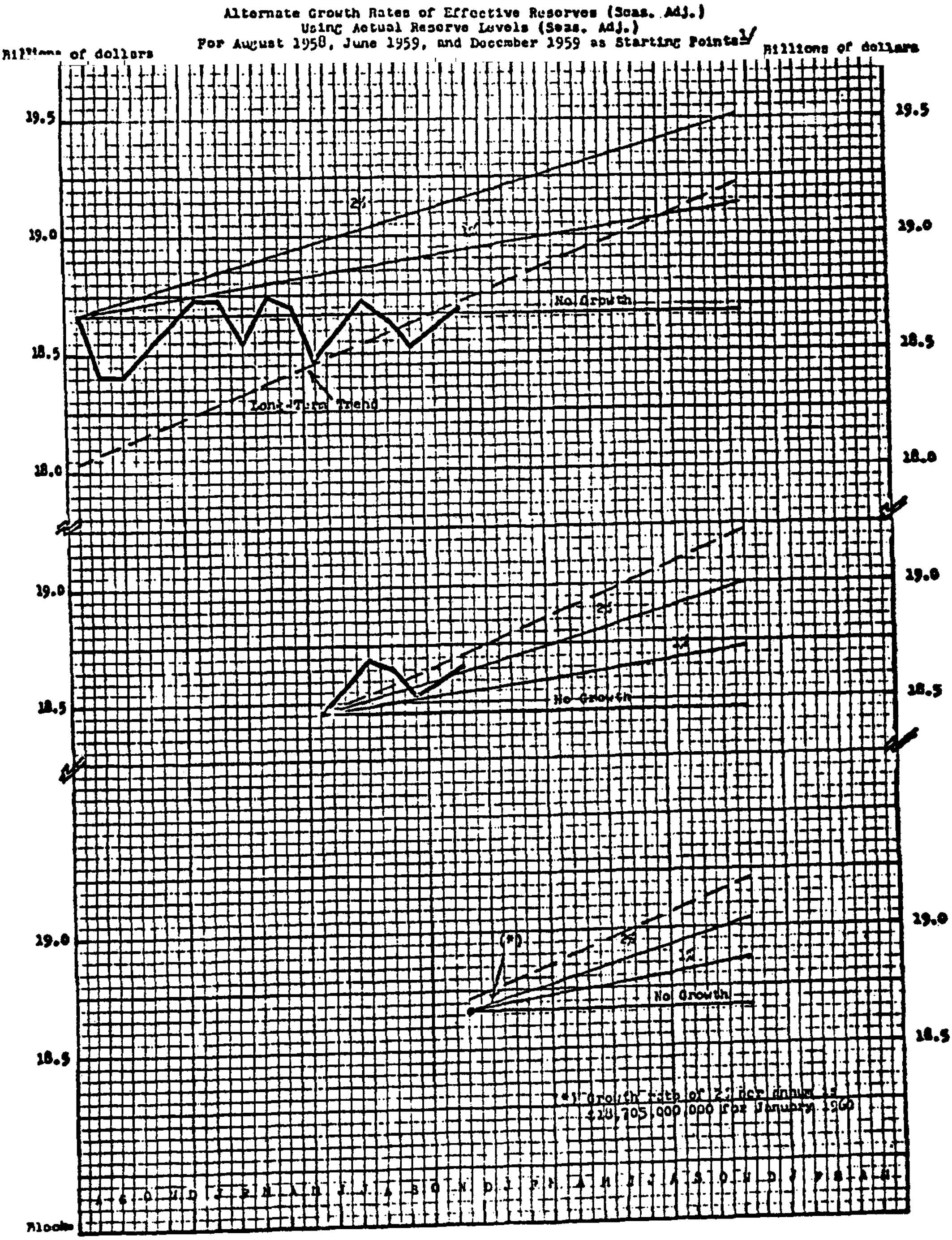


Figure 3 Short-run growth cones. Source: Reprinted from FOMC (1960, 109).

gust 1958 base period also included the postwar trend growth of reserves as a benchmark. December 1959 simply represented the last full month for which data were available.[15] Bryan recognized that requiring the desk to hit a specific target for reserves would be futile. Thus his proposal was to target total reserves to fall within a range of permissible values, providing the desk with some flexibility "to conditions in the money market as they develop" (FOMC 1960, 97).

These charts and accompanying growth rate tables gave Bryan the quantifiable foundation to argue that the FOMC could in fact assess the success or failure of the desk in meeting its policy objectives as stated in the directive. Bryan's introduction of reserve growth cones served to reinforce the attack on the language and intent of the directive. Writing the directive in terms of a reserve target, the FOMC could easily "avoid qualitative terminology as represented by such indefinable terms as tone, feel, ease, tightness, and so on" (FOMC 1960, 98). The record of this meeting indicates that Chairman Martin responded to Bryan's analysis with little discussion, merely noting that Bryan's suggestion and materials "would be taken under study" (98).

A lengthy discussion about the nature of the directive and operating procedures occurred again at the February 9, 1960, meeting of the FOMC. Bryan used his reserve growth cones to propose a February target range of $18,535 million to $18,635 million for total effective reserves. As shown in Table 4, the midpoint of this range put reserves at a level slightly lower than the postwar trend. Bryan felt that this target level provided a needed increase in reserves while maintaining some degree of restraint: policy would provide some monetary stimulus to the economy and avoid any inflationary effect. Even though the midpoint of Bryan's target indicated some restraint, it was appreciably less than what occurred: actual effective reserves for February was $18,203 million, more than $300 million below the lower bound of Bryan's target range.

In Chairman Martin's summary of the discussion at this meeting, he dismissed the usefulness of Bryan's quantitative targets. Martin maintained that such a strategy was too simplistic, a mechanistic approach that would not be wise given the variable nature of the financial markets that must take precedence in policy deliberations and action. Johns defended Bryan's experiment and hoped that "the Committee would not permit proposals such as those advanced . . . to be laughed out of court by attaching a 'mechanistic' approach label to them [and that] such proposals were worthy of serious study" (FOMC 1960, 167). The minutes reveal little additional support, however.

Reserve and money growth continued to decline throughout the early part of 1960. A major reason for the decline was the Fed's dilemma in trying to balance internal and external objectives. The problems stemming from the persistent balance-of-payments deficit called for one type of policy (raising interest rates) while mounting evidence of a domestic slowdown in real output growth called for another (easing). In contrast to the ebullient outlook just a month earlier, the staff report on economic conditions at the February 9, 1960, meeting now discounted the expected rebound in real growth following the end of the steel strike in November 1959. Economic indicators suggested a softening in the economy. Chairman Martin, however, believed that an economic downturn was not likely, even though "he put the possibility forward . . . as an intellectual exercise" (FOMC 1960, 149). By March, however, a recession seemed probable to most observers. For example, Arthur Burns, the chairman of the Council of Economic Advisors and a future Federal Reserve Chairman, was advising President Nixon that a recession was likely (Nixon 1962, 124–28, cited in Meigs 1976). Real output, it turned out, declined at an annual rate of 1.6 percent in the second quarter of 1960.

There persisted a view among most of the FOMC members that the present course of policy and procedures was appropriate. This opinion is exemplified by statements of New York Bank President and Vice Chairman Hayes. Hayes was an ardent supporter of the status quo and an outspoken opponent of Bryan's, or any other, aggregates-based strategy. He considered the daily conference calls, the system of reports, and the frequency of information about money market conditions to be "so extensive that each member [of the FOMC] has ample opportunity to inform the manager if he sees any deviation from the committee's instructions." Arguing that the FOMC "would be giving up a highly advantageous technique, developed over many years, if we were to attempt to couch the instructions in some very exact mathematical terms" (FOMC 1960, 211), Hayes praised the use of net borrowed reserves in conjunction with money market conditions as the best operating guides for policy. Even with evidence of a slowdown in economic activity mounting, Hayes saw no need to alter policy. As he expressed at the March 1 meeting, there simply was "no evidence to suggest that 1960 will be other than a prosperous year, with an upward trend in the economy through most of the year" (210). Unfortunately, Hayes's outlook was shared by the chairman and a majority of the FOMC members.

Reserves in both February and March fell well below the midpoints of Bryan's target levels and below the postwar trend (see Table 4). This development led Governor A. L. Mills to warn the FOMC at the April 12 meeting that, in keeping with Bryan's views, the decline in reserves already was a serious threat to economic growth which, "if not arrested, would in due course lead to serious financial and economic consequences" (FOMC 1960, 347). Why did the FOMC allow the decline in reserves to persist? Bryan focused blame on the committee's continued reliance on free reserves and money market conditions as operating guides. He believed that these measures gave deceptive signals about the degree of

Table 4 Effective Reserves in 1960: Bryan's Target Level, Trend Level, and Actual[a]

Month	Bryan target[b]	Trend	Actual
January	$18,700	$18,748	$18,863
February	18,585	18,790	18,203
March	18,653	18,833	18,027
April	NA	18,876	18,101
May	18,240	18,918	18,236
June	18,419	18,961	18,289
July	18,449	19,004	18,515
August	18,658	19,047	18,499
September	NA	19,089	18,566
October	NA	19,132	18,723
November	18,950	19,175	18,973
December	19,450	19,218	19,270

[a] Figures reported are in millions. The "target" values are midpoints of the ranges specified by Bryan. "Trend" values are determined by assuming a constant 3.6 percent increase in effective reserves. "Actual" effective reserves are based on the first reported values for the month listed.

[b] Bryan did not report a target value for April. Lack of values for September and October reflect his absence from the FOMC during the relevant meetings.

Source: Compiled by the author from information in FOMC 1960, various meetings.

ease or restraint of policy actions. If the FOMC continued down this path, Bryan foresaw "trouble ahead that would be hard to explain" (354).

Martin's consensus view at the April 12 meeting reflects some movement to ease. He stated that the FOMC "should move in the direction of slightly easing the picture as far as reserves were concerned, but with great care on the part of the desk not to do this in an *overt* way" (FOMC 1960, 363; emphasis added). This concession appears as a slightly altered directive from the May 24 meeting: Clause (b) was changed from calling for the desk to "restrain inflationary credit expansion in order to foster sustainable economic growth and expanding employment opportunities" to fostering "sustainable growth in economic activity and employment by providing reserves needed for moderate bank credit expansion" (488). Chairman Martin made it clear, however, that the desk could move to increase reserves only if it did not cause a pronounced change in short-term interest rates. This position again reflects the clash between internal and external policy objectives. While easing reserves would help alleviate the slowing in domestic economic growth, the attendant decline in rates could exacerbate the external imbalances that were becoming politically intolerable.

Table 4 shows that reserves increased slightly following this meeting. Bryan's opinion throughout the summer of 1960 was that policy should aim at increasing the level of effective reserves at a faster pace in order to put them back to the December 1959 level. By June it was clear to all that the economy was in a recession. Bryan urged the FOMC not to "push the panic button" but to undertake immediate actions to reverse the disastrous effects of previous policy. Putting domestic concerns ahead of any external problems, Bryan asserted at the July 6 meeting that it was important to increase effective reserves at a rate that, after seasonal adjustments, would meet the secular needs of the economy without raising inflation. Such a policy "is necessary because of the economic situation and because of the lagged effects of monetary policy" (FOMC 1960, 564). He also noted that his approach would improve economic growth and not "drive short-rates to the ridiculous and obviously unsustainable low levels that have characterized other easing cycles of monetary policy" (565).

Bryan's strategy for getting reserves back on track was based on two key considerations. One was the possible arousal of inflationary expectations, an overarching factor in all of Bryan's policy prescriptions. The other factor was the negative repercussion on the money market if the Fed were to increase reserves immediately and massively.[16] D. C. Johns, however, argued that interest rate movements were of secondary importance. At the July 6 meeting Johns questioned "the System's taking a position of deliberately dampening the downward adjustment of market rates of interest in a period of slack economic growth" (FOMC 1960, 567). Internal and external objectives again gave rise to divergent policy choices.

In light of the deteriorating domestic economy; the FOMC directed the desk to undertake operations that would achieve some ease in reserves by summer's end. Even though the FOMC had decided in favor of easing the reserve position of the banking system, total effective reserves in August and September of 1960 still remained far below their December 1959 level. At the September 13 meeting Bryan pointed out that the August level of effective reserves was lower than in August of the previous year despite the committee's decision to ease. His view was that "the economic situation was deteriorating" and that "[i]n these circumstances, he disagreed with the view of Mr. Hayes that no further monetary ease was required or would be appropriate" (FOMC 1960, 709). The failure of the desk to achieve the desired growth in reserves once again demonstrated the temerity of using free reserves together with money market conditions as operating guides.

Chairman Martin maintained throughout the summer and fall of 1960 that tone and feel was the best policy guide, especially in times of economic uncertainty. At the September 13 meeting he stated that "the System must not let itself be persuaded that if it had expanded the money supply exactly the right amount on a statistical basis, there would not have been any recessions . . . [I]f it got to that point, the only thing necessary would be to keep the levers moving ad infinitum" (FOMC 1960, 736). At the October 4 meeting, tone and feel prevailed over the behavior of any aggregate measure in setting policy. As Robert Rouse, the manager of the open market account, reported, "[A]s the Committee has instructed, we have been operating primarily on the feel of the market rather than on the basis of reserve statistics" (745).

Contrary to the majority opinion of the committee, Bryan advocated a policy of increasing the supply of adjusted reserves in order to provide some monetary stimulation to an economy that was in recession. (Real output decreased at a 0.4 percent rate in the third quarter and at a 3.1 percent rate in the fourth quarter.) As shown in Table 4, Bryan's short-run targets for reserves in November and December 1960 would have put reserves slightly above their postwar trend by the end of 1960. Getting effective reserves back on trend and countering the disastrous policy actions taken during the past year became the principal policy objective for Bryan. The majority of the FOMC took a very different view, however. Chairman Martin, for example, suggested that any overly expansionary monetary policy, an umbrella under which he placed Bryan's suggestions, would be imprudent. In fact, Martin "could not get very pessimistic about the domestic picture . . . [and] continued to feel that the biggest shadow was cast by the balance-of-payments problem" (FOMC 1960, 834). Policy, Martin averred, "would have to be careful that it did not feed fuel to the fires of pessimism by appearing to embark on a cheap money policy" (834). In contrast to Bryan's warnings, at the December 15 meeting Martin voiced the opinion that there had been too much ease recently. He felt that such an easy money policy would do little to affect domestic economic activity and would only exacerbate the critical problem now confronting monetary policy: the persistent balance-of-payments deficit.[17]

Bryan's attempt to convince the FOMC to adopt his short-run reserve growth targets effectively ended with this meeting. Although he never abandoned his conviction that the behavior of monetary aggregates was vital in setting monetary policy, the charts and approach he favored in 1960 would not reappear during his time left on the FOMC.[18] Bryan's experiment with short-term aggregate growth cones swayed a few members of the FOMC. Statements by Johns suggest that he came to appreciate the economic impact that short-run fluctuations in money growth could have, reflecting Bryan's influence and the research of his own staff. Meigs (1976, 450) suggests that Balderston and Mills also were sympathetic. In the final analysis, however, the majority of the FOMC remained unconvinced, relying instead on the dubious tradition of using free reserves and market conditions as operating guides.

CONCLUSION

Bryan's pioneering development and use of aggregate growth targets as a policy alternative to free reserves and money market conditions provides an instructive case study in the early development of postwar U.S. monetary policy. The minutes of the FOMC meetings reveal that money market conditions were of uppermost concern for policy. Bryan's proposals to replace money market activity as the policy operating guide faced a hostile reception in meetings of the FOMC, just as proposals to use monetary growth targets would a decade later. His campaign to change policy challenged not only the convention of maintaining orderly domestic financial markets but also the politically pressing charge to maintain external balance.

Bryan's aggregate-based approach to monetary policy was a dramatic departure for a member of the FOMC at the time. He took new and controversial research results coming out of monetary economics and tried to implement them. Bryan's contributions went beyond merely adopting others' ideas, however. His introduction of short-run aggregate growth targets—growth cones—stands out as a significant and innovative development in monetary policy analysis. Even though his targets and procedures were not adopted by the FOMC at the time, his strategy for monetary policy would resurface as the inflation produced by the policies against which Bryan fought became unacceptable.

The role of monetary aggregates in the formation of policy is as limited today as it was forty years ago. Interest rates remain the primary instrument by which the Fed carries out policy. What then is Bryan's legacy? On an individual basis, Bryan's singular contribution is the development and use of monetary aggregate growth targets—the cones. Although recent events have once again pushed the money supply to secondary importance in its discussions, each year the FOMC must state its annual targets for the aggregates. And discussions at the FOMC during the past year indicate that some members, as in Bryan's time, remain concerned about the behavior of the money supply and its potential effects on economic activity and inflation.

Bryan's contribution also should be considered within a larger perspective. Bryan, like some other bank presidents, pursued a research agenda that resulted in policy prescriptions quite different from that of the Board.[19] His willingness to advocate a controversial view within the FOMC promoted an airing of diverse views and concerns over monetary policy. In his own way, Bryan helped to foster an environment in which alternative theories and approaches to economic analysis could be used for improving monetary policy.

ACKNOWLEDGMENTS

The author thanks Stu Allen, Dave Ault, Michael Bischof, Frederick Deming, Bill Dewald, Jerry Dwyer, Milton Friedman, Gail Heyne Hafer, Joe Haslag, Bob Hetzel, Thomas Mayer, Bill Poole, Anna Schwartz, Frank Steindl, Dave Wheelock, and especially Jim Meigs for helpful comments.

NOTES

1. The FOMC is the policy-making arm of the Federal Reserve System. It is composed of the seven members of the Board of Governors and five of the twelve district bank presidents, four of whom vote on a rotating basis, and the New York Federal Reserve Bank president, who is a permanent voting member. For a recent use of the FOMC transcripts as the basis for analyzing Fed actions, see Edison and Marquez (1998).
2. Meigs (1976) chronicles the contributions of D. C. Johns, the president of the Federal Reserve Bank of St. Louis, and Homer Jones, the director of research at St. Louis. Meigs also provides a brief discussion of Bryan's role in the developing debates that would later be centered on the Federal Reserve Bank of St. Louis.

3. The term cones comes from their construction. For example, supposing that the base period is the average value for the level of an aggregate in the fourth quarter of a year and assuming that this value is \$100, if a 5 percent growth path is the policy objective for the year, then the average value for money in the fourth quarter of the following year would be \$105. Allowing for a growth path somewhat higher (for instance, 7 percent) and lower (for instance, 3 percent) would give quarter-average values of \$107 and \$103, respectively. As shown below, connecting the base period value with the upper and lower ranges creates a cone of possible values.
4. Monetary policy is conducted through the Federal Reserve Bank of New York primarily by buying and selling government securities in the open market. This activity takes place through the Open Market Trading Desk, supervised by the manager of the desk.
5. Atkinson (1969) shows that the FOMC often switched between free reserves and looking to tone and feel during the 1950s and 1960s. Even though the FOMC officially used free reserves as the operating guide, Atkinson's evidence indicates that doing so did not reduce the variance of interest rates or lead to better control over reserves than proposals that used tone and feel as guidelines. For an early analysis of the problems associated with the use of free reserves, see Brunner and Meltzer (1964).
6. Although quotation marks appear, the minutes represent the FOMC Secretary's summary of the discussion and are not necessarily verbatim. Even so, FOMC members had the opportunity to correct the minutes before they entered the permanent record.
7. This notion of flexibility can be found throughout FOMC discussions. For example, compare Roosa's and Martin's comments to those of Chairman Paul Volcker at the December 20–21, 1982, meeting of the FOMC: "I think we're left with what could be termed an eclectic, pragmatic approach. It's going to involve some judgment as to which one of these [aggregate] measures we emphasize, or we may shift from time to time. . . . [W]e're going to have to make some judgments as to which one is more significant at any particular point in time against what nominal GNP is or what the goal is or what the real economy is doing and what prices are doing and all the rest. . . . [T]hat's the way the Federal Reserve used to operate, less elaborately, for years when policy by present standards looked pretty good" (FOMC 1982, 41).
8. The source of the quote is Hayes's testimony before the Joint Economic Committee in 1961.
9. It is likely that Bryan meant "Some Theoretical and Empirical Aspects of the Demand for Money" since there is no reference in the NBER list of publications to the former piece.
10. Thanks to Milton Friedman for making this correspondence available. Bryan's views on monetary policy and the effects of money probably reflect the fact that he received postgraduate training at the University of Chicago (see Box 1). This is the school often associated with Friedman and so-called monetarist economics. To get a feel for the opinion that many economists held of such views, the remarks of Richard Davis, an economist at the Federal Reserve Bank of New York in 1969, can be considered: "[T]he view that 'only money matters' or, perhaps more accurately, that 'mainly money matters' was the province of an obscure sect with headquarters in Chicago. For the most part, economists regarded this group—when they regarded it at all—as a mildly amusing, not quite respectable collection of eccentrics" (1969, 119).
11. In personal discussions, Jim Meigs relates that Bryan and Johns often met outside the FOMC meetings to discuss policy developments. Bryan, the professional economist, is likely to have influenced Johns, a lawyer by training, in matters of monetary policy.
12. The total-effective-reserves measure developed at Atlanta is similar to the St. Louis adjusted monetary base series less currency. Total effective reserves are measured by first calculating the average value for the ratio of required reserves to average deposits beginning in May 1958 through December 1959. May 1958 is used since it is the last time reserve requirements were changed. This ratio is 0.1152. For the period prior to May 1958, this ratio is divided by the monthly ratio of required reserves to deposits and the value of this term multiplied by actual reserves—in other words, $[0.1152/(R_r/D)] \times R$. From May 1958 onward, actual member bank reserves are used. Both reserve measures are then seasonally adjusted. This computation is out-

lined in the appendix to the January 26, 1960, FOMC meeting. This measure can be replicated using the data in Appendix D of Meigs (1962). For example, this author's trend estimate is that reserves increase, on average, $42.7 million per month compared with Bryan's reported estimate of $43 million per month.

13. Meigs (1976, 445) suggests that Bryan introduced the use of total effective reserves at the November 24 meeting of the FOMC. It was at the November 24 meeting that Bryan introduced the charts upon which his policy discussions actually had been based since August.
14. The following draws on Hetzel (1996). See also the related discussions in Meltzer (1991), Schwartz (1997), and Wheelock (1997).
15. The published record does not indicate whether Bryan preferred one base period over another. The fact that Bryan shows little affinity for selecting one base from which to measure changes in reserves causes base drift. For a discussion of this issue and how it influenced monetary policy in the 1970s, see Broaddus and Goodfriend (1984).
16. The notion that any attempt to make up the shortfall in one action could disrupt the market in undesirable ways was used during the 1979–82 period of monetary aggregate targeting. During that period, intermeeting deviations of the aggregates from targets were reduced gradually in order to prevent undue gyrations in interest rates. For an appraisal of policy actions during the 1979–82 period, see, among others, Hetzel (1982) and Poole (1982).
17. The FOMC turned its attention increasingly toward the external balance-of-payments problem as the 1960s unfolded. Bryan believed that monetary policy was not responsible for the problem and could do little to correct it. As he stated at the October 24, 1961, meeting, "for the System to try to correct the balance of payments situation by monetary manipulation [of the Treasury bill rate] struck him as not only absurd but dangerous" (FOMC 1961, 892). For a discussion of how these external events influenced monetary policy during the early 1960s, see, among others, Hetzel (1996), Meltzer (1991), and Schwartz (1997).
18. Bryan's distrust of free reserves deepened over time. By April 1963, for instance, he recognized that the "maintenance of a constant level of free reserves would permit indefinite expansion of the money supply and the financing of inflation" (FOMC 1963, 343). In September 1963 Bryan observed that the "free reserve figure might be a rather dangerous one to use for target purposes, since maintaining free reserves at any selected level would mean supplying all of the reserves demanded" (839). By January 7, 1964, he admonished the committee that policy "had been injecting reserves into the banking system at a rate . . . greater than sustainable in the long run without inflation" (FOMC 1964, 46). The inflationary record of the late 1960s proved his warning to be all too correct.
19. For a related discussion of this issue, see Wheelock (1998).

REFERENCES

Atkinson, T. R. 1969. Tone and Feel of the Market as a Guide for Monetary Management. In *Targets and Indicators of Monetary Policy,* edited by K. Brunner. San Francisco: Chandler Publishing.

Broaddus, A., and Goodfriend, M. 1984. Base Drift and the Longer-Run Growth of M1: Experience from a Decade of Monetary Targeting. Federal Reserve Bank of Richmond *Economic Review* (November/December): 3–14.

Brunner, K. and Meltzer, A. H. 1964. *The Federal Reserve's Attachment to the Free Reserves Concept.* Staff Analysis, House Committee on Banking and Currency, 88th Cong., 2d sess.

Bryan, M. 1938. Limitations of Monetary Management. Speech to the National Catholic Social Action Committee, May 2, Milwaukee, Wisconsin.

Bryan, M. 1948. Some Comments on the Federal Reserve System: Operations and Policy. Speech to the Alabama Bankers Association, August 12, Tuscaloosa, Alabama.

Bryan, M. 1957. The Idea of Creeping Inflation. Speech to the Financial Section, Annual Meeting of the American Life Convention, October 11, Chicago, Illinois.

Bryan, M. 1959. The Sovereign, the Central Bank, and the Monetary Standard. Paper presented to the

Regional Conference on United States Monetary Policy, March 19, cosponsored by Duke University and the American Assembly of Columbia University.

Calomiris, C. W., and Wheelock, D. C. 1998. Was the Great Depression a Watershed for American Monetary Policy? In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century,* edited by M. D. Bordo, C. Goldin, and E. N. White. Chicago: University of Chicago Press.

Davis, R. G. 1969. How Much Does Money Matter? A Look at Some Recent Evidence. Federal Reserve Bank of New York *Monthly Review* (June): 119–31.

Edison, H. J., and Marquez, J. 1998. U.S. Monetary Policy and Econometric Modeling: Tales from the FOMC Transcripts, 1984–1991. *Economic Modeling* 15:411–28.

Federal Open Market Committee. Various years. *Minutes of the Federal Open Market Committee.* (Bound in yearly volumes.) Washington, D.C.: Board of Governors of the Federal Reserve System.

Friedman, M. 1958. The Supply of Money and Changes in Prices and Output. In *The Relationship of Prices to Economic Stability and Growth.* Compendium of papers by panelists appearing before the Joint Economic Committee, March 31, 85th Cong., 2d sess. Joint Committee Print.

Friedman, M. 1982. Monetary Policy: Theory and Practice. *Journal of Money, Credit, and Banking* (February): 98–118.

Goodfriend, M. 1991. Interest Rates and the Conduct of Monetary Policy. *Carnegie-Rochester Conference Series on Public Policy* 34 (Spring): 7–30.

Hetzel, R. L. 1982. The October 1979 Regime of Monetary Control and the Behavior of the Money Supply in 1980. *Journal of Money, Credit, and Banking* (May): 234–51.

Hetzel, R. L. 1995. William McChesney Martin and Monetary Policy in the 1960s. Federal Reserve Bank of Richmond. Photocopy, April.

Hetzel, R. L. 1996. Sterilized Foreign Exchange Intervention: The Fed Debate in the 1960s. Federal Reserve Bank of Richmond *Economic Quarterly* (Spring): 21–46.

Meigs, A. J. 1962. *Free Reserves and the Money Supply.* Chicago: University of Chicago Press.

Meigs, A. J. 1976. Campaigning for Monetary Reform: The Federal Reserve Bank of St. Louis in 1959 and 1960. *Journal of Monetary Economics* 2:439–53.

Meltzer, A. H. 1991. U.S. Policy in the Bretton Woods Era. Federal Reserve Bank of St. Louis *Review* (May/June): 54–83.

Nixon, R. M. 1962. *Six Crises.* Garden City, N.Y.: Doubleday.

Poole, W. 1982. Federal Reserve Operating Procedures: A Survey and Evaluation of the Historical Record Since October 1979. *Journal of Money, Credit and Banking* (November, part 2): 575–96.

Roosa, R. V. 1960. The Change in Money and Credit, 1957–59. *Review of Economics and Statistics* 42 (August): 263–69.

Schwartz, A. J. 1997. From Obscurity to Notoriety: A Biography of the Exchange Stabilization Fund. *Journal of Money, Credit, and Banking* 29 (May): 135–53.

Wheelock, D. C. 1997. Monetary Policy in the Great Depression and Beyond: The Sources of the Fed's Inflation Bias. Federal Reserve Bank of St. Louis Working Paper No. 97-011A, April.

Wheelock, D. C. 1998. National Monetary Policy by Regional Design: The Evolving Role of the Federal Reserve Banks in Federal Reserve Policy. Paper presented at the Center for European Integration Conference "Common Money, Uncommon Regions," July, Bonn, Germany.

27
The Goals of U.S. Monetary Policy

John P. Judd
Vice President and Associate Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

Glenn D. Rudebusch
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

The Federal Reserve has seen its legislative mandate for monetary policy change several times since its founding in 1913, when macroeconomic policy as such was not clearly understood. The most recent revisions were in 1977 and 1978, and they require the Fed to promote both price stability and full employment. The past changes in the mandate appear to reflect both economic events in the U.S. and advances in understanding how the economy functions. In the twenty years since the Fed's mandate was last changed, there have been further important economic developments as well as refinements in economic thought, and these raise the issue of whether to modify the goals for U.S. monetary policy once again. Indeed, a number of other countries—notably those that adopted the Euro as a common currency at the start of this year—have accepted price stability as the new primary goal for their monetary policies.

In this chapter, we spell out the evolution of the legislation governing U.S. monetary policy goals and summarize the debate about whether they could be improved.

THE EVOLUTION OF THE FED'S LEGISLATIVE MANDATE

The Federal Reserve Act of 1913 did not incorporate any macroeconomic goals for monetary policy, but instead required the Fed to "provide an elastic currency." This meant that the Fed should help the economy avoid the financial panics and bank runs that plagued the 19th century by serving as a "lender of last resort," which involved making loans directly to depository institutions through the discount windows of the Reserve Banks. During this early period, most of the actions of monetary policy that affected the macro economy were determined by the U.S. government's adherence to the gold standard.

The trauma of the Great Depression, coupled with the insights of Keynes (1936), led to an acknowledgment of the obligation of the federal government to prevent recessions. The Employment Act of 1946 was the first legislative statement of these macroeconomic

Reprinted from: *FRBSF Economic Letter*, No. 99–04 (Jan. 29, 1999).

policy goals. Although it did not specifically mention the Federal Reserve, it required the federal government in general to foster "conditions under which there will be afforded useful employment opportunities . . . for those able, willing, and seeking to work, and to promote maximum employment, production, and purchasing power."

The Great Inflation of the 1970s was the next major U.S. economic dislocation. This problem was addressed in a 1977 amendment to the Federal Reserve Act, which provided the first explicit recognition of price stability as a national policy goal. The amended Act states that the Fed "shall maintain long run growth of the monetary and credit aggregates commensurate with the economy's long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." The goals of "stable prices" and "moderate long-term interest rates" are related because nominal interest rates are boosted by a premium over real rates equal to expected future inflation. Thus, "stable prices" will typically produce long-term interest rates that are "moderate."

The objective of "maximum" employment remained intact from the 1946 Employment Act; however, the interpretation of this term may have changed during the intervening 30 years. Immediately after World War II, when conscription and price controls had produced a high pressure economy with very low unemployment in the U.S., some perhaps believed that the goal of "maximum" employment could be taken in its mathematical sense to mean the highest *possible* level of employment. However, by the second half of the 1970s, it was well understood that some "frictional" unemployment, which involves the search for new jobs and the transition between occupations, is a necessary accompaniment to the proper functioning of the economy in the long run.

This understanding went hand in hand in the latter half of the 1970s with a general acceptance of the Natural Rate Hypothesis, which implies that if policy were to try to keep employment above its long-run trend permanently or, equivalently, the unemployment rate below its natural rate, then inflation would be pushed higher and higher. Policy can *temporarily* reduce the unemployment rate below its natural rate or, equivalently, boost employment above its long-run trend. However, persistently attempting to maintain "maximum" employment that is above its long-run level would not be consistent with the goal of stable prices.

Thus, in order for maximum employment and stable prices to be mutually consistent goals, maximum employment should be interpreted as meaning maximum *sustainable* employment, referred to also as "full employment." Moreover, although the Fed has little if any influence on the long-run level of employment, it can attempt to smooth out short-run fluctuations. Accordingly, promoting full employment can be interpreted as a countercyclical monetary policy in which the Fed aims to smooth out the amplitude of the business cycle.

This interpretation of the Fed's mandate was later confirmed in the Humphrey-Hawkins legislation. As its official title—the Full Employment and Balanced Growth Act of 1978—clearly implies, this legislation mandates the federal government generally to ". . . promote *full employment and production,* increased real income, balanced growth, a balanced Federal budget, adequate productivity growth, proper attention to national priorities, achievement of an improved trade balance . . . and *reasonable price stability* . . ." (italics added).

Besides clarifying the general goal of full employment, the Humphrey-Hawkins Act also specified numerical definitions or targets. The Act specified two initial goals: an unemployment rate of 4% for full employment and a CPI inflation rate of 3% for price stability. These were only "interim" goals to be achieved by 1983 and followed by a further reduction in inflation to 0% by 1988; however, the disinflation policies during this period were not to impede the achievement of the full-employment goal. Thereafter, the timetable to achieve or maintain price stability and full employment was to be defined by each year's *Economic Report of the President.*

THE DEBATE ABOUT THE FED'S CURRENT MANDATE

The Fed then has two main legislated goals for monetary policy: promoting full employment and promoting stable prices. With this mandate, the Fed has helped foster the exceptional performance of the U.S. economy during the past decade. Still, some have argued that the Fed's mandate could be improved, especially in looking ahead to future attempts to maintain or institutionalize recent low inflation. Much discussion has centered on two topics: the transparency of the goals and their dual nature.

The transparency of goals refers to the extent to which the objectives of monetary policy are clearly defined and can be easily and obviously understood by the public. The goal of full employment will never be very transparent because it is not directly observed but only estimated by economists with limited precision. For example, the 1997 *Economic Report of the President* (which has authority in this matter from the Humphrey-Hawkins Act) gives a range of 5 to 6% for the unemployment rate consistent with full employment, with a midpoint of 5.5%. Research suggests that there is a very wide range of uncertainty around any estimate of the natural rate, with one prominent study finding a 95% probability that it falls in the wide range of 4 to $7\frac{1}{2}$% (see Walsh 1998).

Price stability as a goal is also subject to some ambiguity. Recent economic analysis has uncovered systematic biases, say, on the order of 1 percentage point, in the CPI's measurement of inflation (see Motley 1997). In this case, actual price stability would be consistent with measured inflation of 1%. In addition, at any point in time, different price indexes register different rates of inflation. Over the past year, for example, the CPI has risen about $1\frac{1}{2}$%, while the GDP price index has risen about 1%. Still, a transparent price stability goal could be specified as a precise numerical growth rate (or range) for a particular index (which could take into account any biases). However, economists have also suggested other ways to enhance the transparency of policy. For example, publishing medium-term inflation forecasts might help to clarify the direction of policy (Rudebusch and Walsh 1998). Because the central bank has some control over inflation in the medium term, its forecasts would contain an indication of where it wanted inflation to go.

A second recent proposed modification to the Fed's goals involves focusing to a larger extent on price stability and de-emphasizing business cycle stabilization. Some economists have argued that having dual goals will lead to an inflation bias despite the Fed's best attempts to control inflation. This argument stresses that the temptation to engineer gains in output in the short run will overcome the central bank's desire to control inflation in the long run. As a result of elevated inflation expectations of the public, inflation will end up being higher than the central bank intended, despite its best efforts. This "time-inconsistency" argument, as economists call it, coupled with the pain incurred in the 1970s as inflation skyrocketed and in the early 1980s as inflation was reduced to moderate levels, persuaded many that the primary goal of the central bank should be to stabilize prices.

This view is embodied in the charter for the central bank in the new European Monetary Union: "The primary objective of the European System of Central Banks is to maintain price stability. Without prejudice to the objective of price stability, the ESCB shall support the general economic policies in the Community." Among these latter policies are "a high level of employment" and "a balanced development of economic activities."

Economists remain divided on the importance of the time inconsistency problem and on the need to put primary emphasis on price stability at the expense of output stabilization. Some stress the fact that the central bank is the only entity that can guarantee price stability, and that this goal is not likely to be attained for long unless price stability is designated as the primary goal. Others find the arguments for time inconsistency implausible because policymakers, who are aware of the arguments about an inflationary bias and see the implications for inflation, can conduct policy without an inflationary bias (McCallum 1995). Still others argue that the abdication of other goals is irresponsible (Fuhrer 1997). Also, a good deal of empirical research using simulations of models of the U.S. economy suggests that a focus on dual goals can reduce the variance of real GDP (i.e., smooth the business cycle) while achieving an inflation goal as well (Rudebusch and Svensson 1998).

While these issues are not yet resolved, the experience of the past two decades provides some support to those who think dual goals that lack transparency can function successfully. It is true that some countries around the world have reduced inflation over this period while putting primary emphasis on explicit inflation targeting. But at the same time, with its current legislative mandate, the Fed also has had success in reducing inflation, while maintaining the flexibility of responding to business cycle conditions.

REFERENCES

Fuhrer, J. C. 1997. Central Bank Independence and Inflation Targeting: Monetary Policy Paradigms for the Next Millennium? *New England Economic Review* January/February, pp. 20–36.

Keynes, J. M. 1936. *The General Theory of Employment, Interest, and Money.* Harcourt, Brace, and Company: New York.

McCallum, B. 1995. Two Fallacies Concerning Central Bank Independence. *American Economic Review Papers and Proceedings,* vol. 85, no. 2 (May), pp. 207–211.

Motley, B. 1997. Bias in the CPI: Roughly Wrong or Precisely Wrong. *FRBSF Economic Letter* 97–16 (May 23).

Rudebusch, G. D., and Svensson, L. E. O. 1998. Policy Rules for Inflation Targeting. NBER Working Paper 6512.

Rudebusch, G. D., and Walsh, C. E. 1998. U.S. Inflation Targeting: Pro and Con. *FRBSF Economic Letter* 98–18 (May 29).

Walsh, C. E. 1998. The Natural Rate, NAIRU, and Monetary Policy. *FRBSF Economic Letter* 98–28 (September 18).

28

Seigniorage Revenue and Monetary Policy

Joseph H. Haslag
Senior Economist and Policy Advisor, Federal Reserve Bank of Dallas, Dallas, Texas

Money creation is one potential source of revenue for a government. Seigniorage—government revenue received through creating money—is a relatively inexpensive means of raising funds. Take the United States as an example. It costs just a few pennies to print a $100 bill. The resource costs to the U.S. Treasury are more than offset by the value of the goods that could be purchased with the $100 bill. It is even less expensive for the Federal Reserve to electronically purchase large quantities of Treasury bonds, notes, and bills from traders in New York. It is important to note that the Federal Reserve returns the interest payments on its security holdings (less its expenses) to the U.S. Treasury. Consequently, when the Federal Reserve increases its bond holdings, for example, the U.S. Treasury realizes an effective reduction to its debt expenses.[1] The present value of the reduction in Treasury expenses is equal to the amount of money injected by the Federal Reserve's open market purchase.

The problem is that although money may be cheap to produce, the social costs of money creation are almost certainly greater than what the Federal Reserve pays to create it. Indeed, a large body of empirical evidence suggests that the rate of money creation is closely correlated with inflation. Thus, faster money creation costs society by eroding the purchasing power of money already in circulation, which is the inflation tax. Though tempted by low production costs, governments must balance the benefits with social costs when deciding how much to rely on seigniorage.

The article addresses two questions. First, how much do countries rely on money creation as a source of revenue? The answer to this question gives some idea of the size of the seigniorage revenue "problem." For most of the countries, money creation accounts for less than 2 percent of real GDP. The evidence indicates that seigniorage revenue is not the primary source of revenue for a government, but neither is it quantitatively insignificant.

Second, are monetary policy settings systematically related to a government's reliance on real seigniorage revenue, and, if so, what is the relationship? Such evidence should be a useful guide for economic theories—that is, a good theory should be able to account for a government's reliance on seigniorage revenue versus, say, its reliance on income taxes.

Reprinted from: *Federal Reserve Bank of Dallas Economic Review* (Third Quarter 1998) 10–20.

Sargent (1986) presents some evidence that very rapid money growth does not translate into greater reliance on real seigniorage revenue. He studies monetary policy during four hyperinflation episodes that occurred immediately following World War I. For two countries, Austria and Hungary, Sargent reports data on money growth and the fraction of government spending earned through seigniorage revenue. Austria raised about 67 percent of government expenditures through money creation in the first half of 1919. However, the ratio of money creation to government expenditures fell to about 40 percent of government expenditures by 1922. Between 1919 and 1922, Austrian crowns in circulation went from roughly 4.7 billion to nearly 4.1 *trillion*. For Hungary, money creation accounted for more than 45 percent of its government expenditures in 1921–22, falling to about 33 percent in 1924–25. Between February 1921 and April 1925, Hungary saw its notes in circulation rise from 15 billion kronen to 4.5 trillion kronen.[2] For these two case studies, the evidence suggests that reliance on money creation decreases as the rate of money growth increases. Hyperinflations are rare and probably not good laboratories for studying the relationship between monetary policy and seigniorage revenue. Still, the Austrian and Hungarian data show that dramatic increases in the rate of money growth do not necessarily translate into a government's increased reliance on seigniorage revenue.

In this article, I use data from different countries to identify whether a systematic relationship exists between monetary policy and a country's reliance on seigniorage revenue. Rather than focus on year-to-year realizations, the approach taken in this article is to study the correlation between monetary policy and seigniorage over a longer horizon; specifically, the sample mean is computed from a 30-year period. Both economic theory and problems with statistical inference point to using a sufficient statistic to measure monetary policy. (A sufficient statistic captures changes in the variable that the researcher is studying.) Here, the monetary policy measure is a combination of the money growth rate and the reserve ratio. As such, the evidence bears on whether countries with a high money growth rate–reserve ratio combination also tend, on average, to rely more heavily on seigniorage revenue over these longer horizons than countries with a low money growth rate–reserve ratio combination.

The cross-country evidence indicates a positive association between the monetary policy measure and a country's reliance on seigniorage revenue. Thus, countries with high monetary policy settings tend to rely more on seigniorage revenue than countries with low monetary policy settings. An additional implication follows from the way in which the measure of monetary policy is constructed; specifically, one can infer that the relationship between the reserve ratio, which holds the money growth rate constant, and a country's reliance on seigniorage revenue is concave. The concave relationship also holds between the money growth rate and seigniorage reliance when the reserve ratio is constant. The implied concavity complements Sargent's findings for Austria and Hungary.

It is useful to begin with a brief overview of seigniorage revenue that shows how it fits into a broader picture of government finance.

SEIGNIORAGE REVENUE—AN OVERVIEW

Suppose the government prints new pieces of currency and uses these newly created bills to buy goods and services, such as missiles or computers, or pay workers' salaries.[3] For simplicity, I assume that the economy has a composite commodity (hereafter, the consumption good). The government can buy units of this consumption good with the newly printed money, which is

$$(M_t - M_{t-1})v_t, \tag{1}$$

where M denotes the total quantity of high-powered money in the economy (t denotes time), and v denotes the money's value in terms of the units of the consumption good that can be acquired with one unit of money (that is, the inverse of the price level). Thus, Equation 1 represents the units of the consumption good that can be purchased with newly printed money—in other words, real seigniorage revenue.

Seigniorage revenue is just one part of a larger picture. To see the complete picture, it is necessary to give the government's income statement, or budget constraint. To keep things simple, assume that the government issues only one-period, fully indexed bonds.[4] For this simple economy, the government's budget constraint can be written as

$$g_t + r_t b_{t-1} = \tau_t y_t + b_t + (M_t - M_{t-1})v_t. \tag{2}$$

In Equation 2, g is the total quantity of goods purchased by the government; the product, rb, is the principal and interest payments, measured in units of the good, that the government owes for one-period bonds issued at date t–1; r is the real gross return (principal plus net interest) on government securities worth b goods. Thus, the left-hand side of Equation 2 represents the total expenditures by the government. The right-hand side characterizes the government's total receipts. The product, τy, represents the income tax revenue earned by the government at rate τ, and y is the aggregate level of real income.

Note that in Equation 2 the government has access to an income tax. Representing tax revenue this way is not necessary. However, there is a useful analogy between seigniorage revenue and income tax revenue. The relationship between the income tax rate and tax revenue has been popularized in the Laffer curve. Suppose that income, y, is negatively related to the tax rate. With an increase in the tax rate, for example, people would report less income.[5] The basic supply-side question, therefore, is whether higher tax rates are offset by a lower tax base. Since tax revenues are the product of these two factors, it is impossible to say, *a priori,* whether income tax revenues rise or fall in response to an increase in tax rates.

Seigniorage Revenue and Money Growth

An increase in the money growth rate has an effect on seigniorage that is analogous to the effect that an increase in the tax rate has on income tax revenue. To illustrate this point, I modify the expression for seigniorage revenue to identify a tax rate and tax base. The date t quantity of money in circulation is equal to the product of a growth rate and date t–1 stock. Thus,

$$M_t = \theta_t M_{t-1}, \tag{3}$$

where θ is the gross rate of money supply expansion. With $\theta > 0$, the percentage change in the money supply is θ–1. Use Equation 3 to substitute for M_{t-1} in Equation 1. The resulting government budget constraint is given by

$$g_t + r_t b_{t-1} = \tau_t y_t + b_t + v_t M_t (1 - 1/\theta_t). \tag{2'}$$

The analog to income tax revenue is now more accessible. In Equation 2′, the total revenue from money creation is now the product of a tax base, $v_t M_t$, and a tax rate, (1–1/θ), that is positively related to the rate of money growth.

To complete the analogy to the tax revenue setting, linking the seigniorage tax base to the seigniorage tax rate is necessary. One way to do this is to assume that the real quantity of money—which for seigniorage revenue is the tax base—is a function of its real rate of return. More specifically, let real money balances be positively related to the real return on money. It is straightforward to show that the real rate of return on money is the inverse of the inflation rate; that is, $1/\pi$, where $\pi = p_t/p_{t-1}$.[6] Other things being equal, the rate of inflation is positively related to the rate of money growth. Hence, faster money growth means that the real return on money falls. It follows that faster money growth results in a smaller tax base for real seigniorage revenue.

Overall, faster money growth can lead to either more or less real seigniorage revenue, depending on whether the change in the tax rate or the change in the tax base is quantitatively larger.

Reserve Requirements and the Tax Base

There is another monetary policy tool that could potentially influence real seigniorage revenue. The reserve requirement stipulates that money balances cannot be less than γ percent of bank deposits, where γ denotes the reserve requirement ratio. Consequently, for a given level of deposits, a higher reserve requirement implies that the quantity of real money balances increases; that is, a larger tax base. However, holding the level of deposits constant is unlikely. An increase in the reserve requirement ratio may induce people to decrease their total savings and hence their bank deposits. As a result, people may avoid the inflation tax by reducing their bank deposits.

To illustrate this point, people have two means of saving: government bonds and money. For simplicity, I assume that the real return on the government bonds, r, is constant and that these bonds dominate money in terms of offering a higher rate of return—that is, $1/\pi < r$.

In this economy, banks serve a very simple function. I assume that government bonds are issued in denominations that are too large for any one saver to acquire. The bank costlessly pools the funds to acquire these government bonds. Because the bank maximizes profits in a perfectly competitive market, the rate of return on deposits will also be r. Each person takes the rate of return on deposits as given. The reserve requirement stipulates that the person hold a fraction of these deposits as money balances.[7] Because money is rate of return dominated by government bonds, the person will not hold any fiat money in excess of this reserve requirement. The equilibrium return on a person's savings is

$$q = \frac{\gamma}{1+\gamma} \cdot \frac{1}{\pi} + \frac{r}{1+\gamma}, \tag{4}$$

where q is the gross real return on savings. Note that q is a weighted average of the rate of return on real money balances and on government bonds. With $1/\pi < r$, Equation 4 implies that $q < r$. In other words, the reserve requirement ratio drives a wedge between the return on bonds and the return to savings.

Suppose there is an increase in the reserve requirement ratio. The quantity of real money balances held by people is γd, where d is the quantity of goods deposited with banks. For a given level of deposits, people will hold more money and the tax base rises. Equation 4, however, implies that the real return on savings falls as the reserve requirement ratio increases. It seems reasonable to assume that people's savings are positively related to the real return on savings. Therefore, it follows that a higher reserve requirement ratio will result in a decline in a person's savings. A decline in savings implies a decline in the quantity of bank deposits. As such, γ is increasing and d is falling so that the product—the seigniorage tax base—could either increase or decrease.

The thrust of this section is twofold. First, real seigniorage revenue is formally defined. Second, economic theory offers an ambiguous picture regarding the effects that monetary policy settings have on the size of this revenue. The gist of the economic argument is that people try to avoid taxes, so with higher tax rates, whether it be inflation or income, they have an incentive to reduce the quantity of the good being taxed. The remainder of this article seeks to establish some preliminary observations on the correlation between a country's reliance on seigniorage revenue and its monetary policy settings.

THE DATA

I obtain the data in this article from International Financial Statistics. I use annual observations, spanning the period 1965–94. For each of the variables I examine over this 30-year period, I use the sample mean to measure each country's central tendency. Unfortunately, observations are not available for each country for each year. Each country in the sample has at least fifteen annual observations. The result is a sample of sixty-seven countries.[8]

Following Fischer (1982), I compute the ratio of seigniorage revenue to output, hereafter *S/Y,* for each country.[9] Here, I use high-powered money as the measure of the money stock (*M*). One alternative to computing ratios is to convert each country's seigniorage to a dollar-equivalent value. The chief advantage to using ratios is that no assumptions are required regarding the exchange rate and purchasing power parity.

Before reporting any statistics, it is important to note that the reserve requirement ratio presents a measurement issue. In principal, the average marginal reserve requirement ratio—the ratio that applies to the next dollar deposited—would be measured.[10] In practice, however, measuring this is not so simple. There is a dizzying array of reserve requirements; U.S. banks are currently required to hold reserves equal to 3 percent of the first $49.3 million of checkable deposits and 10 percent of all deposits above the low-reserve tranche. Therefore, it matters whether the deposits are going into small banks or large banks. In other countries, the reserve requirement structures are even more convoluted.[11]

Equations 1 through 4 are built on the notion that there is one reserve requirement ratio that is the marginal reserve requirement. To compute the marginal reserve requirement ratio, one could use the distribution of deposits across the different categories corresponding to the reserve requirement structure. For example, 20 percent of the deposits are in small U.S. banks (with less than $49.3 million in checkable deposits) and 80 percent are in large banks. The average marginal reserve requirement ratio would be $(0.2\cdot0.03) + (0.8\cdot0.1) = 0.086$. Unfortunately, neither the United States nor any other countries report the distribution across deposit categories, which is necessary to construct such a measure. Consequently, I use the reserve-to-deposit ratio, denoted *R/D,* (hereafter reserve ratio) as a proxy for the reserve requirement ratio. Historically, reserve requirements have been applied against deposits included in what is the U.S. counterpart to M2. Accordingly, I use M2 less currency outside the bank as my measure of bank deposits. As it is measured, the reserve ratio ignores any extra information contained in the distribution of deposits across the alternative categories. Instead, different deposit categories are treated as if there is only one type.

Table 1 reports summary statistics for the seigniorage ratio, *S/Y;* as well as a monetary policy measure, *g;* a tax rate measure, *TAX;* and the growth rate of output, y'. On average, seigniorage revenue accounts for a fairly small fraction of total output—about 2 percent.[12] Tax receipts are, on average, about 22 percent of aggregate output. As one would probably expect, seigniorage revenue is not the primary source of government revenue.

Generally, the government budget constraint links the variables in Table 1 together. As such, the statistics describe the central tendencies and average dispersion of monetary policy, fiscal policy, and some aggregate measure of economic activity. The money growth rate, g, is $(M_t/M_{t-1})-1$. *TAX* is the ratio of tax revenue to GNP. Lastly, $y' = (Y_t/Y_{t-1})-1$ is the growth rate of output.

One rather interesting finding is how the reliance on seigniorage revenue is distributed. Approximately three-fourths of the countries collect, on average, less than 2 percent of GNP through money creation. Most of the variation, therefore, occurs among those countries in the top quartile of the distribution. In this sample, Ghana relies most heavily on seigniorage, collecting revenues equal to 10 percent of output, on average, through money creation. Overall, the distribution of *S/Y* ratios is quite skewed toward the low-seigniorage-reliance tail of the distribution.

Table 1 also reports the range of reserve ratios and average money growth rates. The difference between the minimum and maximum values is substantial. Reserve ratios range from a low of 0.6 percent to 64 percent. Money growth rates range from 3.3 percent to nearly 90 percent. This evidence shows that banks hold a substantial fraction of money against deposits in some countries. It also shows that some countries create money at a rapid pace.

Do countries that rely heavily on seigniorage revenue also exhibit large year-to-year volatility in their earnings from money creation? The answer indicates whether countries tend to rely on seigniorage revenue consistently or if there are periods of heavy reliance on seigniorage interspersed with periods in which countries rely less on it. A positive correlation between the seigniorage ratio and volatility would show that countries with large values of *S/Y*, for example, also tend to experience greater year-over-year variability in the *S/Y* ratio. Conversely, if the correlation coefficient is negative, then countries that have relatively high *S/Y* ratios tend to experience less variability in the year-to-year reliance on seigniorage.

The correlation coefficient between a country's average reliance on seigniorage revenue and its sample standard deviation is 0.8462. Thus, the high correlation coefficient suggests that countries with high seigniorage rates have the greatest volatility in year-over-

Table 1 Summary Statistics

Variables	Sample mean	Standard deviation	Minimum	Maximum
S/Y	.0211	.0201	.0025	.0998
R/D	.1712	.1303	.0068	.6402
g	.2085	.1667	.0332	.8981
TAX	.2254	.1008	.0537	.5586
y'	.0407	.0181	−.0018	.0904

S/Y Real seigniorage/real GDP
R/D Bank reserves/deposits (M2 less currency)
g Percentage change in high-powered money
TAX Tax revenue/GDP
y Percentage change in real GDP

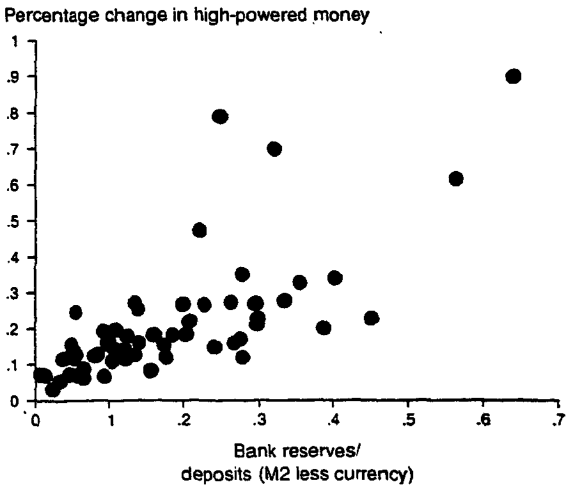


Figure 1 Cross-country plots of reserve requirements versus money growth.

year realizations. In other words, countries that rely, on average, more heavily on money creation as a source of revenue also tend to exhibit the largest variability in reliance from year to year. In contrast, countries that rely relatively little on seigniorage revenue tend to receive about the same fraction of GNP from year to year.[13]

Figure 1 focuses on the two monetary policy variables. Specifically, it plots combinations of the average reserve ratio and the average money growth rate for each country in this sample. The plot suggests that a country with a high average reserve ratio has a high average money growth rate. Formal statistics support the notion that the reserve ratio and money growth are positively related; the correlation coefficient between the reserve ratio and the money growth rate is 0.72.

Thus, three facts emerge from this preliminary review of the data. These facts serve to answer the primary question of how much countries rely on seigniorage revenue. First, for most countries, seigniorage revenue accounts for less than 2 percent of output. Second, countries with the highest average reliance on seigniorage revenue also tend to have the greater year-to-year volatility in the S/Y ratio. Third, the evidence suggests that monetary policy settings are not independent of one another; countries with high money growth rates also tend to have high reserve ratios.

MONETARY POLICY AND SEIGNIORAGE

To determine whether a relationship exists between a country's monetary policy settings and its reliance on seigniorage revenue, I present results from a simple regression. Because of this potentially nonlinear relationship, I use a sufficient statistic, z, to measure monetary policy settings. Specifically, let $z = [R/D/(1 + R/D)][g/(1 + g)]$. The economics motivating this decision is sketched out in the box entitled "A Case for Combining the Money Growth Rate with the Reserve Ratio." Statistical issues also arise, in part, because of the evidence presented in Figure 1. Recall from Figure 1 (and the correlation coefficient) that countries with high average reserve ratios tend to have high average money growth rates. By studying the contribution of each of the monetary policy measures, multicollinearity is a potential problem; that is, if two independent variables are highly correlated, the standard errors of the coefficients are inflated, creating inference problems. In measuring monetary policy settings with a single variable, I am assuming that z is a sufficient statistic for monetary policy. As a sufficient statistic, z is useful because changes in it capture changes in each of the monetary policy variables being studied. As such, z is serving as a measure of the overall thrust of monetary policy as it relates to seigniorage revenue.

One additional property of z is noteworthy. It is straightforward to show that the definition of z implies a concave relationship between it and each of the two monetary policy variables. To illustrate this, consider the effect of a change in the reserve ratio, holding the money growth rate constant. As the reserve ratio increases, z increases also, but the change in z will be smaller as the reserve ratio increases. In other words, for a given increase in the reserve ratio, z will increase at a diminishing rate. The same holds if, for example, the money growth rate increases, holding the reserve ratio constant. With a positive coefficient on z, the relationship between the seigniorage rate and each of the monetary policy variables is concave.[14]

In the benchmark regression, I include the squared value of z and a constant term as additional explanatory variables. In doing so, it is possible to assess whether there are any additional nonlinearities that characterize the relationship between a country's monetary policy settings and its reliance on seigniorage revenue. If the additional quadratic term in the regression is significant, this relationship will vary as z varies. For instance, if the coefficient on z^2 is significantly less than zero, the evidence suggests that the relationship is concave. Conversely, if the coefficient on the squared term is significantly greater than zero, the evidence indicates that the relationship is convex.[15]

The results from the benchmark regression are reported in column 1 of Table 2.[16] The coefficient on z is significantly greater than zero, while the coefficient on z^2 is not significantly different from zero. Thus, the evidence is consistent with the notion that countries with high monetary policy settings (high z values) tend to rely more heavily on seigniorage revenue than do countries with lower monetary policy settings (low z values). As discussed above, the evidence suggests a positive, concave relationship between a country's reserve

Table 2 Regression Results for Seigniorage Ratios and Monetary Policy Variable (Dependent variable: S/Y)

Variable	Benchmark model	Financial sophistication I	Financial sophistication
constant	.011* (.0032)	.0094* (.0034)	.0113* (.0032)
z	.3982† (.1788)	.4273† (.1722)	.3835† (.1808)
z^2	−.0915 (1.157)	−.1554 (1.1854)	−.0072 (1.1545)
$y65$‡	—	−.338E-06 (.416E-06)	—
OECD	—	—	−.006§ (.0033)
OECD*z	—	—	−1.144* (.3909)
OECD*z^2	—	—	165.817* (37.035)
adjusted R^2	.448	.504	.437
Standard error of the estimate	.0148	.0135	.0149

NOTE: Standard errors in parentheses.
* Significant at the 1 percent level.
† Significant at the 5 percent level.
‡ Per capita real GDP in 1965.
§ Significant at the 10 percent level.

Box 1 A Case for Combining the Money Growth Rate with the Reserve Ratio

In this box, I show that the S/Y ratio, in equilibrium, is a nonlinear function of the reserve requirement and money growth rate. This application is a modified version of the economy developed in Champ and Freeman (1994). The chief feature of the model is that a person engages in market activity for two consecutive periods. In other words, N people enter market activity at each date t, stay for two periods, and then exit. It is equivalent to interpret this setup as one in which people are alive for two periods. In this context for a particular date t, those entering the market for the first time are "the young," and those entering the second period of market activity are "the old." Each person receives labor income when young, but nothing when, old. Time is discrete and is indexed by $t = 1, 2, 3$, and so on. I assume there are N people at date $t = 1$ who have only one period in the economy; members of this generation are the "initial old." Preferences are identical for all people born at date t and after.

For simplicity, I focus exclusively on a stationary version of the following economy. All people born at date $t = 1$ and later have identical preferences. Thus, without loss of generality, one can focus on the problem of the representative person, which is characterized by the following equations:

$$max_{c1,c2}\, ln(c_{1t}) + \beta ln(C_{2t+1}) \tag{B.1}$$

$$y \geq c_{1t} + a_t \tag{B.2}$$

$$c_{2t+1} \leq q_{t+1}a_t \tag{B.3}$$

$$a_t = v_t m_t + d_t \tag{B.4}$$

$$v_t m_t \geq \gamma d_t \tag{B.5}$$

where c_{1t} is the young person's consumption at date t; c_{2t+1} is old-age consumption by the person born at date t; β is the person's discount factor; y is the person's labor income; a is the total quantity of goods saved by the young person; q is the gross return on savings carried from date t to date $t + 1$; and d is the quantity of goods stored as bank deposits by the young person. A person can also choose fiat money, which is m. Here, v stands for the value of fiat money—that is, the quantity of the consumption good that can be purchased with one unit of money. The consumption good is perishable.

Equation B.1 is a function that describes the welfare a representative person receives during a market-active period. The person seeks to maximize welfare by consuming as much of the consumption good as possible. Equation B.2 represents the two options—to save or to consume—that the typical young person faces when young, while Equation B.3 indicates that the typical old person can consume up to the value of principal and interest earned on savings. Equation B.4 shows that savings are in the form of either real money balances or bank deposits. Lastly, Equation B.5 is the reserve requirement, dictating that real money balances cannot be less than γ percent of bank deposits.

I assume that deposits offer a greater return than fiat money. Consequently, the typical person will hold the minimum quantity of fiat money balances. Equations B.4 and B.5, therefore, imply that $a_t = (1 + \gamma)d_t$. Substitute for a in Equations B.2 and B.3, and solve Equation B.3 for $(1 + \gamma)d_t$, substituting the result into Equation B.2. After the algebra, the expression is

$$y \leq c_{1t} + c_{2t+1}/q_{t+1}. \tag{B.6}$$

which is the person's lifetime budget.

To maximize lifetime utility, the typical person will choose first- and second-period consumption so that

$$\frac{1}{c_{1t}} = \frac{\beta q}{C_{2t+1}} \tag{B.7}$$

Equation B.7 is an efficiency condition. It says that labor income will be allocated between first- and second-period consumption so that the benefits received from the last good consumed when young (measured by $1/c_{1t}$) are equal to the benefits received from the last good consumed when old (measured by $\beta\ q/c_{2t+1}$). In this economy, the optimizing conditions imply that the typical person will spend all of the labor income. Hence, Equation B.6 holds with equality.

In a stationary equilibrium, $c_{1t} = c_{1t+1}$ and $c_{2t+1} = c_{2t+2}$ at any date t, so that one can drop the time subscripts. For a stationary equilibrium, Equations B.6 and B.7 imply that $c_1 = y/(1 + \beta)$. With $0 < \beta < 1$, the typical person will spend a fixed fraction of labor income on consumption when young.

One might ask why the equilibrium expression for c_1 does not contain q. The answer is that a change in the gross return on assets has two opposing effects on consumption when young: substitution and wealth. With the substitution effect, an increase in q, for example, makes consumption when young more expensive relative to consumption when old. (Note that in the lifetime budget constraint [Equation B.6], $1/q$ can be interpreted as the price of consumption when old). Thus, an increase in the gross return to assets would induce people to consume less when young and more when old. With the wealth effect, when c_2 becomes less expensive, consuming more of both c_1 and c_2 is possible. As such, an increase in q, for instance, will induce people to consume more when young. Clearly, the substitution and wealth effects have opposing impacts on consumption when young. With log utility, these effects exactly offset each other. Consequently, in a stationary equilibrium the value of c_1 is independent of movements in the gross return on assets.

With $c_1 = y/(1 + \beta)$, the level of bank deposits can be represented as

$$d = \frac{\gamma}{1 + \gamma}\left(\frac{\beta}{1 + \beta}\right) \tag{B.8}$$

Next, substituting Equation B.8 Into Equation B.5 yields the expression for the equilibrium value of real money balances; formally,

$$v_t m_t = \frac{\beta}{1 + b} \cdot \frac{\gamma}{1 + \gamma} y. \quad \text{(B.9)}$$

Here one can see the importance of the equilibrium expression for consumption when young. Because of the substitution and wealth effects, neither the quantity of deposits nor the quantity of real money balances is affected by changes in the gross return on assets. The implication is that the tax base for real seigniorage revenue is not affected by changes in real return on assets.

Expressing the equilibrium value of real seigniorage revenue is now possible. Substituing Equation B.9 into the expression for real seigniorage revenue yields $v_t m_t(1 - 1/\theta)$. With $g = \theta - 1$, I divide the expression by y so that the equilibrium reliance on seigniorage revenue per young person is

$$\frac{s}{y} = \frac{\beta}{1+\beta} \cdot \frac{\gamma}{1+\gamma} \cdot \frac{g}{1+g} \quad \text{(B.10)}$$

Equation B.10 indicates that the equilibrium s/y ratio is a nonlinear function of the reserve ratio and the money growth rate.* Indeed, it is straightforward to show that the equilibrium value of s/y is a concave function of both the reserve requirement ratio and the money growth rate (see note 12).

To see how a change in monetary policy affects the equilibrium seigniorage ratio, consider a permanent, unanticipated increase in the reserve requirement ratio. In this model economy, Equation B.5 indicates that the holdings of real money balances will increase. Remember that the equilibrium level of deposits is not affected even though q will decline. Thus, the model economy predicts that the equilibrium seigniorage ratio will rise in response to an increase in the reserve requirement.

Next, consider a permanent, unanticipated increase in the money growth rate. With faster money growth, Equation B.10 indicates that an economy's reliance on real seigniorage revenue will increase. With an increase in g, the tax rate on real seigniorage revenue will rise and the gross return on assets will decline. Because of the utility function, the equilibrium quantity of real money balances is not affected by the decline in the gross real return on assets. Thus, faster money growth translates into an increase in the seigniorage ratio.For this model economy, the s/y ratio increases with respect to an increase in either the reserve requirement ratio or money growth rate, but at a declining rate. For different utility specifications, substitution and wealth effects would not necessarily cancel each other out. Hence, the typical person could respond to an increase in q by increasing consumption when young, thereby saving less. Accordingly, a young person could reduce holdings of real money balances by enough to see a decline in the s/y ratio for a given increase in either the reserve requirement or the money growth rate. The purpose of this box is to illustrate the basic economic trade-offs. Hence, arguing the appropriate utility specification is outside the scope of this article.

Overall, Equation B.10 suggests a particular sufficient statistic for assessing the relationship between monetary policy and a country's reliance on seigniorage revenue. Throughout the statistical analysis, I use the product $z = [\gamma/(1 + \gamma)][g/(1 + g)]$ as the measure of monetary policy.

*Here, the s/y ratio pertains to the ratio of per capita levels, which accounts for the use of lowercase letters.

ratio and money growth rate and its reliance on seigniorage revenue. Further, the adjusted R^2—a measure of the variation in seigniorage that is accounted for by the regression variables—indicates the monetary policy measure accounts for more than 40 percent of the variation in the S/Y ratio, which is a reasonably good fit for a cross-country sample.

Important differences across countries could alter the relationship between monetary policy and the reliance on seigniorage revenue. For example, with only z as an explanatory variable, the regression's constant term captures any differences between countries. Insofar as differences across countries can be measured, additional insight may be gained into the relationship between monetary policy and reliance on seigniorage. Such measurements indicate whether the results obtained in this analysis are robust.

A particular concern is the ability of people to avoid the inflation tax. Such avoidance depends, in part, on a country's financial sophistication. Citizens in countries with more sophisticated financial structure, for instance, can avoid taxation by shifting to nonreservable deposits. They can also dodge the tax by shifting from currency to, say, credit cards as the means of payment. It would seem prudent, therefore, to measure a country's financial sophistication to assess whether this omitted variable affects the relationship between a country's monetary policy settings and its reliance on seigniorage.

The measure of financial sophistication should not depend on the monetary policy settings to get an accurate estimate of the coefficient between monetary policy and seigniorage reliance. In other words, movements in the measure should not reflect behaviors related to monetary policy settings, and yet the variable should be reasonably well correlated with a country's level of financial sophistication. In reality, finding such a measure is quite difficult. Two variables are offered as proxies for financial sophistication: the level of real per capita GDP in 1965 and a dummy variable indicating whether the country is a member of the Organization for Economic Cooperation and Development (OECD).[17] Certainly OECD membership and monetary policy settings are conceivably linked as part of a country's policy package. The more modest claim is that OECD membership and per capita real GDP are less likely to respond to movements in the monetary policy settings than are financial-sophistication measures such as bank deposits.

Figure 2 plots the combination of z and S/Y as well as separate fitted lines for non-OECD and OECD countries. Each line is fitted to a regression of the form $(S/Y)_i = c_0 + \alpha z_i + \beta z_i^2$. These two fitted lines appear quite different. Based on this preliminary look at the data, the evidence suggests that the relationship between a country's monetary policy

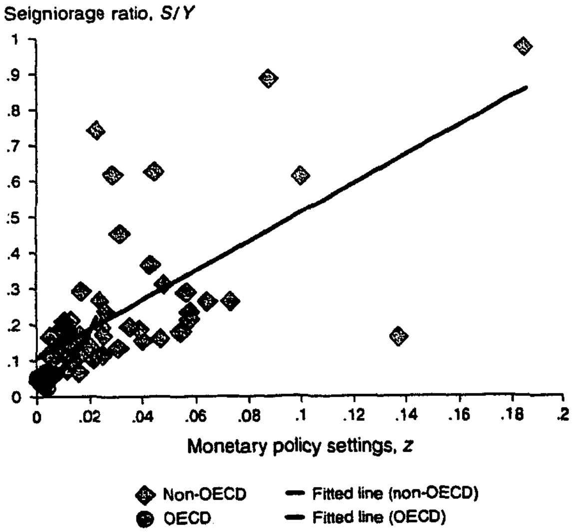


Figure 2 Seigniorage–income ratio and monetary policy settings.

statistic and its reliance on seigniorage revenue is different for developed countries than it is for less developed countries. Indeed, the fitted line for the non-OECD countries is upward sloping, whereas the one for the OECD countries appears to have some curvature.

I report regression results in columns 2 and 3 of Table 2. Here, I use two proxies to measure financial sophistication; one is the OECD membership, and the other is per capita real income in 1965. Two different sets of results emerge. Specifically, with per capita real income as the measure of financial sophistication, the evidence suggests a linear relationship between seigniorage reliance and z. As such, the evidence suggests, as it did when no financial sophistication measures were included, that countries that rely the most heavily on seigniorage revenue have higher monetary policy settings.

Consider, however, the results for a case in which OECD membership is used as a proxy for financial sophistication. These regression results correspond to the evidence presented in Figure 2. The formal statistical analysis supports the eyeball difference presented in the figure; that is, the z-S/Y relationship is significantly different for OECD countries than for non-OECD countries. The coefficient on $OECD \cdot z$ is negative and significant, and the coefficient on $OECD \cdot z^2$ is positive and significant. Thus, the evidence suggests that the relationship between seigniorage rates and monetary policy settings is convex. Indeed, the evidence indicates that an OECD country reaches a minimum reliance on seigniorage revenue at a value of $z = 0.0023$.

To illustrate this result, suppose one is looking at two OECD countries—country A and country B. Each has a different monetary policy setting, with country A always associated with the lower value of z. According to the regression statistics, if $z < 0.0023$ for both countries, then country B would rely less on seigniorage revenue than would country A. In contrast, for $z > 0.0023$, the regression predicts that country B would rely more on seigniorage revenue than would country A.[18]

The convex relationship exhibited by OECD countries is puzzling. In the model economy described in the box, financial sophistication would seem to permit a country's citizens to avoid the inflation tax. Given an increase in z, the equilibrium outcome for the

S/Y ratio would either decline or increase, but at a decreasing rate as people avoid the inflation tax. In other words, it is reasonable to expect that increased tax-avoidance capabilities would result in a more concave relationship between a country's monetary policy settings and its reliance on seigniorage revenue, not a more convex one.

Overall, the evidence suggests that there is a systematic, positive relationship between a country's monetary policy settings and its reliance on seigniorage revenue. Thus, countries with higher monetary policy settings tend to rely more heavily on seigniorage. But compared with less financially sophisticated countries, the more financially sophisticated countries tend to rely on seigniorage revenue at an increasing rate. The findings with respect to financial sophistication are difficult to explain and deserve more attention.

CONCLUDING REMARKS

I present evidence in this article on the importance of seigniorage revenue and its relationship to monetary policy. I use cross-country observations to examine whether the average money growth rate and average reserve ratio are systematically related to a country's reliance on seigniorage revenue. Both economic and statistical considerations suggest that some combination of the money growth rate and the reserve ratio should be used in the empirical analysis. Consequently, a country's monetary policy setting is measured using a combination variable as opposed to investigating two separate relationships—one between seigniorage reliance and the reserve ratio and the other between seigniorage reliance and the money growth rate.

The main finding in this article is that there is a systematic, positive relationship between a country's monetary policy settings and its reliance on seigniorage revenue. Thus, countries that rely most heavily on seigniorage revenue tend to have the highest values of the monetary policy measure. There is some additional evidence that the relationship between the monetary policy variable and the seigniorage rate is nonlinear for OECD countries. Here, OECD membership is used as a proxy for financial sophistication. The evidence suggests that OECD countries rely on seigniorage revenue at an increasing rate for given changes in the monetary policy variable.

The findings in this article constitute a very preliminary investigation of the relationship between seigniorage revenue and monetary policy. There is always a risk of excluding a key variable in a regression, and that risk certainly holds here. One approach would be to control for a host of other environmental factors—for example, a more complete analysis of the depth and structure of financial markets.

The most surprising and, in some ways, the most interesting results are those differentiating between financially developed and less financially developed countries. If these results were to stand up to further scrutiny, economic theory would need to address the puzzle. One possible line of research would be to consider a simple open economy in which two countries differ in terms of financial sophistication and monetary policy rules.

Another avenue for future research would be to recognize that monetary policy variables and seigniorage revenue are jointly determined. While I have tried to describe the correlations without referring to any monetary policy as "causing" movements in seigniorage revenue, the estimated regressions could be interpreted as treating monetary policy as exogenous to the determination of such revenue. Edwards and Tabellini (1991) examine seigniorage revenue as the outcome of various political forces that influence, among other things, monetary policy settings. Thus, future research could attempt to disentangle the relative importance of political factors, controlling for monetary policy explicitly.

NOTES

1. See Cox (1992) for an excellent discussion of the practical relationship between the Federal Reserve and the U.S. Treasury. For an interesting description of seigniorage in medieval times, see Rolnick, Velde, and Weber (1994).
2. For reference, the United States raises, on average, about 2 percent of federal government expenditures through money creation.
3. After all the accounting is consolidated for the government and the central bank, the net change in the government's income state is that money creation amounts to a revenue source to cover various expenses.
4. Bryant and Wallace (1984) offer an explanation for the coexistence of government bonds and money. They argue that the two types of government paper effectively price discriminate between "rich" and "poor" households.

   As far as my assumption about one-period bonds is concerned, I could examine a more complicated maturity structure for government debt. Such generality would not alter the conclusions that I reach about seigniorage revenue, but it would mean that I would have to keep tabs on the entire distribution of government bonds and when each one matures.
5. The reduction in reported income can come either from effective avoidance or from people working less or acquiring less capital. Of course, the discussion describes what happens to the steady-state level of income.
6. There is no explicit interest on money. Consequently, its one-period rate of return is calculated as the ratio of the date t price (the potential selling price) to the date t–1 price (the purchase price). Formally, this is the ratio of v_t/v_{t-1}. With $v_t = 1/p_t$, then simple substitution yields the expression for the gross real return on money.
7. Here, the reserve requirement pins down the fraction of a person's portfolio held in the form of money balances. This approach is qualitatively the same as one in which the reserve requirement pins down the bank's portfolio.
8. The data set is available from the author upon request.
9. Fischer is primarily interested in describing why countries maintain national currencies. Computing the seigniorage-to-GNP ratio demonstrates how important seigniorage is. The ratio represents the command over resources that a government obtains by creating money.
10. The income tax analog is the average marginal tax rate. See, for example, Seater (1985).
11. Historically, the U.S. reserve requirement structure was more convoluted. In the past, for example, it mattered whether the commercial bank was located in a Reserve Bank city or outside.
12. Interestingly, Fischer (1982) presents evidence that several governments have made substantial use of seigniorage. In Fischer's sample, which generally covers the period between 1960 and 1978, Argentina collected, on average, 6.2 percent of GNP through money creation.
13. This result does not bear directly on the relative importance of seigniorage revenue. Rather, it bears on the issue of variability within a country across time. In short, the reader gains a sense of how the countries in the sample rely on seigniorage over time.
14. The effect of a change in the reserve ratio, holding money growth constant, is given by the following derivative: $\partial z/\partial(R/D) = W/(1 + R/D)^2$, where $W = g/(1 + g)$. With $W > 0$, the expression says z is increasing the reserve ratio. In addition, $\partial^2 z/\partial(R/D)^2 = (-2 \cdot W)/(1 + R/D)^3$, which is negative for $W > 0$.
15. To see this relationship, suppose the estimated regression is given by

$$S/Y = c_0 + \alpha z + \beta z^2.$$

For a country with a 1-percentage-point higher average z, an estimate of the change in S/Y is $\alpha + 2\beta z$. Thus, a 1-percentage-point change in z depends on the value of z.

16. In all regressions, the Newey–West procedure is applied to correct any potential bias in standard errors. In this particular application, heteroskedasticity is the chief worry.
17. Per capita real GDP comes from the Summers–Heston Penn World Tables. In addition, regressions are run using per capital real GDP for 1980 and 1994 as alternative measures of financial sophistication in case the 1965 GDP value suffers from some time-specific factors. The regressions are qualitatively the same as those reported in Table 2.
18. Three OECD countries in this sample—France, the Netherlands, and Norway—have z values less than 0.0023. Using the method outlined in Fomby, Hill, and Johnson (1984, 58), one can compute the standard errors for the value of z at which seigniorage reliance is minimized. With 90 percent confidence, the seigniorage-reliance minimizing value of z is between 0.0022 and 0.0024.

REFERENCES

Bryant, J., and Wallace, N. 1984. A Price Discrimination Analysis of Monetary Policy. *Review of Economic Studies* 51 (April): 279–88.

Champ, B., and Freeman, S. 1994. *Modelling Monetary Economies.* New York: John Wiley and Sons Inc.

Cox, W. M. 1992. Two Types of Paper: The Case for Federal Reserve Independence. Federal Reserve Bank of Dallas *Southwest Economy,* Issue 6, November/December, 3–8.

Edwards, S., and Tabellini, G. 1991. Explaining Fiscal Policies and Inflation in Developing Countries. *Journal of International Money and Finance* 10 (March): S16–S48.

Fischer, S. 1982. Seigniorage and the Case for a National Money. *Journal of Political Economy* 90 (April): 295–313.

Fomby, T. B., Hill, R. C., and Johnson, S. 1984. *Advanced Econometric Methods.* New York: Springer-Verlag Inc.

Rolnick, A. J., Velde, F. R. and Weber, W. E. 1994. The Debasement Puzzle: An Essay on Medieval Monetary Policy. Federal Reserve Bank of Minneapolis Working Paper no. 536 (October).

Sargent, T. J. 1986. *Rational Expectations and Inflation.* New York: Harper & Row.

Seater, J. 1985. On the Construction of Marginal Federal Personal and Social Security Tax Rates in the U.S. *Journal of Monetary Economics* 15 (January): 121–35.

29

Monetary Policy Comes of Age
A Twentieth Century Odyssey

Marvin Goodfriend
Senior Vice President and Director of Research, Federal Reserve Bank of Richmond, Richmond, Virginia

In the early 1960s the Federal Reserve (Fed) was little known outside of the financial services industry and university economics departments. Twenty years later Fed Chairman Paul Volcker was one of the most recognized names in American public life. Now hardly a week goes by when the Fed is not featured prominently in the business news. The Fed was thrust into the limelight in the intervening years when the public came to associate it with inflation-fighting policy actions that raised interest rates and weakened economic activity. Even though inflation has been held in check since the mid-1980s, the public remains acutely aware of Fed policy today.

Monetary economists and central bankers alike now understand that effective monetary policy must be built on a consistent commitment to low inflation. That is why in recent years the Fed has made low inflation a particularly high priority. The large fraction of the public having first-hand experience with high inflation naturally supports the view that inflation must be contained. As the collective memory of inflation fades, however, public support for low inflation will become increasingly difficult to sustain. A permanent national commitment to price stability requires that citizens personally unfamiliar with the trauma of high inflation understand the rationale for price stability and the tactical policy actions needed to maintain it.

This article reviews the history of U.S. monetary policy in the 20th century with the aim of providing that understanding. It identifies mistakes that led to high and volatile inflation, lessons learned from the experience, and principles applied in the pursuit of low inflation today. U.S. monetary policy came of age in the 20th century in the sense that the country left the strict rules of the gold standard for the freedom of an inconvertible paper standard, which the Fed only slowly and painfully learned to manage. What follows is the story of that 20th century odyssey.

Reprinted from: *Federal Reserve Bank of Richmond Economic Quarterly,* v. 83, n. 1 (Winter 1997) 1–22.

MONETARY POLICY UNDER THE GOLD STANDARD

When the Federal Reserve was established in 1913, inflation was not the problem it was to become in the latter part of the century. The nation was on a gold standard and the purchasing power of money in 1913 was about what it had been 30 years before, or for that matter, 100 years before. The gold standard sharply restricted inflation by requiring that money created by the U.S. Treasury be backed by gold.[1]

The classical gold standard yielded price stability only to the extent that the Treasury's stock of monetary gold happened to expand at a rate sufficient to satisfy the economy's demand for money at stable prices. For instance, slow growth in the gold supply caused the price level to decline at over 1 percent per year from 1879 to 1897, and gold discoveries and new mining techniques caused inflation to average over 2 percent per year between 1897 and 1914. Nevertheless, by the standard of what was to come, the variation of inflation was very small.

Although the economy grew rapidly throughout the gold standard years, the period was marked by a number of recessions associated with temporary deflations and substantial interest rate movements. Sudden sustained short-term interest rate spikes of over 10 percentage points occurred on eight occasions between the Civil War and the founding of the Federal Reserve. Five of these spikes were associated with bank runs characterized by a demand to convert bank deposits into currency that could not be satisfied by the fractional cash reserves held by banks.[2]

Finally, in response to the banking panic of 1907 and the ensuing recession, the nation was no longer willing to run monetary policy entirely according to the classical gold standard rules. The Federal Reserve was established in the United States with the power to create currency and bank reserves at least somewhat independently of the nation's monetary gold. The Fed was given authority to create currency and reserves by making loans to banks through its discount window or by acquiring securities in the money market. The Fed's mission was to provide an elastic supply of money to smooth short-term interest rates against liquidity disturbances, while preserving the link between money and gold in the long run in order to restrain inflation.[3]

Through its dominant presence in the market for currency and bank reserves, the Fed easily gained control of short-term interest rates and eliminated the kind of interest rate spikes seen earlier.[4] By smoothing short-term interest rates, however, the Fed was obliged to substitute its discretionary management of short rates for the impersonal market forces that had determined rates previously. The Bank of England had successfully managed short rates for decades in the context of the classical gold standard.[5] And the Fed could have followed similar gold standard operating procedures. However, the classical gold standard collapsed with World War I, and the nation was never willing to support Fed procedures geared to defending the gold standard. The Fed was left without clear operational procedures for positioning short-term interest rates to stabilize economic activity around full employment with stable prices.

THE OBSTACLES TO UNDERSTANDING MONETARY POLICY

Improvements in monetary policy that seemed within reach after the founding of the Fed proved elusive. The 1930s saw the sharpest deflation, the worst banking crisis, and the longest and deepest economic depression in American history.[6] Then, beginning in the mid-1960s there were two decades of unprecedented peacetime inflation that tripled the general price level by the early 1980s.[7]

Why has it taken so long for the Fed to give price stability pride of place?[8] Initially, there was a tendency to underestimate the disruptive potential of inflation and a willingness to be tolerant of each new burst of inflation in the hope that it would soon die down. Such hope seemed reasonable since protracted peacetime inflation had never before been a problem in the United States. Another difficulty was that it took some time for economists to develop a framework capable of understanding monetary policy in the absence of a link to gold. Prior to the 20th century the world had little practical experience with monetary regimes in which money was unbacked by a commodity such as gold or silver. With some exceptions, mainly during wartime, there was little empirical evidence on such regimes and little interest in analyzing them.

The main problem was confusion within the economics profession about the determination of the general price level and the control of inflation in a regime of inconvertible paper money.[9] There was also little understanding of the role played by inflation expectations in the wage- and price-setting process and in the determination of interest rates. Finally, the relationship between unemployment and inflation was seriously misunderstood. The resolution of these disputes provided the foundation for today's monetary policy success.

INFLATIONARY MONETARY POLICY AT MID-CENTURY

Largely as a result of the nation's unfortunate experience with inflation in the period from the mid-1960s through the early 1980s, monetary economists and central bankers now understand that the costs of inflationary monetary policy are significant and varied. First are the costs that even a steady, perfectly anticipated inflation imposes on society. Then there are the disruptive and destabilizing costs of unstable inflation, more difficult to quantify but substantial nonetheless. These latter costs stemmed from alternating expansionary and contractionary policy actions. Specifically, there was a tendency—known as go-stop monetary policy—for the Fed to exacerbate the cyclical volatility of inflation and unemployment. And there was a related tendency to produce rising inflation and increasingly volatile inflation expectations over time. The forces giving rise to these tendencies are identified and described below together with their disruptive consequences.

The Cost of Steady Inflation

The cost of steady inflation begins with the fact that a steadily falling purchasing power of money causes people to hold less cash than they would if prices were stable. Attempts to economize on money holdings manifest themselves in several ways. Banks invest in teller machines, people visit banks or teller machines more frequently, businesses devote more time and effort to managing their cash balances, etc.[10] Even more important, individuals and firms take steps to protect the value of their savings and investments against loss due to inflation. The effort and resources devoted to dealing with inflation are wasted from society's point of view in the sense that they could be better employed in producing goods and services.

Another major cost of steady inflation stems from the incomplete indexation of the tax system. The biggest problem in this regard results because taxes are assessed on *nomi-*

nal interest earnings and *nominal* capital gains, that is, on investment returns in dollars. Inflation causes nominal returns to rise because investors demand compensation for the declining purchasing power of money. For instance, long-term bond rates contain a premium for expected inflation over the life of the bond. Since nominal returns are taxed as income, however, inflation reduces the after-tax return to saving and investment and thereby tends to inhibit capital accumulation and economic growth.[11]

Go-Stop Monetary Policy[12]

A central bank such as the Fed that is charged with conducting monetary policy on a discretionary basis is naturally inclined to give considerable weight to the public's mood. Go-stop monetary policy was, in good part, a consequence of the Fed's inclination to be responsive to the shifting balance of public concerns between inflation and unemployment. Of course, difficulties in judging the strength of the economy and in gauging inflationary pressures compounded the problem, as did ignorance of the lags in the effect of policy.

For the most part of the public tolerated inflation as long as it was low, steady, and predictable. When labor markets were slack, the public was even willing to risk higher inflation in order to stimulate additional economic activity. Only when economic activity was strong and inflation moved well above the prevailing trend did inflation top the list of public concerns.

It is easy to understand why inflation need not greatly concern the public when it is steady and predictable. Individuals and firms are inconvenienced only slightly by steady inflation. As long as wages, prices, and asset values move up in tandem, there are no big financial consequences, especially when inflation is low. Likewise, a temporary and modest increase of inflation around a low, well-established trend need not immediately arouse concerns.

However, a persistent departure of inflation above trend causes anxieties because people wonder where a new trend might be established. Investors worry about how much of an inflation premium to demand in interest rates; businesses worry about how aggressively to price in order to cover rising costs; and workers worry about maintaining the purchasing power of their wages.

In marked contrast to inflation, which affects all, unemployment actually affects relatively few at a given time. The unemployment rate in recent decades has risen at most to only about 10 percent of the labor force. The public is concerned about unemployment not so much because of those who are currently unemployed but because people are afraid of becoming unemployed. It follows that the public is generally more concerned about unemployment when the unemployment rate is rising, even if it is still low, than when it is falling, even if it is already high.

The above-mentioned reasoning helps explain why the Fed produced go-stop monetary policy in the 1960s and '70s. In retrospect, one observes the following pattern of events.

First, because inflation became a major concern only after it clearly moved above its previous trend, the Fed did not tighten policy early enough to preempt inflationary outbursts before they became a problem.

Second, by the time the public became sufficiently concerned about inflation for the Fed to act, pricing decisions had already begun to embody higher inflation expectations. Thus delayed, a given degree of restraint on inflation required a more aggressive increase in short-term interest rates with greater risk of recession.

Third, in any cyclical episode there was a relatively narrow window of broad public support for the Fed to tighten monetary policy. The window opened after inflation was widely recognized as the major concern and closed when tighter monetary policy caused the unemployment rate to begin to rise. Often the Fed did not take full advantage of a window of opportunity to raise short rates, because it wanted more confirmation that higher short-term rates were required.

Fourth, it was probably easier for the Fed to maintain public support for fighting inflation with prolonged rather than preemptive tightening. A more gradual lowering of interest rates in the later stage of a recession was a less visible means of fighting inflation than raising rates more sharply earlier. Once unemployment peaked and began to fall, the public's anxiety about it diminished. Prolonged tightening was attractive as an inflation-fighting measure in spite of the fact that it probably lengthened the "stop" phase of the policy cycle.

Rising Inflation and Unstable Inflation Expectations

Over time, deliberately expansionary monetary policy in the "go" phase of the policy cycle came to be anticipated by workers and firms. Workers learned to take advantage of tight labor markets to make higher wage demands, and firms took advantage of tight product markets to pass along higher costs in higher prices. Increasingly aggressive wage- and price-setting behavior tended to neutralize the favorable employment effects of expansionary policy. And the Fed became evermore expansionary on average in its pursuit of low unemployment, causing correspondingly higher inflation and inflation expectations. Lenders demanded unprecedented inflation premia in long-term bond rates. And the absence of a long-run anchor for inflation caused inflation expectations and long bond rates to fluctuate widely.[13]

The breakdown of mutual understanding between the markets and the Fed greatly inhibited the conduct of monetary policy. The Fed continued to manage closely short-term nominal interest rates.[14] But the result of an interest rate policy action is largely determined by its effect on the real interest rate, which is the nominal rate minus the public's expected rate of inflation. And the Fed found it increasingly difficult to estimate the public's inflation expectations and to predict how its policy actions might influence those expectations. Compounding the problem, enormous increases in short-term interest rates were required by the early 1980s to stabilize the economy. Stabilization policy became more difficult because the public could not predict what a given policy action implied for the future, and consequently, the Fed could not predict how the economy would respond to its policy actions.

THE CONTROL OF INFLATION: DISINFLATION IN THE 1980S

By the late 1970s, policymakers and monetary economists had come to understand the costly and disruptive features of inflation discussed above. With considerable public support, the Fed under the leadership of Chairman Paul Volcker initiated the great disinflation in October 1979, marking the beginning of the period in which the Fed would make lowering inflation a priority. What followed was a tightening of monetary policy that succeeded in bringing the inflation rate down permanently for the first time in the post–Korean War period, first from over 10 percent to around 4 percent by 1983, and then to around 3 percent by the mid-1990s.

This section reviews three developments that paved the way for the Fed to take responsibility for price stability. Most important was the progress that economists made in understanding money demand and supply. Next was the failure of nonmonetary approaches to controlling inflation. Finally, and to a lesser extent, was the idea advanced by monetary economists that the unemployment cost of disinflation might be minimized if the disinflation were credible.

The Central Bank's Responsibility for Inflation

The consensus among monetary economists that central banks are responsible for inflation is built on both theory and evidence. Above all, there is the substantial body of evidence from the inflationary experiences of a great many nations, including the widespread inflation in the industrialized world during the 1960s and '70s, showing that sustained inflation is always associated with excessive money growth. The evidence also clearly indicates that inflation is stopped by slowing the growth of the money supply.[15]

The theory of money demand and supply supports the cross-country evidence by illuminating the mechanics of the link between monetary policy and inflation. The theory of money demand implies that control of the money supply is necessary and sufficient to control the trend rate of inflation. And the theory of money supply implies that a central bank can control the trend rate of money growth. As will become clear below, money demand may be thought of as the fulcrum by which a central bank controls inflation, and the money supply may be thought of as the lever by which it does so.

Money Demand

The theory of money demand asserts that individuals and businesses choose to hold a target stock of money that is proportional to their expenditures, a target that balances the convenience of holding money against the foregone interest earnings.[16] The key implication of money demand theory for monetary policy is that there is a reasonably stable long-run relationship between a nation's demand for money and its production and exchange of goods and services.

It follows that sustained inflation results when the growth of the nation's money stock exceeds the rate of growth of the nation's physical product.[17] Prices must rise in this case because otherwise individuals and firms would spend their growing excess money balances. Since one person's expenditure is another person's receipts, the spending would put upward pressure on prices until the inflation rate matched the rate of money growth in excess of the growth of output. Only then would the ongoing increase in the stock of money be willingly absorbed by the public.

The theory of money demand also implies that the overall price level cannot move very much over the long run if the stock of money grows in tandem with the growth of output.[18] If an inflation were to start, it would reduce the purchasing power of a given nominal stock of money and cause individuals and businesses to cut their spending in an effort to maintain their inventory of monetary purchasing power. With no additional money balances forthcoming in the aggregate, the downward pressure on spending would stop the inflation.

Money Supply

The nation's basic money supply consists of currency and checkable deposits held by households and businesses. A central bank can control the former because it has a monopoly on the creation of currency.[19] Checkable deposits are created by banks. A central bank also has the power to control checkable deposits because banks must hold reserves to service their deposits, and a central bank controls the aggregate stock of bank reserves.[20]

The financial services industry has long been creating new instruments in which the public can hold liquid balances, e.g., certificates of deposit and money market mutual funds. New financial instruments usually do not add to the basic money supply since they are only imperfect substitutes for currency or checkable deposits.[21] Nevertheless, the introduction of money substitutes has adversely affected the predictability of money demand in the short run. In practice, however, money demand is sufficiently stable and money supply sufficiently controllable over time, so that financial innovations do not fundamentally alter a central bank's power over inflation.[22]

Failed Approaches to Controlling Inflation

A variety of nonmonetary approaches to controlling inflation were tried in the 1960s and '70s. In the United States, for example, the federal government published voluntary wage-price guidelines at various times to persuade firms and workers to forego price and wage increases deemed excessive.[23] Actual controls were imposed for a few years in the early '70s but for the most part they were lifted by the mid-'70s.[24] By the end of the period, both controls and guidelines came to be regarded as arbitrary, unfair, and ineffective. Moreover, where they were effective they often created allocative disruptions, e.g., price controls in the energy sector created shortages and long lines at gas stations.

In the early 1960s economists believed that budget policy might play a key role in fighting inflation. In the United States, however, it quickly became clear in the Vietnam War period that political concerns would immobilize fiscal policy as a practical economic policy tool. Moreover, it later became clear that the inflation of the 1970s was not closely related to the government's fiscal situation.[25]

Even after the Fed under Chairman Volcker had begun its momentous disinflation, the Carter administration imposed credit controls in early 1980 in an effort to foster the process. The credit control program caused a sharp recession with little impact on inflation and was phased out at midyear.[26]

Thus did policymakers learn the hard way that policies for stopping inflation other than monetary control didn't work. As much as anything else, the failure of nonmonetary approaches to disinflation set the stage for the Fed to take responsibility for bringing inflation down.

Credibility for Low Inflation and the Unemployment Cost of Disinflation

In the early 1960s policymakers were inclined to accept the inflationary consequences of policy actions taken to stimulate aggregate demand and employment. That inclination was based to a great extent on evidence of a century-long negative Phillips curve correlation between unemployment and (wage) inflation in the United Kingdom that appeared to offer a trade-off in which the benefits of lower inflation would have to be balanced against the costs of higher unemployment.[27]

When stimulative policy succeeded in driving down the unemployment rate in the '60s, the resulting increase in inflation at first seemed consistent with a stable Phillips curve trade-off; and the rising inflation was tolerated as a necessary evil.[28] In the 1970s, however, the Phillips curve correlation broke down as inflation and unemployment both moved higher, and it became clear that high inflation could not buy permanently low unemployment.[29]

Even though protracted inflation was widely understood by the late 1970s to have costs with no offsetting benefits, it was recognized that bringing inflation down would be costly too. Previous experience with go-stop policy made it clear that there was a short-run trade-off between unemployment and inflation.[30] Policymakers expected the temporary unemployment cost of a large permanent disinflation to exceed the costs of earlier disinflations that the Fed had produced in the "stop" phase of its policy cycles.

To some degree a view then emerging in the academic community might have encouraged the Fed to pursue the disinflation. The view holds that the unemployment cost of disinflation can be minimized if a disinflation policy is credible. The idea that credibility would govern the costliness of disinflation has since become widely accepted in theory.[31] The acquisition and maintenance of credibility for low inflation have become major practical concerns of Fed policymakers and central bankers around the world.

The idea underlying the role of credibility is that wage- and price-setting behavior is geared to expectations of money growth. The Fed supports the ongoing inflation as long as money grows in excess of output. If the Fed's disinflation is credible, the Fed slows money growth and wage and price inflation come down, too, with little effect on employment. On the other hand, if the disinflation is not credible, then wage and price inflation continues as before. If the Fed persists in slowing money growth anyway, a deficiency of aggregate demand causes unemployment as households and businesses cut spending in an attempt to maintain their targeted monetary purchasing power.[32]

In practice, however, disinflation is nearly always costly because credibility for low inflation is hard to acquire after it has been compromised. Moreover, a central bank's commitment to low inflation is only as credible as the public's support for it. The Fed probably embarked on the disinflation in 1979, in part, because the public finally seemed ready to accept it.

Although its discount rate changes often made the headlines prior to 1979, the Fed rarely sought publicity for its monetary policy actions. Chairman Volcker broke sharply with tradition by initiating the period of disinflationary policy with a high-profile announcement signaling that the Fed would take responsibility for inflation and bring it down.[33] In so doing, Chairman Volcker built credibility by staking his own reputation and the Fed's on achieving the low inflation objective. The unprecedented increases in short-term interest rates that followed further demonstrated the Fed's commitment to reducing inflation.[34]

Nevertheless after two decades of rising inflation, a widespread skepticism worked against Fed credibility.[35] Wage and price setters doubted that there would be sufficiently widespread public support for the Fed's disinflation. Indeed, the inflation was not broken until a sustained slowing of money growth beginning in 1981 created a serious recession that tested the Fed's determination and the public's support.[36] Although the recession was the worst since the 1930s, it was less severe than might have been expected considering the size of the accompanying disinflation. Most remarkable is that the roughly 6 percentage point disinflation occurred in just two years: 1981 and 1982. The size and speed of the disinflation suggests that the acquisition of credibility played a key role in making it happen.

MONETARY POLICY AT THE CLOSE OF THE CENTURY: MAINTAINING LOW INFLATION

The Fed has succeeded in maintaining low inflation for almost 15 years now. With luck the United States should enter the 21st century with inflation near what it was under the gold standard at the opening of the 20th century. Macroeconomic performance during the low inflation period has been good, especially when compared to the inflationary period preceding it. The only recession during the period, in 1990–91, was mild by recent standards. Over the period as a whole, employment growth has been strong and productivity growth may have picked up somewhat. Moreover, both short- and long-term interest rates are around one-third of what they averaged in the early 1980s and are much less volatile too.

The promise of low inflation is being fulfilled. The challenge today is for the Fed to understand the secret of its success. In that regard the low inflation period has as much to teach as the traumatic period that preceded it. In reviewing below the lessons learned and principles applied, we shall see that the best way of assuring our continued monetary policy success would be for Congress to give the Fed a legislative mandate for low inflation.

Lessons Learned and Principles Applied

One of the most important lessons learned from the last four decades is that credibility for low inflation is the foundation of effective monetary policy. The Fed has acquired credibility since the early 1980s by consistently taking policy actions to hold inflation in check. In effect, the Fed has reestablished a mutual understanding between itself and the markets. From this perspective, wage and price setters keep their part of an implicit bargain by not inflating as long as the Fed demonstrates its commitment to low inflation. Ironically, the Fed has learned from nearly a century of experience to pursue rule-like behavior in order to fully achieve the gains from moving away from the gold standard.

Experience shows that the guiding principle for monetary policy is to preempt rising inflation. The go-stop policy experience teaches that waiting until the public acknowledges rising inflation to be a problem is to wait too long. At that point, the higher inflation becomes entrenched and must be counteracted by corrective policy actions more likely to depress economic activity.

The main tactical problem for the Fed is to decide when preemptive policy actions are necessary and how aggressive they should be. In this regard, the Fed must be careful to consider any adverse effect that a poorly timed policy tightening could have on employment and output. For that matter, the Fed must be prepared to ease monetary policy when a weakening economy calls for it. The central bank's credibility depends not only on its inflation-fighting credentials but also on its perceived competence.

A natural starting point to balance these concerns is to use a policy rule-of-thumb based on historical data to benchmark Fed policy. The stance and direction of monetary policy can then be chosen in light of historical experience conditioned on any special current circumstances. The most relevant historical experience is, of course, the relatively brief low inflation period since the mid-1980s. As the Fed extends low inflation over time, the nation will build up a richer relevant history against which to benchmark policy.[37]

However, even our brief experience with low inflation contains useful insights such as this. In some years, such as 1994, inflationary pressures might be judged to call for a particularly aggressive preemptive tightening. At other times, such as in 1996, there might be some concern about the potential for rising inflation but enough doubt to adopt a wait-and-see attitude. The Fed's success in 1994 and 1996 suggests that the key to effective management of short-term interest rates over the business cycle is to move rates up decisively and preemptively when warranted in order to build credibility for low inflation. With credibility "in the bank," so to speak, the Fed can hold rates steady or move them down out of concern for unemployment at other times.[38] The lesson is that credibility enhances flexibility.

A Legislative Mandate for Price Stability

Largely as a result of the common understanding of the theory and history of monetary policy reviewed above, there is today a consensus among monetary economists and central bankers that maintaining low inflation is the foundation of effective monetary policy. Moreover, there is an *emerging* consensus that a central bank's commitment to price stability should be strengthened by legislation making low inflation the primary goal of monetary policy.[39]

The recommended priority for price stability derives not from any belief in its intrinsic value relative to other goals such as full employment and economic growth. Price stability should take priority for two reasons: first, the Fed actually has the power to guarantee it over the long run, and second, monetary policy encourages employment and economic growth in the long run mostly by controlling inflation.[40] Also, and this is very important, a mandate for price stability would not prevent the Fed from taking the kinds of policy actions it takes today to stabilize employment and output in the short run. What it *would* do is discipline the Fed to justify these actions against a commitment to protect the purchasing power of money.

Two often-repeated objections to a mandate for low inflation deserve mention here. One is the notion that low inflation targeting is largely irrelevant because two enormous oil price increases in the 1970s—in 1973/74 and 1979/80—were responsible for the worst inflation of that period.[41] The claim continues that our success in controlling inflation will be determined by whether we have large oil price shocks in the future or not. Clearly, oil price increases create a problem for the economy: the higher price of oil diverts expenditure to oil products and raises real costs throughout the economy, with adverse consequences for demand and employment in non-oil sectors.

The economy must adjust to the higher real cost of oil in any case. The problem for a central bank is to make sure that the adjustment problem is not compounded with monetary instability. A central bank with a mandate for low inflation is more likely to resist excessive monetary accommodation than one with a weaker commitment to price stability. This is because an oil price shock will be less likely to set in motion wage and price increases that the central bank will be inclined to accommodate. The Fed was in just this predicament when the 1970s oil price shocks hit, since rising inflation trends were already well established before each oil shock. The destabilizing effects on inflation, inflation expectations, and employment and output would almost surely have been less troublesome in a climate of stable inflation.

A second objection to a mandate for low inflation is that it would hold back economic growth. In fact, the opposite is more nearly true. In terms of the earlier discussion of money demand and supply, trend growth of national output continually raises the demand for money, and the Fed accommodates the growing demand for money at stable prices.

Would monetary policy prevent the economy from growing faster if labor productivity unexpectedly surged? Not for long, because unemployment would begin to rise as businesses found that they could meet demand with less labor input. And the Fed would resist rising unemployment by easing monetary policy to encourage faster growth in aggregate demand. In short, the Fed's policy procedures do not "target growth." A mandate for price stability would allow the Fed to naturally and automatically accommodate an increase in productivity growth over time.

Ultimately the Fed can only secure full credibility for low inflation with the backing of the public. The public's misunderstanding of the tactics of monetary policy is particularly troublesome. For instance, accusations that the Fed was "busting ghosts" when it ran short-term interest rates up in 1994 threatened to undermine support for policy actions that were clearly called for.[42] Preemptive policy actions in 1994 laid the foundation for continued economic expansion. The task ahead must be to broaden and deepen the public's understanding and support for the strategy and tactics of monetary policy and to lock in credibility for low inflation with a legislative mandate.

CONCLUSION

American monetary policy has come full circle in the 20th century. Early in the century the nation overcame a long-standing distrust of government intervention in the monetary system to establish a central bank. The Federal Reserve embodied the idea that discretionary monetary policy could improve on the rules of the gold standard, rules that were seen as unduly restrictive. We now know that the faith then placed in discretion over rules was somewhat misplaced. Today, monetary economists and central bankers alike understand that effective monetary policy must be built on a consistent commitment to low inflation.

Numerous lessons were learned on the 20th century odyssey. The most important is that the Federal Reserve, through its management of monetary policy, is responsible for inflation. This became clear partly as a result of advances in monetary theory and partly as a result of evidence on money demand and money supply. It was also the result of a learning process in which nonmonetary approaches to controlling inflation were seen to fail, and the monetary approach succeeded.

Discretionary monetary policy actions can be invaluable in fighting a financial crisis or a weak economy. But we learned that the promise of discretion can be realized fully only in the context of a monetary policy that makes price stability a priority. Otherwise discretion leads inexorably to go-stop policy that brings rising and unstable inflation and inflation expectations, with adverse consequences for interest rates and employment.

The go-stop experience taught that the Fed should fight inflation by tightening monetary policy before price pressures break out into the open. Waiting until inflation has begun to rise may better assure public support for higher short-term interest rates. But delayed tightening allows higher inflation to become more firmly established, requiring even higher rates to choke it off, with a greater risk of recession.

An emerging consensus among monetary economists and central bankers supports the need for a legislative mandate to make low inflation the primary goal of monetary policy. That recommendation has broad backing for three reasons. A central bank can guarantee low inflation over time. Monetary policy most effectively stabilizes employment over the business cycle when it has credibility for low inflation. And full credibility for low inflation needs the support of a legislative mandate.

Monetary policy has come of age in the 20th century in the sense that monetary economists and central bankers have come to terms with the past—lessons have been learned and principles have been applied successfully. The country should build on that professional consensus to broaden the public's understanding and support for price stability and the preemptive policy procedures to sustain low inflation. The nation has the opportunity to bring a tumultuous chapter in its monetary history to a close. It should grasp that opportunity and enjoy the benefits that sustained price stability would bring.

NOTES

1. Technically, the United States was on a bimetallic (gold and silver) standard until 1900. Though it is true that the money supply was limited by the size of the Treasury's gold and silver holdings, there was considerable short-run variability in the money multiplier. See Cagan (1965) and Freidman and Schwartz (1963).
2. Major banking panics occurred in 1873, 1884, 1890, 1893, and 1907.
3. This latter understanding was viewed as part of the Fed's mission, although it is only implicitly, not explicitly, stated in the Federal Reserve Act of 1913 itself.
4. See Goodfriend (1988).
5. See Hawtrey (1938).
6. According to Friedman and Schwartz (1963) U.S. real net national product fell by more than one-third from 1929 to 1933, implicit prices of goods and services fell by more than one-quarter, and wholesale prices by more than one-third. More than one-fifth of the commercial banks in the United States holding nearly one-tenth of the deposits closed because of financial difficulties. As a result of the sharp contraction in economic activity, the unemployment rate peaked at over 20 percent in 1932–33, and remained above 10 percent for the remainder of the decade.
7. The Fed had already recognized inflation as a problem on three occasions prior to the mid-1960s: in the aftermath of World War II, during the Korean War and the period of the 1951 Fed-Treasury Accord, and again in the mid-1950s. See footnote 12.
8. Under the leadership of Benjamin Strong, Governor of the Federal Reserve Bank of New York, the Fed made price stability a priority briefly in the 1920s. See Hetzel (1985).
9. See, for example, Bronfenbrenner and Holzman (1963) and Friedman (1987).
10. Estimates in Lucas (1994) imply that the economization of money balances that occurred at a rate of inflation of 5 percent per year (associated with a short-term nominal interest rate of about 6 percent) wasted about 1 percent of U.S. GDP. The payment of interest on transactions deposits in recent years raises money balances and reduces this welfare cost somewhat. The bulk of the welfare gain to reducing inflation is probably realized at a slightly positive inflation rate. See Wolman (1997).
11. Feldstein (1996) reports that the net present value of the welfare gain of shifting from 2 percent inflation per year to price stability forever is about 30 percent of the current level of GDP.
12. Friedman (1964, 1972, and 1984) discusses go-stop policy. Romer and Romer (1989) document that since World War II the Fed tightened monetary policy decisively to fight inflation on six occasions beginning, respectively, in October 1947, September 1955, December 1968, April 1974, August 1978, and October 1979. The unemployment rate rose sharply after each policy shock. Only two significant increases in unemployment were not preceded by Fed action to fight inflation. One occurred in 1954 after the Korean War and the second occurred in 1961, after the Fed tightened monetary policy to improve the international balance of payments.
13. The monthly average 30-year bond rate rose from around 8 percent in early 1978 to peak above 14 percent in the fall of 1981. The long bond rate was near 13 percent as late as the summer of 1984.
14. See Cook (1989).
15. See, for instance, Friedman (1987), Poole (1978), and Sargent (1986).
16. See McCallum and Goodfriend (1987).
17. The public's target ratio of money to expenditure may exhibit a trend at times in response to, say, rising interest rates or technical progress in the payments system. For instance, the ratio of money to expenditure will trend downward if money provides transaction services more efficiently over time. In that case, the money growth rate consistent with price stability will be below the growth of physical product.
18. See the preceding note.
19. Electronic private substitutes for government currency have become feasible recently. See Lacker (1996).

20. See Cagan (1965).
21. There have been exceptions, however. For instance, a new deposit type known as the negotiable order of withdrawal (NOW) account was introduced in the late '70s and early '80s as part of the deregulation of the prohibition of interest on checkable deposits. NOW accounts were interest-earning substitutes for demand deposits and so were immediately included in the Fed's M1 measure of the basic money supply for purposes of targeting and control. See Broaddus and Goodfriend (1984).
22. For instance, see Lucas (1988) and Meltzer (1963) on the long-run stability of the demand for M1.
23. See Heller (1966) and Shultz and Aliber (1966).
24. See Kosters (1975).
25. Government fiscal concerns are the driving force behind high inflations. See Sargent (1986).
26. See Schreft (1990).
27. See Phillips (1958).
28. See Heller (1966) and Tobin (1972).
29. See Fischer (1994), pp. 267–68.
30. King and Watson (1994), for example, report a significant negative correlation between unemployment and inflation over the business cycle.
31. Barro and Gordon (1983), Fellner (1976), Sargent (1986), and Taylor (1982) contain early discussions of credibility as it relates to monetary policy. Persson and Tabellini (1994) contains a recent survey of research on the role of credibility in monetary and fiscal policy. The new large-scale Federal Reserve Board macroeconomic model is designed to take account of different degrees of credibility in policy simulations. See "A Guide to FRB/US" (1996).
32. What happens is this: In the first instance households and businesses attempt to exchange financial assets for money. Such actions, however, cannot satisfy the aggregate excess demand for money directly. They drive asset prices down and interest rates up until the interest sensitive components of aggregate expenditure grow slowly enough to eliminate the excess demand for money. As the disinflation gains credibility, wage and price inflation slows, and real aggregate demand rebounds until the higher unemployment is eliminated.

    Ball (1994) shows that a perfectly credible disinflation need have no adverse effects on employment even in a model with considerable contractual inertia in the price level.
33. The Fed did not explicitly assert its responsibility for inflation in the initial announcements of its disinflationary policy. However, by emphasizing the key role played by money growth in the inflation process, and by announcing a change in operating procedures to emphasize the control of money, the Fed *implicitly* acknowledged its responsibility for inflation. See *Federal Reserve Bulletin* (November 1979), pp. 830–32.
34. The Fed took short-term rates from around 11 percent in September 1979 to around 17 percent in April 1980. This was the most aggressive series of actions the Fed has ever taken in so short a time, although the roughly 5 percent increase in short rates from January to September of 1973 was almost as large. See Goodfriend (1993).
35. The collapse of confidence in U.S. monetary policy in 1979 and 1980 was extraordinary. The price of gold rose from around $275 per ounce in June 1979 to peak at about $850 per ounce in January 1980, and it averaged over $600 per ounce as late as November 1980. Evidence of a weakening economy caused the Fed to pause in its aggressive tightening in early 1980. But with short rates relatively steady, the 30-year rate jumped sharply by around 2 percentage points between December and February, signaling a huge jump in long-term inflation expectations. The collapse of confidence in early 1980 was caused in part by the ongoing oil price shock and the Soviet invasion of Afghanistan in December 1979. But the Fed's hesitation to proceed with its tightening at the first sign of a weakening economy probably also played a role. In any case, the Fed responded with an unprecedented 3 percentage point increase in short rates in March, taking them to around 17 percent. See Goodfriend (1993).
36. After making its disinflationary policy commitment in October 1979, the Fed let the growth of

effective M1 overshoot its target range in 1980 and the inflation rate continued to rise, peaking at over 10 percent in the fourth quarter. Then, in sharp contrast to the preceding four years, effective M1 actually undershot its target range in 1981. Effective M1 grew around 4.6 percentage points slower in 1981 than its average annual growth over the preceding five years. Further, the actual 2 percent shortfall in M1 from the midpoint of its 1981 target was built into the 1982 target path. See Broaddus and Goodfriend (1984).

The unemployment rate rose from around 6 percent in 1978 to average nearly 10 percent in the recession year of 1982.

37. Simple policy rule specifications studied with models estimated on historical data can be of great practical value in benchmarking actual policy decisions. McCallum (1988) and Taylor (1993) present two rules, respectively, that are particularly useful in this regard. McCallum models the monetary base (currency plus bank reserves) as the Fed's policy instrument, and has it responding to a moving average of base velocity and departures of nominal GDP from a target path. Taylor models the real short-term interest rate (the market interest rate minus expected inflation) as the policy instrument, and has it responding to inflation and the gap between actual and potential GDP.

    Each specification has advantages and disadvantages. Taylor's rule matches more closely the way the Fed thinks of itself as operating. But McCallum's rule makes clear that the ultimate power of the Fed over the economy derives from its monopoly on the monetary base. McCallum's rule has the advantage that it could still be used if disinflation happened to push the market short rate to zero, or if inflation expectations became excessively volatile. In either situation the Fed might be unable to use the real short rate as its policy instrument.
38. See Board of Governors "Monetary Policy Report to Congress" (1994, 1995, and 1996).
39. In 1995, Senator Connie Mack introduced a bill that would make low inflation the primary goal of monetary policy. In 1989, Fed Chairman Alan Greenspan testified in favor of a prior resolution that would have mandated a price stability objective for the Fed. Academics as diverse as Blinder (1995), Fischer (1994), and Friedman (1962) all agree that the Fed should be given some sort of mandate for low inflation. The remarkable convergence of professional thinking in favor of a mandate was evident at the Federal Reserve Bank of Kansas City's August 1996 conference on price stability. See *Achieving Price Stability* (1996).

    Inflation targeting is employed by a number of central banks around the world. See Leiderman and Svensson (1995).
40. Rudebusch and Wilcox (1994) report empirical evidence on inflation and productivity growth. Dotsey and Ireland (1996) study the question in a quantitative, theoretical model.
41. Oil prices rose from around $3 to $12 a barrel during the 1973/74 oil price shock, and from about $15 to over $35 in 1979/80.
42. See Thurow (1994). By successfully keeping inflation in check, preemptive policy actions *necessarily* appear to be busting ghosts. So the appearance of ghost busting is a consequence of good monetary policy.

REFERENCES

Achieving Price Stability. A Symposium Sponsored by the Federal Reserve Bank of Kansas City. Jackson Hole, Wyoming, 1996.

A Guide to FRB/US: A Macroeconomic Model of the United States. Manuscript. Macroeconomic and Quantitative Studies, Division of Research and Statistics, Federal Reserve Board: July 1996.

Ball, L. Credible Disinflation with Staggered Price-Setting. *American Economic Review,* vol. 84 (March 1994), pp. 282–89.

Barro, R. J., and Gordon, D. B. Rules, Discretion and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics,* vol. 12 (July 1983), pp. 101–21.

Blinder, A. Central Banking in Theory and Practice. The Marshall Lecture, Cambridge University, Federal Reserve Board, processed May 1995.

Board of Governors of the Federal Reserve System. Monetary Policy Report to Congress. *Federal Reserve Bulletin,* July 1994, February 1995, and August 1996.

Board of Governors of the Federal Reserve System. Announcements. *Federal Reserve Bulletin,* vol. 65 (November 1979), pp. 830–32.

Broaddus, A., and Goodfriend, M. Base Drift and the Longer Run Growth of M1: Experience from a Decade of Monetary Targeting. Federal Reserve Bank of Richmond *Economic Review,* vol. 70 (November/December 1984), pp. 3–14.

Bronfenbrenner, M., and Holzman, F. D. Survey of Inflation Theory. *American Economic Review,* vol. 53 (September 1963), pp. 593–661.

Cagan, P. *Determinants and Effects of Changes in the Stock of Money, 1875–1960.* New York: Columbia University Press, 1965.

Cook, T. Determinants of the Federal Funds Rate: 1979–1982. Federal Reserve Bank of Richmond *Economic Review,* vol. 75 (January/February 1989), pp. 3–19.

Dotsey, M., and Ireland, P. The Welfare Cost of Inflation in General Equilibrium. *Journal of Monetary Economics,* vol. 37 (February 1996), pp. 29–47.

Economic Report of the President. Washington: Government Printing Office, 1996.

Feldstein, M. The Costs and Benefits of Going from Low Inflation to Price Stability. *NBER* Working Paper 5469. Forthcoming in C. Romer and D. Romer, eds., *Monetary Policy and Inflation.* Chicago: University of Chicago Press, 1997.

Fellner, W. *Towards a Reconstruction of Macroeconomics: Problems of Theory and Policy.* Washington, D.C.: American Enterprise Institute for Public Policy Research, 1976.

Fischer, S. Modern Central Banking. In Forrest Capie, Stanley Fischer, Charles Goodhart, and Norbert Schnadt, eds. *The Future of Central Banking: The Tercentenary Symposium of the Bank of England.* Cambridge: Cambridge University Press, 1994.

Friedman, M. The Quantity Theory of Money. In John Eatwell, Peter Newman, and Murray Milgate, eds. *The New Palgrave Dictionary of Money and Finance.* New York: The Stockton Press, 1987.

Friedman, M. Monetary Policy for the 1980s. In John H. Moore, ed. *To Promote Prosperity: U.S. Domestic Policy in the Mid-1980s.* Stanford: Hoover Institution Press, 1984.

Friedman, M. *An Economist's Protest: Columns in Political Economy.* New Jersey: Thomas Horton and Company, 1972.

Friedman, M. Statement before U.S. Congress. House of Representatives, Committee on Banking and Currency in *The Federal Reserve System After Fifty Years.* Subcommittee on Domestic Finance. Hearings, 88 Cong. 2 Sess. Washington: Government Printing Office, 1964.

Friedman, M. Should There Be an Independent Monetary Authority? In Leland Yeager, ed. *In Search of a Monetary Constitution.* Cambridge, Mass.: Harvard University Press, 1962.

Friedman, M., and Schwartz, A. J. *A Monetary History of the United States, 1867–1960.* Princeton, N.J.: Princeton University Press, 1963.

Goodfriend, M. Interest Rate Policy and the Inflation Scare Problem: 1979–1992. Federal Reserve Bank of Richmond *Economic Quarterly,* vol. 79 (Winter 1993), pp. 1–24.

Goodfriend, M. Central Banking Under the Gold Standard. *Carnegie-Rochester Conference Series on Public Policy,* vol. 29 (Autumn 1988), pp. 85–124.

Goodfriend, M., and Whelpley, W. Federal Funds: Instrument of Federal Reserve Policy. Federal Reserve Bank of Richmond *Economic Review,* vol. 72 (September/October 1986), pp. 3–11.

Greenspan, A. Statement before the U.S. Congress, House of Representatives, Subcommittee on Domestic Monetary Policy of the Committee on Banking, Finance and Urban Affairs, *Zero Inflation.* Hearing, 101 Cong. 1 Sess. Washington: Government Printing Office, 1989.

Hawtrey, R. G. *A Century of Bank Rate.* London: Longran, 1938.

Heller, W. W. *New Dimensions of Political Economy.* Cambridge, Mass.: Harvard University Press, 1966.

Hetzel, R. L. The Rules versus Discretion Debate over Monetary Policy in the 1920s. Federal Reserve Bank of Richmond *Economic Review,* vol. 71 (November/December 1985), pp. 3–14.

King, R. G., and Watson, M. W. The Post-War U.S. Phillips Curve: A Revisionist Econometric History. *Carnegie-Rochester Conference Series on Public Policy,* vol. 41 (December 1994), pp. 157–219.

Kosters, M. H. *Controls and Inflation: The Economic Stabilization Program in Retrospect.* Washington, D.C.: American Enterprise Institute for Public Policy Research, 1975.

Lacker, J. M. Stored Value Cards: Costly Private Substitutes for Government Currency. Federal Reserve Bank of Richmond *Economic Quarterly,* vol. 82 (Summer 1996), pp. 1–25.

Leiderman, L., and Svensson, L. eds. *Inflation Targets.* London: Center for Economic Policy Research, 1995.

Lucas, R. E., Jr. On the Welfare Cost of Inflation. Center for Economic Policy Research, Stanford University, February 1994.

Lucas, R. E., Jr. Money Demand in the United States: A Quantitative Review. *Carnegie-Rochester Conference Series on Public Policy,* vol. 29 (Autumn 1988), pp. 137–68.

McCallum, B. T. Robustness Properties of a Rule for Monetary Policy. *Carnegie-Rochester Conference Series on Public Policy,* vol. 29 (Autumn 1988), pp. 173–204.

McCallum, B. T. and Goodfriend, M. S. Demand for Money: Theoretical Analysis. In *The New Palgrave: A Dictionary of Economics.* London: Macmillan Press, 1987. Reprinted in Federal Reserve Bank of Richmond *Economic Review,* vol. 74 (January/February 1988), pp. 16–24.

Meltzer, A. The Demand for Money: The Evidence from the Time Series. *Journal of Political Economy,* vol. 71 (June 1963), pp. 219–46.

Persson, T. and Tabellini, G. eds. *Monetary and Fiscal Policy, Vol. 1: Credibility.* Cambridge, Mass.: MIT Press, 1994.

Phillips, A. W. The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. *Economica,* vol. 25 (November 1958), pp. 283–99.

Poole, W. *Money and the Economy: A Monetarist View.* United States: Addison-Wesley Publishing Company, 1978.

Romer, C. D., and Romer, D. H. Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. *NBER Macroeconomics Annual,* vol. 4 (1989), pp. 121–70.

Rudebusch, G., and Wilcox, D. Productivity and Inflation: Evidence and Interpretations. Federal Reserve Board, Washington, D.C., 1994.

Sargent, T. J. *Rational Expectations and Inflation.* New York: Harper and Row, 1986.

Schreft, S. L. Credit Controls: 1980. Federal Reserve Bank of Richmond *Economic Review,* vol. 76 (November/December 1990), pp. 25–55.

Schultz, G. P., and Aliber, R. Z. eds. *Guidelines: Informal Controls and the Market Place.* Chicago: University of Chicago Press, 1966.

Taylor, J. B. Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy,* vol. 39 (December 1993), pp. 195–214.

Taylor, J. B. Establishing Credibility: A Rational Expectations Viewpoint. *American Economic Review Papers and Proceedings,* vol. 72 (May 1982), pp. 81–85.

Thurow, L. C. The Fed Goes Ghostbusting. *The New York Times,* May 6, 1994.

Tobin, J. The Cruel Dilemma. In A. M. Okun, ed., *Problems of the Modern Economy: The Battle Against Unemployment.* Revised. New York: W. W. Norton and Company, Inc., 1972.

Wolman, A. Zero Inflation and the Friedman Rule: A Welfare Analysis. Federal Reserve Bank of Richmond *Economic Quarterly,* vol. 83, no. 4 (Fall 1997) 1–21.

30

Lessons on Monetary Policy from the 1980s

Benjamin M. Friedman
Professor of Economics, Harvard University, Cambridge, Massachusetts

The half-decade running from mid-1982 to mid-1987 was a pretty good era for U.S. monetary policy, as these things go. A sharp easing of policy, beginning some time around midyear 1982, helped set in motion a recovery from the most severe business downturn the United States had experienced since the 1930s—a downturn that tight monetary policy earlier on had deliberately brought about in order to slow the alarming acceleration of prices. The recovery that ensued developed into a sustained expansion that continued without interruption through the end of 1987, thereby setting a new record for the longest recorded business expansion in U.S. peacetime experience. A fiscal policy based on unprecedentedly large (and continually growing) structural budget deficits was a major factor underlying this record-length expansion, but at least monetary policy did not stand in the way. Just as importantly, the severe 1981–82 recession served its intended purpose of substantially restoring price stability, and even after five years of expansion inflation remained modest by recent historical standards. At least as judged by these outcomes for the standard objectives of macroeconomic policy, U.S. monetary policy was a distinct success during these years.

Economists hoping to say something useful about monetary policy have had a tougher time. The quantitative relationships connecting income and price movements to the growth of familiar monetary aggregates, including especially the *M*1 measure of the money stock that had been the chief focus of monetary policy during 1979–82, utterly fell apart during this period.[1] Moreover, the collapse of these long-standing empirical regularities was not merely a matter of larger than usual quarter-to-quarter or year-to-year variances around longer-run benchmarks that otherwise continued to be reliable. Double-digit *M*1 growth, sustained on average over fully five years, repeatedly led prominent economists who had relied on these relationships in the past to offer widely publicized warnings of an imminent re-acceleration of prices. Yet the inflation rate fell dramatically, and then remained low. The presumption that "inflation is always and everywhere a monetary phenomenon" became progressively less compelling as a substantive rather than merely tautological description of the determination of prices.

Reprinted from: *Journal of Economic Perspectives,* v. 2, n. 3 (Summer 1988) 51–72.

Economists who preferred to think about monetary policy in different terms had no more success in fitting the major developments of this period into some alternative conceptual framework. Relationships connecting income and prices to the monetary base, or to measures of credit, fell apart just as visibly as did those centered on *M*1. Although real interest rates declined somewhat from the record levels posted in 1981 and early 1982, they nevertheless remained historically high throughout the next five years. Yet the economy's rate of growth in after-inflation dollars was about average for post war business cycle expansions, and in this case the expansion continued on for five years without even a single negative quarter.

In the meanwhile, the rapidly changing structure of the U.S. international economic position brought importance of a whole new magnitude to complications that most economists addressing questions about U.S. monetary policy had acknowledged often enough in the past, but nonetheless had usually ignored in substance. The U.S. current account balance, which had frequently changed sign since World War II but had always remained trivially small in comparison to the economy's size, suddenly became large enough to matter in a domestic macroeconomic context. The dollar exchange rate therefore emerged as a primary focus of U.S. macroeconomic policy. In addition, just within these few years the steady string of mounting current account deficits transformed the United States from the world's largest creditor country to the largest debtor. The exchange rate therefore acquired a new importance in the U.S. financial markets as well.

In the eyes of many economists, the Federal Reserve System has been steering without a rudder ever since it effectively abandoned its commitment to monetary growth targets in 1982. The visible success of monetary policy during the past half-decade is therefore all the more puzzling. In fact, over the course of this period the Federal Reserve's conduct of monetary policy appears to have centered ever more closely on controlling short-term interest rates. Whether this development really means that U.S. monetary policy has now returned to the conceptual basis of a quarter-century ago, when short-term nominal interest rates (or their equivalent) were practically the only focus of the policy process, remains unclear. At the very least, it raises the issue of whether the blatant flaws that crippled this policy strategy in the past were inherent and unavoidable. The other side of the same question is what will happen if the Federal Reserve continues to pursue what is basically an interest rate strategy, if and when inflation again becomes a major problem.

THE 1979–1982 EXPERIMENT AND ITS AFTERMATH

The basic facts describing the conduct of U.S. monetary policy in recent years are well known, although their interpretation has been the subject of much disagreement. In October 1979, the Federal Reserve System announced that it was adopting a new policy strategy placing primary emphasis on reducing over time the growth of the money stock. Further, to gain better control over money growth the Federal Reserve would be implementing new operating procedures, based in the first instance on the stock of nonborrowed reserves rather than the federal funds rate or some other short-term interest rate. The bands within which the federal funds rate would be free to fluctuate in the interval between meetings of the Federal Open Market Committee were accordingly widened by more than an order of magnitude. Several more narrowly technical measures, designed in principle to enhance control over money growth, were also part of the overall package.

The immediate motivation underlying this dramatic move was the rapidly deteriorating inflation situation, together with growing concerns about the dollar exchange rate. The overall U.S. inflation rate, as measured by the GNP deflator, had risen from a post-recession low of 6.4 percent in 1976 to 8.9 percent in 1979. The comparable rise measured by consumer prices was from 4.8 percent to 13.3 percent, the highest rate since 1946. At the same time, dollar exchange rates declined between 1976 and 1979 by 16.4 percent on a trade-weighted basis (14.5 percent after correcting for differential inflation). Especially once the OPEC cartel announced yet another major hike in crude petroleum prices, fears of an uncontrollable inflation spiral or a precipitous decline in the dollar, or both, began to spread.

Although the Federal Open Market Committee had at least mentioned money growth targets in its formal policy directives ever since 1970, and had formally reported money growth targets to Congress since 1975, in practice there was little correspondence between the stated targets and actual money growth. Evidence from the 1970s shows that the Federal Reserve did systematically adjust the federal funds rate in the direction required to offset deviations of actual money growth from the targets, but that the magnitude of these adjustments was far too small to be effective for plausible estimates of the interest elasticity of money demand.[2] Perhaps more importantly, the Committee typically did not set the coming year's targets so as to make up for past deviations, but instead let bygones be bygones and so treated each year as independent of the past. As the years rolled on, the difference between actual and targeted money growth was usually positive, and the average difference was positive. As a result, the actual rate of *M*1 growth gradually drifted upward, from 4.9 percent per annum during 1965–69 to 6.1 percent per annum during 1970–74, and then 7.1 percent per annum during 1975–79 (although the fastest growth for any year during this period, 8.4 percent, occurred during 1972).

The effects of the new combination of policy strategy and policy tactics implemented in October 1979 were immediately visible, and they continued to be so for the next several years, although in some aspects they ran counter to the new policy's declared intent. At the most basic level, the Federal Reserve did carry through on its commitment to contain, and then reduce, the growth rate of the *M*1 money stock. The rate of *M*1 growth, measured from the fourth quarter of one year to the fourth quarter of the next (the same basis used for officially reporting money growth targets to Congress) had been 8.2 percent in 1978. The "gradualist" objective of reducing the money growth rate by one percent each year, until it reached a level consistent with price stability, would imply targets of roughly 7 percent for 1979, 6 percent for 1980, 5 percent for 1981, and so on.[3] Although *M*1 had grown at a 9.0 percent per annum rate in the first three quarters of 1979, with the sharp policy shift in the final quarter the growth for 1979 as a whole was 7.9 percent. A continuation of the new policy delivered 7.3 percent *M*1 growth in 1980, and 5.1 percent in 1981 (see Figure 1). For this three-year period viewed as a whole, the new policy did manage to achieve results roughly consistent with the objective of reducing the money growth rate by one percent per year.

Not surprisingly, delivering on this objective involved interest rates that were both higher on average and also more volatile.[4] Nominal interest rates immediately rose to record highs, and then declined sharply as the economy entered a recession and the Federal Reserve Board also imposed credit controls as authorized by President Carter under the Credit Control Act of 1969.[5] Nominal interest rates then rose to yet new record highs in 1981 after business began to recover and credit controls were no longer in effect. Interest rates were high in real terms as well. The difference between the three-month Treasury bill

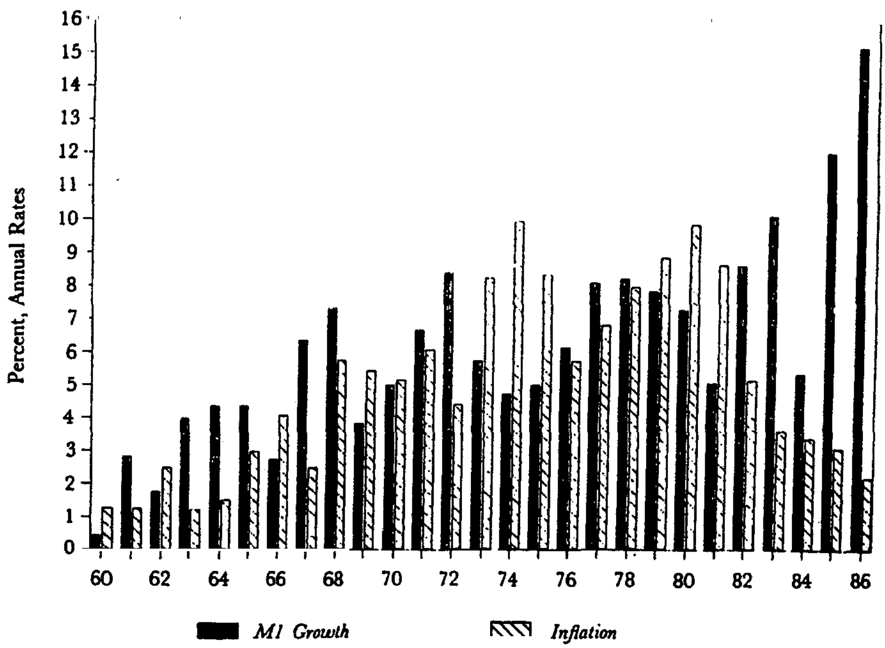


Figure 1 Money growth and price inflation.

rate and the next quarter's inflation rate, which had averaged close to zero during the post World War II period up until then, fluctuated in the 4–8 percent per annum rage throughout 1981 and 1982. Short-run volatility of interest rates—month-to-month, day-to-day, and even within the trading day—increased by what for most measures were large multiples.

The most striking respect in which the results of the new monetary policy did not correspond to its declared intent was that the short-run volatility of money growth increased sharply too. The Federal Reserve's success in gradually reducing the yearly *M*1 growth rate—8.2 percent in 1978, 7.9 percent in 1979, 7.3 percent in 1980, 5.1 percent in 1981—did not carry over to the quarterly growth rate, which varied in the immediate post-1979 period from a high of 21.1 percent per annum in 1980:*Q*3 to a low of *minus* 5.8 percent per annum in 1980:*Q*2, nor to any finer time calibration. The standard deviation of *M*1 growth from one quarter to the next jumped from 2.37 percent per annum during 1970:*Q*1–1979:*Q*3 to 6.64 percent per annum during 1979:*Q*4–1982:*Q*2 (although omitting the two back-to-back extreme quarters during the 1980 credit control episode, the increase was only to 3.00 percent per annum). Although the Federal Reserve had never formally acknowledged any commitment to stabilize money growth on a within-year basis, critics of monetary policy, both at the time and subsequently, focused on this increase in short-run volatility as casting doubt on the strength of the central bank's commitment to money growth targets in a broader perspective.

The monetary policy experiment that commenced in October 1979 ended almost as abruptly as it had begun. By the summer of 1982 the slowdown in U.S. business activity had developed into what was, by many measures, the most severe recession since the great depression of the 1930s. Unemployed labor, idle industrial capacity and business bankruptcies were all at post war record levels. At the same time, bank failures had also reached record levels, and there were increasing signs of fragility throughout the financial structure. In August 1982 an impasse over Mexico's ability to meet its dollar-denominated external obligations called widespread attention for the first item to the debt problems not just of Mexico but of developing countries throughout Latin America and sub-Saharan Africa.

Yet the demand for money balances continued to be strong. Following the 5.1 percent growth in 1981, the Federal Reserve had set a 2.5–5.5 percent target range for *M*1 growth in 1982. The revised data now available show that actual *M*1 growth was within this range (4.6 percent per annum) during the first half of 1982; but the data available at that time showed *M*1 growth running consistently above the stated target range, despite continuing high interest rates and contracting real economic activity. In August the Federal Reserve allowed short-term interest rates to drop by almost three percentage points notwithstanding the emergence of still more rapid money growth. In October, within less than a week of the third anniversary of the announcement that had proclaimed the new dedication to money growth targets, Federal Reserve Chairman Paul Volcker publicly acknowledged that the *M*1 growth target was no longer in effect. With a further speed-up in the fourth quarter, *M*1 growth for the year 1982 came to 8.6 percent—above the target range, above the 1978 growth rate from which the program of cutting back on money growth had originally begun, and even above the postwar record growth rate set back in 1972.

The experience of the next five years largely continued the course set in the latter half of 1982. The yearly *M*1 growth rate was in double digits in each of 1983, 1985 and 1986 (see again Figure 1). For the five-year span ending at mid-1987, the average *M*1 growth was 10.8 percent per annum. Only in 1984 did actual money growth fall within the stated target range. In both 1983 and 1985 the Federal Reserve officially changed the target range at midyear, once the difference between actual and targeted *M*1 growth became obvious. In 1986 the Federal Reserve suspended its *M*1 target range without setting a new one. In 1987 the Federal Reserve simply set no target range for *M*1 growth at all.

Throughout the post-1982 period the Federal Reserve's official pronouncements continued to emphasize targets for broader monetary aggregates in place of *M*1, but it is not clear to what extent these measures genuinely guided monetary policy. Actual *M*2 growth did fall within the stated target range in every year between 1983 and 1986.[6] Actual *M*3 growth exceeded the stated target range in 1983 and 1984, but fell within it in 1985 and 1986. In 1987, until the October drop in stock prices, the Federal Reserve either instigated or accepted (depending upon one's perspective) rising market interest rates, and also raised the discount rate, despite the fact that both *M*2 growth and *M*3 growth were falling short of the stated target range; and the decline of market interest rates immediately after the stock market crash appeared to have little to do with money growth patterns in any direct way.

By contrast, short-term nominal interest rates since 1982 have resumed the smooth pattern characteristic of the pre-1979 era, thereby suggesting a renewed role for interest rates—as before—at the center of the monetary policymaking process. The standard deviation of the month-to-month change in the three-month U.S. Treasury bill rate, for example, had risen from .42 percent per annum between January 1970 and September 1979, to 1.54 percent per annum between October 1979 and September 1982. It dropped back to .32 percent per annum between October 1982 and June 1987. The standard deviation of the month-to-month change in the federal funds rate rose from .50 percent to 1.92 percent, and then fell to .38 percent, over the same three periods.

Although some of this return to interest rate smoothness on a month-to-month basis may simply have reflected the continuity of the business expansion and the absence of a reacceleration of prices during this period, the pronounced stability of short-term interest rates over a substantial span of time more likely indicates the return to a monetary policy approach based on closely controlling interest rate movements. At the least, it presents a strong contrast to the behavior that the Federal Reserve has accepted for the monetary aggregates.

COLLAPSE OF THE MONEY-INCOME AND MONEY-PRICE RELATIONSHIPS

What makes this unusual record of monetary policy actions look so successful in retrospect is that the bizarre behavior of money growth in no way corresponded to the behavior of income or prices. The familiar relationships that had characterized prior experience simply disappeared.

Instabilities in the money-income relationship—or, in more sophisticated forms, the money demand function—had actually begun to become more pronounced as early as the mid-1970s, and their appearance had already spawned a substantial new body of empirical literature even before the new monetary policy experiment commenced in October 1979.[7] By 1980 the Federal Reserve System had already adopted a whole new set of definitions of the monetary aggregates, designed in part to overcome just such difficulties. A survey paper bearing the suggestive title "The Search for a Stable Money Demand Function: A Survey of the Post-1973 Literature," and including more than eighty references, was already in print in the *Journal of Economic Literature* before Paul Volcker acknowledged in October 1982 that the Federal Reserve was suspending its $M1$ growth target (Judd and Scadding, 1982).

As Figure 2 makes clear, however, the instability that generated so much concern and research in the pre-1982 period was small stuff in comparison to what followed. The figure plots the ratio of the $M1$ money stock to GNP for each quarter since the start of the rede-

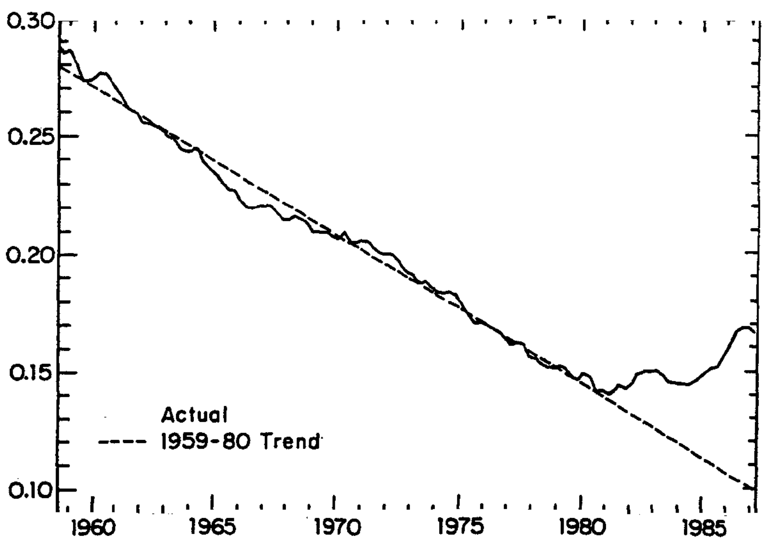


Figure 2 Ratio of money to GNP. *Note:* The series plotted is the ratio of $M1$, as a quarterly average of monthly data (source: Board of Governors of the Federal Reserve System) to quarterly GNP at annual rates (source: U.S. Department of Commerce). Both series are seasonally adjusted.

fined $M1$ series in 1959:$Q1$. Through the end of 1980, the $M1$-to-GNP ratio displayed the familiar downward trend of roughly 3 percent per annum that most students of the money-income relationship had come to see as inevitable in the post-war period, with a standard deviation around this trend of only .0044 (in comparison to a 1980:$Q4$ value of .1466). After 1980 the $M1$-to-GNP ratio not only experienced wider fluctuations but even reversed course. A simple extrapolation of the 1959–80 trend implies a ratio of .1007 by 1987:$Q2$ (the last quarter plotted). The actual value in 1987:$Q2$ was .1686, different from the trend extrapolation by more than 15 times the 1959–80 standard deviation.

Discussion of this phenomenon at the popular level has typically offered as an explanation the fact that "velocity" has declined. Because the so-called income velocity of money is nothing other than the ratio of GNP to money (the reciprocal of the ratio plotted in Figure 2), however, such explanations are completely empty of content. Given the definition of "velocity" in this context, the fact that velocity declined is simply identical to the fact that money grew rapidly while income did not. Saying that money growth outpaced income growth because velocity declined is like saying that the sun rose because it was morning.[8]

The mere fact of instability in the simple money-income ratio need not, of course, imply instability in more fully specified behavioral representations of the money-income relationship. The impression that stands out on a first glance at Figure 2 is representative of the results that researchers employing a variety of statistical strategies have found, however. A standard Goldfeld-type money demand function, estimated for quarterly data spanning 1952:$Q3$–1979:$Q3$, indicates a standard error of .42 percent. Extending the sample to 1986:$Q4$ raises the standard error to .61 percent. Deleting the earlier data, so that the sample is 1974:$Q2$–1986:$Q4$, further raises the standard error to .84 percent. Dynamic out-of-sample simulations of such equations deliver cumulative errors with root mean squares in the range of 4–8 percent for different parts of the post-1982 period, in comparison with 0.5–1.5 percent for different parts of the pre-1974 period. Attempts to do better with alternative specifications have met at best only very limited success.[9]

The story is approximately the same for efforts to investigate the money-income relationship from the perspective of determining income rather than money. A "St. Louis" type equation relating nominal GNP to four-quarter lags on both $M1$ and high-employment government expenditures, estimated in logarithmic differences for quarterly data spanning 1960:$Q1$–1979:$Q3$, indicates an adjusted coefficient of determination ($\bar{R}^2$) of .32. Extending the sample to 1986:$Q4$ reduces the $\bar{R}^2$ to .11. Deleting the earlier data, so that the sample is 1970:$Q3$–1986:$Q4$, further reduces the $\bar{R}^2$ to just .02. More sophisticated autoregression methods testing for a significant role of money in "causing" either nominal or real income, in the sense of accounting for income fluctuations not already accounted for by prior fluctuations in income itself, have produced results that are sufficiently varied to generate more skepticism than confidence in any strong conclusion on the subject, either positive or negative.[10]

Finally, in considering the money-*price* relationship it is even necessary to be on guard against results that are strongly statistically significant but with the wrong sign to make any sense in economic terms. The double-digit average growth rate maintained for five years following mid-1982 represents the most rapid sustained money growth the United States has experienced since World War II, yet these same years also saw the strongest sustained deceleration of prices in the postwar period (see again Figure 1). Price inflation as measured by the GNP deflator peaked at 9.7 percent in 1981 and declined in each of the next five years, reaching 2.6 percent in 1986. The rate of increase of consumer prices peaked at 13.3 percent in 1979 and declined in all but one of the next seven years, reaching 1.1 percent in 1986. Given the role that high-variance observations play in dominating results based on the least-squares methodology, as of the late 1980s it is necessary to take care not to find results indicating that *faster* money growth implies *slower* inflation.

Two examples, both drawn from the same paper, readily illustrate the pitfalls that confronted anyone who continued to rely closely on straightforward money-income and money-price relationships during this period. First, Figure 3 is an expansion, both backward in time and forward, of a figure included in a paper by Milton Friedman (1984) in the *American Economic Review*. The figure plots the respective annualized quarter-to-quarter growth rates of GNP and, with a one-quarter lag, *M*1. The figure covers 1960:*Q*1–1987:*Q*2 and distinguishes three time intervals. The middle one, 1979:*Q*4–1983:*Q*4, is identical to that plotted by Friedman. It spans the period from the October 1979 inception of the new monetary policy experiment through what was presumably the most recent observation available as of his time of writing.

After pointing out that the correlation between these two series during 1979:*Q*4–1983:*Q*4 was .46, or .71 after eliminating the two quarters affected by the credit control episode, Friedman wrote, "Two things are notable about the relation between money and income in these years: first, the lag is both shorter on the average and less variable than in earlier years, second, the relation is unusually close. I believe that both are a consequence of the exceptionally large fluctuations in *M*1 growth. The effect was to enhance the importance of the monetary changes relative to the numerous other factors affecting nominal income and thereby to speed up and render more consistent the reaction."

Table 1 summarizes the record of the GNP-to-lagged-*M*1 growth correlation and the variability of *M*1 growth for the three intervals shown in Figure 3. Money growth on a quar-

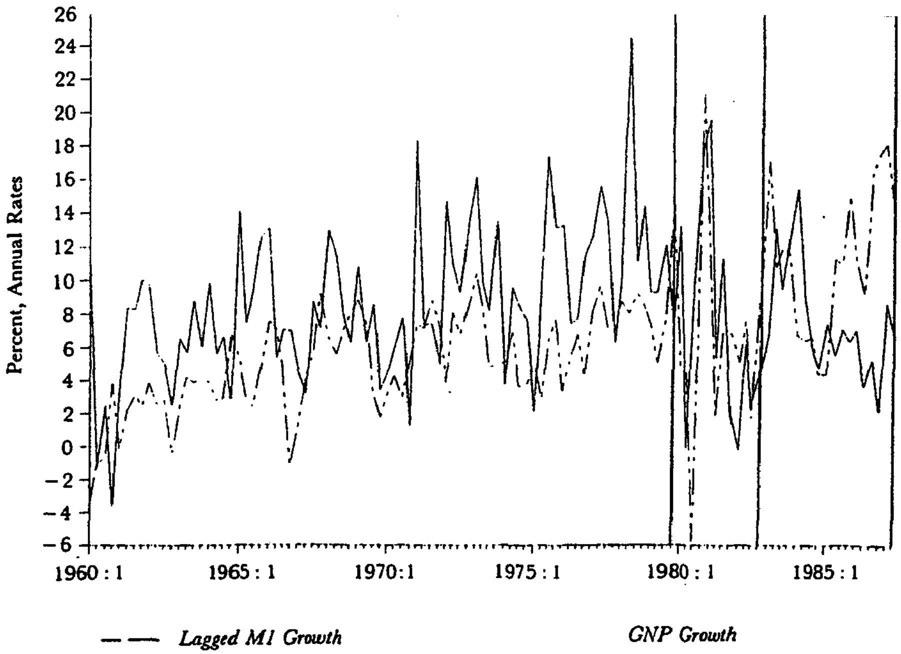


Figure 3 Growth rates of GNP and lagged money.

Table 1 Money Growth Volatility and the GNP-to-Lagged-*M*1 Correlation, 1960–1987

Sample	Standard deviation of M1 growth	Correlation between GNP growth and lagged M1 growth
1960: *Q*1–1979: *Q*3	2.87%	.47
1979: *Q*4–1983: *Q*4	6.16	.45
1979: *Q*4–1980: *Q*1, 1980: *Q*4–1983: *Q*4	4.18	.47
1984: *Q*1–1987: *Q*2	4.80	−.10

Data are seasonally adjusted at annual rates. Money data are quarterly averages.

ter-to-quarter basis (as used by Friedman in his paper) was certainly more variable during 1979:*Q*4–1983:*Q*4 than it had been during the prior two decades. The GNP-to-lagged-*M*1 correlation was not "unusually close" during 1979:*Q*4–1983:*Q*4 compared to the past, however. The correlation of .45 computed over these eleven quarters (Friedman reported .46) is essentially identical to that for the previous 79 quarters. Excluding 1980:*Q*2 and 1980:*Q*3 reduces the variability of money growth, but does not materially affect the GNP-to-lagged money correlation. (Subsequent data revisions have reduced the .71 correlation reported by Friedman to .47 as shown in Table 1—identical to the correlation for the earlier period.)

More importantly, what stands out in both Table 1 and Figure 3 is the changes that occurred after 1983. Although the variability of money growth remained high, the positive GNP-to-lagged-*M*1 correlation disappeared entirely. In its place is a small *negative* correlation.

Table 2, focusing on the money-price relationship, is simply an updated version of a table that Friedman presented in the same 1984 paper. The horizontal line in each column indicates entries not included in the original version. In describing the data shown above

Table 2 Rates of Change in Money and in Inflation Eight Quarters Later

	Annual rate of change over eight quarters		
Period for money	M1	Deflator eight quarters later	Period for deflator
1971: *Q*3–1973: *Q*3	6.9%	9.5%	1973: *Q*3–1975: *Q*3
1973: *Q*3–1975: *Q*3	5.2	6.3	1975: *Q*3–1977: *Q*3
1975: *Q*3–1977: *Q*3	6.4	8.3	1977: *Q*3–1979: *Q*3
1977: *Q*3–1979: *Q*3	8.6	9.4	1979: *Q*3–1981: *Q*3
1979: *Q*3–1981: *Q*3	6.1	4.8	1981: *Q*3–1983: *Q*3
1981: *Q*3–1983: *Q*3	9.2	3.3	1983: *Q*3–1985: *Q*3
1983: *Q*3–1985: *Q*3	8.1	2.8	1985: *Q*3–1987: *Q*3

Data are seasonally adjusted.
Source: Friedman (1984), updated. (The entries above the lines differ from Friedman's becauseof subsequent data revisions, but the differences are slight.)

the two lines, Friedman wrote, "The long-period evidence suggests that inflation has much inertia and that the lag between money and inflation is of the order of two years. Table [2] shows that this relation has held in recent years as well. There is a one-to-one relation between movements in monetary growth, and in the GNP deflator two years later over successive two-year periods since 1971. . . . The increased rate of monetary growth in the 1981–83 biennium suggests that we have passed the trough in inflation and that inflation will be decidedly higher from 1983 to 1985 than it was from 1981 to 1983."[11]

As the below-the-line entries in Table 2 show, quite the opposite happened. Growth of *M*1 during 1981:*Q*3–1983:*Q*3 was the fastest for any of the six biennia in Friedman's sample, but inflation in 1983:*Q*3–1985:*Q*3 turned out to be the lowest. Rapid money growth continued in 1983:*Q*3–1985:*Q*3, but inflation slowed still further in 1985:*Q*3–1987:*Q*3. The simple correlation between the two time series shown, calculated for the first five observations only, is .70. Calculated for all seven observations, the correlation is *minus* .23.

OTHER MONEY AND CREDIT AGGREGATES

The breakdown of long-standing relationships to income and prices has not been confined to the *M*1 money measure. Neither *M*2 nor *M*3, nor the monetary base, nor the total debt of domestic nonfinancial borrowers has displayed a consistent relationship to nominal income growth or to inflation during this period. On a quarter-to-quarter basis, standard relationships like Goldfeld-type equations fitting movements in these aggregates to movements of income and interest rates, or St. Louis-type equations fitting movements of nominal income to movements of an aggregate and a measure of fiscal policy, showed pronounced deterioration for each of these aggregates. (The *largest* $\bar{R}^2$ for any of these St. Louis equations, estimated for quarterly data spanning 1970:*Q*3–1986:*Q*4, is .09.) On a longer-term basis, the average growth rate for each of these aggregates during the half-decade from mid-1982 to mid-1987 was in excess of any prior postwar experience; yet inflation lessened substantially, and the average growth of nominal income was hardly extraordinary. By mid-1987 the ratio of each aggregate to nominal GNP was above the level implied by an extrapolation of the corresponding pre-1980 trend by an amount ranging from three standard deviations (for *M*2) to twenty-three (for the credit aggregate), based on the pre-1980 variability. It is difficult to imagine how anyone could have successfully predicted the behavior of either income or prices during this period on the basis of foreknowledge of the path of any—or, for that matter, all—of these aggregates.

The manifest failure of the credit aggregate to perform satisfactorily in this context perhaps merits a special comment. I had earlier advocated the use of a broad credit aggregate in conjunction with one or more monetary aggregates in formulating U.S. monetary policy, precisely on the ground that credit might provide some safeguard against false signals given by the monetary aggregates under conditions of instability affecting the public's demand for money (see, for example, Friedman, 1982, 1983). Because credit is a measure of activity on the liability side of the public's balance sheet, while the monetary aggregates are various measures of the non-bank public's assets, it seemed (and to me still seems) reasonable to think that expanding the information base explicitly underlying the monetary policy process to encompass both money and credit measures would provide potentially useful diversification in the context of portfolio behavior that is at best imperfectly understood, and inevitably subject to a multiplicity of shocks. In addition, empirical investiga-

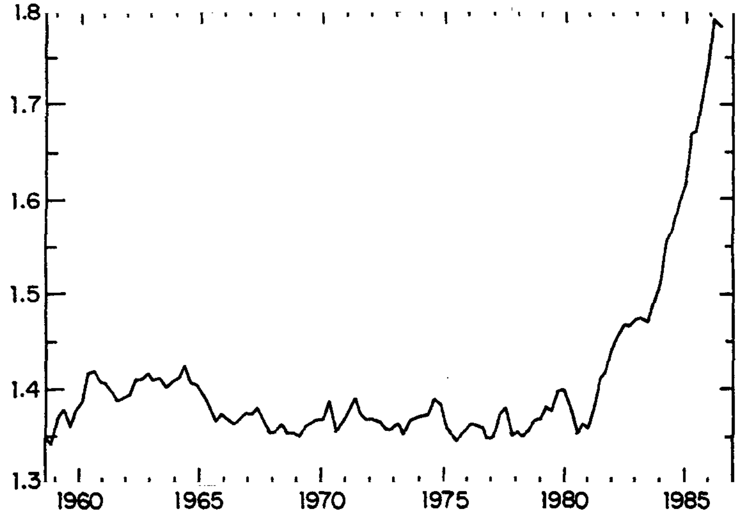


Figure 4 Ratio of credit to GNP. *Note:* The data are analogous to those in Figure 2 except that the numerator is end of quarter outstanding debt of domestic nonfinancial borrowers. *Source:* Board of Governors of the Federal Reserve System.

tions relying on a variety of statistical methods indicated little basis for concluding that the total debt of movements of income or prices than was any of the standard monetary aggregates.[12]

In any event, the movement of credit during the post-1982 period bore no more relation to income or prices than did that of any of the monetary aggregates. Worse still, the false signals provided by the growth of credit were in the same direction as those provided by the growth of money. Figure 4 plots the credit-to-GNP ratio for the same sample for which Figure 2 shows the $M1$-to-GNP ratio. After decades of trendless stability, the credit ratio began an unprecedented climb in 1982 which has not stopped as of the time of writing. Moreover, disaggregated data show that essentially all categories of domestic nonfinancial borrowers—including the federal government, state and local governments, individuals, and businesses—have played major roles in this extraordinary surge of indebtedness.[13] Anyone who had relied on prior credit-based relationships to predict the behavior of income or prices during this period would have made forecasts just as incorrect as those derived from money-based relationships. Anyone who had derived additional confidence in such predictions because the respective signals given by both money and credit confirmed one another would have found that confidence misplaced.

I have speculated elsewhere on the causes of the breakdown of the relationship between credit and income in the 1980s (Friedman, 1987). For purposes of the monetary policy issues under discussion here, it is sufficient to say that attempts to "fix up" this relationship in any simple way are likely to be no more successful than such efforts directed at parallel relationships for the monetary aggregates have been.

A RELATIONSHIP THAT DID HOLD UP

In sharp contrast to the collapse of relationships connecting the ultimate objectives of monetary policy to standard quantity measures of policy actions, the long-standing relationship between the two most prominent macroeconomic policy objectives—inflation and unemployment—remained intact during this period. The point is of some interest because one of the principal supposed merits widely claimed in favor of the use of publicly announced money growth targets for monetary policy was a potential lessening of the real costs of disinflation. The idea was that public knowledge of such targets would affect expectations in such a way as to minimize (according to some models, to eliminate altogether) the usual negative impact of disinflationary monetary policy on employment, output, incomes, and profits.

What this idea should have led one to predict about the real costs of the major disinflation achieved in the United States in the 1980s is far from clear. For almost three years beginning in October 1979, the Federal Reserve did approximately carry through on its widely publicized intention to reduce the yearly rate of money growth by one percent per year—that is, as long as everyone understood that "money" meant *M1*. Because the familiar claim that most measures of money move roughly together over time turned out to be false, however, anyone who watched *M2* or *M3* instead of *M1* would have seen no evidence of monetary decelaration.[14] Moreover, on a within-year basis even *M1* growth became not more regular but less after October 1979 (see again Table 1). After mid-1982 there was no reason for anyone to find the Federal Reserve's commitment to its stated money growth targets credible.

Regardless of whether any part of the policy experience of the 1980s constituted a good test for an effect of preannounced money growth targets in reducing the real costs of disinflation, it is clear that no such reduction occurred. Ten years ago, Arthur Okun (1978) surveyed a variety of econometric estimates of these costs, none of which incorporated any expectations effects due to reliance on preannounced money growth targets. Okun's survey indicated that the cost of each percentage point reduction in inflation achieved by monetary policy would be between two and six "point-years" of unemployment, with a median estimate of three point-years.[15] Table 3 shows the annual rate of change of the GNP deflator and the annual average unemployment rate beginning in 1978. The table also shows, for years beginning in 1980, the cumulative excess of the unemployment rate above 6 percent (the approximate average for the two prior years and a standard "full employment" benchmark).

Table 3 Percentage Rates of Inflation and Unemployment, 1978–1987

	Inflation rate	Unemployment rate	Cumulative excess unemployment
1978	7.3%	6.1%	—
1979	8.9	5.8	—
1980	9.0	7.1	1.1%
1981	9.7	7.6	2.7
1982	6.4	9.7	6.4
1983	3.9	9.6	10.0
1984	3.7	7.5	11.5
1985	3.2	7.2	12.7
1986	2.6	7.0	13.7
1987	3.2	6.5	14.2

Data for 1987 based on first half year only.

The slowing of inflation from nearly 10 percent per annum in 1980–81 to roughly 3 percent per annum a half-decade later required approximately 14 point-years of unemployment—right at the lower end of the range implied by the econometric models Okun surveyed. Especially in light of the evidence suggesting that something like a third (and perhaps as much as a half) of this disinflation was a product of the 74 percent appreciation of the dollar's trade-weighted exchange rate between 1980 and early 1985—an appreciation which has been almost entirely reversed as of the time of writing—this result seems fully consistent with Okun's prediction.[16]

Everyone had always known that sufficiently tight monetary policy, maintained for a sufficiently long time, could halt even the most deeply rooted inflation. The reluctance to proceed in that fashion lay not in disbelief that such a policy would do its job, but in concern for the resulting real costs. What was new beginning in October 1979 was the willingness to bear those costs.

INTERNATIONAL COMPLICATIONS

The United States is not a small open economy. U.S. production and U.S. incomes bulk sufficiently large in comparison to the combined economic scale of all countries participating in the current free international trading system that independent forces originating in this country can and do affect economic conditions on a worldwide basis. The U.S. financial markets bulk sufficiently large in comparison to world capital markets to exert analogous effects.

The United States is an open economy, however, and the increasing practical importance of this fact is perhaps the most interesting single development of the 1980s from a monetary policy perspective. There are at least two major aspects of the U.S. economy's openness that now matter, far more than earlier on, for purposes of monetary policy. The more straightforward of the two is simply the demonstrated ability of U.S. exports and U.S. imports to differ by a margin large enough to affect the economy's aggregate output and growth.

The possibility of such an outcome had always been present, of course; but throughout the postwar era, until the 1980s, the U.S. current account was always approximately in balance over periods long enough to even out the ups and downs of the business cycle. The current account was in surplus, albeit by less than $1 billion per annum (0.2 percent of GNP), on average during the 1950s. In the 1960s the average annual surplus was $3 billion (0.4 percent of GNP). In the 1970s the current account was in deficit on average, but by less than $1 billion per year (not even 0.1 percent of GNP). In 1980 there was a $2 billion surplus, compared to overall GNP of $2.7 trillion. It is little wonder that many analysts of U.S. monetary policy during these decades practically ignored potential effects on real economic activity via exchange rate channels.

The 1980s have been starkly different. With a 74 percent appreciation of the dollar on a trade-weighted basis between the end of 1980 and early 1985 (64 percent after allowing for differential inflation rates), together with a host of other factors that may have further compounded the problem, the ability of many U.S. industries to compete in world markets all but collapsed. By 1986 the current account deficit was $141 billion, or 3.3 percent of GNP. On an overall basis, including the trade balance, real GNP in 1986 was up over the 1980 level by 16.5 percent. Excluding the trade balance—that is, focusing on domestic absorption rather than production—the increase over the six years was 23.3 percent. Not surprisingly, exchange rate effects on export and import flows have now emerged as a genuine focus of monetary policy concern, not just a subject for obligatory mention.

The more complex aspect of the newly enhanced importance of the U.S. economy's openness from a monetary policy perspective is an outgrowth of the first. A current account deficit means that a country is borrowing from abroad to finance the excess of its imports over its exports (where both are broadly defined)—or, equivalently, to finance the excess of its domestic investment and government deficit over its domestic saving. Since 1980 U.S. fiscal policy has delivered an unbroken string of unprecedentedly large federal government budget deficits, notwithstanding five years of business expansion beginning in 1983. At the same time, domestic saving has been below average despite the introduction of a variety of supposed saving incentives. As a result, the United States has been borrowing from abroad in record volume even though the share of GNP devoted to net investment in both business and residential capital has been unusually small. Especially since 1984, the United States has borrowed from abroad on a larger scale in relation to the economy's size than the country did at any of the four peak periods of reliance on foreign capital to finance its canals, then its railroads, then its steel and other industries, and then its public utilities, during the nineteenth century.[17]

This massive borrowing from abroad transformed the United States from the world's largest creditor country to its largest debtor in just four years. The U.S. net foreign asset position peaked at $141 billion at the end of 1981. It was still $136 billion at the end of 1982. The United States officially became a net debtor again (for the first time since before World War I) in early 1985. By the end of 1986 the U.S. net foreign debt was $265 billion—more than that of Mexico, Brazil and Argentina combined.[18] On a gross basis (that is, without any balancing of U.S. holdings abroad) foreign holdings in the United States as of year-end 1986 came to $1.3 trillion, including such readily liquefiable assets as $267 billion of U.S. Treasury securities, $149 billion of corporate bonds and other debt securities, $167 of corporate equities, and $477 billion of bank deposits.[19]

As a result of this surge in foreign holdings during the 1980s, foreign investors have assumed an importance in the U.S. financial markets that was previously unknown in modern times. Whether or not foreign investors will bid in volume on any given day's auction of U.S. Treasury securities is now often the paramount focus of attention among participants not just in the U.S. bond market but in the stock market too. Popular discussions frequently raise the fear that foreign investors might decide to "get out" of dollar holdings, with disastrous consequences for the dollar exchange rate as well as for the prices of dollar denominated securities. The notion that foreign investors in the aggregate can sell their dollar holdings to anyone but each other is mistaken, of course, but their attempts to do so can move both exchange rates and securities prices. In particular, given the size of foreign holdings, a widespread move to liquidate bond and equity portfolios as a preliminary to selling dollars (even if only to other foreign buyers) could easily move securities prices by a substantial margin. As a result, managing dollar exchange rates to prevent such occurrences has added a whole new dimension of U.S. monetary policy.

The contrast between the rigid hands-off attitude toward the foreign exchange markets that the Reagan administration espoused during its first four years in office, and the series of ad hoc international agreements that it has orchestrated beginning with the widely publicized meeting at New York's Plaza Hotel in September 1985, is just the most visible part of the increased importance of the exchange rate for the United States. Although there is little hard evidence as yet, it appears that the change has affected U.S. monetary policy in more ongoing, and more subtle, ways as well. For example, even at the simplest level it is no longer so obvious that an open market purchase which lowers short-term interest rates will necessarily lower long-term interest rates as well. If dollar exchange rates fall as a result of lower U.S. short-term rates, foreign investors may sell their long-term U.S. assets as a preliminary to reducing (again, not collectively) their overall dollar holdings, and domestic market participants may seek to do the same in anticipation of the action by foreign investors.[20]

If the net result of these developments amounts to an erosion of sovereignty over U.S. monetary policy, that loss is no more than the inevitable price of becoming a debtor country—and the world's largest, at that. Losing control over one's affairs is part of what a mounting debt level is all about, no less for a borrowing nation than for business or individual borrowers. The policy actions and public discussions that ensued after the October 1987 stock market crash dramatically illustrated the tension that can arise under such circumstances between a monetary policy designed to support the currency and a monetary policy designed to prop up the domestic economy. At least thus far the United States' situation in this regard is not yet so constraining as, for example, Britain's was during the 1950s and 1960s, when repeated ad hoc concessions were necessary to avert the eruption of a "sterling crisis." Even so, the problem of formulating and implementing U.S. monetary policy is now more complicated by one additional dimension.

VACUUM AT THE CENTER?

It is difficult to escape the conclusion that there is now a conceptual vacum at the center of the U.S. monetary policymaking process. The seemingly endless quest from various quarters to impose some kind of simple rule on the conduct of monetary policy is, at least for now, no longer a going concern. The interactions among money, income and prices during the 1980s—including the half-decade that followed the monetary targets experiment, even more than what happened during 1979–82—have undermined it both intellectually and practically. The claim that reliance on some kind of simple rule would open the way to costless disinflation is at best unproven, and for practical purposes doubtful. The notion that some kind of simple rule would adequately encompass the new complications due to the increased practical importance of the U.S. economy's openness is dubious at best. In retrospect, the contemptuous dismissal by some economists of the Federal Reserve's reluctance to adopt the monetary policy rules they had proposed—and the readily voiced assumption that that reluctance could have stemmed only from ignorance or a faulty set of objectives on the part of Federal Reserve officials, or perhaps even their self-aggrandizement at the public expense—stands as a sorry reminder that economists outside government can also make each of these errors.

The most powerful element of what has happened in this regard in the 1980s is the collapse of the longer-run relationship between money on the one side and income and prices on the other. In this decade the main event has been very different from the quarter-to-quarter or even year-to-year irregularity that was always the focus of debate about whether these relationships were stable before. Proponents of simple monetary policy rules in the past could and did claim that such rules failed to offset short-run economic fluctuations that policy could probably not hope to eliminate anyway, but that over longer horizons the anchor they provided would keep the economy on a steadier course than an alternative policy that attempted to achieve "fine tuning." With nominal GNP by mid-1987 more than 40 percent below the value implied by the long-run relationship to $M1$

which prevailed during 1959–80, the problem is no longer in the fine tuning but in the anchor itself.

The relevant issue here for policy purposes is not whether for any time interval there exists some abstract notion of "money" that conceptually bears a stable and reliable relation to income and prices, and that statisticians can seek to uncover after that interval ends, but whether policymakers can identify and measure that quantity substantially in advance of their need to base planning and operating decisions on it. Appeals to the tradition of the "quantity theory" are of no use in this context in the absence of a clear statement of what is the quantity and what is the theory. Especially in a world of institutions that increasingly blur the distinction between transactions balances and saving balances, being precise about either the theory or the quantity is ever more difficult. The continuing (indeed increasing) interest within the economics profession in some kind of constitutional constraint on monetary policy jars harshly against the likelihood of what such constraints suggested (but not enacted) in years past would have meant in the 1980s.

Nor is the problem merely one of money (or credit) "targets." The more flexible idea that such variables as money and credit, which are endogenous to the monetary policy process in the short run, should be used as "information variables" to guide the initial setting and ultimate readjustment of whatever genuinely exogenous instrument the central bank is using, has always seemed highly attractive—at least to me. The information-variable approach to monetary policy makes no sense, however, in the absence of a reasonably compact set of variables that reliably provide information about the macroeconomic outcomes monetary policy is seeking to affect. After the experience of the 1980s, it is difficult to foresee any ready consensus on what that compact set of variables should be. Subsequent experience may provide a new basis for such a consensus, and new data to substantiate it; but that prospect remains a matter for the future, if not later still.

In the meanwhile, the Federal Reserve System has not ceased operations. Nor should it be inclined to do so, in light of the performance of both income and prices during the past half-decade. Five years of fairly steady economic growth, with inflation consistently lower than at any time since before the Vietnam War, represents no small achievement by today's standards. In the world of practical affairs, it is difficult to argue with success.

Notwithstanding the Federal Reserve's continuing formulation of money growth targets that it reports to Congress, as current law requires, and even notwithstanding the relatively high success rate in meeting the target for *M*2, it seems clear enough that the Federal Reserve System since mid-1982 has centered its monetary policy actions primarily around controlling short-term nominal interest rates.[21] In so doing, Federal Reserve decision makers have no doubt taken account of the movements of money (and perhaps credit, too); but they have also taken account of many other potential information sources, including longer-term asset prices and yields, dollar exchange rates, and numerous aspects of nonfinancial economic activity. More to the point, they have apparently proceeded in the absence of any well-articulated conceptual framework linking the interest rate as the chief policy instrument to the main macroeconomic policy objectives, or linking the associated large and diverse information base to either the policy instrument or the policy objectives. Although procedures differ in various details, the overall approach is strongly reminiscent of the practice of the 1950s and 1960s.[22]

It is therefore useful to ask why the policy approach followed at that time failed. The voluminous investigation of this question, both at the time and subsequently, supported three general conclusions: First, Federal Reserve officials systematically confused the level of interest rates as the instrument of monetary policy with the level of interest rates as an ultimate objective of monetary policy. As a result, they usually delayed too long before raising or lowering interest rate levels, and even then made changes of insufficient magnitude. Second, with no nominal quantity at the center of the policy process, the overall approach lacked an anchor to provide price stability. Although inflation was not therefore inevitable, there was little protection against it when various inflationary pressures arose. Third, once inflation did emerge, Federal Reserve officials (and many other people too) often failed to distinguish nominal from real interest rates. As a result, they often associated higher observed interest rates with a tighter policy stance even when the increase in nominal interest rates merely kept pace with, or even fell short of, rising inflationary expectations.

Are these three flaws inherent in the approach to monetary policy that the Federal Reserve System followed a quarter-century ago, and that it has apparently been following again since mid-1982? Or is it possible to design and implement monetary policy along these lines, albeit in a way that has learned from the still relatively recent past? Were the familiar failures of monetary policy under this approach in the past inevitable? Or does the experience of the last half-decade show that this kind of monetary policy can work, and work well? Research on these questions may be the best contribution economists concerned with U.S. monetary policy can now make.

ACKNOWLEDGMENTS

I am grateful to Kenneth Kuttner for research assistance; to Stanley Fischer, William Poole, Joseph Stiglitz and Timothy Taylor for helpful comments on a previous draft; and to the National Science Foundation and the Harvard Program for Financial Research for research support.

NOTES

1. The *M*1 money stock consists of currency, checkable deposits (including both non-interest-bearing demand deposits and interest-bearing NOW accounts) and travelers checks; its value as of June 1987 was $747 billion. *M*2 consists of *M*1 plus a much larger quantity of savings-type accounts, including ordinary passbook accounts and certificates of deposit (in amounts up to $100,000), money market deposit accounts and money market mutual funds (both of which can have limited checking facilities), and overnight repurchase agreements and Eurodollars; its June 1987 value was $2.8 trillion. *M*3 consists of *M*2 plus institutionally oriented instruments like certificates of deposit in amounts over $10,000 and money market mutual funds used by institutional investors, as well as repurchase agreements and Eurodollars extending beyond overnight; its June 1987 value was $3.6 trillion.
2. See, for example, De Rosa and Stern (1977) and Lombra and Moran (1980).
3. The official *M*1 target range for 1979 was 5–8 percent.
4. Increased short-run volatility of short-term interest rates, as a result of no longer accommodating temporary disturbances affecting money demand, is a straightforward implication of Poole's (1970) analysis of the money growth target strategy for monetary policy. Whether long-term interest rates should be expected to be more or less volatile is a more complicated question, however, involving changing risk factors and expectations of future inflation and interest rates.
5. The 1969 legislation under which the Board acted was quite far reaching, empowering the Federal Reserve Board, whenever explicitly authorized by the President, to "prohibit or limit any extensions of credit under any circumstances the Board deems appropriate." In 1980 the Board

proceeded under this authority to impose special reserve-type requirements on increases in certain kinds of consumer credit by all lenders (including non-banks), on increases in deposits at money market mutual funds, and on increases in non-deposit liabilities at banks that were not members of the Federal Reserve System.

6. In 1983 the stated target range for *M2* growth covered only part of the year.
7. The standard reference to state first is the contrast between the findings in Goldfeld (1973) and Goldfeld (1976). The most widely read studies done at the time by the Federal Reserve's own staff include Enzler et al. (1976), Porter et al. (1979), and Simpson and Porter (1980).
8. I owe the analogy to William Bennett.
9. The specific results cited here are from Goldfeld and Sichel (forthcoming), which also provides an extensive survey. Roley (1985) also showed the results of experimenting with a wide variety of alternative specifications.
10. See, for example, the differing results reported in Friedman (1986), Eichenbaum and Singleton (1986), and Stock and Watson (1987).
11. Friedman made the some prediction more forcefully in writings directed at broader audiences. In a column in the September 26, 1983 issue of *Newsweek,* for example, Friedman wrote, "Inflation has not yet accelerated. That will come next year, since it generally takes about two years for monetary acceleration to work its way through to inflation . . . The monetary explosion from July 1982 to July 1983 leaves no satisfactory way out of our present situation . . . The result is bound to be renewed stagflation—recession accompanied by *rising* inflation and high interest rates." A lengthy interview in the March 19, 1984, issue of *Fortune* indicated that Friedman ". . . also sees a strong possibility that by the end of [1984] inflation could reach an annual rate as high as 9%."
12. There are several obvious problems with attempting to measure the relevant concept of credit in this way. One is simply that the available data measure long-term debts at par value rather than at market prices (or some equivalent for nonmarketable debts). Another is that, although the category of "nonfinancial" borrowers excludes any entity explicitly set up as a financial intermediary, there is inevitably some degree of double-counting due to what amounts to financial intermediation carried out by ordinary businesses and even individuals. Whether this problem is more or less severe than comparable problems affecting the monetary aggregates—for example, the apparently widespread use of U.S. currency in black markets around the world, or even in the United States for a variety of purposes not related to familiar theories of demand for money—is an empirical question.
13. The one exception is the farm sector.
14. For the five years 1978–82, the simple correlations among the fourth-quarter-over-fourth-quarter growth rates of the major M's were each *negative:* –.53 between *M1* and *M2*, –.57 between *M1* and *M3*, and –.12 between *M2* and *M3*.
15. A "point year" of unemployment is one percentage point of unemployment in excess of the rate that corresponds to "full employment," maintained for one year. Some writers—for example, Fischer (1985)—have focused on real output rather than unemployment, and have argued on that basis that the post-1980 disinflation involved smaller costs than Okun's survey implied. The focus of the evidence that Okun surveyed was the inflation-unemployment relationship, however. His translation of the cost estimate into foregone real output simply relied on the usual three-for-one "Okun's Law" relation, which has not held up during the 1980s.
16. See, for example, Sachs (1985) for an analysis of the importance of the dollar's appreciation in the U.S. disinflation.
17. See Solomon (1986) for a review of the nineteenth century experience.
18. The official accounts include numerous obvious mismeasurements, but there is no ground for claiming that their sum is very different from zero. The largest adjustments in favor of the United States in a set of true accounts would be the revaluation of U.S. gold stocks, and of the net of U.S. direct investment abroad and foreign direct investment in the United States, to current market values. The largest adjustments against the United States would be the revaluation

to market value of U.S. banks' loans to developing countries, and allowance for the accumulation over time of each year's "errors and omissions" flow.

19. These data are from Scholl (1987), Table 1.
20. The effect is analogous to what is sometimes claimed along the lines that open market purchases would lead investors to sell long-term bonds out of fear that the resulting increase in money growth would bring higher inflation. Market experience in the United States has not borne out this earlier line of reasoning, but there appears to be more evidence to support the effect operating via exchange rates.
21. See Wallich (1984) for a description in different but equivalent terms.
22. See, for example, the descriptions given by Brunner and Meltzer (1964) and Guttentag (1966).

REFERENCES

Brunner, K. and Meltzer, A. H. *The Federal Reserve's Attachment to the Free Reserve Concept.* U.S. Congress, House of Representatives, Committee on Banking and Currency, Subcommittee on Domestic Finance, 88th Congress, 2nd Session. Washington: U.S. G.P.O., 1964.

DeRosa, P. and Stern, G. H. Monetary Control and the Federal Funds Rate," *Journal of Monetary Economics,* April 1977, *3,* 217–231.

Eichenbaum, M. and Singleton, K. J. Do Equilibrium Real Business Cycle Theories Explain Postwar U.S. Business Cycles? In Fischer, ed., *NBER Macroeconomics Annual.* Cambridge: The MIT Press, 1986.

Enzler, J., Johnson, L., and Paulus, J. Some Problems of Money Demand. *Brookings Papers on Economic Activity,* 1976, No. 1, 261–278.

Fischer, S. Contracts, Credibility, and Disinflation. Mimeo, National Bureau of Economic Research, 1985.

Friedman, B. M. Money, Credit, and Interest Rates in the Business Cycle. In Gordon, ed., *The American Business Cycle: Continuity and Change.* Chicago: University of Chicago Press, 1986.

Friedman, B. M. New Directions in the Relationship Between Public and Private Debt. *Science,* April 24, 1987, *236,* 397–403.

Friedman, B. M. The Roles of Money and Credit in Macroeconomic Analysis. In Tobin, ed., *Macroeconomics Prices and Quantities: Essays in Memory of Arthur M. Okun.* Washington: The Brookings Institution, 1983.

Friedman, B. M. Time to Reexamine the Monetary Targets Framework. *New England Economic Review,* March/April 1982, 15–23.

Friedman, M. Lessons from the 1979–82 Monetary Policy Experiment. *American Economic Review,* May 1984, *74,* 397–400.

Goldfeld, S. M. The Case of the Missing Money. *Brookings Papers on Economic Activity,* 1976, No. 3, 683–730.

Goldfeld, S. M. The Demand for Money Revisited. *Brookings Papers on Economic Activity,* 1973, No. 3, 577–638.

Goldfeld, S. M., and Sichel, D. E. The Demand for Money. In Friedman and Hahn, eds., *Handbook of Monetary Economics.* Amsterdam: North-Holland Publishing Company, 1990, 299–356.

Guttentag, J. M. The Strategy of Open Market Operations. *Quarterly Journal of Economics,* February 1966, *80,* 1–30.

Judd, J. J., and Scadding, J. L. The Search for a Stable Money Demand Function: A Survey of the Post-1973 Literature. *Journal of Economic Literature,* September 1982, *20,* 993–1023.

Lombra, R., and Moran, M. Policy Advice and Policymaking at the Federal Reserve. In Brunner and Meltzer, eds., *Monetary Institutions and the Policy Process.* Amsterdam: North-Holland Publishing Company, 1980.

Okun, A. M. Efficient Disinflationary Rules. *American Economic Review,* May 1978, *68,* 348–352.

Poole, W. Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model. *Quarterly Journal of Economics,* May 1970, *84,* 197–216.

Porter, R. D., Simpson, T. D. and Mauskopf, E. Financial Innovation and the Monetary Aggregates. *Brookings Papers on Economic Activity,* 1979, No. 1, 213–237.

Roley, V. V. Money Demand Predictability. *Journal of Money, Credit and Banking,* November 1985, *17,* Part 2, 611–641.

Sachs, J. D. The Dollar and the Policy Mix: 1985. *Brookings Papers on Economic Activity,* 1985, No. 1, 117–185.

Scholl, R. B. The International Investment Position of the United States in 1986. *Survey of Current Business,* June 1987, *67,* 38–45.

Simpson, T. D., and Porter, R. D. Some Issues Involving the Definition and Interpretation of the Monetary Aggregates. In *Controlling Monetary Aggregates III.* Boston: Federal Reserve Bank of Boston, 1980.

Solomon, R. The United States as a Debtor in the Nineteenth Century. Mimeo, The Brookings Institution, 1986.

Stock, J. H., and Watson, M. W. Interpreting the Evidence on Money-Income Causality. Mimeo, National Bureau of Economic Research, 1987.

Wallich, H. C. Recent Techniques of Monetary Policy. Federal Reserve Bank of Kansas City, *Economic Review,* May 1984, 21–30.

31

Monetary Policy in the 1990s

Robert T. Parry

President and Chief Executive Officer, Federal Reserve Bank of San Francisco, San Francisco, California

This chapter discusses challenges in conducting monetary policy in the 1990s—now that the monetary aggregates have proved unreliable. The source of these challenges is financial change—that is, the sweeping financial deregulation and innovation that began 20 years ago. As these developments spread through the financial markets, they undermined the reliability of the monetary aggregates. As a result, the aggregates were no longer reliable indicators of monetary policy. Moreover, they have been confusing to the public who may watch them to figure out the stance of monetary policy. M1, which used to be our main indicator, has been soaring for three years. But the growth of M2, which replaced M1 as our prime indicator, has been feeble. And contrary to *either* indicator, we've had moderate growth and well-behaved inflation. My focus will be on how the Fed has handled policy without reliable aggregates, and on a couple of options under discussion as indicators or targets—real interest rates and nominal GDP.

THE ECONOMIC SETTING IN THE UNITED STATES AND ABROAD

Not surprisingly, a good deal of attention has been focused on the rapid real GDP growth in the latter half of last year, and especially on the nearly 6 percent growth rate registered in the fourth quarter. But a longer-term perspective shows the current expansion so far has been moderate by postwar standards. Over the 11 quarters since the business cycle trough in early 1991, real GDP has risen at an average annual rate of $2\frac{3}{4}$ percent—well below the 5 percent average growth rate in previous expansions that have lasted at least this long.

Why did the economy "creep" out of the recession, rather than "boom" back, as it usually does? Basically because the United States, as well as many of our major trading partners, are in a stage of transition—a stage marked by disinflation and fiscal restraint. The defense cuts and other deficit-reducing measures in the U.S. are important factors that have

Reprinted from: *FRBSF Weekly Letter, No. 94–09 (Mar. 4, 1994).* This chapter is adapted from a speech given by Robert T. Parry, President and Chief Executive Officer of the Federal Reserve Bank of San Francisco, on January 14, 1994, to the Los Angeles chapter of the National Association of Business Economists.

restrained aggregate demand and slowed economic growth. And for the last few years, economic activity in our major trading partners has been lackluster, or worse. In the other G-7 countries—Canada, France, Germany, Italy, the U.K., and Japan—output grew on average by only $1\frac{3}{4}$ percent in 1991, not at all in 1992, and (based on incomplete data) probably by less than 1 percent last year. Although there are a variety of special factors operating in individual countries, one common reason behind these developments is policies designed to cut rates of inflation. All of these countries have made significant progress in reducing inflation in recent years.

The Fed's role in countering these forces was to lower interest rates—down to about a third what they were in 1989. But we lowered them gradually and cautiously because of our concerns about inflation. Like many other central banks, we want to bring inflation down and keep it close to zero where it will not distort economic decisionmaking. Thus, in February, we raised the funds rate by 25 basis points against a background of a sharp acceleration in the pace of economic activity and declines in the amount of slack in labor and product markets.

Although a policy of lowering inflation has its costs in the short run, I believe it *is* worth it, because, *in the long run,* inflation reduces economic well-being (see Motley 1993). For one thing, inflation often is associated with uncertainty about future inflation, which fosters higher long-term real interest rates. Uncertainty also complicates the planning and contracting businesses do that are so essential to capital formation, and it drives people to wasteful hedging activities. Finally, inflation heightens the distortionary effects of our tax system.

MONETARY AGGREGATES

Now comes the problem of *implementing* a low-inflation policy without relying on the monetary aggregates. The beauty of the aggregates was that they helped us solve the "lag problem"—that is, the classic "long and variable lag" between policy actions and inflation—probably $1\frac{1}{2}$ to 2 years. The aggregates were easily measured, we could control them reasonably well in the short run, and they had a fairly stable relationship to long-run inflation.

What happened to them? To summarize some 20 years in a single phrase, deregulation and innovation swept through financial markets. Interest rate ceilings on deposits were eliminated, new substitutes for deposits in M1 and M2 cropped up, and shifting funds from one instrument to another got a lot less expensive. Of course, this innovation and deregulation has been great for the overall economy: It has brought us more choices than ever to manage our financial affairs, and it has made financial markets far more dynamic and efficient.

But it has created problems for monetary policy. Growth rates of M1 and M2 no longer give us dependable information about future inflation—they often just reflect portfolio substitutions. For example, over the past two years, M2 growth has slowed dramatically—to an average of only $1\frac{1}{2}$ percent; if M2 *were* a reliable indicator of future inflation, it would imply outright *deflation* in 1994. With inflation currently a little below 3 percent, that's clearly wide of the mark. Why did M2 growth slow so dramatically? One important reason is the steep yield curve of the last few years. Households switched *out of* short-term, low-yielding M2 holdings and *into* long-term, higher-yielding stock and bond mutual funds.

Now, I do not mean to imply that because we have lost the aggregates as reliable indicators, we are helpless. We have *always* looked at a number of real and financial variables and have based our decisions on a good deal of judgment. And I think we have done fairly well. Real GDP growth has been respectable, and inflation has come down. The core inflation rate is now only a little above 3 percent—better than the 4 to $4\frac{1}{2}$ percent rates we saw around the turn of the decade. And I think we are in a good position to make further gradual progress on inflation. But I would certainly be more comfortable about it if I could look at a reliable leading indicator of inflation. Several indicators or targets to replace the aggregates have been suggested in recent years. I will focus on two: real interest rates and nominal income targeting.

REAL INTEREST RATES

The real interest rate is appealing because it has a direct effect on business and household spending decisions. But it also has problems. Real interest rates are hard to measure because they depend on unobservable expectations of future inflation. And the Fed cannot target real interest rates beyond the short run because they are determined by market forces. Finally, real interest rates are meaningful indicators only compared with a benchmark—an equilibrium real rate—that would be consistent with full employment. That equilibrium rate is not directly observable, and it is difficult to estimate, because it is affected by factors like productivity, government spending, and income tax rates. So the real interest rate does not appear to be a good candidate for the Fed's main inflation indicator (Cogley 1993, and Trehan 1993).

But that is not to say that real interest rates are *never* useful. If real rates stay very high or very low, that can be a warning sign. Look at the 1970s, for instance. Real rates were persistently *negative,* and that meant significant inflationary pressures were building up. More recently, in the past year or so, short-term real rates have been close to zero (Figure 1). The recent rise in the funds rate has raised real short-term rates somewhat, but they still are at fairly low levels. Is this an early warning? Well, I would say this situation *does* bear watching.

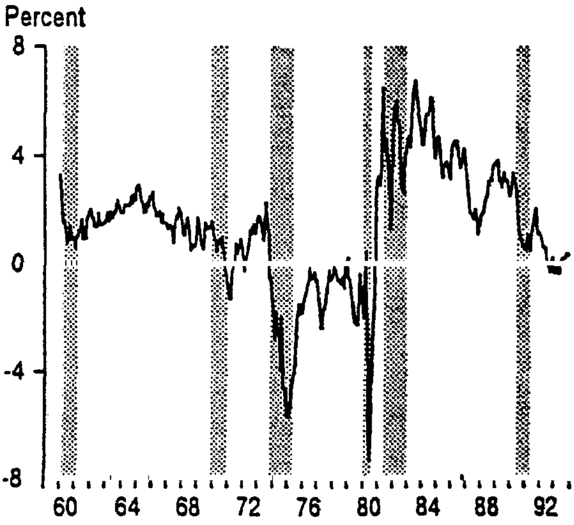


Figure 1 Real 3-Month T-bill rate (3-Month T-bill rate less CPI inflation over prior 12 months).

NOMINAL GDP

The second approach uses targets for aggregate demand, or nominal GDP. Nominal GDP is appealing as a monetary policy target because its long-run relationship with inflation is relatively stable. Furthermore, it will *remain* stable unless there's a sudden dramatic change in the trend growth of real GDP. So it is clearly immune to the effects of financial change that have undermined the monetary aggregates.

The *problem* with nominal GDP is that it does not respond to policy actions as quickly as money did, though the lag *is* shorter than the inflation lag. Some recent research on so-called "feedback rules" for nominal GDP suggests a way around this lag problem (McCallum 1990, Judd and Motley 1993, and Taylor 1993). Feedback rules provide "recommendations" for policy in the short run that are designed to control nominal GDP—and therefore inflation—in the long run. The policymaker sets a target for nominal GDP that is consistent with the inflation goal. Then, if the latest quarter's *actual* data are outside the target, the formula indicates by how much the funds rate should be raised or lowered.

Consider an example of an inflationary episode using one version of the rule the staff at the San Francisco Fed has explored. Suppose the inflation target is 1 percent. To allow for trend growth in real GDP of about 3 percent, a nominal GDP growth target would be set at $3\frac{1}{2}$ percent. Now suppose *actual* nominal GDP growth in one quarter comes in at $4\frac{1}{2}$ percent. That feedback rule would call for raising the funds rate by 20 basis points. If nominal GDP continued to exceed its target, the rule would likewise call for further increases until nominal GDP hit its target.

So with this approach, policymakers would have a guide for responding to actual *recent* data on aggregate demand and have more confidence that they would hit their inflation target in the long run. Of course, this approach is still in the research stage. And, I personally wouldn't be comfortable with *strictly* following *any* formula. But I think this approach merits consideration. The policy recommendations it generates might be a useful input that provides a benchmark in making judgmental policy decisions.

CONCLUSION

This chapter addressed some of the issues involved in conducting monetary policy in the 1990s—a time of worldwide disinflation, fiscal restraint, and continuing dynamism in financial markets. Replacing the monetary aggregates as targets for policy is not going to be easy. They not only served as a guide for monetary policymakers, but they also gave useful signals to everyone else about the future effects of policy. Even without useful guidance from the aggregates, though, we have managed to lower inflation. So let me conclude by assuring you that the erosion of the aggregates as reliable inflation indicators has *not* eroded our commitment to moving gradually toward zero inflation, which I believe is the best way the Fed can help the U.S. economy achieve its maximum growth potential.

REFERENCES

Cogley, T. 1993. Monetary Policy and Long-Term Real Interest Rates. FRBSF *Weekly Letter* 93–42 (December 3).

Judd, J. P., and Motley, B. 1993. Using a Nominal GDP Rule to Guide a Discretionary Monetary Policy. Federal Reserve Bank of San Francisco *Economic Review* (3) pp. 3–11.

McCallum, B. 1990. Targets, Indicators, and Instruments of Monetary Policy. *NBER Reprint* No. 1550.

Motley, B. 1993. Inflation and Growth. FRBSF *Weekly Letter* 93–44 (December 31).

Taylor, J. B. 1992. Discretion Versus Policy Rules in Practice. *CEPR Publication* No. 327. Center for Economic Policy Research, Stanford University.

Trehan, B. 1993. Real Interest Rates. FRBSF *Weekly Letter* 93–38 (November 5).

This Page Intentionally Left Blank

32

Price Stability
Is a Tough Central Bank Enough?

Lawrence J. Christiano
Professor of Economics, Northwestern University, Evanston, Illinois

Terry J. Fitzgerald
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

The objective of monetary policy, laid out in the Federal Reserve Act, is to "promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." Many believe the Federal Reserve's primary focus should be achieving stable prices, arguing that price stability is essential to attaining the companion goals of maximum employment and moderate interest rates.

How can price stability be achieved? Conventional wisdom offers a simple and straightforward answer. Make sure there is a tough, independent central bank with an unwavering commitment to price stability. According to this view, a tough central bank is all that is required.

Recently, however, an alternative view has challenged the conventional wisdom. In this new view, a tough, independent central bank is not enough to ensure price stability; an appropriate fiscal policy is also required, *no matter how tough the central bank.* Because fiscal policy plays such a prominent role in the theory underlying this view, it has been called the *fiscal theory of the price level.* Thus, we refer to it throughout this *Commentary* as the *fiscal theory view.*[1]

Whether one accepts the conventional or the fiscal theory view has significant implications for the way central banks should conduct business. The traditional view has fostered the notion that central banks with a mandate to foster price stability should stay away from the fiscal authorities, to reduce their likelihood of being pressured into making poor decisions.

The fiscal theory view implies that central bankers must do more than just make sure their own house is in order: They must also ensure the fiscal authority adopts an appropriate policy. If this view is correct, then governments that seek to achieve price stability by focusing exclusively on establishing tough central banks may find price stability *impossible* to achieve—unless an appropriate fiscal policy is, fortuitously, in place.

Reprinted from: *Federal Reserve Bank of Cleveland Economic Commentary,* August 1, 2000.

This *Economic Commentary* describes the conventional and the fiscal theory views of price stability, discusses why their conclusions are so different, and examines U.S. inflation behavior in light of the two views.[2]

THE CONVENTIONAL WISDOM[3]

The conventional wisdom asserts that the price level is determined by the quantity of money supplied relative to the quantity of money demanded. If supply grows faster than demand, the price level must rise—that is, inflation results. Furthermore, inflation can occur only when the monetary authority expands the money supply too rapidly. But there is more to the story.

The Role of Fiscal Policy

At the heart of both the conventional and the fiscal theory views lies the simple observation that today's government debt must be paid off with future government surpluses. Specifically, the current value of the government debt must equal the expected present value of future government surpluses. As we will discuss, the views differ as to how this condition gets satisfied.

Under the conventional view, the current real (that is, inflation-adjusted) value of the government debt—the nominal value divided by the price level—is a fixed quantity, because the nominal debt is given and the price level is determined in the money market. The government has two ways of generating surpluses to pay off the debt. It can collect more tax revenues than it spends—referred to as fiscal surpluses—or it can print more money, referred to as seignorage. Increases in seignorage are accompanied by increases in inflation.[4]

A Game of Chicken

The key observation of the conventional wisdom is that the fiscal authority controls one source of surplus, while the monetary authority controls the other. Therefore, a tension exists between the authorities over the amount each will contribute toward paying off the debt. If one authority announces it will collect very low surpluses, then the other has no choice but to collect large surpluses to ensure fiscal solvency (that is, to pay off the debt). This situation is sometimes compared to the game of "chicken." If one authority can convince the other that it will collect little surplus no matter what (it will not swerve), then the other authority is forced to relent and collect high surpluses (it must swerve).

Suppose the fiscal authority adopts a "loose" fiscal policy, reflecting a tax cut or an increase in expenditures, so that fiscal surpluses are reduced. Because today's real government debt is fixed, simple arithmetic dictates the monetary authority must, sooner or later, increase seignorage to counteract the decline in fiscal surpluses. This increase in seignorage is accomplished by increasing the growth rate of the money supply, and thus it is accompanied by an increase in inflation.

Although the monetary authority may wish to maintain low money growth and low inflation, it has no choice but to increase money growth. It does have some discretion over the timing of the increase, so that inflation might rise sooner or later, or it may be spread out over time. But whatever the timing, if the fiscal authority reduces its surpluses, money growth—and thus inflation—must go up at some point.

The Conventional Answer—Yes (A Tough Central Bank Is Enough)

The same arithmetic suggests a solution to the inflation problem: Design central banks so they can credibly commit to not "swerving" when an irresponsible fiscal authority threatens to reduce future surpluses. Governments around the world have sought to implement this solution by making central banks independent and directing them to assign a high priority to inflation. With the monetary authority completely committed to a fixed value for seignorage (it will not swerve), the arithmetic forces the fiscal authority to adopt a consistent fiscal policy (it must swerve). This is the basis of the conventional view.

THE FISCAL THEORY VIEW

According to the fiscal theory view, the conventional framework just described is not always appropriate. In particular, the price level is not always determined in the money market. Instead, the fiscal theory view asserts that the price level is sometimes determined in much the same way as the price of Microsoft stock.[5] This alternative view has dramatic implications for achieving price stability.

An Analogy to Microsoft

How is the price of Microsoft stock determined? According to standard theory, Microsoft behaves in a way that will maximize its profits, and the stock price reflects how much profit Microsoft is expected to generate for its stockholders. The key point is that the price of the stock responds so that it always equals the expected present value of future profits per share of outstanding stock.

According to the fiscal theory, the aggregate price level is determined similarly. Government policies imply a stream of future surpluses, and the price level reflects the expected value of this stream to bondholders. In particular, the price level is the level at which the current real value of the government debt equals the expected present value of future real surpluses.

Recall that under the conventional view, it is also true that the real government debt must equal the expected present value of future surpluses. The key distinction is that under that view, the price level (and thus the value of real debt) is determined elsewhere, so future surpluses must be adjusted with changes in the price level to ensure fiscal solvency. Under the fiscal theory view, the causation between the price level and surpluses runs in the opposite direction. Future surpluses are not calibrated to ensure fiscal solvency, just as Microsoft does not adjust its profits to justify changes in its stock price. Instead, the price level adjusts so that the real debt (stock price) is equal to whatever expected future surpluses (profits) are implied by the current policies (firm decisions).

Suppose the expected present value of real future government surpluses is represented as 10 billion apples. Further suppose the total nominal value of current government debt equals $1 billion. According to the fiscal theory, the price level today must be 10 cents per apple, so that the real value of current government debt is 10 billion apples (1 billion divided by 0.1).

If the price level were not 0.1, but were 0.05 instead (5 cents per apple), then the current value of the government debt, in terms of apples, would be 20 billion. However, the government would generate surpluses of only 10 billion. People could get more apples by selling government debt and buying apples today (they could buy 20 apples for $1) than they could by holding the government debt and waiting to collect the government surpluses ($1 of debt would entitle them to 10 apples). People would choose the former, thereby driving up the price of apples.

Game Over

Consider again the game of chicken. Suppose that of the 10 billion apples owed by the government, the fiscal authority initially agrees to raise 9 billion in future fiscal surpluses, while the monetary authority agrees to raise 1 billion. Now suppose the fiscal authority adopts a loose fiscal policy that reduces future fiscal surpluses to 4 billion apples. According to conventional wisdom, this policy change, combined with the requirement of fiscal solvency, leaves the central bank with no choice but to swerve and increase seignorage to 6 billion apples, thereby increasing inflation.

But the situation is quite different under the fiscal theory. Now there is a third option for achieving fiscal solvency that is not envisioned in the game of chicken—the value of outstanding government debt can be reduced. This is accomplished through a jump in the price level. Rather than increasing seignorage to 6 billion apples, suppose the monetary authority hangs tough and refuses to increase seignorage. What would happen? Total expected future surpluses would fall from 10 billion apples to 5 billion. Nominal government debt would remain $1 billion. But, unlike under the conventional view, the price level is free to adjust to equate the real value of the debt with expected future surpluses. In particular, this would occur if the price level jumped to 20 cents per apple, so that the current real debt would be 5 billion apples. Fiscal solvency does not require either authority to change its policy (that is, to swerve).

Under the conventional view, the change to a loose fiscal policy means the monetary authority has no choice but to increase money supply growth and, thus, inflation. This game of chicken is not played out in the fiscal theory—the monetary authority is not forced by fiscal solvency to adjust its policy.

Despite the fact that the monetary authority refused to budge in our example, the price of apples doubled! This increase in the price level occurs without the complicity of the monetary authority. That is the crucial distinction from the conventional wisdom—a tough central bank does not prevent the price increase.[6]

The Fiscal Theory Answer—No (A Tough Central Bank Is Not Enough)

Ironically, the disappearance of the game of chicken prevents a tough central bank from being able to achieve price stability on its own. Just as the monetary authority is not forced to swerve by a change in fiscal policy, a tough central bank cannot force the fiscal authority to swerve and increase fiscal surpluses.

In fact, the fiscal theory produces a striking, more general, result. When fiscal surpluses fluctuate unpredictably through time, the theory implies that it is impossible for the central bank to perfectly stabilize the variability of inflation, even though it can control the average rate of inflation.[7] That is, the central bank cannot insulate the price level from unexpected movements in tax revenues and expenditures, no matter how tough and independent it may be.

How can price stability be achieved, according to the fiscal theory view? One possibility is to design fiscal policy to minimize fluctuations in the present value of future surpluses. Such a policy would limit the amount the price level must adjust from period to period.

A second way to achieve price stability, the one we focus on, makes use of the fact that with an appropriately chosen fiscal policy, a tough central bank can achieve price stability. What is an appropriate fiscal policy? It is a policy that ensures government debt will not grow too rapidly. The key point here is that a tough central bank is not enough for any fiscal policy, but it is enough for some policies. Thus, the fiscal theory view asserts it is crucial to be aware of the nature of fiscal policy.

WHICH VIEW TO BELIEVE?

The key policy assumption of the fiscal theory is that it allows for fiscal policies such that if the real value of the government debt were to grow explosively, no policy adjustments would be made to keep it in line.[8] In contrast, the conventional wisdom assumes government policy will always adjust to keep the debt in line. We can gain insight into the plausibility of each assumption by observing policies that have actually been in place when government debt has grown rapidly.

It seems clear that governments often stand ready to adjust fiscal policy when the debt gets large. For example, the Maastricht Treaty records the intention of European Union members to adjust fiscal policy in the event their debt grows too rapidly. Likewise, the International Monetary Fund uses an array of sanctions and rewards to encourage its member countries to keep their debt in line by suitably adjusting fiscal policy.

This may seem like the end of the story—toss out fiscal theory. But fiscal theory advocates do not claim that fiscal policies never respond to exploding debt. Instead, they argue that debt limitations such as those imposed by the Maastricht Treaty and the IMF are precisely the types of fiscal policies that allow the conventional wisdom to hold. The fiscal theory view provides a rationale for such policies—because without them, countries would run the risk of substantial price instability regardless of the central bank's toughness and commitment.

WHICH VIEW EXPLAINS U.S. INFLATION?

Which view explains the historical and recent behavior of inflation in the United States? Inflation increased steadily through the 1960s and 1970s, peaked in the early 1980s, and has declined fairly steadily since. In fact, it has remained relatively low and stable since the mid-1980s. Conventional wisdom argues that the slowdown in inflation can be attributed to the policies of a tough, committed Federal Reserve—initially led by the unquestionably tough Paul Volcker and currently led by the equally committed Alan Greenspan.

The fiscal theory view does not dispute the Federal Reserve's commitment to price stability over this period, nor its important contribution to achieving price stability; rather, it asserts that developments in fiscal policy have also played an important role in stabilizing prices. Some research supports the notion that the fiscal theory characterizes the behavior of inflation in the 1960s and 1970s.[9] During the 1980s, however, fiscal policy changed dramatically as government debt increased rapidly and pressure mounted for tax increases or expenditure cuts to bring the debt back in line. The fiscal theory view can be used to argue that this change represented a shift to an appropriate fiscal policy under which a tough central bank could achieve price stability. Without the change in fiscal policy, price stability may not have been possible under any monetary policy.

CONCLUSION

Is a tough central bank enough to achieve price stability? Both the conventional wisdom and the fiscal theory view agree that it is necessary ingredient. They differ in their analyses of whether a tough central bank alone is sufficient to achieve price stability, as in the conventional wisdom, or whether an appropriate fiscal policy is also necessary, as in the fiscal theory view. Resolving this disagreement will be of primary importance as nations worldwide seek to achieve and maintain price stability.

As for the United States, the fiscal theory view provides a warning that the current environment of low and stable inflation may not have been achieved by a tough central bank alone. If the fiscal theory view has merit, a renewed emphasis should be placed on "locking in" an appropriate fiscal policy.

NOTES

1. A more thorough discussion of the issues presented here can be found in Lawrence J. Christiano and Terry J. Fitzgerald, "Understanding the Fiscal Theory of the Price Level," Federal Reserve Bank of Cleveland, *Economic Review,* vol. 36, no. 2 (2000 Quarter 2), pp. 3–38.
2. In this *Commentary,* we assume that complete price stability is desirable. Some research suggests this may not be the case—some price variability may be desirable. In fact, one argument in the fiscal theory literature is that it may generate an optimal degree of price instability. See Christiano and Fitzgerald (2000) for a discussion of this point.
3. This discussion is based on the classic analysis in Thomas Sargent and Neil Wallace, "Some Unpleasant Monetarist Arithmetic," Federal Reserve Bank of Minneapolis, *Quarterly Review,* vol. 5, no. 3 (1981), pp. 1–17.
4. Here we assume the economy is on the "right" side of the Laffer curve, where seignorage is an increasing function of the inflation rate.
5. The Microsoft example is taken from John Cochrane, "Money as Stock: Price Level Determination with No Money Demand," National Bureau of Economic Research, Working Paper no. 7498, January 2000.
6. We follow Sargent and Wallace in thinking of the game of chicken as reflecting that the actions of the fiscal authority can force, as a matter of feasibility, the monetary authority to increase the money supply. Under the fiscal theory, the fiscal authority's actions may also affect the actions of the monetary authority, but this is a matter of the monetary authority's preferences and not feasibility. For example, the monetary authority may want to swerve in our example, even though it is feasible not to, because it may not like the outcomes when it does not swerve. Fleshing this out requires specifying the preferences and objectives of the monetary and fiscal authorities.
7. This result is due to Michael Woodford.
8. Exploding debt is not envisioned under the fiscal theory. The idea is that as long as there is absolutely no doubt about the government's commitment to not adjusting policy in the face of exploding debt, prices will respond so that the debt does not explode in the first place.
9. John Cochrane argues that the fiscal theory can explain the behavior of inflation over the entire postwar period in "A Frictionless View of U.S. Inflation," in Ben S. Bernanke and Julio Rotemberg, eds., *NBER Macroeconomics Annual,* Cambridge, Mass.: MIT Press, 1998. In a comment on that article, Michael Woodford indicates the fiscal theory may provide a good explanation for the 1970s, but he is more skeptical that it characterizes the 1980s and 1990s, in part for the reasons given in this *Commentary.*

33

A Hitchhiker's Guide to Understanding Exchange Rates

Owen F. Humpage
Economic Advisor, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Each day, more than $1.2 billion worth of foreign exchange changes hands around the globe, an amount that far exceeds the daily value of world trade. Approximately 83 percent of these transactions involve U.S. dollars, but not all involve U.S. citizens.[1]

Relatively small changes in the prices at which these trades occur—exchange rates—can have immediate and profound effects on economic events, ranging from family vacations to corporate profits. Large changes can shake governments, as recently demonstrated in Southeast Asia. Yet, despite the importance of exchange rates, most people find their behavior unfathomable.

Economists often view the *nominal* exchange rate (the foreign currency price of a dollar typically quoted in *The Wall Street Journal*) as the product of the real exchange rate and a component reflecting the difference between domestic and foreign inflation. Unlike their nominal counterparts, real exchange rates are not directly observable, but economists estimate them because of their influence on international competitiveness (see Figure 1). This dichotomy between a real exchange rate and an inflation differential has proved useful for understanding the complex connections between economic fundamentals and nominal exchange rates, and especially for appreciating the role of monetary policy in determining exchange rates.

This chapter offers a nontechnical guide for those seeking a quick tour of exchange rates. The first part considers the role of inflation and monetary policy in determining exchange-rate movements. The next section utilizes balance-of-payments concepts to illustrate the economic role of the real exchange rate. The final, but perhaps most important, part of the narrative introduces the crucial role of expectations.

INFLATION DIFFERENTIALS

An exchange rate is the relative price of one nation's money versus another's. Currently, for example, 1.82 German marks exchange for one U.S. dollar. If the Federal Reserve creates excessive amounts of dollars (that is, more than people currently wish to hold) at a

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (Jan. 1, 1998).

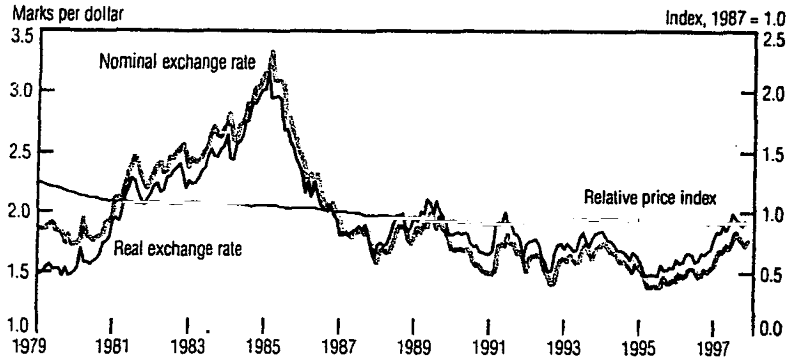


Figure 1 Mark/dollar exchange rates. Sources: DRI/McGraw-Hill; and author's calculations.

faster pace than the Bundesbank issues excessive amounts of marks, the value of the dollar will fall relative to the mark, say to 1.75.

Ignoring the difficulties associated with expectations and perceptions of monetary policy, the dynamics underlying the dollar's depreciation might proceed as follows: Faster money growth creates inflationary pressures in the United States, causing people to shift their purchases away from U.S. goods toward the now relatively less expensive German goods. To acquire German goods, however, people must first convert their dollars to marks. The increased demand for marks (and the greater supply of dollars) will bid up the value of the mark relative to the dollar in the foreign exchange market; that is, the dollar will depreciate against the mark. Holding others things constant, this dollar depreciation will continue as long as the U.S. inflation rate exceeds the German inflation rate, and will tend to match the inflation differential between the two countries. If, for example, Germany's inflation rate is 2 percent per year and the U.S. inflation rate is 3 percent per year, the dollar will depreciate 1 percent per year against the mark, other things being equal.

The explanation above contains two important implications for monetary policy: First, because monetary policy ultimately determines only the domestic inflation rate, a central bank that wants to engineer a depreciation of its currency must create more money than its trading partners and thereby generate a higher inflation rate. Second, because any resulting exchange-rate depreciation will ultimately offset the inflation differential, a monetary-induced depreciation cannot secure a competitive trade advantage. The real exchange rate, which I discuss below, will remain unaffected in the long term. Any trading gains from engineering a dollar depreciation are purely transitory, lasting only until prices fully adjust.

Economists refer to the relationship linking exchange-rate movements and inflation rates across countries as relative purchasing power parity (PPP). Recent estimates suggest that once disturbed, PPP takes an average of eight years to become reestablished.[2]
One interpretation of this finding is that goods prices adjust slowly to monetary shocks, implying that monetary policy may be able to affect the real exchange rate in the interim. A second interpretation is that nonmonetary events, such as productivity shocks or changes in preferences for domestic versus foreign goods, are important determinants of exchange-rate movements. The latter perspective draws attention to the determination of real exchange rates.[3]

REAL EXCHANGE RATES AND THE BALANCE OF PAYMENTS

By assuming that world inflation rates and expectations are constant, we can focus on the real component of the nominal exchange rate. This will change in response to any economic event that affects the real (inflation-constant) demand for, or supply of, traded goods and international investments. To understand the connection, however, one must first understand balance-of-payments accounting.

The balance of payments records all transactions between residents of the United States and residents of the rest of the world.[4] Anything that creates a debit item in the U.S. balance of payments creates a supply of dollars (and a demand for foreign currency). When we import a German car or invest in German stocks—both debits in balance-of-payments accounting—we must first acquire marks in the foreign exchange market. Likewise, anything that creates a credit item in the U.S. balance of payments, such as the sale of domestic wheat or U.S. Treasury bonds, generates a demand for dollars (and a supply of foreign currency).

A country that incurs a current account deficit—like the United States—is consuming more of the world's output than it is producing. Its imports are a debit item in its current account, creating a supply of its own currency. Such a country must pay for its extra current consumption (that is, its surfeit of imports) by giving foreigners financial claims on its future output (stocks, bonds, bank deposits, and so on). The resulting foreign capital inflows are credit items in the balance of payments and represent a demand for dollars. As Table 1 indicates, current account and capital account balances must offset each other exactly (assuming that everything is measured correctly). In sum, then, neither an excess supply nor an excess demand for dollars exists.

The capital account does not passively respond to the current account, as the previous paragraph implies. Myriad individuals make separate decisions about importing, exporting, and investing abroad, and each of these decisions affects the balance of payments independently. If at any time the collective intentions are not consistent with equilibrium in the accounts, attempts to enact these plans will cause the real exchange rate to change. The exchange-rate adjustment in turn forces people to reevaluate their plans in such a way as to pull the current and capital accounts into balance. (Other economic variables, like real interest rates, might also shift and contribute to the process.)

Contrary to popular belief, the mere existence of a current account surplus or deficit implies nothing about how dollar exchange rates will behave. A country may incur a cur-

Table 1 U.S. Balance of Payments (Billions of Dollars)

	1995	1996	1997[a]
Current account balance	–129	–148	–160
Trade in goods and services	–102	–111	–113
Investment income	7	3	–11
Unilateral transfers	–34	–40	–36
Capital account balance	129	148	160
Official capital flows	101	129	66
Private capital flows	43	66	166
Statistical discrepancy	–15	–47	–72

[a] Based on three quarters of data expressed at an annual rate.
Source: U.S. Department of Commerce, Bureau of Economic Analysis.

How to Create a Real Mark/Dollar Exchange Rate

First, divide the German Consumer Price Index (CPI) by the U.S. CPI. Second, construct an index number for this ratio. If possible, choose a base year for the index that represents an equilibrium. (This is difficult and arbitrary; I chose February 1987 only because German and U.S. officials expressed satisfaction with the nominal exchange rate at the time.) The index will fall when U.S. inflation exceeds German inflation. Finally, multiply the nominal exchange rate by this index of consumer prices to obtain the real exchange rate. The real exchange rate equals the nominal exchange rate in the base year.

rent account deficit through various routes, each with different implications for its exchange rates. If, for example, domestic demand for foreign goods initially increases, imports will expand, the current account deficit will grow, and the dollar will depreciate. The dollar depreciation will encourage a counterbalancing inflow of foreign capital by making U.S. financial securities more attractive to foreigners. This—the standard view—holds that the dollar depreciates when the U.S. trade deficit widens. Alternatively, suppose that an improved domestic investment climate draws an inflow of foreign capital. The dollar will appreciate, making domestic goods expensive relative to foreign goods and striking a balance between increased capital inflows and the larger trade deficit. This second case connects a dollar appreciation to a U.S. trade deficit.

All economic events that affect the real demand for, or supply of, traded goods and financial investments can potentially determine the level of real exchange rates. While almost any economic variable would seem a possible candidate, real interest-rate differentials, productivity differentials, trade restraints, tax rates, and relative preferences for domestic versus foreign goods seem key.

I have discussed nominal exchange rates as the product of a real component and an inflation differential, identifying or alluding to fundamental economic variables that most economists regard as important to the determination of exchange rates. Nevertheless, empirical models have typically failed to explain past exchange-rate movements or to predict future exchange-rate paths in terms of combinations of these fundamentals. Although exchange rates bear some long-term correspondence to fundamentals, the relationship is not close in the short or medium term. Instead, research suggests that the best guess of tomorrow's (or next week's or next month's) exchange rate is today's (or this week's or this month's) exchange rate, and even this projection is not very good. This peculiarity has never been fully explained, but expectations seem to play a crucial role.

EXPECTATIONS

Economists typically view exchange rates as very similar to asset prices. The current price of an asset reflects the present discounted value of the income stream that investors expect it to generate over its lifetime. Similarly, an exchange rate reflects the present discounted value of all relevant fundamentals, including current fundamentals and their expected future values. Foreign exchange traders face strong incentives to acquire every piece of information about current and anticipated economic developments that could possibly influence exchange rates. If these dealers are successful, their current quotations will incorporate all available, relevant data, and only new information that causes revisions in traders' expectations will influence exchange rates. This implies, for example, that previously anticipated changes in monetary policies or other fundamentals will not affect current exchange rates; only unanticipated changes will.

One might expect profit-seeking exchange dealers to formulate their expectations, and therefore their exchange quotations, without making systematic errors. To the extent that they can do so, revisions to their quotes will be fairly random and will impart a zigzag pattern to exchange-rate movements. Over time, a net change in one direction or the other may emerge as exchange rates adjust to persistent shifts in underlying fundamentals. On a day-to-day basis, the exchange rate will bounce—in a seemingly random manner—around any such path.

Figure 2 illustrates such a pattern for daily movements in the mark/dollar exchange rate. Over the period displayed (January 1, 1995 to December 31, 1997), the average daily change in the exchange rate was 0.02 percent, suggesting a slight trend toward dollar appreciation. Daily movements about this average seem fairly random, with appreciations or depreciations of almost 1 percent being quite common. The largest single-day appreciation was 2.7 percent, and the largest single-day depreciation was 2.5 percent.

The foregoing discussion is based on the implicit assumption that expectations are fairly uniform and remain firmly anchored to a set of generally recognized fundamentals. The seemingly random pattern of Figure 2, together with the small size of day-to-day changes, appears consistent with that view. Nevertheless, other evidence conflicts with it. The aforementioned poor performance of exchange-rate models, for example, undermines the validity of this assumption. Particularly damaging is the fact that since the inception of floating exchange rates, both nominal and real exchange rates have increased substantially, but the volatility of their underlying fundamentals has changed little.[5]

Although exchange traders are highly effective users of information, they probably are not perfectly efficient. Indeed, why would trade occur, especially in such large volumes, if all traders had identical information at all times? For one thing, information about economic fundamentals (both its acquisition and its interpretation) is costly, which may explain why many foreign exchange traders generate profits from technical trading rules—essentially rules of thumb—instead of from models based on economic fundamentals. Many

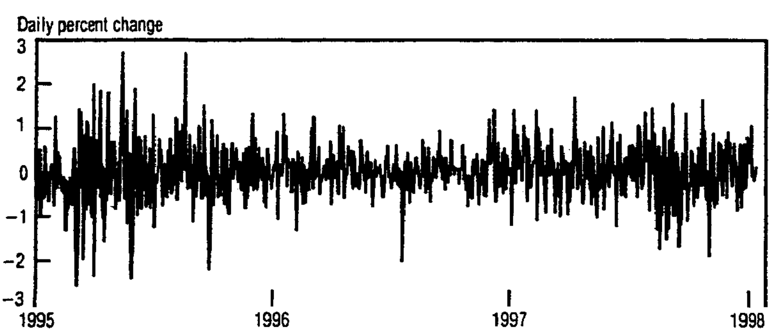


Figure 2 Daily mark/dollar exchange rates, 1995–97. **SOURCE:** Federal Reserve Bank of New York.

of these rules project past trends into the future. One, for example, requires that a trader buy when the exchange rate rises by some fixed percentage above its past trough, and sell when it falls some fixed amount below its past peak.

Such trading rules could increase short-term volatility. As time passes and as information becomes freely accessible, traders may increasingly respond to fundamentals; initially, however, traders may not be linked to fundamentals in a fixed or even consistent way.[6] Figure 2 demonstrates striking periods of high and low exchange-rate volatility. The standard deviation of exchange-rate changes in 1995, for example, is almost twice as large as in 1996, suggesting more market uncertainty about exchange rates in the former period.

CONCLUSION

Over the 25 years since dollar exchange rates began to float, economists have learned to garnish their exchange-rate predictions with humility. Most probably feel secure in betting that a country with a relatively high inflation rate will eventually see its currency depreciate in foreign exchange markets. Few, however, would venture a guess about the time path of the adjustment, or gamble with equal confidence on the long-run exchange-rate implications of most other economic variables, or even speculate about the short run. Nevertheless, international trade and capital flows continue to grow, despite the periodic trepidation about movements in dollar exchange rates.

NOTES

1. See Bank for International Settlements, *Central Bank Survey of Foreign Exchange and Derivatives Market Activity.* Basle: BIS, May 1996; and International Monetary Fund, *Direction of Trade Statistics Yearbook, 1997.* Washington, D.C.: IMF, 1997, p. 2.
2. See Kenneth A. Froot and Kenneth Rogoff, "Perspectives on PPP and Long-Run Real Exchange Rates," in Gene M. Grossman and Kenneth Rogoff, eds., *Handbook of International Economics,* vol. 3, part 2. Amsterdam: North-Holland, 1995, pp. 1647–88.
3. The price indexes used to calculate PPP usually include traded and nontraded goods. Deviations between the prices of these goods within a single country can affect calculated values of the real exchange rate.
4. The balance of payments, as presented in table 1, records international transactions on a nominal basis. I assume throughout this section that all transactions are measured on a real (constant-price) basis.
5. See Jeffrey A. Frankel and Andrew K. Rose, "Empirical Research on Nominal Exchange Rates," in Gene M. Grossman and Kenneth Rogoff, eds., *Handbook of International Economics,* vol. 3, part 2. Amsterdam: North-Holland, 1995, pp. 1689–729.
6. See Gregory P. Hopper, "What Determines the Exchange Rate: Economic Factors or Market Sentiment?" Federal Reserve Bank of Philadelphia, *Business Review* (September/October 1997), pp. 17–29.

34

Output and Inflation A 100-Year Perspective

Kevin Lansing
Economist, Federal Reserve Bank of San Francisco, San Francisco, California

Jeffrey Thalhammer
Research Associate, Federal Reserve Bank of San Francisco, San Francisco, California

While economists generally accept that monetary policy can influence nominal variables such as the price level and inflation, they continue to debate the relationship between monetary policy and *real* variables such as the unemployment rate and real GDP. During the early 1960s, many economists and policymakers believed that policy could exploit a stable trade-off between inflation and real economic activity. One version of the hypothesized trade-off, originally described by A. W. Phillips (1958) using U.K. data from 1861–1957, implied that policymakers could permanently lower unemployment by generating higher inflation. Some years later, economists Edmund Phelps (1967) and Milton Friedman (1968) argued persuasively that any such trade-off was bound to be short-lived: once people came to expect the higher inflation, monetary policy could not keep unemployment below its long-run equilibrium or "natural" level. This claim was later borne out by the experience of the 1970s, when rising U.S. inflation did not bring about the lower unemployment rates promised by the Phillips curve. On the contrary, higher inflation coincided with higher unemployment—a combination that became known as "stagflation."

In the late 1990s, the situation is precisely the reverse—the U.S. economy exhibits low inflation combined with low unemployment. Some commentators seem to view this combination as a puzzle or breakdown in the traditional relationship between inflation and real economic activity. This chapter challenges such a view by putting the recent data into a 100-year historical perspective.

STABILIZATION POLICY

Validation of the Phelps-Friedman argument by the experience of the 1970s still left open the possibility that policymakers might exploit a *short-term* or transitory link between in-

Reprinted from: *FRBSF Economic Letter,* No. 99-18 (May 28, 1999).

flation and real activity to smooth business cycle fluctuations. For example, the Fed might pursue an expansionary monetary policy to stimulate aggregate demand when real activity was "too low" and adopt a contractionary monetary policy to reduce aggregate demand when real activity was "too high"—all the time recognizing that no permanent trade-off between inflation and real activity exists. Indeed, one can see how this principle might be used nowadays to guide the Fed in pursuit of its legally mandated goals to promote "maximum" employment and "stable" prices (see Federal Reserve Bank of San Francisco, 1999). Efforts to smooth business cycles using monetary policy face some difficult problems. First, history has shown that monetary policy affects the economy only through long and variable lags. Second, a policy of "leaning against the wind" requires policymakers to decide in real time (using preliminary data) whether economic activity is "too low" or "too high" relative to some benchmark. Benchmarks such as the natural rate of unemployment or the long-run trend of real GDP cannot be observed directly, however. These must be estimated from available data by combining economic theory, statistical analysis, and sound judgment.

ESTIMATING THE LONG-RUN TREND IN REAL OUTPUT

One very simple way to estimate the long-run trend of real GDP per person is to fit a straight line through the data. This would be appropriate if one believes that technological change occurs at a constant rate, while transitory shocks generate fluctuations around the long-run trend. Many economists believe, however, that technological change occurs unevenly and is influenced by many factors, such as the training and education of the workforce, the amount of resources devoted to inventive activity, and tax policy, to name a few. If this were true, then the assumption of an invariant long-run trend could lead one to the mistaken conclusion that observed movements in real GDP are due to transitory shocks pushing the economy away from trend (thus calling for a monetary policy response) when in fact the long-run trend itself has shifted. To avoid such a mistake, it is desirable to employ a trend measure that is capable of accounting for shifts in the long-run trend of real GDP. One way this might be done, for example, would be to fit a piecewise straight line through successive midpoints of business cycle expansions. This approach has a serious drawback, however, because the midpoint date is not known until after the expansion has ended. Policy decisions must be made in real time without the benefit of such hindsight.

John Cochrane (1994) has suggested a way of using consumption data to help account for shifts in the long-run trend of real GDP. Cochrane's idea is based on the permanent income hypothesis, which says that people's consumption decisions depend primarily on their "permanent" income, i.e., income that is expected to persist into the future. According to this theory, transitory changes in income do not have much impact on consumption; people use saving and borrowing to maintain a smooth pattern of consumption when hit by transitory income shocks. If we observe a change in income that is not accompanied by a change in consumption, then we can infer that people view the income change as transitory. On the other hand, if we observe simultaneous changes in income and consumption, then we can infer that people view the income change as permanent. Individual consumers have first-hand knowledge regarding the future prospects of their employer (and their employer's industry) and hence are in good position to judge whether changes in their own income are likely to be permanent or transitory.

A shift in the long-run trend of real GDP will affect people's permanent income and should thus be accompanied by a change in consumption. By regressing output data on con-

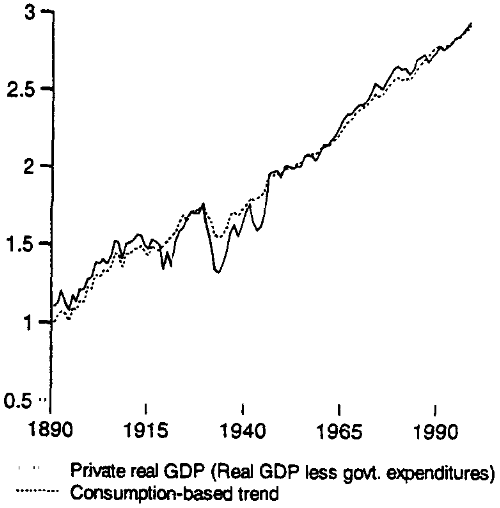


Figure 1 Private real GDP/person: 87$bil (logs). Data source: Farmer (1999).

sumption data (both of which are observable in real time), these shifts can be automatically incorporated into a measure of trend output. The resulting "output gap," i.e., the deviation of real GDP from trend, should provide an informative signal about where the U.S. economy is operating relative to its long-run potential (for additional discussion, see Cogley and Schaan 1995).

Figure 1 plots private real GDP per person (real GDP less government expenditures) for the period 1890–1998 together with a consumption-based trend. This approach places the U.S. economy close to its long-run trend during the late 1990s. The strong growth of real GDP over this period could thus be interpreted as reflecting an upward shift in the economy's ability to produce goods and services—a shift that people view as permanent based on the strong growth of their consumption. Given this interpretation of the data, the current U.S. output gap is actually quite small in comparison to some other periods of fast economic growth, such as the early to mid-1960s.

ESTIMATING THE LONG-RUN TREND IN INFLATION

Figure 2 plots U.S. inflation over the period 1890–1998. As with real GDP, it is useful to construct a long-run trend of inflation so that short-term or transitory movements can be defined as deviations from the trend. Here we adopt a trailing five-year moving average as our measure of trend inflation, which approximates an exponential smoothing technique. As noted by Cogley (1998), this technique accounts for periodic shifts in what might be viewed as the Fed's long-run inflation target and has the advantage that the trend can be computed in real time (unlike a centered moving average trend).

The trailing five-year moving average also can be interpreted as a measure of expected inflation. In particular, people are likely to use observations about past inflation to help make predictions about future inflation. According to the Phelps-Friedman argument, only the unexpected component of inflation (the deviation from trend) should influence real economic activity. Figure 2 shows that U.S. inflation is slightly below trend in the late

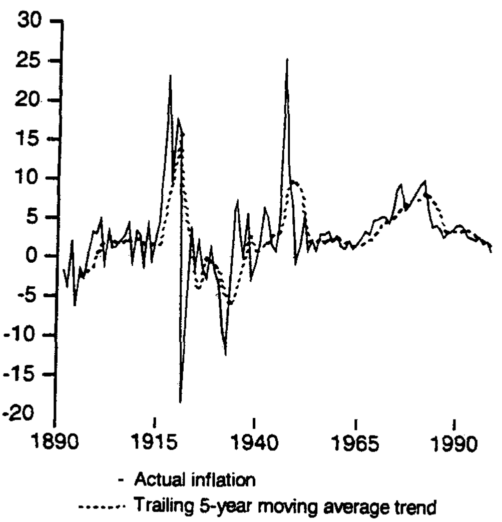


Figure 2 Inflation: percent change in GDP deflator. Data source: Farmer (1999).

1990s—a situation that has been labeled a puzzle by some commentators, given the strong growth of real GDP over the same period.

SHORT-TERM MOVEMENTS IN OUTPUT AND INFLATION

Figures 3 and 4 plot short-term movements in output and inflation. The "output gap" (private real GDP minus the consumption-based trend) consistently falls during recessions. The "inflation gap" (inflation minus the moving average trend) usually falls during recessions but is sometimes observed to rise during these periods (Figure 3). An example is the

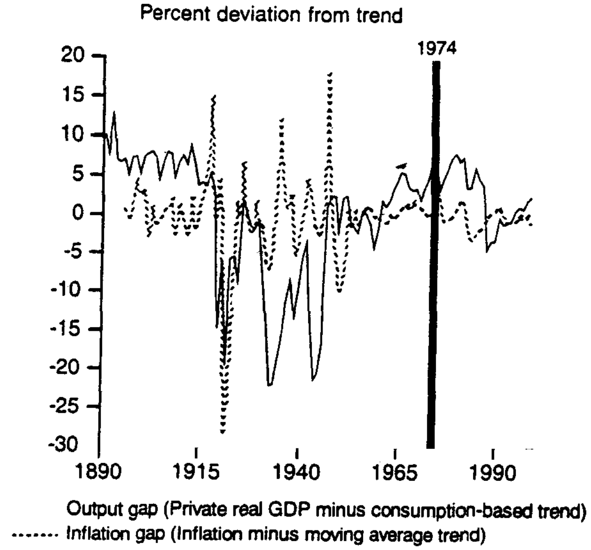


Figure 3 Short-term movements in output and inflation.

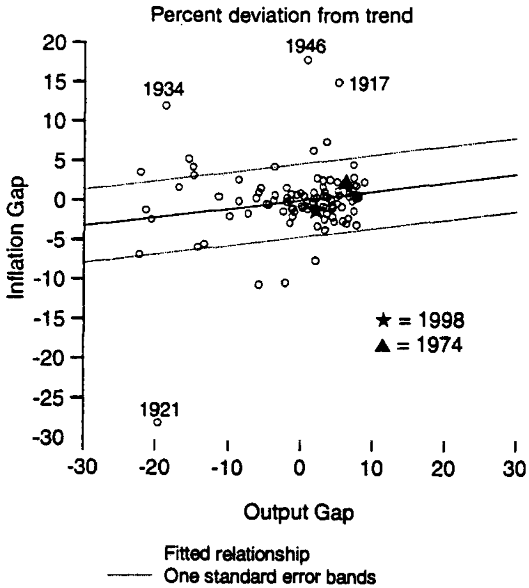


Figure 4 Scatterplot of short-term movements.

1974 recession when OPEC production cutbacks led to a 68 percent increase in the price of crude oil. Thus, the presence of transitory supply-shocks (which by definition should be separated from shifts in the long-run trend) help to explain why the output and inflation gaps do not always move in the same direction. Notice that the two gap measures appear to be going in opposite directions at the end of our data sample in 1998 (Figure 3). It is worth noting that the price of crude oil dropped by nearly 40 percent in 1998—an event that can be viewed as a favorable transitory supply shock.

Despite periods when the two gap measures move in opposite directions, the correlation between short-term movements in output and inflation is positive for the whole sample period (correlation coefficient of 0.18), the post-WWII sample period (correlation coefficient of 0.20), and the 1917–1946 sample period, which includes two world wars and the Great Depression (correlation coefficient of 0.34). This positive correlation accounts for the "traditional" view of the short-term output-inflation relationship. It also suggests that transitory demand shocks (which cause the two gap measures to move in the same direction) are an important feature of the data.

Data for 1998 place the U.S. economy close to the center of the scatter diagram with an inflation rate that is slightly below the best-fit regression line (Figure 4). Given the statistical uncertainty in the position of the best-fit line and the possibility of a transitory supply shock from falling oil prices, the 1998 output-inflation combination should not be viewed as "puzzle" or a breakdown in the historical relationship.

REFERENCES

Cochrane, J. H. 1994. Permanent and Transitory Components of GNP and Stock Prices. *Quarterly Journal of Economics* 109(1), pp. 241–265.

Cogley, T. 1998. A Simple Adaptive Measure of Core Inflation. Federal Reserve Bank of San Francisco Working Paper 98-06.

Cogley, T. and Schaan, D. 1995. Using Consumption to Track Movements in Trend GDP. *FRBSF Weekly Letter* 95-28 (September 1).

Farmer, R. E. A. 1999. *Macroeconomics.* South-Western College Publishing.

Federal Reserve Bank of San Francisco. 1999. U.S. Monetary Policy: An Introduction. *FRBSF Economic Letter* 99-01 (January 1).

Friedman, M. 1968. The Role of Monetary Policy. *American Economic Review* 58, pp. 1–17.

Phelps, E. S. 1967. Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time. *Economica* 34, pp. 254–281.

Phillips, A. W. 1958. The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. *Economica* 25, pp. 283–289.

35

Inflation and Growth

Robert J. Barro

Professor of Economics, Harvard University, Cambridge, Massachusetts

In recent years, many central banks have placed increased emphasis on price stability. Monetary policy—whether expressed in terms of interest rates or growth of monetary aggregates—has been increasingly geared toward the achievement of low and stable inflation. Central bankers and most other observers view price stability as a worthy objective because they think that inflation is costly. Some of these costs involve the average rate of inflation, and others relate to the variability and uncertainty of inflation. But the general idea is that businesses and households are thought to perform poorly when inflation is high and unpredictable.

The academic literature contains a lot of theoretical work on the costs of inflation, as reviewed recently by Briault (1995). This analysis provides a presumption that inflation is a bad idea, but the case is not decisive without supporting empirical findings. Although some empirical results (also surveyed by Briault) suggest that inflation is harmful, the evidence is not overwhelming. It is therefore important to carry out additional empirical research on the relation between inflation and economic performance. This article explores this relation in a large sample of countries over the last 30 years.

DATA

The data set covers over 100 countries from 1960 to 1990. Table 1 provides information about the behavior of inflation in this sample. Annual inflation rates were computed in most cases from consumer price indexes. (The deflator for the gross domestic product was used in a few instances, when the data on consumer prices were unavailable.) Table 1 shows the mean and median across the countries of the inflation rates in three decades: 1960–70, 1970–80, and 1980–90. The median inflation rate was 3.3 percent per year in the 1960s (117 countries), 10.1 percent in the 1970s (122 countries), and 8.9 percent in the 1980s (119 countries). The upper panel of Figure 1 provides a histogram for the inflation rates observed over the three decades. The bottom panel applies to the 44 observations for which the inflation rate exceeded 20 percent per year.[1]

The annual data were used for each country over each decade to compute a measure

Reprinted from: *Federal Reserve Bank of St. Louis Review* (May/June 1996) 153–169.

of inflation variability, the standard deviation of the inflation rate around its decadal mean. Table 1 shows the mean and median of these standard deviations for the three decades. The median was 2.4 percent per year in the 1960s, 5.4 percent in the 1970s, and 4.9 percent in the 1980s. Thus, a rise in inflation variability accompanied the increase in the average inflation rate since the 1960s.

Figure 2 confirms the well-known view that a higher variability of inflation tends to accompany a higher average rate of inflation.[2] These charts provide scatter plots of the standard deviation of inflation (measured for each country around its own decadal mean) against the average inflation rate (the mean of each country's inflation rate over the decade). The upper panel considers only inflation rates below 15 percent per year, the middle panel includes values above 15 percent per year, and the lower panel covers the entire range. The positive, but imperfect, relation between variability and mean is apparent throughout.

Table 1 also gives the means and medians of the growth rate of real per capita gross domestic product (GDP) and the ratio of investment to GDP for the three decades. The median growth rate fell from 3.1 percent in the 1960s (118 countries) to 2.5 percent in the 1970s (123 countries) and 0.4 percent in the 1980s (121 countries). The median investment

Table 1 Descriptive Statistics on Inflation, Growth, and Investment*

Variable	Mean	Median	Number of countries
1960–70			
Inflation rate	.054	.033	117
Standard deviation of inflation rate	.039	.024	117
Growth rate of real per capita GDP	.028	.031	118
Ratio of investment to GDP	.168	.156	119
1970–80			
Inflation rate	.133	.101	122
Standard deviation of inflation rate	.075	.054	122
Growth rate of real per capita GDP	.023	.025	123
Ratio of investment to GDP	.191	.193	123
1980–90			
Inflation rate	.191	.089	119
Standard deviation of inflation rate	.134	.049	119
Growth rate of real per capita GDP	.003	.004	121
Ratio of investment to GDP	.174	.173	128

* The inflation rate is computed on an annual basis for each country from data on consumer price indexes (from the World Bank; STARS databank; issues of *World Tables*; International Monetary Fund; *International Financial Statistics*—yearbook issues; and individual country sources). In a few cases, figures on the GDP deflater were used. The average inflation rate for each country in each decade is the mean of the annual rates. The standard deviation for each country in each decade is the square root of the average squared difference of the annual inflation rate from the decadal mean. The values shown for inflation in this table are the mean or median across the countries of the decade-average inflation rates. Similarly, the figures for standard deviations are the mean or median across the countries of the standard deviations for each decade. The growth rates of real per capita GDP are based on the purchasing power–adjusted GDP values complied by Summers and Heston (1993). For the 1985–90 period, some of the figures come from the World Bank (and are based on market exchange rates rather than purchasing-power comparisons). The ratios of real investment (private plus public) to real GDP come from Summers and Heston (1993). These values are averages for 1960–69, 1970–79, and 1980–89.

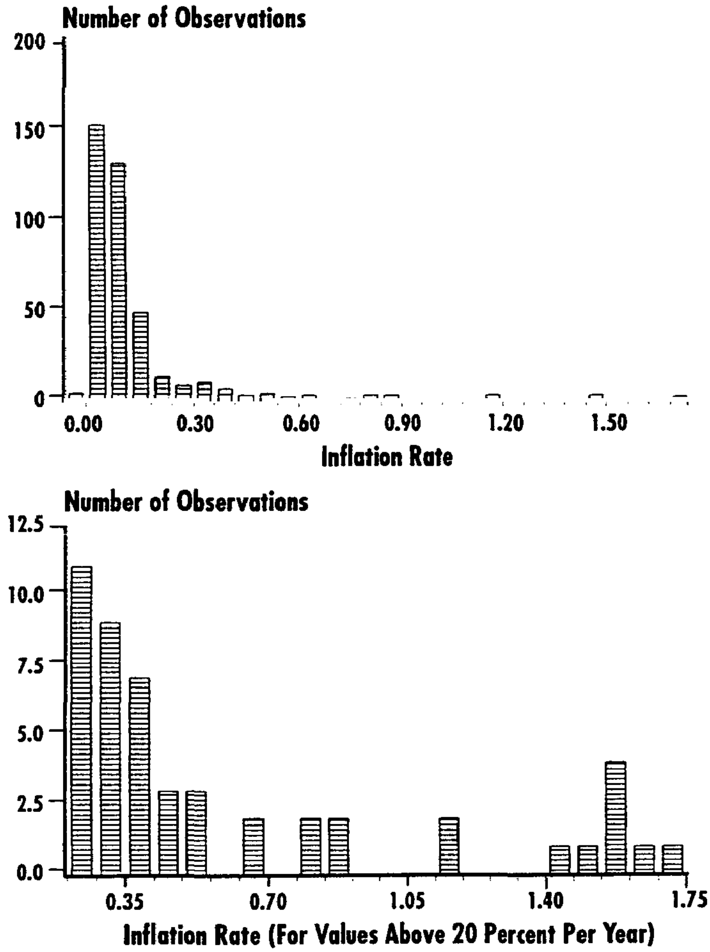


Figure 1 Histograms for inflation rate.

ratio went from 16 percent in the 1960s to 19 percent in the 1970s and 17 percent in the 1980s. In contrast to inflation rates, the growth rates and investment ratios tend to be symmetrically distributed around the median.

FRAMEWORK FOR THE ANALYSIS OF GROWTH

To assess the effect of inflation on economic growth, I use a system of regression equations in which many other determinants of growth are held constant. The framework is based on an extended view of the neoclassical growth model, as described in Barro and Sala-i-Martin (1995, Chapters 1 and 2). My empirical implementations of this approach include Barro (1991 and 1996).

A general notion in the framework is that an array of government policies and private-sector choices determine where an economy will go in the long run. For example, favorable public policies—including better maintenance of the rule of law and property rights, fewer distortions of private markets, less nonproductive government consumption,

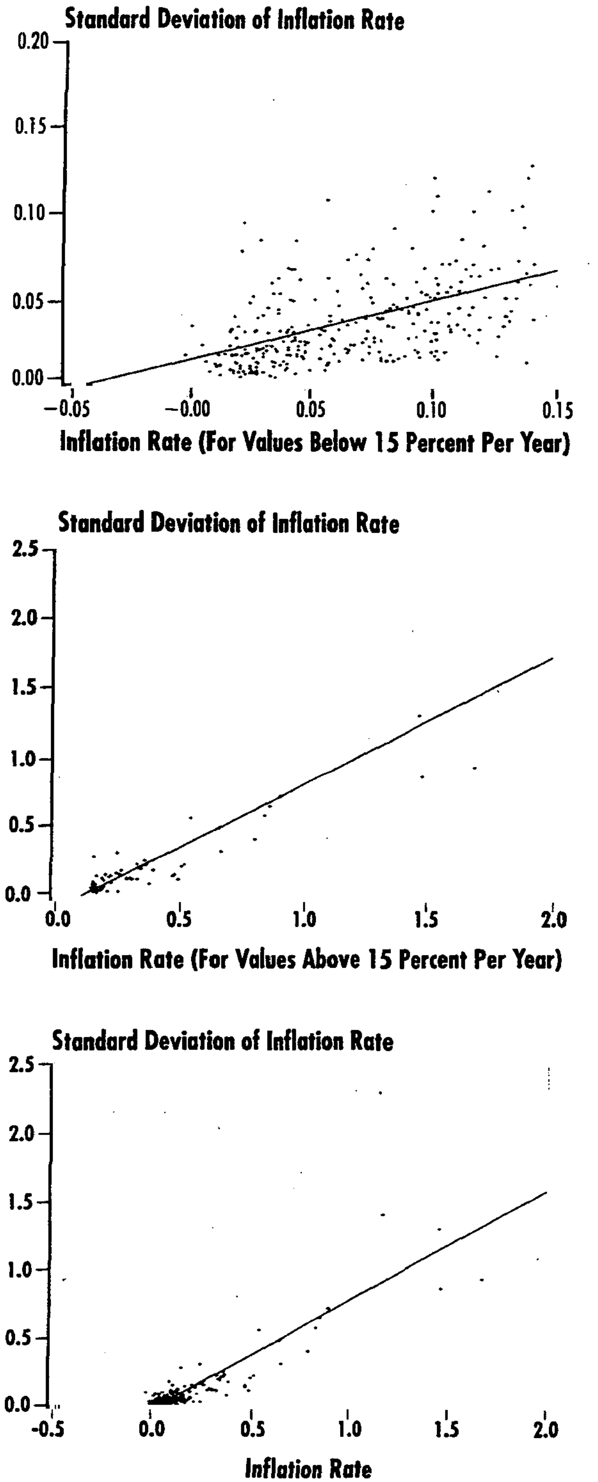


Figure 2 Standard deviation of inflation versus mean inflation.

and greater public investment in high-return areas—lead in the long run to higher levels of real per capita GDP. (Henceforth, the term *GDP* will be used as a shorthand to denote real per capita GDP.) Similarly, a greater willingness of the private sector to save and a reduced tendency to expend resources on child rearing (lower fertility and population growth) tend to raise standards of living in the long run.

Given the determinants of the long-run position, an economy tends currently to grow faster the lower its GDP. In other words, an economy's per capita growth rate is increasing in the gap between its long-term prospective GDP and its current GDP. This force generates a convergence tendency in which poor countries grow faster than rich countries and tend thereby to catch up in a proportional sense to the rich places. However, poor countries grow fast only if they have favorable settings for government policies and private-sector choices. If a poor country selects unfavorable policies—a choice that likely explains why the country is currently observed to be poor—then its growth rate will not be high and it will not tend to catch up to the richer places.

Another important element is a country's human capital in the forms of education and health. For given values of prospective and actual GDP, a country grows faster—that is, approaches its long run position more rapidly—the greater its current level of human capital. This effect arises because, first, physical capital tends to expand rapidly to match a high endowment of human capital, and, second, a country with more human capital is better equipped to acquire and adapt the efficient technologies that have been developed in the leading countries.

PANEL ESTIMATES OF GROWTH EQUATIONS

Overview of the Results

Table 2 lists the explanatory variables used as determinants of the growth rate of real per capita GDP. The details for a similar setup are in Barro (1996). The results apply to growth rates and the other variables observed for 78 countries from 1965 to 1975, 89 countries for 1975 to 1985, and 84 countries from 1985 to 1990. This sample reflects the availability of the necessary data. The first period starts in 1965, rather than 1960, so that five-year lags of the explanatory variables are available.

The estimation is by instrumental variables, where the instruments consist mainly of prior values of the regressors. For example, the 1965–75 equation includes the log of 1965 GDP on the right-hand side and uses the log of 1960 GDP as an instrument. This procedure should lessen the estimation problems caused by temporary measurement error in GDP. The right-hand side also contains period averages of several variables—government spending ratios, fertility rates, black-market premia, and investment ratios—and uses five-year earlier values of these variables as instruments.

The use of lagged variables as instruments is problematic, although better alternatives are not obvious. One favorable element here is that the residuals from the growth equations turn out to be virtually uncorrelated over the time periods. In most respects, the instrumental results do not differ greatly from OLS estimates. The largest difference turns out to be for the estimated effect of the investment ratio on the growth rate.

Since the general pattern of results has been considered elsewhere,[3] I will provide only a brief sketch here and will focus the main discussion on the effects of inflation. One familiar finding in Table 2 is that the estimated coefficient on initial log (GDP) is significantly negative with a magnitude of around 2.5 percent. Thus, conditional on the other vari-

Table 2 Regressions for Per Capita Growth Rate *

Variable	1	2	3	4	5
Log(GDP)	–.0241 (.0030)	–.0242 (.0030)	–.0246 (.0029)	–.0242 (.0030)	–.0231 (.0029)
Male schooling	.0144 (.0036)	.0145 (.0036)	.0146 (.0036)	.0136 (.0037)	.0116 (.0038)
Female schooling	–.0100 (.0049)	–.0100 (.0049)	–.0104 (.0050)	–.0087 (.0052)	–.0063 (.0053)
Log(life expectancy)	.0359 (.0122)	.0354 (.0121)	.0381 (.0124)	.0333 (.0120)	.0407 (.0130)
Log(GDP)*human capital	–.44 (.17)	–.44 (.17)	–.41 (.17)	–.47 (.17)	–.45 (.16)
Log(fertility rate)	–.0175 (.0053)	–.0175 (.0053)	–.0173 (.0053)	–.0176 (.0053)	–.0146 (.0055)
Government consumption ratio	–.117 (.027)	–.116 (.027)	–.118 (.027)	–.120 (.027)	–.115 (.026)
Public education spending ratio	.114 (.090)	.112 (.089)	.146 (.090)	.081 (.091)	.057 (.091)
Black-market premium	–.0127 (.0052)	–.0125 (.0051)	–.0150 (.0050)	–.0109 (.0055)	–.0137 (.0054)
Rule-of-law index	.00426 (.00093)	.00424 (.00093)	.00426 (.00093)	.00418 (.00095)	.00404 (.00093)
Terms-of-trade change	.126 (.028)	.127 (.028)	.129 (.028)	.123 (.028)	.117 (.028)
Investment ratio	.019 (.022)	.020 (.022)	.018 (.022)	.024 (.022)	.013 (.022)
Democracy index	.063 (.025)	.063 (.025)	.066 (.025)	.059 (.026)	.066 (.026)
Democracy index squared	–.064 (.022)	–.063 (.022)	–.067 (.023)	–.060 (.023)	–.066 (.023)
Inflation rate	–.0236 (.0048)	–.0209 (.0082)	–.0197 (.0069)	–.0306 (.0083)	–.0254 (.0086)
Standard deviation of inflation rate	—	–.0036 (.0086)	—	—	—
Latin America dummy	—	—	—	—	–.0060
R^2	.63, .60, .48	.63, .60, .49	.64, .60, .46	.63, .59, .48	.63, .61, .49
Number of observations	78, 89, 84	78, 89, 84	78, 89, 84	78, 89, 84	78, 89, 84

* Standard errors of the coefficient estimates are shown in parentheses. The R^2 values and numbers of observations apply to each period individually.

The systems have three equations, where the dependent variables are the growth rate of real per capita GDP for 1965–75, 1975–85, and 1985–90. The variables GDP (real per capita GDP) and schooling (years of attainment at the secondary and higher levels) refer to 1965, 1975, and 1985. Life expectancy at birth is for 1960–64, 1970–74, and 1980–84 The rule-of-law index applies to the early 1980s (one observation for each country). The terms-of-trade variable is the growth rate over each period of the ratio of export to import prices. The variable log(GDP)* human capital is the product of log(GDP) (expressed as a deviation from the sample mean) and the estimated effect of the schooling and life-expectancy variables (also expressed as deviations from sample means). Variables measured as averages over each period are the log of the total fertility rate, the ratio of government consumption (exclusive of defense and education) to GDP, the ratio of public educational spending to GDP, the black-market premium on foreign exchange, the ratio of gross investment (public plus private) to GDP, the Gastil/Bollen indexes of political rights (where 0 indicates the fewest rights and 1 the most), and the CPI inflation rate. The standard deviation of the inflation rate is measured from annual observations in relation to the mean inflation rate for each period. The Latin American dummy equals 1 for countries in Latin America and 0 otherwise. Individual constants (not shown) are estimated for each period. Estimation is by instrumental variables. Columns 1 and 2 include actual inflation as an instrument. Column 2 also uses the standard deviation of inflation as an instrument. Column 3 uses the five-year earlier values of inflation as instruments. Columns 4 and 5 use prior colonia status as instruments. Column 5 also uses the Latin America dummy as an instrument. The other instruments for all columns are the five-year earlier value of log(GDP) (for example, for 1960 in the 1965–75 equation); the schooling, life-expectancy, rule-of-law, and terms-of-trade variables; and earlier values of the other variables. For example, the 1965–75 equation uses the averages of the black-market premium and the government-spending and investment ratios for 1960–64. The estimation allows for different error variances in each period and for correlation across these errors. The estimated correlation of the errors for column 1 is—0.13 between the 1965–75 and 1975–85 equations, 0.10 between the 1965–75 and 1985–90 equations, and 0.07 between the 1975–85 and 1985–90 equations. The pattern is similar for the other columns. The estimates are virtually the same if the errors are assumed to be independent over the time periods.

ables, convergence in real per capita GDP occurs at roughly 2.5 percent per year.[4] Growth tends also to be increasing in the initial levels of human capital in the forms of education (average years of school attainment at the secondary and higher levels) and health (proxied by the log of life expectancy at birth). The negative coefficient on the interaction term between initial GDP and human capital[5] means that the rate of convergence is higher in a place that starts with more human capital.

For given starting values of the state variables (represented by initial human capital and GDP), growth is estimated to fall with higher fertility (the average woman's total fertility rate), higher government consumption (the ratio to GDP of government consumption, exclusive of spending on education and defense), and a larger black-market premium on foreign exchange (intended as a proxy for market distortions more broadly).

Growth is enhanced by greater maintenance of the rule of law, as measured by Knack and Keefer's (1995) subjective index. One problem here is that this variable is observed only in the early 1980s (and is included among the instruments). Growth also rises in response to a contemporaneous improvement in the terms of trade, measured by the growth rate of the ratio of export prices to import prices. (The contemporaneous terms-of-trade change is included with the instruments.)

The estimated coefficients on the ratio of public educational spending to GDP and on the ratio of total real investment to real GDP are positive, but insignificant. The estimated coefficient on investment becomes higher and significant if the contemporaneous investment ratio is included with the instruments. (The timing in the data indicates that much of the positive association between investment and growth represents the reverse response of investment to growth.) The estimate becomes even larger and resembles that reported in other studies, such as Mankiw, Romer, and Weil (1992), if life expectancy is deleted as a regressor.

Finally, an increase in democracy—measured by indexes of political rights from Gastil (1982–83) and Bollen (1990)—have a nonlinear effect (which I did not find for log[GDP] or the human-capital variables). At low levels of democracy, more freedom is estimated to raise growth. But once a moderate level of democracy is attained (corresponding roughly to "half" the way toward full representative democracy), further liberalization is estimated to reduce growth. These effects are discussed at length in Barro (1996).

Preliminary Results on Inflation

To get a first-pass estimate of the effect of inflation on economic growth, I included the inflation rate over each period as an explanatory variable along with the other growth determinants listed in Table 2. If contemporaneous inflation is also included with the instruments, then column 1 of Table 2 indicates that the estimated coefficient of inflation is –0.024 (s.e. = 0.005). Thus, an increase by 10 percentage points in the annual inflation rate is associated on impact with a decline by 0.24 percentage points in the annual growth rate of GDP. Since the "t-statistic" for the estimated coefficient is 4.9, this result is statistically significant.[6]

Figure 3 depicts graphically the relation between growth and inflation. The horizontal axis plots the inflation rate; each observation corresponds to the average rate for a particular country over one of the time periods considered (1965–75, 1975–85, and 1985–90). The top panel in the chart considers inflation rates below 15 percent per year, the middle panel includes values above 20 percent per year, and the bottom panel covers the full range of inflation. The vertical axis plots the growth rate of GDP, *net of the part of the growth rate that is explained by all of the explanatory variables aside from the inflation rate.*[7] Thus, the panels illustrate the relation between growth and inflation after all of the other growth determinants have been held constant.

The panels of Figure 3 show downward-sloping regression lines (least-squares lines) through the scatter plots. The slope of the line in the lower panel corresponds approximately to the significantly negative coefficient shown in column 1 of Table 2. The panels show, however, that the fit is dominated by the inverse relation between growth and inflation at high rates of inflation. For inflation rates below 20 percent per year, as shown in the upper panel, the relation between growth and inflation is not statistically significant.

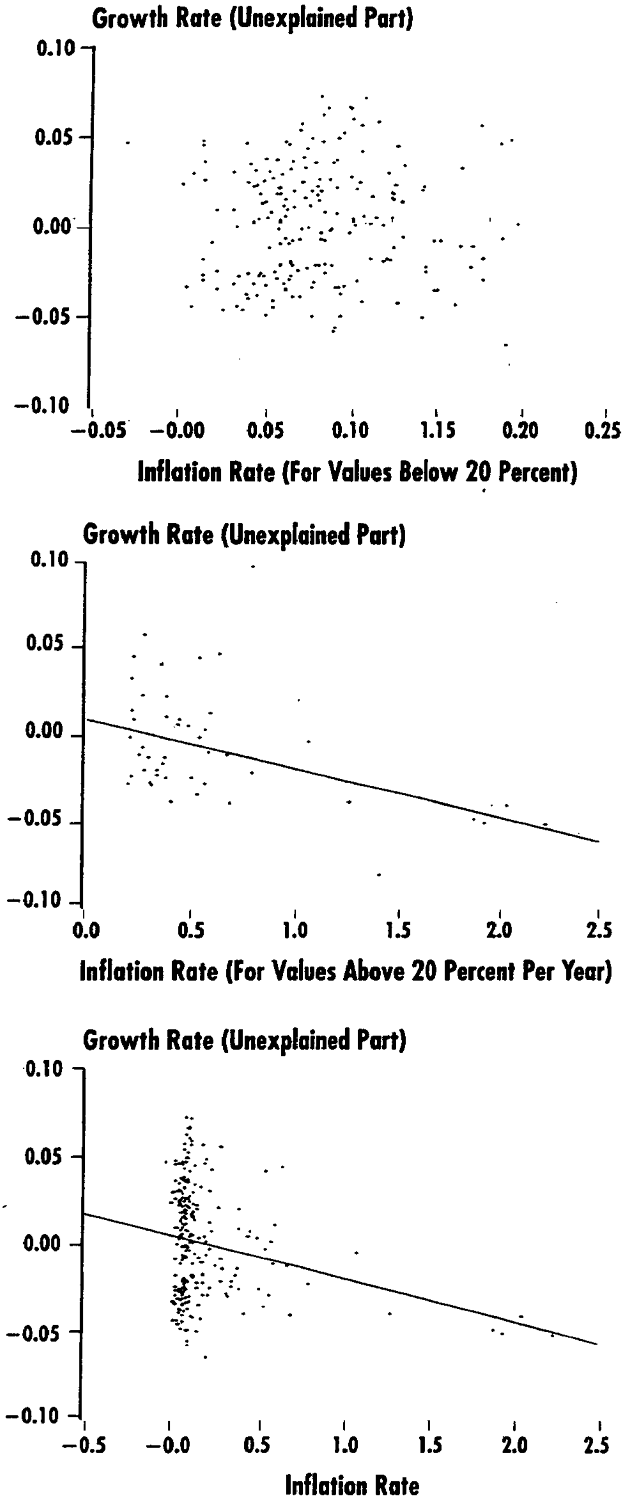


Figure 3 Growth rate versus inflation rate.

To put it another way, one can reestimate the panel while restricting the observations to those for which the inflation rate is less than some cutoff value, x. To get a statistically significant estimate for the inflation coefficient, x has to be raised to roughly 50 percent per year. With an inflation cutoff of 50 percent, the estimated coefficient is –0.029 (0.015). For lower values of the cutoff, the estimated coefficient tends to be negative but insignificant; some results are $x = 40$ percent, coefficient = –0.023 (0.018); $x = 25$ percent, coefficient = –0.011 (0.027); $x = 15$ percent, coefficient = –0.032 (0.042).

The results indicate that there is not enough information in the low-inflation experiences to isolate precisely the effect of inflation on growth, but do not necessarily mean that this effect is small at low rates of inflation. To check for linearity of the relation between growth and inflation, I reestimated the system on the whole sample with separate coefficients for inflation in three ranges: up to 15 percent, between 15 percent and 40 percent, and over 40 percent. The estimated coefficients on inflation in this form are –0.016 (0.035) in the low range, –0.037 (0.017) in the middle range, and –0.023 (0.005) in the upper range. Thus, the clear evidence for the negative relation between growth and inflation comes from the middle and upper intervals. However, since the three estimated coefficients do not differ significantly from each other (p-value = 0.65), the data are consistent with a linear relationship. In particular, even at low rates of inflation, the data would not reject the hypothesis that growth is negatively related to inflation.

Although statistically significant effects arise only when the high-inflation experiences are included, the results are not sensitive to a few outlier observations. Table 3 shows the 27 cases of inflation in excess of 40 percent per year for one of the time periods (1965–75, 1975–85, and 1985–90). Note that Uruguay appears three times (although it is by no means the overall champion for high inflation), and Argentina, Brazil, Peru, Uganda, and Zaire show up twice each. The other countries, with one observation each, are Chile, Indonesia, Bolivia, Haiti, Israel, Guinea-Bissau, Mexico, Mozambique, Nicaragua, Poland, Sierra Leone, Turkey, Yugoslavia, and Zambia. (Guinea-Bissau, Mozambique, and Poland are excluded from the regression sample because of missing data on other variables.) The exclusion of any small number of these high-inflation observations—Nicaragua and Zaire had been suggested to me—has a negligible effect on the results.

The estimates are also reasonably stable over time. If different coefficients for inflation are allowed for each period, but the system is otherwise the same as in column 1 of Table 2, then the resulting estimates are –0.019 (0.015) for 1965–75, –0.029 (0.010) for 1975–85, and –0.023 (0.005) for 1985–90. These values do not differ significantly from each other (p-value = 0.20). (The higher significance of the estimated coefficients in the two later periods reflects the larger number of high-inflation observations.)

The standard deviation of inflation can be added to the system to see whether inflation variability has a relation with growth when the average inflation rate is held constant. The strong positive correlation between the mean and variability of inflation (Figure 2) suggests that it would be difficult to distinguish the influences of these two aspects of inflation. However, when the two variables are entered jointly into the regression system in column 2 of Table 2, the estimated coefficient on inflation remains similar to that found before (–0.021 [0.008]), and the estimated coefficient on the standard deviation of inflation is virtually zero (–0.004 [0.009]).[8] Thus, for a given average rate of inflation, the variability of inflation has no significant relation with growth. One possible interpretation of this result is that the realized variability of inflation over each period does not adequately measure the uncertainty of inflation, the variable that one would have expected to be negatively related to growth. This issue is worth further investigation.

Table 3 High-Inflation Observations—Period-Average Inflation Rate Exceeds 40 Percent Per Year

	Country	Inflation rate
1965–75		
	Chile	0.68
	Indonesia	0.53
	Uruguay	0.50
1975–85		
	Argentina	1.26
	Bolivia	1.06
	Brazil	0.66
	Haiti	0.48
	Israel	0.78
	Peru	0.56
	Uganda	0.53
	Uruguay	0.41
	Zaire	0.44
1985–90		
	Argentina	1.92
	Brazil	2.04
	Guinea-Bissau*	0.53
	Mexico	0.53
	Mozambique*	0.48
	Nicaragua	1.87
	Peru	2.22
	Poland*	0.81
	Sierra Leone	0.63
	Turkey	0.43
	Uganda	0.78
	Uruguay	0.58
	Yugoslavia	1.41
	Zaire	0.59
	Zambia	0.56

* Not included in regression sample because of missing data on other variables.

The Endogeneity of Inflation

A key problem in the interpretation of the results is that they need not reflect causation from inflation to growth. Inflation is an endogenous variable, which may respond to growth or to other variables that are related to growth. For example, an inverse relation between growth and inflation would arise if an exogenous slowing of the growth rate tended to generate higher inflation. This increase in inflation could result if monetary authorities reacted to economic slowdowns with expansionary policies. Moreover, if the path of monetary aggregates did not change, then the equality between money supply and demand at each point in time implies that a reduction in the growth rate of output would tend automatically to raise the inflation rate.

Narayana Kocherlakota's commentary uses a sophisticated cash-in-advance theory to focus on this last source of endogeneity bias. To fix ideas, suppose as he does that the money growth rate, μ_t, is determined exogenously. The relation between the inflation rate, π_t, and μ_t is given from the money-supply-equals-money-demand condition by

$$\pi_t = \mu_t - g_t + v_t, \tag{1}$$

where g_t is the growth rate of output and v_t is an independent shock to velocity (which Kocherlakota takes to be nil). Suppose that the effect of inflation on the growth rate is described by

$$g_t = -\alpha\pi_t + \varepsilon_t, \tag{2}$$

where e_t is an independent shock and α is the coefficient that we want to estimate.

In this framework, the OLS regression coefficient, $\hat{\beta}$, of g_t on, π_t can be shown to be given by

$$\hat{\beta} = -\left\{\frac{\alpha\cdot[VAR(\mu) + VAR(v)] + VAR(\varepsilon)}{VAR(\mu) + VAR(v) + VAR(\varepsilon)}\right\} \tag{3}$$

Thus, $\hat{\beta}$ will be close to α if the variances of the shocks to money growth and velocity are much greater than those to output growth.

Kocherlakota uses data from my macroeconomics textbook, normally an impeccable source, to estimate the variances. However, these figures are inappropriate for my panel estimation, which applies to different time periods. Using averages over the three periods in the panel and measuring money by either M1 or M2, I find that VAR(μ) $\approx$ .032, VAR(v) $\approx$ .004, and VAR(e) $\approx$ .0002. Then a value $\alpha = 0$ corresponds to $\hat{\beta} \approx -.006$, and a value $\alpha = -.020$ corresponds to $\hat{\beta} \approx -.026$. That is, the bias would be small and could not account for the empirical findings. Moreover, if inflation, rather than money growth, were determined exogenously, then the bias would be nil.

The Kocherlakota argument also implies that the results would be very different if one included money growth, rather than inflation, in the growth regressions. If the panel estimation from column 1 of Table 2 is redone with M1 growth replacing the inflation rate, then the estimated coefficient is –0.0225 (0.0057). If M2 growth is used instead, the result is –0.0191 (0.0054). Thus, the estimated coefficient is about the same with money growth as with inflation. Basically, the results reveal an inverse relation between the growth rate of GDP and the growth rate of prices *or* money. The distinction between inflation and money growth has nothing to do with the findings.

It is also possible that the endogeneity of inflation would produce a positive relation between inflation and growth. This pattern tends to emerge if output fluctuations are driven primarily by shocks to money or to the aggregate demand for goods.

Another possibility is that some omitted third variable is correlated with growth and inflation. For example, better enforcement of property rights is likely to spur investment and growth and is also likely to accompany a rules-based setup in which the monetary authority generates a lower average rate of inflation. The idea is that a committed monetary policy represents the application of the rule of law to the behavior of the monetary authority. Some of the explanatory variables in the system attempt to capture the degree of maintenance of the rule of law. However, to the extent that these measures are imperfect, the inflation rate may proxy inversely for the rule of law and thereby show up as a negative influence on growth. The estimated coefficient on the inflation rate could therefore reflect an effect on growth that has nothing to do with inflation per se.

Some researchers like to handle this type of problem by using some variant of fixed-effects estimation; that is, by allowing for an individual constant for each country. This procedure basically eliminates cross-sectional information from the sample and therefore relies on effects within countries from changes over time in inflation and other variables. It is not apparent that problems of correlation of inflation with omitted variables would be less serious in this time-series context than in cross-sections. (If a country is undergoing an inflation crisis or implementing a monetary reform, then it is likely to be experiencing other crises or reforms at the same time.) Moreover, the problems with measurement error and timing of relationships would be more substantial in the time series. The one thing that is clear is that fixed-effects procedures eliminate a lot of information.

Another way to proceed is to find satisfactory instrumental variables—reasonably exogenous variables that are themselves significantly related to inflation. My search along these lines proceeded along the sequence now described.

Central Bank Independence

One promising source of instruments for inflation involves legal provisions that more or less guarantee central bank independence. Recent literature argues that a greater degree of independence leads to lower average rates of money growth and inflation and to greater monetary stability.[9] The idea is that independence enhances the ability of the central bank to commit to price stability and, hence, to deliver low and stable inflation. Alesina and Summers (1993, Figures 1a and b) find striking negative relationships among 16 developed countries from 1955 to 1988 between an index of the degree of central bank independence and the mean and variance of inflation. Thus, in their context, the measure of central bank independence satisfies one condition needed for a good inflation instrument; it has substantial explanatory power for inflation.

Because of the difficulty of enacting changes in laws, it is plausible that a good deal of the cross-country differences in legal provisions that influence central bank independence can be treated as exogenous. Problems arise, however, if the legal framework changes in response to inflation (although the sign of this interaction is unclear). In addition, exogeneity would be violated if alterations in a country's legal environment for monetary policy are correlated with changes in unmeasured institutional features—such as structures that maintain property rights—that influence growth rates. This problem is, however, mitigated by the inclusion of other explanatory variables, notably the index of the rule of law, in the regression framework.

Cukierman (1992, chapter 19) argues that the legal provisions that govern central bank action differ substantially from the way that the banks actually operate. In particular, he distinguishes the legal term of office of the central bank governor from the observed turnover. The latter variable would be more closely related to bank performance (and, hence, to inflation), but cannot be treated as exogenous to growth or omitted third variables. Thus, for the purpose of constructing instruments for inflation, the preferred strategy is to focus on the extent to which inflation can be explained by differences in legal provisions for the central bank.

Table 4 shows an index of central bank independence for 67 countries, based on the information compiled by Cukierman (1992, chapter 19, Appendix A) over time periods that correspond roughly to the four decades from the 1950s to the 1980s. The index is an average over the time periods and for numerous categories of legal provisions contained in the charters of the central banks. (See the notes to Table 4.) The details of construction differ somewhat from those used by Cukierman, but the values shown in Table 4 are similar to those reported in his Table 19.3 for the 1980s.

Table 4 also contains the average inflation rate from 1960 to 1990 for the 67 countries in my sample that have data on the index of central bank independence. A comparison

Table 4 Inflation Rates and Central Bank Independence*

Country	Index of bank independence	Inflation rate 1960–90	Country	Index of bank independence	Inflation rate 1960–90
West Germany	0.71	0.037	Zambia	0.34	0.174
Switzerland	0.65	0.038	South Africa	0.33	0.099
Austria	0.65	0.043	Nigeria	0.33	0.125
Egypt	0.57	0.094	Malaysia	0.32	0.034
Denmark	0.53	0.069	Uganda	0.32	0.353
Costa Rica	0.52	0.117	Italy	0.31	0.088
Greece	0.52	0.109	Finland	0.30	0.073
United States	0.51	0.049	Sweden	0.30	0.067
Ethiopia	0.50	0.058	Singapore	0.30	0.034
Ireland	0.50	0.083	India	0.30	0.074
Philippines	0.49	0.107	United Kingdom	0.30	0.077
Bahamas	0.48	0.063†	South Korea	0.29	0.11
Tanzania	0.48	0.133	China	0.29	0.039
Nicaragua	0.47	0.436	Bolivia	0.29	0.466
Israel	0.47	0.350	Uruguay	0.29	0.441
Netherlands	0.47	0.045	Brazil	0.28	0.723
Canada	0.47	0.054	Australia	0.27	0.067
Venezuela	0.45	0.100	Thailand	0.27	0.052
Barbados	0.44	0.075	Western Samoa	0.26	0.112‡
Argentina	0.44	0.891	New Zealand	0.25	0.085
Honduras	0.44	0.058	Nepal	0.23	0.084
Peru	0.44	0.606	Panama	0.23	0.033
Chile	0.43	0.416	Zimbabwe	0.22	0.074
Turkey	0.42	0.235	Hungary	0.21	0.047
Malta	0.42	0.035	Japan	0.20	0.054
Iceland	0.42	0.229	Pakistan	0.19	0.072
Kenya	0.40	0.082	Colombia	0.19	0.170
Luxembourg	0.40	0.044	Spain	0.16	0.096
Zaire	0.39	0.357	Morocco	0.15	0.055
Mexico	0.37	0.227	Belgium	0.13	0.048
Indonesia	0.36	0.366	Yugoslavia	0.12	0.395
Botswana	0.36	0.076	Poland	0.12	0.293†
Ghana	0.35	0.256	Norway	0.12	0.066
France	0.34	0.064			

* The index of central bank independence is computed from data in Cukierman (1992, chapter 19, Appendix A). The index is a weighted average of the available data from 1950 to 1989 of legal provisions regarding (1) appointment and dismissal of the governor (weight 1/6), (2) procedures for the formulation of monetary policy (weight 1/6), (3) objectives of central bank policy (weight 1/6), and (4) limitations on lending by the central bank (weight 1/2). The first category is an unweighted average of three underlying variables that involve the governor's term of office and the procedures for appointment and dismissal. The second category is an unweighted average of two variables, one indicating the location of the authority for setting monetary policy and the other specifying methods for resolving conflicts about policy. The third category relates to the prominence attached to price stability in the bank's charter. The fourth category is an unweighted average of four variables: limitations on advances, limitations on securitized lending, an indicator for the location of the authority that prescribes lending terms, and the circle of potential borrowers from the central bank. For each underlying variable, Cukierman defines a scale from 0 to 1, where 0 indicates least favorable to central bank independence and 1 indicates most favorable. The overall index shown in Table 4 runs correspondingly from 0 to 1. See Table 1 for a discussion of the inflation data.

‡ 1975–90

† 1970–90

between the index and the inflation rate reveals a crucial problem; the correlation between the two variables is essentially zero, as in clear from Figure 4. This verdict is also maintained if one looks separately over the three decades from the 1960s to the 1980s and if one holds constant other possible determinants of inflation. In this broad sample of countries, differences in legal provisions that ought to affect central bank independence have no explanatory power for inflation.[10] This negative finding is of considerable interest, because it suggests that low inflation cannot be attained merely by instituting legal changes that appear to promote a more independent central bank. However, the result also means that we have to search further for instruments to clarify the relation between growth and inflation.[11]

Lagged Inflation

Earlier values of a country's inflation rate have substantial explanatory power for inflation.[12] Lagged inflation would also be exogenous with respect to innovations in subsequent growth rates. Hence, if lagged inflation is used as an instrument, then the estimated relation between growth and inflation would not tend to reflect the short-run reverse effect of growth on inflation.

One problem, however, is that lagged inflation would reflect persistent characteristics of a country's monetary institutions (such as the extent to which policymakers have credibility), and these characteristics could be correlated with omitted variables that are relevant to growth (such as the extent to which political institutions support the maintenance of property rights). The use of lagged inflation as an instrument would therefore not rule out the problems of interpretation that derive from omitted third variables. However, the inclusion of the other explanatory variables in the regression framework lessens this problem. Another favorable element is that the residuals from the growth equations are not significantly correlated over the time periods.

Column 3 of Table 2 shows the estimated effect of inflation on the growth rate when lagged inflation (over the five years prior to each sample period) is used as an instrument. The estimated coefficient is–0.020 (0.007), similar to that found in column 1 when contemporaneous inflation is included as an instrument. Thus, it seems that most of the esti-

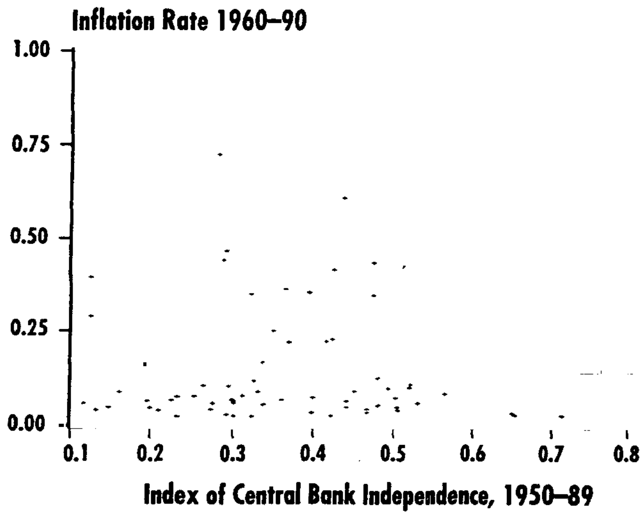


Figure 4 Inflation and central bank independence.

Table 5 Inflation Rates and Prior Colonial Status*

Period	All countries	Non-colony	British colony	French colony	Spanish or Portugese colony	Other colony	Latin America not Spanish or Portugese colony
1960–70	.054	.045	.033	.030	.089	.194	.031
	(121)	(31)	(43)	(21)	(19)	(7)	(7)
1970–80	.131	.110	.120	.093	.218	.147	.109
	(131)	(32)	(50)	(20)	(21)	(8)	(11)
1980–90	.182	.124	.139	.074	.523	.136	.097
	(132)	(31)	(51)	(22)	(20)	(8)	(11)
1960–90	.126	.089	.104	.066	.294	.161	.090
	(117)	(30)	(42)	(20)	(18)	(7)	(7)

* The numbers shown in parentheses are the numbers of countries with available data that fall into each category. See Table 1 for a discussion of the inflation data. Countries that were independent before 1776 are treated as non-colonies. Otherwise, the colonial status refers to the most recent outside power; for example, the Philippines is attributed to the United States rather than to Spain; Rwanda and Burundi are attributed to Belgium rather than to Germany; and the Dominican Republic is attributed to France rather than to Spain. Some countries that were dominated by other countries for some periods are treated as non-colonies. Examples are Hungary, Poland, South Korea, and Taiwan. The only present-day colony in the sample is Hong Kong. The last column refers to countries that are located in Latin America but are not former Spanish or Portugese colonies.

mated negative relation between growth and inflation does not represent reverse short-term (negative) effects of growth on inflation.

The significant negative influence of inflation on growth still shows up only when the high-inflation observations are included. The results are, however, again consistent with a linear relation and with stability over the time periods. The standard deviation of inflation also remains insignificant if it is added to the regressions (with lagged values of this standard deviation included as instruments).

Prior Colonial Status

Another possible instrument for inflation comes from the observation that prior colonial status has substantial explanatory power for inflation. Table 5 breaks down averages of inflation rates from 1960 to 1990 by groups of countries classified as non-colonies (defined as those that were independent prior to U.S. independence in 1776) and former colonies of Britain, France, Spain, or Portugal, and other countries (in this sample, Australia, Belgium, the Netherlands, New Zealand, and the United States).

Table 5 indicates that the average inflation rate for all 117 countries from 1960 to 1990 is 12.6 percent per year. The average for the 30 non-colonies of 8.9 percent is similar to that of 10.4 percent for the 42 British colonies and 6.6 percent for the 20 French colonies. However, the rates are strikingly higher for the 18 Spanish or Portuguese colonies—29.4 percent—and somewhat higher for the 7 other colonies—16.1 percent.

A key reason for the low average inflation rate for the former French colonies is the participation of most of the sub-Saharan African states in the fixed–exchange rate regime of the CFA Franc.[13] This type of reasonably exogenous commitment to relatively low inflation is exactly the kind of experiment that provides for a good instrument for inflation.

For many of the former British colonies, a significant element may be their prior experience with British organized currency boards, another system that tends to generate low inflation.[14] These boards involved, at one time or another before independence, most of the British colonies in Africa, the Caribbean, southeast Asia, and the Middle East.

The high average inflation rate for the 16 former Spanish colonies in the sample does not reflect, per se, their presence in Latin America. For seven Latin American countries that are not former Spanish or Portuguese colonies,[15] the average inflation rate for 1960–90 is only 9.0 percent, virtually the same as that for the non-colonies (see Table 5). Also, four former Portuguese colonies in Africa experienced the relatively high average inflation rate of around 20 percent.[16] For Portugal and Spain themselves, the average inflation rate of 10.9 percent for 1960–90 is well below the rate of 29.4 percent experienced by their former colonies. However, 10.9 percent inflation is substantially higher than that experienced by France (6.4 percent) and the United Kingdom (7.7 percent).

Column 4 of Table 2 shows the estimated effect of inflation on the growth rate of GDP when the instruments exclude contemporaneous or lagged inflation but include indicators of prior colonial status. The two variables used are a dummy for whether the country is a former Spanish or Portuguese colony and a dummy for whether the country is a former colony of a country other than Britain, France, Spain, or Portugal.[17] The estimated coefficient on the inflation rate is now–0.031 (0.008), somewhat higher in magnitude than that found when contemporaneous or lagged inflation is used as an instrument. The significantly negative relation between growth and inflation again arises only when the high-inflation experiences are included in the sample. The results also continue to be stable over the time periods.

One question about the procedure is whether prior colonial status works in the growth regressions only because it serves as an imperfect proxy for Latin America, a region that is known to have experienced surprisingly weak economic growth.[18] However, column 5 of Table 2 shows that if a dummy variable for Latin America is included in the system (and the indicators of prior colonial status and the Latin America dummy are used as instruments), then the estimated coefficient of inflation remains negative and significant, –0.025 (0.009). Moreover, the estimated coefficient on the Latin America dummy is only marginally significant, –0.0060 (0.0034). The results are basically the same if the Latin America dummy is added to the system from column 1 of Table 2, in which contemporaneous inflation is used as an instrument. It therefore appears that much of the estimated effect of a Latin America dummy on growth rates in previous research reflected a proxying of this dummy for high inflation. In particular, the negative effect of inflation on growth does not just reflect the tendency for many high-inflation countries to be in Latin America.

ESTIMATED EFFECTS OF INFLATION ON INVESTMENT

A likely channel by which inflation decreases growth is through a reduction in the propensity to invest. (This effect is already held constant by the presence of the investment ratio in the growth regressions.) I have investigated the determination of the ratio of investment to GDP within a framework that parallels the one set out in Table 2. The results are in Table 6.

In the case of the investment ratio, the use of instruments turns out to be crucial for isolating a negative effect of inflation. In column 1 of Table 6, which uses contemporaneous inflation as an instrument, the estimated coefficient on the inflation rate is virtually zero, –0.001 (0.011). In contrast, the result in column 2 with lagged inflation used as an instrument is –0.059 (0.017). Similarly, the result in column 3 with the indicators of prior colonial status used as instruments is –0.044 (0.022). The last two estimates imply that an

Table 6 Regressions for Investment Ratio*

Variable	1	2	3
Log(GDP)	–.008 (.010)	–.011 (.011)	–.011 (.010)
Male schooling	.016 (.011)	.010 (.012)	.013 (.011)
Female schooling	–.018 (.012)	–.012 (.013)	–.016 (.013)
Log(life expectancy)	.228 (.045)	.242 (.047)	.231 (.046)
Log(fertility rate)	–.010 (.018)	–.010 (.019)	–.013 (.019)
Government consumption ratio	–.172 (.083)	–.215 (.088)	–.220 (.087)
Public education spending ratio	.18 (.27)	–.06 (.29)	.09 (.28)
Black-market premium	–.017 (.013)	.001 (.014)	.000 (.015)
Rule-of-law index	.0150 (.0034)	.0151 (.0036)	.0146 (.0035)
Terms-of-trade change	.047 (.062)	.060 (.070)	.059 (.067)
Democracy index	.092 (.059)	.111 (.066)	.103 (.065)
Democracy index squared	–.096 (.052)	–.108 (.059)	–.097 (.058)
Inflation rate	–.001 (.011)	–.059 (.017)	–.044 (.022)
R^2	.64, .62, .67	.64, .61, .62	.65, .62, .66
Number of observations	78, 89, 84	78, 89, 84	78, 89, 84

* The systems have three equations, where the dependent variables are the ratios of real gross investment to real GDP for 1965–75, 1975–85, and 1985–90. See the notes to Table 2 for definitions of the variables. Estimation is by instrumental variables. Column 1 includes inflation as an instrument. Column 2 uses inflation over the previous five years as an instrument. Column 3 uses prior colonial status as instruments. See the notes to Table 2 for descriptions of the other instrument.

increase in average inflation by 10 percentage points per year would lower the investment ratio on impact by 0.4 percentage point to 0.6 percentage point.

Even when the instruments are used, the adverse effect of inflation on investment shows up clearly only when the high-inflation observations are included in the sample. This finding accords with the results for growth rates.

CONCLUDING OBSERVATIONS

A major finding from the empirical analysis is that the estimated effects of inflation on growth and investment are significantly negative when some plausible instruments are used in the statistical procedures. Thus, there is some reason to believe that the relations reflect causation from higher long-term inflation to reduced growth and investment.

It should be stressed that the clear evidence for adverse effects of inflation comes from the experiences of high inflation. The magnitudes of effects are also not that large. For example, an increase in the average inflation rate by 10 percentage points per year is estimated to lower the growth rate of real per capita GDP (on impact) by 0.2 percentage point to 0.3 percentage point per year.

Some people have reacted to these kinds of findings by expressing skepticism about the value of cross-country empirical work. In fact, the wide differences in inflation experiences offered by the cross-section provide the best opportunity for ascertaining the long-term effects of inflation and other variables on economic performance. If the effects cannot be detected accurately in this kind of sample, then they probably cannot be pinpointed anywhere else. In particular, the usual focus on annual or quarterly time series of 30 to 40 years for one or a few countries is much less promising.

In any event, the apparently small estimated effects of inflation on growth are misleading. Over long periods, these changes in growth rates have dramatic effects on standards of living. For example, a reduction in the growth rate by 0.2 percentage point to 0.3 percentage point per year (produced on impact by 10 percentage points more of average inflation) means that the level of real GDP would be lowered after 30 years by 4 percent to 7 percent.[19] In mid-1995, the U.S. GDP was over $7 trillion; 4 percent to 7 percent of this amount is $300 billion to $500 billion, more than enough to justify a keen interest in price stability.

NOTES

1. Table 1 shows that the cross-country mean of inflation exceeded the median for each decade. This property reflects the skewing of inflation rates to the right, as shown in Figure 1. That is, there are a number of outliers with positive inflation rates of large magnitude, but none with negative inflation rates of high magnitude. Because this skewness increased in the 1980s, the mean inflation rate rose from the 1970s to the 1980s, although the median rate declined.
2. See, for example, Okun (1971) and Logue and Willett (1976).
3. See, for example, Barro (1996).
4. The actual rate is slightly higher because the observed growth rates are averages over periods of 10 or 5 years. See Barro and Sala+Martin (1995, p. 81).
5. Human capital is measured as the overall estimated effect from the levels of school attainment and the log of life expectancy.
6. This estimate is similar to that reported by Fischer (1993, Table 9). For earlier estimates of inflation variables in cross-country regressions, see Kormendi and Meguire (1985) and Grier and Tullock (1989).
7. The residual is computed from the regression system that includes all of the variables, including the inflation rate. But the contribution from the inflation rate is left out to compute the variable on the vertical axis in the scatter diagram. The residual has also been normalized to have a zero mean.
8. This system includes on the right-hand side standard deviations of inflation measured for the periods 1965–75, 1975–85, and 1985–90. These variables are also included with the instruments.
9. Bade and Parkin (1982); Grilli, Masciandaro, and Tabellini (1991); Cukierman (1992); and Alesina and Summers (1993).
10. Cukierman's (1992, chapter 20) results concur with this finding, especially for samples that go beyond a small number of developed countries, the kind of sample used in most of the literature on central bank independence.
11. Cukierman et al. (1993) use as instruments the turnover rate of bank governors and the average number of changes in bank leadership that occur within six months of a change in government. These measures of actual bank independence have substantial explanatory power for inflation but would not tend to be exogenous with respect to growth.
12. I have carried out SUR estimation of a panel system with the inflation rate as the dependent variable (for 1965–75, 1975–85, and 1985–90), where the independent variables are lagged inflation and the other instrumental variables used in Table 2. The estimated coefficient of lagged inflation is 0.74 (0.06). The only other coefficients that reach marginal significance are for log (GDP), 0.037 (0.019); the blackmarket premium, 0.059 (0.033); the change in the terms of trade, –0.40 (0.22); and the rule-of-law index, –0.009 (0.005). The R^2 values for the three periods are 0.55, 0.24, and 0.37.
13. For discussions of the CFA Franc zone, see Boughton (forthcoming) and Clement (1994). The zone maintained a fixed exchange rate with the French Franc for 45 years until the devaluation

from 50 to 100 CFA Francs per French Franc in January 1994. At the time of the devaluation, the zone covered 14 African countries grouped around three central banks: the West African Monetary Union of Benin. Burkina Faso. Ivory Coast. Mali. Niger. Senegal. and logo; a group of central African countries consisting of Cameroon, Central African Republic, Chad, Congo, Equatorial Guinea, and Gabon; and the Comoros. Some original members of the zone left to establish independent currencies—Djibouti in 1949, Guinea in 1958, Mali in 1962 (until it rejoined in 1984), Madagascar in 1963, Mauritania in 1973, and the Comoros in 1981 (to set up its own form of CFA franc). Equatorial Guinea, which joined in 1985, is the only member that is not a former colony of France (and not French speaking).

14. See Schwartz (1993).
15. The seven in the sample are Barbados, Dominican Republic (attributed to France rather than to Spain; see the notes to Table 5), Guyana, Haiti, Jamaica, Suriname, and Trinidad and Tobago. Five other former British colonies in Latin America that are not in this sample—Bahamas, Belize, Grenada, St. Lucia, and St. Vincent—experienced the relatively low average inflation rate of 6.9 percent from 1970 to 1990.
16. These four are Angola, Cape Verde, Guinea-Bissau, and Mozambique. Data are unavailable for Cape Verde and Guinea-Bissau in the 1960s (prior to independence). The figures for Angola in the 1980s are rough estimates.
17. I have carried out SUR estimation of a panel system with the inflation rate as the dependent variable (for 1965–75, 1975–85, and 1985–90), where the independent variables are the two colony dummies and the other instrumental variables—mainly lagged variables—used in Table 2. This system excludes lagged inflation (see footnote 12). The estimated coefficient on the Spain-Portugal colonial dummy is 0.14 (0.03) and that on the dummy for other colonies is 0.11 (0.05). The R^2 values are 0.38 for 1965–75, 0.14 for 1975–85, and 0.10 for 1985–90. Thus, inflation is difficult to explain, especially if most contemporaneous variables and lagged inflation are excluded as regressors. Two variables that are sometimes suggested as determinants of inflation—trade openness (measured by lagged ratios of exports and imports to GDP) and country size (measured by log of population)—are insignificant if added to the system. Years since independence also has no explanatory power for inflation. This result may arise because the former colonies of Spain and Portugal in Latin America became independent at roughly the same time.
18. See, for example, the results in Barro 1991.
19. In the model, the fall in the growth rate by 0.2 percent to 0.3 percent per year applies on impact in response to a permanent increase in the inflation rate. The growth rate would also decrease for a long time thereafter, but the magnitude of this decrease diminishes toward zero as the economy converges back to its (unchanged) long-run growth rate. Hence, in the very long run, the effect of higher inflation is a path with a permanently lower level of output, not a reduced growth rate. The numerical estimates for the reduced level of output after 30 years take account of these dynamic effects. The calculation depends on the economy's rate of convergence to its long-term growth rate (assumed, based on the empirical estimates, to be 2 percent to 3 percent per year). Also, the computations unrealistically neglect any responses of the other explanatory variables, such as the human-capital measures and the fertility rate.

REFERENCES

Alesina, A. and Summers, L. H. Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence. *Journal of Money, Credit and Banking* (May 1993), pp. 151–62.

Bade, R. and Parkin, J. M. Central Bank Laws and Monetary Policy. Unpublished manuscript, University of Western Ontario, 1982.

Barro, R. J. Democracy and Growth. *Journal of Economic Growth* (March 1996), pp. 1–27.

Barro, R. J. Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics* (May 1991), pp. 407–33.

Barro, R. J. and Sala-i-Martin, X. *Economic Growth.* McGraw Hill, 1995.

Bollen, K. A. Political Democracy: Conceptual and Measurement Traps. *Studies in Comparative International Development* (Spring 1990), pp. 7–24.

Boughton, J. M. The CFA Franc Zone: Currency Union and Monetary Standard. In *Monetary Integration,* Anthony Courakis and George Tavlas, eds. Cambridge University Press, forthcoming.

Briault, C. The Costs of Inflation. *Bank of England Quarterly Bulletin* (February 1995), pp. 33–45.

Clement, J. A. P. Striving for Stability: CFA Franc Realignment. *Finance & Development* (June 1994), pp. 10–13.

Cukierman, A. *Central Bank Strategy, Credibility, and Independence: Theory and Evidence.* MIT Press, 1992.

Cukierman, A., Kalaitzidakis, P., Summers, L. H., et al. Central Bank Independence, Growth, Investment, and Real Rates. *Carnegie-Rochester Conference Series on Public Policy* (Vol. 39, 1993), pp. 95–140.

Fischer, S. The Role of Macroeconomic Factors in Growth. *Journal of Monetary Economics* (Vol. 32, no. 3, 1993), pp. 485–512.

Gastil, R. D. and followers. *Freedom in the World: Political Rights and Civil Liberties.* Greenwood Press, 1982–83 and other years.

Grier, K. B., and Tullock, G. An Empirical Analysis of Cross-National Economic Growth, 1951–80. *Journal of Monetary Economics* (Vol. 24, 1989), pp. 259–76.

Grilli, V., Masciandaro D., and Tabellini, G. Political and Monetary Institutions and Public Financial Policies in the Industrial Countries. *Economic Policy: A European Forum* (October 1991), pp. 341–92.

Knack, S., and Keefer, P. Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures. *Economics and Politics* (November 1995), pp. 207–27.

Kormendi, R. C., and Meguire, P. G. Macroeconomic Determinants of Growth: Cross-Country Evidence. *Journal of Monetary Economics* (Vol. 16, 1985), pp. 141–63.

Logue, D. E., and Willett, T. D. A Note on the Relation Between the Rate and Variability of Inflation. *Economica* (Vol. 43, no. 17, 1976), pp. 151–58.

Mankiw, N. G., Romer, D., and Weil, D. N. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* (May 1992), pp. 407–37.

Okun, A. M. The Mirage of Steady Inflation. *Brookings Papers on Economic Activity* (Vol. 2, 1971), pp. 485–98.

Schwartz, A. J. Currency Boards: Their Past, Present, and Possible Future Role. *Carnegie-Rochester Conference Series on Public Policy* (Vol. 39, 1993), pp. 147–88.

Summers, R., and Heston, A. Penn World Tables, Version 5.5. National Bureau of Economic Research, 1993.

36
Economic Activity and Inflation

Bharat Trehan
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

This chapter reviews four papers on the relationship between inflation and economic activity that were presented at a recent macroeconomics workshop organized by the Federal Reserve Bank of San Francisco and the Stanford Institute of Economic Policy Research. The papers focused on the Phillips curve, named for A. W. Phillips (1958); he showed that a plot of (wage) inflation against unemployment for the U.K. produced a downward-sloping curve, indicating that higher unemployment was accompanied by lower inflation. The nature of this relationship has been studied and debated extensively, reflecting its importance for numerous policy issues. For instance, policymakers would like to know how much unemployment would increase if they tried to reduce inflation by some amount. This relationship also plays a key role in forecasting inflation in many Keynesian models.

DOES THE PHILLIPS CURVE ALWAYS SLOPE DOWNWARDS?

Haldane and Quah begin by showing that 50 years of postwar U.K. data do—at first glance—exhibit a negative relationship between unemployment and inflation, much as Phillips found in data through 1957. However, they then demonstrate that this relationship conceals two quite different relationships that held during the pre-1980 and post-1980 subperiods. Before 1980, the unemployment rate moved relatively little and was generally below 4%, while the inflation rate was fairly volatile and appears to have been centered around 10%. A plot of inflation against unemployment over that period would produce a vertical Phillips curve. After 1980, unemployment moved around more, ranging between 4% and 12%, while inflation varied less than the prior period, remaining below 10% most of the time. Here, the Phillips curve appears horizontal. Note that unemployment is higher on average while inflation is lower in the second subsample than in the first, which is why we get an apparently downward-sloping Phillips curve for the overall postwar period.

Other data sets Haldane and Quah examined also reveal differing relationships between inflation and unemployment. For instance, over periods that approximate the length of the business cycle, U.S. data reveal a negative relationship between the two variables, while U.K. data exhibit no such relationship.

Reprinted from: *FRBSF Economic Letter,* No. 99-09 (Mar. 12, 1999).

Finding different relationships between unemployment and inflation in different data sets suggests that the conventional downward-sloping Phillips curve is not a fundamental economic relationship and prompts a search for the features of the economy responsible for the changing patterns observed in the data. In their theoretical work, the authors show how differences in policymakers' beliefs about the economy can lead to differences in the observed Phillips curve. First, they present a model economy where it is optimal for fully informed policymakers to reduce the rate of inflation when unemployment is low and to raise the rate of inflation when unemployment is high. Intuitively, periods of low unemployment (for instance) make it less costly to reduce inflation. However, while such behavior tends to stabilize both output and inflation, the result is a Phillips curve that is upward sloping! The authors then show that this prediction can change if policymakers do not know exactly how the economy functions. For instance, a conventionally sloped Phillips curve can emerge if policymakers are mistaken about the nature of the tradeoff between inflation and unemployment.

According to the authors, the "mistakes" they focus on are meant to describe the changing beliefs of the British authorities in the postwar era. At first, monetary authorities probably believed that there was a long-run tradeoff between inflation and unemployment. Later, they realized that this belief was incorrect, but still believed there was a fixed short-run tradeoff between inflation and unemployment which they could exploit. More recently, they have abandoned this belief as well. The authors show that these beliefs have different implications for the behavior of output and inflation generated by their model, and that the predictions from their model are consistent with some (but not all) of the patterns observed in the data.

LEARNING ABOUT THE TRADEOFF

Sargent's paper also emphasizes the role of changing beliefs and uses these changes to explain variations in the U.S. inflation rate since the 1960s. A key feature of his paper is that these changes in policymakers' beliefs occur because of their analysis of incoming data about the economy.

As in the Haldane and Quah paper, Sargent assumes that monetary policymakers attempt to minimize fluctuations in inflation and unemployment around some target levels. He further assumes that policymakers can control the rate of inflation reasonably well, but they do not know exactly how variations in the rate of inflation will affect the unemployment rate. More formally, they do not know the "true" model. They try to learn the "truth" by statistically estimating the relationship between inflation and unemployment—a Phillips curve—using all the available data. They reestimate this relationship each period when new data come in. Policymakers also believe that this relationship tends to shift over time, so they employ estimation techniques that de-emphasize older data. Finally, each period they use the latest estimates of the Phillips curve to determine the optimal rate of inflation.

Typically (but not always), the Phillips curve estimated by policymakers show a tradeoff between inflation and unemployment, which tends to reflect the tradeoff between surprise inflation and unemployment inherent in the "true" model. (This kind of model will be described in more detail in the discussion of the paper by Ireland below.) In an attempt to exploit this tradeoff, policymakers tend to push the inflation rate to a relatively high level.

However, Sargent's model economy does not settle down at this point. The economy is subject to random shocks, so the shape of the estimated Phillips curve keeps changing. (The assumption that older data get de-emphasized is important here.) Eventually, the estimated Phillips curve appears vertical, that is, the data suggest that there is no tradeoff between inflation and unemployment. Optimizing policymakers then reduce inflation to near zero (the target level). Sargent interprets the reduction in U.S. inflation during the 1980s as just such an event: Sometime in the late 1970s, incoming data began to suggest that the U.S. Phillips curve was vertical; policymakers responded by lowering the inflation rate.

Sargent's analysis suggests that, unfortunately, the reduction in inflation is not likely to be permanent. This is because policymakers' belief in a vertical Phillips curve does not last. At some point, incoming data again begin to suggest a tradeoff between unemployment and inflation, and policymakers respond by trying to exploit the perceived tradeoff once again. (Recall that the perceived tradeoff reflects the tradeoff between unemployment and surprise inflation in the "true" model.) The whole cycle then repeats itself. Simulations of his model show the economy cycling between periods of high and low inflation, which mirrors changes in policymakers' beliefs about the Phillips curve.

DECLINING INFLATION: A LUCKY BREAK?

The two papers discussed so far attribute changes in inflation—and in the correlation between economic activity and inflation—to changes in policymakers' understanding of the economy. In sharp contrast, Ireland suggests that the observed variation in inflation originated in events that had little to do with monetary policy.

Ireland uses a well-known model of inflation due to Robert Barro and David Gordon. In the Barro-Gordon model, surprise inflation leads to reduction in the unemployment rate. Policymakers have an incentive to exploit this tradeoff, that is, to create surprise inflation in order to push the unemployment rate below the level where market forces would take it—this level is referred to as the natural rate of unemployment. However, the public recognizes the temptation faced by the government and adjusts its expectations of inflation accordingly. In equilibrium, then, the economy ends up with higher than optimal inflation and unemployment at the natural rate.

An important assumption of this model is that the gap between the natural rate and the unemployment rate desired by policymakers varies positively with the natural rate. This means, for instance, that the higher the natural rate, the greater the amount by which policymakers want to reduce unemployment. Since a greater reduction in unemployment can be obtained only by greater (surprise) inflation, the end result is that inflation varies positively with the natural rate.

Using data for the 1960–1990 period, Ireland shows that the long-run relationship between U.S. inflation and unemployment is consistent with the predictions of the Barro-Gordon model. While the short-run relationship found in the data is not consistent with the model, Ireland believes that this is not fatal for the theory, since the model being used is very simple. Thus, the model can be used to explain the behavior of inflation as follows: A series of negative shocks pushed the natural rate up during the 1970s. Policymakers responded by pushing the inflation rate up, as implied by the model. Positive shocks during the 1980s and 1990s then pushed the natural rate down; policymakers responded by lowering the inflation rate. Thus, Ireland argues that the rise in inflation during the 1970s reflected bad luck rather than unusually bad policy, while the reduction in inflation over the last two decades reflects good luck rather than good policy.

A BETTER WAY TO PREDICT INFLATION

Gali and Gertler focus on the role firms play in determining the behavior of the inflation rate. Their model is a variant of recent "New Keynesian" models of the macroeconomy. In these models, firms are assumed to have some monopoly power in the market, which allows them a degree of freedom in setting prices for their products. These firms determine the price they will charge by looking not only at current costs of production, but also at anticipated costs, since prices cannot be adjusted instantaneously. Because changes in costs are related to changes in output over the business cycle, these models imply that prices should change before output does. However, U.S. data show that changes in prices typically lag changes in output over the business cycle.

Gali and Gertler make several changes to the basic New Keynesian model to improve its ability to predict inflation. First, they assume that some firms are backward-looking and use simple rules of thumb to set prices. In addition, they employ a measure of the firms' marginal cost in order to predict inflation, on the grounds that this is what matters for inflation in the New Keynesian models, rather than more common measures of the output gap used in earlier empirical work.

Gali and Gertler show that their model does quite well in explaining the behavior of U.S. inflation over the 1960–1996 period. Further, their estimates suggest that relatively few U.S. firms are backward-looking. This result is important because the existence of backward-looking firms often has been proposed as the reason for the empirical finding that prices tend to be "sticky," that is, for the empirical finding that the price level tends to change more slowly than is easily explained a priori. Their finding implies that the observed stickiness in prices results from the fact that real marginal costs are sticky. Stated more generally, sticky prices reflect the fact that the variables that firms look at in order to set prices themselves change slowly in response to changes in economic conditions.

SUMMING UP

Though the underlying issues are far from settled, the papers presented at this workshop provide some interesting insights about the relationship between economic activity and inflation. Haldane and Quah argue that the conventional, downward-sloping Phillips curve does not represent a fundamental relationship, and that the relationship we observe will change as policymakers' beliefs about the economy change. Sargent's message is consistent, though his paper focuses on the role of learning. Policymakers learn about the economy by reestimating the Phillips curve as new data come in and vary the rate of inflation in response to these changing estimates. Ireland offers a plausible, alternative explanation for the observed variations in U.S. inflation: Inflation went up during the 1970s and came down in the 1980s because of the nature of the shocks hitting the economy and not because policymakers learned something about how the economy functions. Finally, the Gali and Gertler paper suggests that attempts to use economic activity to forecast inflation will do better if one accounts for the relationship between firms' marginal costs and inflation.

REFERENCES

Barro, R. J., and Gordon, D. B. 1983. A Positive Theory of Monetary Policy in a Natural Rate Model. *Journal of Political Economy* 91 pp. 589–610.

Gali, J., and Gertler, M. 1998. Inflation Dynamics: A Structural Econometric Analysis. Mimeo. New York University.

Haldane, A., and Quah, D. 1998. U.K. Phillips Curves and Monetary Policy. Mimeo. Bank of England.

Ireland, P. 1998. Does the Time Consistency Problem Explain the Behavior of U.S. Inflation? Mimeo. Boston College.

Phillips, A. W. 1958. The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. *Economica* 25 pp. 283–299.

Sargent, T. 1999. The Temporary (?) Conquest of American Inflation. Mimeo. Stanford University.

This Page Intentionally Left Blank

37

Inflation, Financial Markets, and Capital Formation

Sangmok Choi
Economist, Korean Ministry of Finance, Seoul, Korea

Bruce D. Smith
Professor of Economics, University of Texas, Austin, Texas

John H. Boyd
Professor of Finance, University of Minnesota, Minneapolis, Minnesota

A consensus among economists seems to be that high rates of inflation cause "problems," not just for some individuals, but for aggregate economic performance. There is much less agreement about what these problems are and how they arise. We propose to explain how inflation adversely affects an economy by arguing that high inflation rates tend to exacerbate a number of financial market frictions. In doing so, inflation interferes with the provision of investment capital, as well as its allocation.[1] Such interference is then detrimental to long-run capital formation and to real activity. Moreover, high enough rates of inflation are typically accompanied by highly variable inflation and by variability in rates of return to saving on all kinds of financial instruments. We argue that, by exacerbating various financial market frictions, high enough rates of inflation force investors' returns to display this kind of variability. It seems difficult then to prevent the resulting variability in returns from being transmitted into real activity.

Unfortunately, for our understanding of these phenomena, the effects of permanent increases in the inflation rate for long-run activity seem to be quite complicated and to depend strongly on the initial level of the inflation rate. For example, Bullard and Keating (forthcoming) find that a permanent, policy-induced increase in the rate of inflation *raises* the long-run level of real activity for economies whose initial rate of inflation is relatively low. For economies experiencing moderate initial rates of inflation, the same kind of change in inflation seems to have no significant effect on long-run real activity. However, for economies whose initial inflation rates are fairly high, further increases in inflation significantly *reduce* the long-run level of output. Any successful theory of how inflation affects real activity must account for these nonmonotonicities.

Along the same lines, Bruno and Easterly (1995) demonstrate that a number of economies have experienced sustained inflations of 20 percent to 30 percent without suf-

Reprinted from: *Federal Reserve Bank of St. Louis Review* (May/June 1996) 9–29.

fering any apparently major adverse consequences. However, once the rate of inflation exceeds some critical level (which Bruno and Easterly estimate to be about 40 percent), significant declines occur in the level of real activity. This seems consistent with the results of Bullard and Keating.

Evidence is also accumulating that inflation adversely affects the allocative function of capital markets, depressing the level of activity in those markets and reducing investors' rates of return. Again, however, these effects seem highly nonlinear. In a cross-sectional analysis, for example, Boyd, Levine, and Smith (1995) divide countries into quartiles according to their average rates of inflation. The lowest inflation quartile has the highest level of financial market activity, and the highest inflation quartile has the lowest level of financial market activity. However, the two middle quartiles display only very minor differences. Thus for the financial system, as for real activity, there seem to be threshold effects associated with the inflation rate.

Moreover, as we will show, high rates of inflation tend to depress the real returns equity-holders receive and to increase their variability. In Korea and Taiwan, there were fairly pronounced jumps in the rate of inflation in 1988 and 1989, respectively. In each country, before those dates, inflation's effects on rates of return to equity, rate of return volatility, and transactions volume appear to be insignificant. After the dates in question, these effects are generally highly significant. Thus it seems possible that—to adversely affect the financial system—inflation must be "high enough."

Why does inflation affect financial markets and real activity this way? We produce a theoretical model in which—consistent with the evidence—higher inflation reduces the rate of return received by savers in all financial markets. By itself this effect might be enough to reduce savings and hence the availability of investment capital. However, we do not believe that this explanation by itself is very plausible, for two reasons. First, to explain the nonmonotonicities we have noted, the savings function would have to bend backward. Little or no empirical evidence exists to support this notion. Second, almost all empirical evidence suggests that savings is not sufficiently sensitive to rates of return to make this a plausible mechanism for inflation to have large effects. Thus an alternative mechanism is needed.

We present a model in which inflation reduces real returns to savings and, via this mechanism, exacerbates an informational friction afflicting the financial system. The particular friction modeled is an adverse selection problem in capital markets. However, the specific friction seems not to be central to the results we obtain.[2] What is central is that the severity of the financial market friction is endogenous and varies positively with the rate of inflation.

In this specific model, higher rates of inflation reduce savers' real rates of return and lower the real rates of interest that borrowers pay. By itself, this effect makes more people want to be borrowers and fewer people want to be savers. However, people who were not initially getting credit represent "lower quality borrowers" or, in other words, higher default risks. Investors will be uninterested in making more loans to lower quality borrowers at lower rates of interest and therefore must do something to keep them from seeking external finance. The specific response here is that markets ration credit, and more severe rationing accompanies higher inflation. This rationing then limits the availability of investment capital and reduces the long-run level of real activity. In addition, when credit rationing is sufficiently severe, it induces endogenously arising volatility in rates of return to savings. This volatility must be transmitted to real activity and, hence, to the rate of inflation. Variable inflation therefore necessarily accompanies high enough rates of inflation, as we observe in practice.

This story, of course, does not explain why these effects are strongest at high—and not at low—rates of inflation. The explanation for this lies in the fact that—at low rates of inflation—our analysis suggests that credit market frictions are potentially innocuous. Thus at low rates of inflation, credit rationing might not emerge at all, and none of the mechanisms mentioned in the previous paragraph would be operative. In this case our economy would act as if it had no financial market frictions. When this occurs, our model possesses a standard Mundell-Tobin effect that makes higher inflation lead to *higher* long-run levels of real activity.[3] However, once inflation exceeds a certain critical level, credit rationing must be observed, and higher rates of inflation can have the adverse consequences noted above.

Finally, our analysis suggests that a certain kind of "development trap" phenomenon is ubiquitous, particularly at relatively high rates of inflation.[4] We often observe that economies whose performance looks fairly similar at some point in time—like Argentina and Canada circa 1940—strongly diverge in terms of their subsequent development. Although this is clearly often because of differences in government policies, presumably many governments confront similar policy options. One would thus like to know whether intrinsically similar economies can experience divergent economic performance for purely endogenous reasons. The answer in models with financial market frictions is that this can occur fairly easily: When the severity of an economy's financial market frictions is endogenous, it is possible that—for endogenous reasons—the friction is perceived to be more or less severe. If it is perceived to be more (less) severe, financial markets provide less (more) investment capital. The result is a reduced (enhanced) level of real economic performance. This validates the original perception that the friction was (was not) severe. Thus, as we show, development traps should be expected to be quite common.

The remainder of the article proceeds as follows: In the first section we lay out a theoretical model that illustrates the arguments just given, while in the second section we describe an equilibrium of the model. In the third section we discuss how inflation affects the level of real activity when the financial market friction is not operative, while in the fourth section we take up the same issue when it is. In the fifth section we examine when the friction will or will not be operative and derive the theoretical implications we have already discussed. In the sixth section we show that an array of empirical evidence supports these implications. In the final section we offer our conclusions.

A SIMPLE ILLUSTRATIVE MODEL

The purpose of this section is to present a model that illustrates how inflation interacts with a particular financial market friction. This friction is purposely kept very simple in order to highlight the economic mechanisms at work. Later we will argue that these mechanisms are operative very generally in economies where financial markets are characterized by informational asymmetries.

The Environment

The economy is populated by an infinite sequence of two period lived, overlapping generations. Each generation is identical in its size and composition. We describe the latter below and index time periods by $t = 0, 1, \ldots$.

At each date, a single final commodity is produced via a technology that utilizes homogeneous physical capital and labor as inputs. An individual producer employing K_t units of capital and N_t units of labor at t produces $F(K_t,N_t)$ units of final output. For purposes of exposition, we will assume that F has the constant elasticity of substitution form

$$F(K,N) = [aK^{\rho} + bN^{\rho}]^{1/\rho}, \tag{1}$$

and we will assume throughout that $\rho < 0$ holds.[5] Defining $k \equiv K/N$ to be the capital-labor ratio, it will often be convenient to work with the intensive production function $f(k) \equiv F(k,1)$. Clearly, here

$$f(k) = [ak^{\rho} + b]. \tag{1'}$$

Finally, to keep matters notationally simple, we assume that capital depreciates completely in one period.[6]

Each generation consists of two types of agents. Type 1 agents—who constitute a fraction $\lambda \in (0,1)$ of the population—are endowed with one unit of labor when young and are retired when old. We assume that all young-period labor is supplied inelastically. In addition, type 1 agents have access to a linear technology for storing consumption goods whereby one unit stored at t yields $x > 0$ units of consumption at $t + 1$.

Type 2 agents represent a fraction $1 - \lambda$ of each generation. These agents supply one unit of labor inelastically when old and have no young-period labor endowment.[7] In addition, type 2 agents have no access to the technology for storing goods. They do, on the other hand, have access to a technology that converts one unit of final output at t into one unit of capital at $t + 1$. Only type 2 agents have access to this technology.

We imagine that any agent who owns capital at t can operate the final goods production process at that date. Thus type 2 agents are producers in old age. It entails no loss of generality to assume that all such agents run the production process and work for themselves in their second period.

With respect to agents' objective functions, it is simplest to assume that all agents care only about old-age consumption and that they are risk neutral.[8] These assumptions are easily relaxed.

The central feature of the analysis is the presence of an informational friction affecting the financing of capital investments. In particular, we assume that each agent is privately informed about his own type. We also assume that nonmarket activities, such as goods storage, are unobservable, while all market transactions are publicly observed. Thus, to emphasize, an agent's type and storage activity are private information, while all market transactions—in both labor and credit markets—are common knowledge. This set of assumptions is intended to keep the informational asymmetry in our model very simple: Since type 2 agents cannot work when young, they cannot credibly claim to be type 1. However, type 1 agents might claim to be type 2 when young. We now describe what happens if they do so.

If a type 1 agent wishes to claim to be type 2, he cannot work when young, and he must borrow the same amount as type 2 agents do. Since type 1 agents are incapable of producing physical capital, it will ultimately be discovered that they have misrepresented their type. To avoid punishment, we assume that a dissembling type 1 agent absconds with his loan, becoming autarkic and financing old-age consumption by storing the proceeds of his borrowing. Dissembling type 1 agents never repay loans. Notice, however, that since type 2 agents cannot store goods, they will never choose to abscond, and hence they always repay their loans.[9] Obviously, lenders will want to avoid making loans to dissembling type 1 agents. How they do so is the subject of the section on equilibrium conditions in financial markets.

It remains to describe the initial conditions of our economy. At $t = 0$ there is an initial old generation where each agent is endowed with one unit of labor (supplied inelastically) and with $K_0 > 0$ units of capital. No other agents are endowed at any date with capital or consumption goods.

Trading

Three kinds of transactions occur in this economy. First, final goods and services are bought and sold competitively. We let p_t denote the dollar price at t of a unit of final output. Second, producers hire the labor of young type 1 agents in a competitive labor market, paying the real wage rate w_t at t. And third, young (nondissembling) type 1 workers save their entire labor income, which they supply inelastically in capital markets, thereby acquiring claims on type 2 agents—and possibly on some dissembling type 1 agents—and claims on the government, such as money or national debt. The model we present here is not rich enough to capture any distinction between different types of financial claims, such as debt or equity.[10] We thus think of young agents as simply acquiring a generalized claim against producers of capital. It entails no loss of generality to think of financial market activity as being intermediated, say through banks, mutual funds, or pension funds. We assume there is free entry into the activity of intermediation. We also let R_{t+1} be the real gross rate of return earned by intermediaries between t and $t + 1$ on (nondefaulted) investments, and we let r_{t+1} be the real gross rate of return earned by young savers. After describing government policy, we return to a description of equilibrium conditions in these markets.

The Government

We let M_t denote the outstanding per capita money supply at t. At $t = 0$ the initial old agents are endowed with the initial per capita money supply, $M_{-1} > 0$. Thereafter, the money supply evolves according to

$$M_{t+1} = \sigma M_t, \tag{2}$$

where $\sigma > 0$ is the exogenously given gross rate of money creation. We assume that the government makes a once and for all choice of σ at $t = 0$: In steady-state equilibria the (gross) rate of inflation will equal σ.

Our ultimate purpose is to examine how different choices of σ affect financial markets and, through this channel, capital formation. To make our results as stark as possible, we assume that the government uses the proceeds of money creation to finance a subsidy to private capital formation. It should then be transparent that any adverse effects of inflation are a result of the presence of inflation alone and not what the revenue from the inflation tax is used for. More specifically, then, we assume that any monetary injections (withdrawals) occur via lump-sum transfers to young agents claiming to be type 2. Genuine type 2 agents will use these transfers entirely to invest in capital; hence government policy here consists of a capital subsidy program financed by printing money. If we let τ_t denote the real value of the transfer received by young type 2 agents at t, and we let $\mu_t \in [0,1]$ denote the fraction of dissembling type 1 agents in the time t population, then the government budget constraint implies that the real value of transfers, per capita, equals the real value, per capita, of seigniorage revenue. Thus

$$[(1-\lambda)+\lambda\mu_t]\,\tau_t=(M_t-M_{t-1})/p_t \tag{3}$$

must hold at all dates. If we let $m_t \equiv M_t/p_t$ denote time t real balances, equations 2 and 3 imply that

$$[(1-\lambda)+\lambda\mu_t]\,\tau_t=[(\sigma-1)/\sigma]m_t. \tag{3'}$$

EQUILIBRIUM CONDITIONS

Factor Markets

Let b_t denote the real value of borrowing by young type 2 agents at t. These agents also receive a transfer τ_t. All resources obtained by these individuals are used to fund capital investments at t. Each old producer at $t+1$ will hence have the capital stock

$$K_{t+1}=b_t+\tau_t. \tag{4}$$

Let L_t denote the quantity of young labor hired by a representative producer at t. Each such producer combines this with his own unit of labor to obtain $N_t=L_t+1$ units of labor services. Then the producer's profits, net of loan repayments, are $F(K_t,L_t+1)-w_tL_t-R_tb_{t-1}$ since an interest obligation of R_tb_{t-1} was incurred at $t-1$. Producers wish to maximize old-period income. At t, b_{t-1} is given by past credit extensions, so that the only remaining choice variable is L_t. Profits are maximized when

$$\begin{aligned} w_t &= F_2(K_t,L_t+1)=F_2(K_t/N_t,1)\\ &= f(k_t)-k_tf'(k_t)=b(ak_t^{\rho}+b)^{(1-\rho)/\rho}\\ &\equiv w(k_t) \end{aligned} \tag{5}$$

where $k_t \equiv K_t/N_t$ is the capital labor ratio. Equation 5 asserts the standard result that labor earns its marginal product.

For future reference, it will be useful to have an expression for the consumption, c2\t, of old type 2 agents at t. Clearly

$$\begin{aligned} c_t^2 &= F_1(\cdot)K_t+F_2(\cdot)(L_t+1)-w_tL_t-R_tb_{t-1}\\ &= [F_1(\cdot)-R_t]b_{t-1}+w_t+F_1(\cdot)\tau_{t-1}. \end{aligned} \tag{6}$$

The first equality in equation 6 follows from Euler's law, while the second follows from equations 3, 4, and 5. Equation 6 asserts that old producers have income equal to the marginal product of their own labor, plus the value of the capital paid for through transfer payments $[F_1(\cdot)\tau_{t-1}]$, plus the net income obtained from capital attained through borrowing $[(F_1(\cdot)-R_t)b_{t-1}]$.

Financial Markets

Intermediaries face a fairly standard adverse selection problem in financial markets.[11] If they lend to a dissembling type 1 agent, the loan will not be repaid.[12] Hence it is desirable not to lend to these agents, but at the same time such agents cannot be identified ex ante. Intermediaries will hence structure financial contracts to deter type 1 agents from dissembling or, in other words, to induce self-selection (only type 2 agents choose to accept funding).

Using typical assumptions in economies with adverse selection, we assume that intermediaries announce financial contracts consisting of a loan quantity b_t, and an interest rate (or return to the intermediary) of R_{t+1}. Each intermediary announces such a contract, taking the contracts offered by other intermediaries as given. Hence we seek a Nash equilibrium set of financial contracts. On the deposit side we assume that intermediaries behave competitively (that is, each intermediary assumes it can raise all the funds it wants at the going rate of return on savings r_{t+1}).

One objective of intermediaries is to induce self-selection. This requires that type 1 agents prefer to work when young and save their young-period income rather than to borrow b_t, receive the transfer τ_t, and abscond. If they work when young and save the proceeds, their utility is $r_{t+1}w_t$. If they borrow b_t, obtain the transfer τ_t, and abscond, their utility is $x(b_t + \tau_t)$. Hence self-selection requires that

$$r_{t+1}w_t \geq x(b_t + \tau_t). \tag{7}$$

Standard arguments[13] establish that equation 7 holds in any Nash equilibrium and that all type 1 agents are deterred from dissembling. Hence $\mu_t = 0$ holds at all dates.

In addition, since there is free entry into intermediation, all intermediaries must earn zero profits in equilibrium. Since $\mu_t = 0$, this simply requires that

$$R_{t+1} = r_{t+1}. \tag{8}$$

An equilibrium in financial markets now requires that five conditions be satisfied. First, given r_{t+1} and τ_t, the quantity of funds obtained in the marketplace by each type 2 agent must satisfy equation 7. Second, equation 8 must hold. Third, sources and uses of funds must be equal. Sources of funds at each date are simply the savings of young type 1 agents, which in per capita terms are λw_t. Uses of funds are loans to borrowers [$(1 - \lambda)b_t$ per capita], plus real balances (m_t per capita), plus per capita storage (s_t). Thus equality between sources and uses of funds obtains if and only if

$$\lambda w(k_t) = (1 - \lambda)b_t + m_t + s_t. \tag{9}$$

The fourth condition is that type 2 agents will be willing to borrow if and only if

$$F_1(K_t,N_t) = F_1(K_t/N_t,1) = f'(k_t)$$
$$= a[a + bk^{-\rho}]^{(1-\rho)/\rho} \geq R_{t+1} = r_{t+1} \tag{10}$$

holds[14]. Equation 10 implies that type 2 agents perceive nonnegative profits from borrowing. And finally, type 1 agents are willing to supply funds to intermediaries if and only if the return they receive is at least as large as the return available on alternative savings instruments (money and storage). This requires that

$$r_{t+1} \geq p_t/p_{t+1} \tag{11a}$$

$$r_{t+1} \geq x. \tag{11b}$$

We will want agents to hold money in equilibrium. Hence equation 11a must always hold with equality. We will assume that equation 11b is a strict inequality; hence in equilibrium $s_t = 0$ (storage is dominated in rate of return). Equation 11b is validated, at least near steady states, by the assumption that

$$1/x > \sigma. \tag{a1}$$

We will henceforth impose equation a1.[15]

Some Implications

We now know that in equilibrium all young type 1 agents supply their labor to producers. Hence labor market clearing requires that the per capita labor demand of producers $[(1 - \lambda)L_t]$ equals the per capita labor supply of young type 1 agents (λ). Therefore,

$$L_t = \lambda/(1 - \lambda). \tag{12}$$

It is an immediate implication that $N_t = 1 + L_t = 1/(1 - \lambda)$ and that

$$k_t \equiv K_t/N_t = (1 - \lambda)K_t. \tag{13}$$

In addition, under equation a1, $s_t = 0$ holds, so that equation 9 becomes

$$(1 - \lambda)b_t = \lambda w(k_t) - m_t. \tag{9'}$$

Now note that $k_{t+1} = (1 - \lambda)\, K_{t+1} = (1 - \lambda)(b_t + \tau_t)$. Using this fact in equation 9′, we obtain

$$k_{t+1} = \lambda w(k_t) - m_t - (1 - \lambda)\tau_t. \tag{9''}$$

Finally, we know that $\mu_t = 0$. Using this fact in equation 1′ and substituting the result into equation 9″ yields

$$k_{t+1} = \lambda w(k_t) - m_t/\sigma. \tag{14}$$

It is also possible to derive some further implications from the preceding discussion. Equations 10 and 11a at equality, imply that

$$f'(k_{t+1}) \geq r_{t+1} = p_t/p_{t+1}. \tag{15}$$

We can now use the identity $p_t/p_{t+1} \equiv (M_{t+1}/p_{t+1})(p_t/M_t)(M_t/M_{t+1}) \equiv m_{t+1}/\sigma m_t$ to write equation 15 as

$$f'(k_{t+1}) \geq m_{t+1}/\sigma m_t. \tag{15'}$$

Finally, equation 7 must hold in equilibrium. Substituting equation 4 into equation 7, and using $K_{t+1} = k_{t+1}/(1 - \lambda)$, we obtain the equivalent condition

$$r_{t+1}w(k_t) \geq xk_{t+1}/(1 - \lambda). \tag{16}$$

Equation 11a also implies an alternative form of equation 16:

$$[m_{t+1}/\sigma m_t]w(k_t) \geq xk_{t+1}/(1 - \lambda). \tag{16'}$$

We can now reduce our search for an equilibrium to the problem of finding a sequence $\{k_t, m_t\}$ that satisfies equations 14, 15′, and 16′ at all dates, with $k_0 > 0$ given as an initial condition. We now make an additional comment. If equation 15′ is a strict inequality at any date, young type 2 agents perceive positive profits to be made from borrowing and hence will want to borrow an arbitrarily large amount. Because this is not possible, if equation 15′ is a strict inequality, their borrowing must be constrained. The relevant constraint is equation 7. In this case equation 7 at equality determines b_t, and equation 16′ will hold with equality. In equilibrium, at least one of the conditions (equations 15′ or 16′) must thus hold with equality. If equation 15′ is an equality, the equilibrium coincides with standard equilibria that obtain in similar economies with no informational asymmetries.[16] In this case we say the equilibrium is Walrasian. If equation 15′ holds as a strict inequality, then equation 16′ is an equality. We refer to this situation as credit rationing.

WALRASIAN EQUILIBRIA

We now describe sequences that satisfy equations 14 and 15′ at equality. For the present we do not impose equation 16′: This amounts to assuming that agents' types are publicly observed. In the section on the endogeneity of financial market frictions, we ask when such sequences will also satisfy equation 16 or, in other words, when Walrasian resource allocations can be sustained even in the presence of the informational asymmetry. We begin with steady-state equilibria, and then briefly describe the nature of equilibrium paths that approach the steady state. Because the material in this section is quite standard,[17] we attempt to present it fairly concisely.

Steady States

In a steady state k_t and m_t are constant. Hence equation 15′ at equality reduces to

$$f'(k) = 1/\sigma = p_t/p_{t+1}, \tag{17}$$

while equation 14 becomes

$$m = \sigma[\lambda w(k) - k]. \tag{18}$$

It is immediately apparent from equation 17 that *increases* in the rate of money growth (and inflation), σ, increase the steady-state capital-labor ratio, per capita output, and productivity of labor. This is true for all rates of money growth satisfying equation a1. Because the empirical evidence cited in the introduction strongly suggests that higher inflation can lead to higher long-run levels of real activity *only* if initial rates of inflation are relatively low, it is clear that our model cannot confront the whole array of empirical experience in the absence of the informational asymmetry.

For future reference, it will be convenient to give an explicit form for the capital stock (or variables related to it) as a function of the money growth rate. To this end we define the variable

$$z_t \equiv (b/a)k_t^{-\rho} \equiv w(k_t)/k_t f'(k_t). \tag{19}$$

It is readily verified that z_t is simply the ratio of labor's share to capital's share: The assumption that $\rho < 0$ implies that z_t is an increasing function of k_t. Hence movements in z_t simply reflect similar movements in k_t.

It is easy to check that $f'(k_t) = a^{1/\rho}[1 + (b/a)k_t^{-\rho}]^{(1-\rho)/\rho} \equiv a^{1/\rho}(1 + z_t)^{(1-\rho)/\rho}$. Then, if we let $z^*(\sigma)$ denote the value of z satisfying equation 17 for each σ, we have that

$$z^*(\sigma) = [a^{-1/\rho}(1/\sigma)]^{\rho/(1-\rho)} - 1. \tag{20}$$

Equations 19 and 20 give the capital stock in a Walrasian steady state.

Dynamics

Equations 14 and 15′ at equality describe how the economy evolves given k_0 and m_0: the initial capital-labor ratio and initial real balances. The initial price level is endogenous, and so $m_0 \equiv M_0/p_0$ is endogenous.

It is easy to show that the monetary steady state is a saddle or, in other words, that there is only one choice of m_0 that averts a hyperinflation where money asymptotically loses all value. Thus nonhyperinflationary equilibria are determinate (only one possible equilibrium path approaching the monetary steady state exists), and it is possible to show that the steady state is necessarily approached monotonically. Walrasian equilibria therefore *cannot* display economic fluctuations in output, real returns to investors, or the rate of inflation.

Summary

Walrasian equilibria are unique. Growth traps are therefore impossible. Moreover, Walrasian equilibria do not display economic fluctuations. Finally, Walrasian equilibria have the feature that increases in the long-run rate of inflation lead to higher long-run levels of real activity and productivity.

EQUILIBRIA WITH CREDIT RATIONING

In this section we investigate sequences $\{k_t,m_t\}$ that satisfy equations 14 and 16′ at equality at all dates. In the section on the endogeneity of financial market frictions, we then examine when a Walrasian equilibrium or an equilibrium with credit rationing will actually obtain. As before, we begin with steady-state equilibria.

Steady States

When k_t and m_t are constant, equation 16′ at equality implies that the steady-state capital-labor ratio satisfies

$$w(k)/k = x\sigma/(1-\lambda). \tag{21}$$

Equation 21 says that the capital stock is determined by how financial markets control borrowing to induce self-selection. The rate of inflation matters because it affects the rate of return that nondissembling type 1 agents receive on their savings. As inflation rises, this return falls,[18] with the consequence that the utility of working and saving declines. To prevent type 1 agents from dissembling, the utility of doing so must also fall. Equation 21 describes the consequences for the per capita capital stock.

It will be convenient to transform equation 21 as follows: First note that it can be written as

$$[w(k)/kf'(k)]f'(k) = x\sigma/(1-\lambda). \tag{21'}$$

Second, given equation 19 and our previous observations about $f'(k)$, it is easy to verify that $[w(k)/kf'(k)]f'(k) = a^{1/\rho}z(1+z)^{(1-\rho)/\rho}$. This observation allows us to rewrite the equilibrium condition equation 21′ as

$$a^{-1/\rho}[x\sigma/(1-\lambda)] = z(1+z)^{(1-\rho)/\rho} \equiv H(z). \tag{22}$$

Equation 22 determines the steady-state equilibrium value(s) of z as a function of the long-run inflation rate σ. Equation 19 then gives the steady-state per capita capital stock. Steady-state real balances are determined from equation 14 with k and m constant:

$$m = \sigma[\lambda w(k) - k]. \tag{23}$$

Equation 21 permits us to rewrite equation 23 as

$$m = \sigma k\{[x\lambda\sigma/(1 - \lambda)] - 1\}. \tag{23'}$$

For future reference, it will be convenient to define the function $A(\sigma)$ by

$$A(\sigma) \equiv [x\lambda/(1 - \lambda)]\sigma. \tag{24}$$

We can now state our first result.

RESULT 1. *Define $\hat{\sigma}$ by*

$\underline{z}(\sigma)$

Then if $\sigma \leq \hat{\sigma}$, there exists a solution to equation 22. If, in addition,

$$A(\sigma) > 1 \tag{a2}$$

all solutions to equation 22 yield positive levels of real balances. (Result 1 is proved in the Appendix.)

As the Appendix establishes, the function $H(z)$ defined in equation 22 has the configuration depicted in Figure 1. In particular,

$$H(0) = 0 = \lim_{z \to \infty} H(z)$$

and H attains a unique maximum at $z = -\rho$. Thus, if

$$H(-\rho) \geq a^{-1/\rho}[x\sigma/(1 - \lambda)] \tag{26}$$

equation 22 has a solution, which is depicted in Figure 1. If $\sigma < \hat{\sigma}$ where

$$\hat{\sigma} = [(1 - \lambda)/x]a^{1/\rho}H(-\rho),$$

there will be exactly two solutions to equation 22 that are denoted by $\underline{z}(\sigma)$ and $\bar{z}(\sigma)$ in Figure 1.

The conditions $A(\sigma) > 1$ and $\sigma \leq \hat{\sigma}$ are equivalent to

$$(1 - \lambda)/x\lambda < \sigma \leq \hat{\sigma}. \tag{a3}$$

We henceforth assume that equation a3 holds. We also assume that

$$1/x \geq \hat{\sigma} \tag{a4}$$

so that equation a3 implies satisfaction of equation a1.[19]

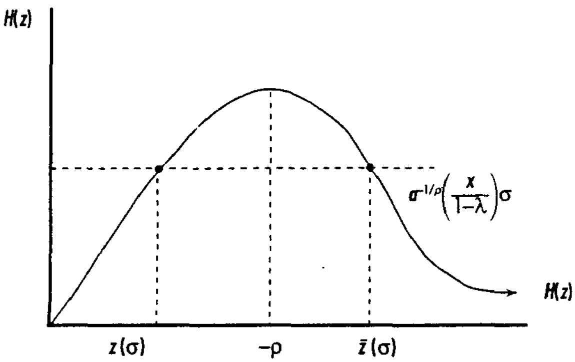


Figure 1 Determination of z under credit Rationing.

Evidently, when $\sigma \leq \hat{\sigma}$, there are two solutions to equation 22. This multiplicity of candidate equilibria derives from the way that financial markets respond to the presence of the adverse selection problem. To induce self-selection at any given value of σ, $w(k)$ and $(b + \tau)$ must be linked. One way that self-selection can occur is for $w(k)$ and $(b + \tau)$ both to be low; this requires that z be low. Alternatively, $w(k)$ and $(b + \tau)$ can both be relatively high; this requires that z be high. The possibility that there is more than one way for financial markets to address an informational asymmetry has a generality beyond this particular model, as shown by Boyd and Smith (forthcoming) or Schreft and Smith (forthcoming and 1994).

The Effects of Higher Inflation

The consequences of an increase in the steady-state inflation rate are depicted in Figure 2. Evidently, an increase in σ raises $\underline{z}(\sigma)$ and reduces $\bar{z}(\sigma)$ or, in other words

$$\underline{z}'(\sigma) > 0 > \bar{z}'(\sigma) \tag{27}$$

holds. The same statements apply to k. Hence, in the low- (high-) capital-stock steady state, an increase in the inflation rate raises (lowers) the steady-state capital stock. These effects occur because an increase in σ reduces the steady-state return on savings. Other things equal, this lowers the utility of honest type 1 agents and would cause them to misrepresent their type. To preserve self-selection $w(k)$ must rise relative to $(b + \tau) = k/(1 - \lambda)$. In the low- (high-) capital-stock steady state, this requires that k rise (fall). Thus higher inflation exacerbates informational asymmetries, with implications for the capital stock that are adverse in the high-capital-stock steady state.

Figure 3 depicts the solutions to equation 22 as a function of σ, where we denote by $\hat{z}(\sigma)$ any solutions to that equation. Evidently, there can be no solution to equation 22 if the government sets σ above $\hat{\sigma}$. For σ satisfying equation a3, clearly we have

$$\underline{z}(\sigma) < -p < z(\sigma). \tag{28}$$

Of particular interest in this context is the possibility that an increase in the long-run rate of inflation can reduce the long-run capital stock, real activity, and productivity. Such consequences are often observed empirically when inflation increases, particularly when

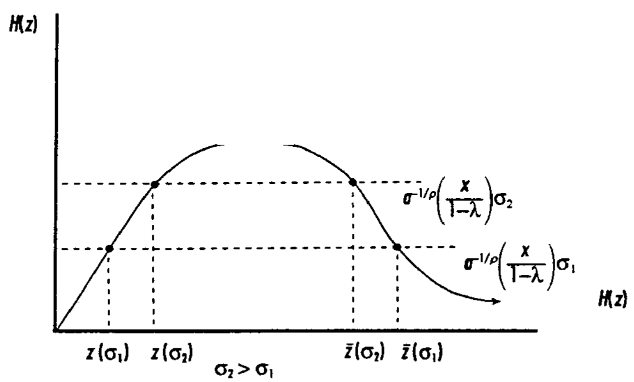


Figure 2 The consequences of higher inflation.

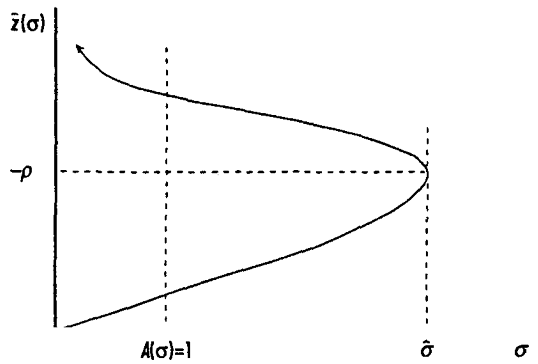


Figure 3 Inflation and its consequences under credit rationing.

the initial rate of inflation is relatively high. This outcome is observed in the high-capital-stock steady state. We now want to know which, if either, steady state can be approached under credit rationing.

Dynamics

Given an initial capital-labor ratio, k_0, and an initial level of real balances, m_0, equations 14 and 16′ at equality govern the subsequent evolution of k_t and m_t. The Appendix establishes our second result.

RESULT 2. *(a) The low-capital-stock steady state is a saddle. All* $\{k_t, m_t\}$ sequences approaching it do so monotonically. (b) The high-capital-stock steady state is a sink if $\hat{\sigma}$ *is not too large.*

Result 2a implies that both the high- and the low-capital-stock steady states can potentially be approached. To approach the low-capital-stock steady state, m_0 must be chosen to lie on a "saddle path;" that is, there is a unique choice of m_0 that allows the economy to approach the low-capital-stock steady state.

Result 2b implies that, for some open set of values of k_0, there is a whole interval of choices of m_0 that allow the high-capital-stock steady state to be approached. The requirement of avoiding a hyperinflation thus no longer implies what m_0 must be. Monetary equilibria have become *indeterminate.* A continuum of possible equilibrium values of m_0 exist and hence so does a continuum of possible equilibrium paths approaching the high-capital-stock steady state. This is a consequence of the informational friction afflicting capital markets.

Not only is the informational asymmetry a source of indeterminacy, it is a potential source of endogenous economic volatility as well. We now establish that such volatility *must* be observed near the high-capital-stock steady state whenever the rate of inflation is sufficiently high. At high rates of inflation, the economy must thus pay a price to avoid the low-capital-stock steady state: This price is the existence of endogenous volatility in real activity, inflation, and asset returns.

RESULT 3. *Suppose that* σ *is sufficiently close to* $\hat{\sigma}$*. Then all paths approaching the high-capital-stock steady state do so nonmonotonically.* (Result 3 is proved in the Appendix.)

Summary

When financial market frictions bind, there can be two steady-state equilibria differing in their levels of real development. Both steady states can potentially be approached. A continuum of paths approaching the high-capital-stock steady state exists so that the operation of financial markets creates an indeterminacy. If the steady-state inflation rate is high enough, all such paths display endogenously arising volatility in real activity, real returns, and inflation. In this sense high inflation also engenders variable inflation.

THE ENDOGENEITY OF FINANCIAL MARKET FRICTIONS

In the section on Walrasian equilibria, we described equilibria under the assumption that information about borrower type is publicly available. In the section on equilibria with credit rationing, we described candidate equilibria under the assumption that equation 16′ holds as an equality. In this section we ask when equation 16′ will and will not be an equality in equilibrium. When it is, credit rationing will occur. When it is not, self-selection occurs even with Walrasian allocations. In the former situation, financial market frictions are severe enough to affect the allocation of investment capital for entirely endogenous reasons. In the latter situation, it transpires—again for entirely endogenous reasons—that financial market frictions are not severe enough to affect allocations. One of our main results is that when the steady-state inflation rate is high enough, financial market frictions *must* matter and credit rationing *must* occur. Thus high enough rates of inflation imply that market frictions must adversely affect the extension of credit and capital formation as well.

When Are Walrasian Allocations Consistent With Self-Selection?

When do candidate Walrasian equilibria (sequences $\{k_t, m_t\}$ satisfying equations 14 and 15′ at equality) also satisfy equation 16′? For simplicity of exposition, we focus our discussion on steady states.

Walrasian steady states satisfy equation 16′ when equation 17 holds and when the implied value of k satisfies

$$[w(k)/kf'(k)]f'(k) \geq x\sigma/(1-\lambda). \tag{29}$$

We have already established that $[w(k)/kf'(k)] = a^{1/\rho}z(1+z)^{(1-\rho)/\rho}$; hence equation 29 is equivalent to

$$H[z^*(\sigma)] \geq a^{-1/\rho}[x\sigma/(1-\lambda)]. \tag{30}$$

We now demonstrate our fourth result.

RESULT 4. *Equation 30 is satisfied if and only if* $\underline{z}(\sigma) \leq z^*(\sigma) \leq \bar{z}(\sigma)$ *holds.*

Result 4 is proved in the Appendix. The result asserts that Walrasian allocations are consistent with self-selection if and only if the steady-state value of z under full information lies between the values of z solving equation 16′ at equality. When this condition is satisfied, the Walrasian allocation continues to constitute an equilibrium, even in the presence of the informational asymmetry. Endogenous factors allow the friction to be sufficiently mild that it does not affect the allocation of investment capital. Thus, when $z^*(\sigma) \in [\underline{z}(\sigma).\bar{z}(\sigma)]$. Walrasian allocations are equilibrium allocations. When $z^*(\sigma) \notin [\underline{z}(\sigma)\bar{z}(\sigma)]$. Walrasian allocations are inconsistent with self-selection and do not constitute legitimate equilibria.

Credit Rationing

We now ask the opposite question: When do solutions to equation 16′ at equality satisfy equation 15′? Since $f'(k) = a^{1/\rho}(1 + z)^{(1-\rho)/\rho}$, clearly they do so if and only if

$$\hat{z}(\sigma) \leq z^*(\sigma). \tag{31}$$

In particular, equation 31 asserts that credit can be rationed if and only if the solution to equation 22 yields a lower capital stock than would obtain under a Walrasian allocation. This observation has the following implication: The smaller (larger) solution to equation 16′ at equality is an equilibrium if and only if $\underline{z}(\sigma)[\bar{z}(\sigma)] \leq z^*(\sigma)$. We now put all these facts together.

The Steady-State Equilibrium Correspondence

Here we describe the full set of steady-state equilibria for each potential choice of σ. We begin by depicting $z^*(\sigma)$ and $\hat{z}(\sigma)$ simultaneously in Figure 4a and b. It is easy to check that $z^*(\sigma)$ is an increasing function, and that $z^*(a^{-1/\rho}) = 0$. Combining this with our previous results about the correspondence $\hat{z}(\sigma)$, it follows that there are three possible configurations of the steady-state equilibrium correspondence. We now briefly discuss each case. The first case is the one of primary interest to us.

Case 1. Suppose that

$$z^*(\hat{\sigma}) > -\rho. \tag{32}$$

Then we have the configuration depicted in Figure 4a.[20] The loci $z^*(\sigma)$ and $\hat{z}(\sigma)$ intersect twice at $\underline{\sigma}$ and $\bar{\sigma}$.[21]

For $\sigma < \underline{\sigma}$, $z^*(\sigma) < \underline{z}$ holds. Hence neither the Walrasian situation nor the credit rationing situation constitutes a legitimate equilibrium. Then if $\sigma < \underline{\sigma}$, there are *no* monetary steady states.

For $\sigma \in [\underline{\sigma},\bar{\sigma}]$, $z^*(\sigma) \in [\underline{z}(\sigma),\bar{z}(\sigma)]$ holds. It follows that the Walrasian steady state is consistent with self-selection whenever $\sigma \in [\underline{\sigma},\bar{\sigma}]$ and hence is a true steady-state equilibrium. At the same time, $\underline{z}(\sigma) \leq z^*(\sigma)$ also holds. Thus $\underline{z}(\sigma)$ is a legitimate steady state with credit rationing. Clearly, $\bar{z}(\sigma) > z^*(\sigma)$ holds for all $\sigma \in [\underline{\sigma},\bar{\sigma}]$ and hence $\bar{z}(\sigma)$ is *not* a legitimate steady state for $\sigma < \bar{\sigma}$. Thus, for $\sigma \in [\underline{\sigma},\bar{\sigma}]$ exactly two steady-state equilibria exist: one with credit rationing and one without. Our previous results imply that both steady states

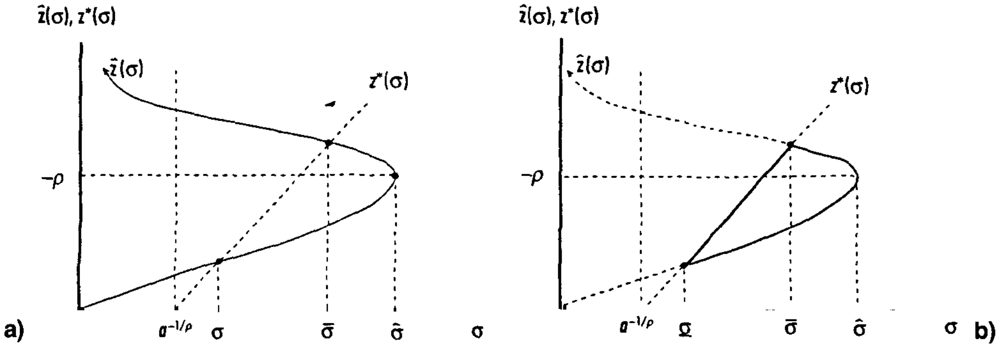


Figure 4 (a) Multiple steady states; (b) the steady-state equilibrium correspondence (case 1).

are saddles and hence that both can potentially be approached.[22] If credit rationing arises, the result will be that the capital stock is depressed. The capital stock is low, and therefore $w(k)$ must be low relative to $(b + \tau) = k/(1 - \lambda)$. This forces intermediaries to ration credit to induce self-selection. Credit rationing can thus arise for fully endogenous reasons.

Suppose that two intrinsically identical economies[23] with $\sigma \in [\underline{\sigma},\bar{\sigma}]$ land in different steady states. The economy with a low capital stock will experience credit rationing, while that with a high capital stock does not. Thus the better-developed economy will appear to have a better functioning financial system, as in fact it does. However, the inefficient functioning of capital markets in the poorer economy is a purely endogenous outcome.

When $\sigma > \bar{\sigma}$ holds, $z^*(\sigma) < \bar{z}(\sigma)$ holds as well. Hence Walrasian outcomes are no longer consistent with self-selection and they cannot be equilibria. Thus, when steady-state inflation exceeds a critical level ($\bar{\sigma}$) informational frictions *must* interfere with the operation of capital markets.

Since $z^*(\sigma) > \bar{z}(\sigma)$ for all $\sigma > \bar{\sigma}$, both $\bar{z}(\sigma)$ and $\underline{z}(\sigma)$ constitute legitimate equilibria with credit rationing. For $\sigma \in (\bar{\sigma},\hat{\sigma})$, there are thus again two steady-state equilibria. Our previous results indicate that one is a sink and one a saddle; hence both can potentially be approached. In the high-(low-) capital-stock steady state, credit rationing appears to be less (more) severe.

To summarize, in this case for $\sigma \in (\underline{\sigma},\sigma\hat{})$ potentially two steady-state equilibria exist. In one credit market frictions are relatively severe; in the other they are less so.

We have thus far not insisted that a steady-state equilibrium have a positive level of real balances. Keeping this condition in mind, we present our fifth result.

RESULT 5. *Suppose that* $A(\sigma) > 1$ *holds. Then any steady state has positive real balances.* (Result 5 is proved in the Appendix.[24])

In this case, then, the steady-state equilibrium correspondence is given by the solid locus in Figure 4b. For $\sigma \leq \bar{\sigma}$, the steady-state equilibrium value of z, and hence of k, increases with σ. Thus, *for low initial rates of inflation,* increases in σ result in higher steady-state capital stocks and output levels (*unless* increases in σ result in a shift from a Walrasian equilibrium to an equilibrium with credit rationing). However, for $\sigma > \bar{\sigma}$, equilibria lying along the upper branch of this locus will have z (and hence k) *decreasing* as σ increases. Thus, at high initial inflation rates, increases in σ can reduce long-run output levels. This situation is very consistent with the empirical evidence reviewed in the introduction.[25]

Case 2. (Figure 5a).

In this case $z^*(\sigma)$ and $\hat{z}(\sigma)$ (generically) have two intersections, as they did previously. In addition, for $\sigma < \underline{\sigma}$ there are *no* steady-state equilibria, as in Case 1. Similarly, for $\sigma \in [\underline{\sigma},\bar{\sigma}]$ there are two steady-state equilibria, just as in Case 1. However, here for $\sigma \in [\bar{\sigma},\hat{\sigma}]$, $z(\sigma) > z^*(\sigma)$ holds, so that neither the Walrasian nor the credit rationing allocations are legitimate steady states. Hence steady-state equilibria exist if and only if $\sigma \in [\sigma,\bar{\sigma}]$.

The steady-state equilibrium correspondence for Case 2 is depicted in Figure 5b. In this case no branch of the correspondence exists for which z (and k) are decreasing in σ. Thus this case cannot easily capture the empirical observations cited in the introduction.

Case 3.

Here $\underline{z}(\sigma) > z^*(\sigma)$ holds for all σ. It follows that there are *no* steady-state equilibria for any value of σ.

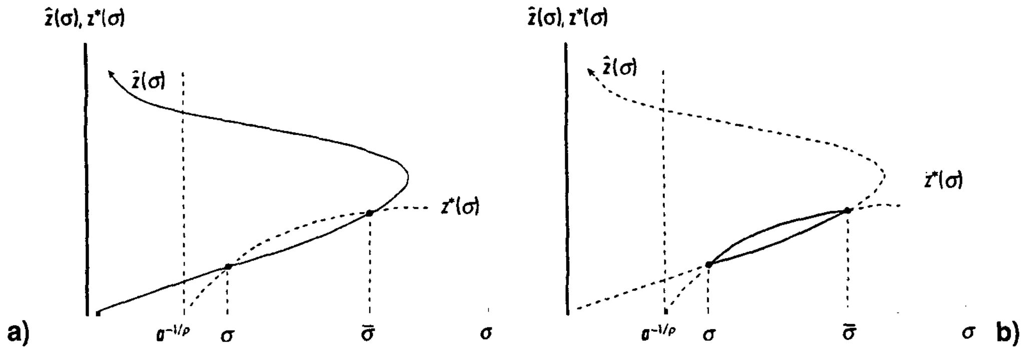


Figure 5 (a) Multiple steady states; (b) the steady-state equilibrium correspondence (case 2).

Discussion

Of the various possible configurations of the steady-state equilibrium correspondence, only that in Case 1 seems like it can easily confront empirical findings like those of Bullard and Keating (1995) and Bruno and Easterly (1995). We therefore regard this as the most interesting case and explore it somewhat further.

As shown in Result 3, some critical value (σ_c) of the money growth rate exists, with $\sigma_c < \hat{\sigma}$ such that for all $\sigma > \max\{\bar{\sigma}, \sigma_c\}$, equilibrium paths approaching the high-capital-stock steady state necessarily display endogenous oscillation. Then, in particular, if $(1 - \lambda)/x\lambda < \bar{\sigma}$ (see the Appendix), there are three distinct possibilities:

- $\sigma \in [\max\{\underline{\sigma}, (1 - \lambda)/x\lambda\}, \bar{\sigma}]$. Here one steady-state equilibrium exists with credit rationing and one exists without. Paths approaching both steady states do so monotonically. Increases in σ (within this interval) raise the capital stock in each steady state.[26]
- $\sigma \in (\bar{\sigma}, \sigma_c)$.[27] Here there are two steady-state equilibria, each displaying credit rationing. Dynamical equilibrium paths approaching each steady state may do so monotonically. In the higher of the steady states, increases in the steady-state inflation rate are detrimental to capital formation and the long-run level of real activity.
- $\sigma \in (\sigma_c, \hat{\sigma})$. Here there continue to be two steady-state equilibria with credit rationing (*if* $\sigma < \hat{\sigma}$). Now equilibrium paths approaching the high-capital-stock steady state necessarily display endogenous fluctuations. This is the price paid for avoiding convergence to the low-capital-stock steady state. Moreover, if low levels of real activity are to be avoided, high rates of money growth induce volatility in all economic variables, including the inflation rate. High rates of inflation are then associated with variable rates of inflation.

An Example

We now present a set of parameter values satisfying equations a4, A19 (implying that we have a Case 1 economy), A24′, A26′ (implying that intermediaries have no incentive to pool different agent types in any steady-state equilibrium), and A27 [implying that $(1 - \lambda)/\lambda x < \bar{\sigma}$]. One set of parameter values satisfying these conditions is given by $\hat{\sigma} = 2$, $\rho = -1$, $x = 1/32$, $\lambda = 63/64$, and $a = 1/16$. For these parameter values, equation a3 reduces to $\sigma \in (0.508, 2)$. These parameter values imply, parenthetically, that the government can allow the money supply to grow as rapidly as 100 percent per period, or could contract the money supply by as much as 49 percent per period. They also imply an empirically plausible elasticity of substitution between capital and labor of 0.5. It is also easy to check that, for all $\sigma > (1 - \lambda)/\lambda x$, the high-capital-stock steady state has labor's share exceeding capital's share, as is true empirically.

SOME EMPIRICAL EVIDENCE

The theoretical analysis of the previous sections yields several predictions that can be tested empirically.

1. Increases in the steady-state rate of inflation reduce the real returns investors receive.
2. Such increases can lead to greater inflation variability and also to greater variability in the returns on all assets.
3. Higher long-run rates of inflation raise steady-state output levels for economies whose rate of inflation is initially low enough.[28] For economies with initially high rates of inflation ($\sigma \geq \bar{\sigma}$) further increases in inflation must reduce long-run output levels, unless the economy is in a development trap.
4. When higher inflation is detrimental to long-run output levels, inflation adversely affects the level of activity in financial markets.

As we have noted, many of these results are empirically well-supported in the existing literature. For example, it is well-known that higher rates of inflation are typically accompanied by greater inflation variability, as shown in Friedman (1992). Similarly, the third implication listed above is consistent with the empirical evidence presented by Bullard and Keating (forthcoming) and Bruno and Easterly (1995), which we summarized in the introduction. We now address evidence for the remaining propositions.

Table 1 presents the results of four regressions using stock market data for the United States over the period 1958–93.[29] The dependent variables are the growth rate of the real value of transactions on the New York Stock Exchange (RV); real returns on the Standard & Poor's 500 Index, inclusive of dividends (RR); nominal returns on the Standard & Poor's Index, inclusive of dividends (NR); and the standard deviation of daily returns on the Standard & Poor's Index (V). The explanatory variable of interest is the rate of inflation in the Consumer Price Index (INF). Other explanatory variables are also employed. However, the results appear to be quite robust to the inclusion of other explanatory variables. These regressions were selected as being representative of a much larger set that we estimated. Finally, all data are reported as deviations from their sample means,[30] and pass standard stationarity tests in that form. In some regressions we corrected for serial correlation using a Cochrane-Orcutt procedure.

As is apparent from Table 1, higher rates of inflation significantly reduce the growth rate of stock market transactions. As predicted by theory, higher inflation thus attenuates financial market activity. In addition, as the inflation rate rises, the real return received by investors falls significantly.[31] Indeed, over this time period even nominal returns to investors appear to be negatively associated with inflation. Finally, higher inflation increases the volatility of stock returns. All of these results are consistent with the predictions of our model.

Figure 6 depicts the ratio of the value of stock market transactions (on the New York Stock Exchange) to gross domestic product (GDP), plotted against the rate of inflation. Apparently, higher inflation rates also tend to reduce the level of financial market activity using this particular measure.

Table 1 Regressions from Stock Market Data*: United States

(1)

$$RV_{\dagger}^{\dagger} = .00 + .01\ V(\dagger) + 2.9\ GIP(\dagger) - .05\ INF(\dagger)$$
$$\quad (.01)\quad (.003)^{\ddagger}\qquad (3.8)\qquad\qquad (.02)^{\ddagger}$$

$R^2 = .03$, DW $= 1.97$, $Q(60) = 77.7$

(2)

$$RR_{\dagger} = -\ .01 - .01\ RR(\dagger - 1) - .35\ V(\dagger) + .10\ RRAT(\dagger) - 2.9\ INF(\dagger)$$
$$\quad (.20)\quad (.05)\qquad\qquad (.10)^{\ddagger}\qquad (.12)\qquad\qquad (1.1)^{\ddagger}$$

$R^2 = .09$, $Q(60) = 56.2$

(3)

$$NR_{\dagger} = -\ .01 - .01\ NR(\dagger - 1) - .34\ V(\dagger) + .10\ RRAT(\dagger) - 1.9\ INF(\dagger)$$
$$\quad (.20)\quad (.05)\qquad\qquad (.10)^{\ddagger}\qquad (.13)\qquad\qquad (1.1)^{\ddagger}$$

$R^2 = .06$, $Q(60) = 56.4$

(4)

$$V_{\dagger} = .01 + .19\ RRAT(\dagger) - 2.87\ GIP(\dagger) + 2.26\ INF(\dagger)$$
$$\quad (.1)\quad (.08)^{\ddagger}\qquad (37.6)\qquad\qquad (.73)^{\ddagger}$$

$R^2 = .02$, DW $= 1.99$, $Q(60) = 71.5$

* Monthly, 1958–93.

† Denotes that a Cochrane-Orcutt procedure has been employed.

‡ Denotes significance at the 5 percent level or higher.

Standard errors are in parentheses.

DW: Durbin-Watson statistic.

Q: Ljung-Box Q statistic.

Table 2 reports the results of estimating similar regressions using stock market data from Chile. Here RV represents the growth rate of the real value of stock market transactions on the Santiago Stock Exchange, and RRAT is the real interest rate on 30–89 day bank deposits. A lack of daily data prevents us from examining the volatility of stock market returns. As in the case of the United States, we see that higher rates of inflation significantly

Figure 6 Ratio of the value of stock market transactions to GDP: United States.

Table 2 Regressions from Stock Market Data*: Chile

(1)

$$RV_{\dagger} = .00 + .75\,GIP(\dagger) - .30\,INF(\dagger)$$
$$\quad (.02)\quad (.20)^{\dagger} \qquad (1.5)$$
$$R^2 = .10,\ DW = 2.24,\ Q(33) = 25.5$$

(2)

$$RR_{\dagger} = .00 + .22\,RR(\dagger - 1) - .02\,RRAT(\dagger) - 2.56\,INF(\dagger)$$
$$\quad (.01)\quad (.08)^{\dagger} \qquad (.007)^{\dagger} \qquad (.68)^{\dagger}$$
$$R^2 = .19,\ Q(33) = 19.9$$

(3)

$$NR_{\dagger} = .00 + .17\,NR(\dagger - 1) - .02\,RRAT(\dagger) - 1.30\,INF(\dagger)$$
$$\quad (.01)\quad (.09)^{\dagger} \qquad (.007)^{\dagger} \qquad (.66)^{\dagger}$$
$$R^2 = .09,\ Q(33) = 19.7$$

* Monthly, 1981–91.

† Denotes significance at the 5 percent level or higher.

Standard errors are in parentheses.

DW: Durbin-Watson statistic.

Q: Ljung-Box Q statistic.

reduce investors' real and nominal rates of return on the stock exchange. The point estimate suggests that higher inflation also depresses the growth rate of market transactions, although here the point estimate is not significantly different from zero.

Figure 7 depicts the ratio of the value of stock market transactions to GDP for Chile, plotted against its rate of inflation. Again we perceive a negative relationship, particularly if the one single-digit inflation year (1982) is excluded as an outlier.

Tables 3 and 4 report analogous regression results for Korea and Taiwan. Here we proceed somewhat differently, since both Korea and Taiwan experienced fairly pronounced

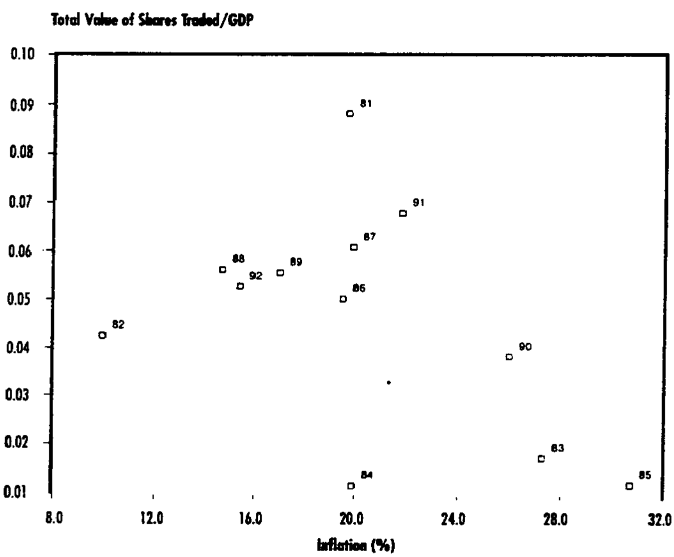


Figure 7 Ratio of the value of stock market transactions to GDP: Chile.

Table 3 Regressions from Stock Market Data*: Korea

A. 1982–87

(1)

$$RV_{\dagger} = .01 + .12\,V(\dagger) - .008\,GIP(\dagger) - .12\,INF(\dagger)$$
$$(.05)\ (.13)\quad (.008)\quad (.1)$$

$R^2 = .04$, DW = 2.3, Q(24) = 33.2

(2)

$$RR_{\dagger} = .00 - .17\,RR(\dagger - 1) + .25\,V(\dagger) - 1.69\,RRAT(\dagger) - 2.32\,INF(\dagger)$$
$$(.72)\ (.12)\quad (1.9)\quad (4.50)\quad (4.50)$$

$R_2 = .03$, Q(24) = 17.6

(3)

$$V^{\dagger}_{\dagger} = .04 + .29\,RRAT(\dagger) - .01\,GIP(\dagger) - .11\,INF(\dagger)$$
$$(.11)\ (.42)\quad (.005)^{\ddagger}\quad (.42)$$

$R^2 = .39$, DW = 2.34, Q(24) = 15.6

B. 1988–94

(1)

$$RV_{\dagger} = .00 + .34\,V(\dagger) + .02\,GIP(\dagger) - .28\,INF(\dagger)$$
$$(.07)\ (.16)^{\ddagger}\quad (.01)^{\ddagger}\quad (.13)^{\ddagger}$$

$R^2 = .11$, DW = 2.15, Q(27) = 29.4

(2)

$$RR_{\dagger} = -.19 - .18\,RR(\dagger - 1) + 1.9\,V(\dagger) - 6.5\,RRAT(\dagger) - 9.77\,INF(\dagger)$$
$$(.77)\ (.11)\quad (1.74)\quad (3.55)^{\ddagger}\quad (3.86)^{\ddagger}$$

$R^2 = .11$, Q(27) = 34.8

(3)

$$NR_{\dagger} = -.23 - .23\,NR(\dagger - 1) + 2.5\,V(\dagger) - 5.43\,RRAT(\dagger) - 7.11\,INF(\dagger)$$
$$(.73)\ (.11)^{\ddagger}\quad (1.72)\quad (3.36)\quad (3.63)^{\ddagger}$$

$R^2 = .10$, Q(27) = 32.1

(4)

$$V^{\dagger}_{\dagger} = -.01 + .33\,RRAT(\dagger) - .01\,GIP(\dagger) + .40\,INF(\dagger)$$
$$(.04)\ (.36)\quad (.006)^{\ddagger}\quad (.38)$$

$R^2 = .08$, DW = 1.95, Q(24) = 33.0

* Monthly.
† Denotes that a Cochrane-Orcutt procedure has been employed.
‡ Denotes significance at the 5 percent level or higher.
Standard errors are in parentheses.
DW: Durbin-Watson statistic.
Q: Ljung-Box Q statistic.

jumps in their rates of inflation in 1988 and 1989, respectively. In particular, in Korea the average monthly inflation rate was 0.27 percent over the period 1982–87 while from 1988–94 it was 0.54 percent. In Taiwan, the average monthly rate of inflation over the period 1983–88 was 0.07 percent, but jumped to 0.33 percent from 1989–93. These increases are apparent in Figures 8 and 9, respectively.

In view of these marked changes in the inflation rate, we proceeded as follows. For each country we divided the sample and ran regressions analogous to those reported above. For Korea the results are reported in Table 3. Over the low inflation period (1982–87), inflation has no significant effects on the real return on equity, its volatility, or on the growth rate of stock market transactions. However, during the period of higher inflation, increases

Table 4 Regressions from Stock Market Data*: Taiwan

A. 1983–88

(1)

$$RV_{\dagger} = -.02 + 1.1\ GIP(\dagger) + 9.1\ INF(\dagger)$$
$$\quad (.05) \quad (.52)^{\dagger} \quad (6.5)$$

$R^2 = .08$, DW = 1.8, Q(24) = 13.4

(2)

$$RR_{\dagger} = .00 + .23\ RR(\dagger - 1) - .01\ RRAT(\dagger) - 7.3\ INF(\dagger)$$
$$\quad (.02) \quad (.13)^{\dagger} \quad (.01) \quad (8.5)$$

$R^2 = .18$, Q(24) = 19.6

B. 1988–94

(1)

$$RV_{\dagger}^{\ddagger} = -.03 + .5\ GIP(\dagger) - 7.1\ INF(\dagger)$$
$$\quad (.05) \quad (.50) \quad (4.0)^{\dagger}$$

$R^2 = .11$, DW = 1.88, Q(21) = 14.1

(2)

$$RR_{\dagger} = .00 + .36\ RR(\dagger - 1) - .01\ RRAT(\dagger) - 18.7\ INF(\dagger)$$
$$\quad (.01) \quad (.11)^{\dagger} \quad (.01) \quad (10.1)^{\dagger}$$

$R^2 = .21$, Q(24) = 26.8

(3)

$$NR_{\dagger} = .00 + .35\ NR(\dagger - 1) - .01\ RRAT(\dagger) - 17.8\ INF(\dagger)$$
$$\quad (.01) \quad (.11)^{\dagger} \quad (.01) \quad (10.1)^{\dagger}$$

$R^2 = .17$, Q(24) = 25.4

* Monthly.
† Denotes significance at the 5 percent level or higher.
‡ Denotes that a Cochrane-Orcutt procedure has been employed.
Standard errors are in parentheses.
DW: Durbin-Watson statistic
Q: Ljung-Box Q statistic.

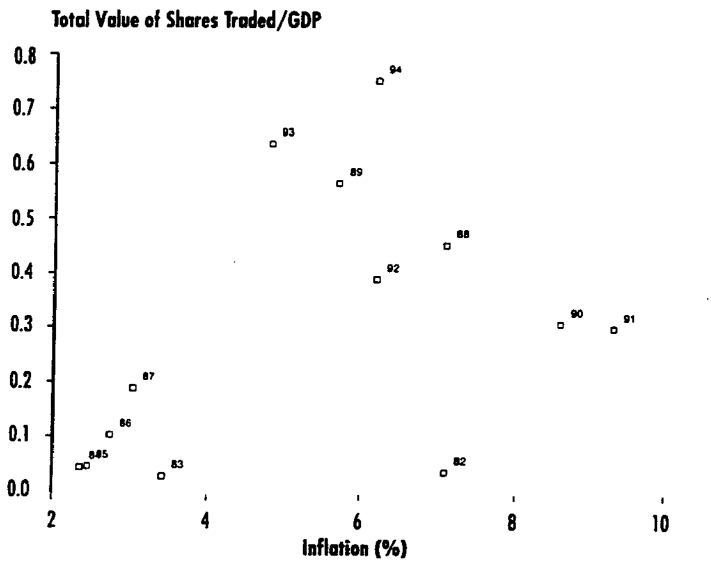


Figure 8 Ratio of the value of stock market transactions to GDP: Korea.

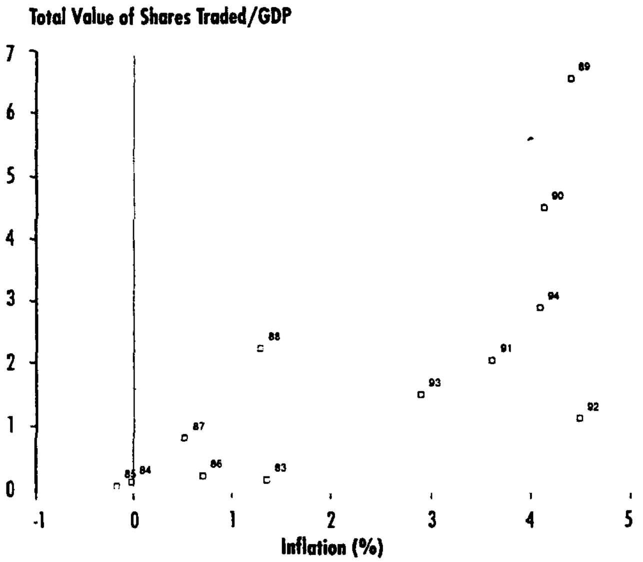


Figure 9 Ratio of the value of stock market transactions to GDP: Taiwan.

in the rate of inflation lead to statistically significant reductions in the growth rate of transactions and the real and nominal return on equity. With respect to the volatility of market returns, our point estimates again suggest that inflation leads to higher volatility, but the inflation coefficient is not significantly different from zero.

Figure 8 represents Korea's ratio of the value of stock market transactions to GDP and its rate of inflation. Clearly, in the higher inflation period of 1988–93, the negative relationship between market activity and the rate of inflation is highly pronounced. This is not the case for the low-inflation period 1982–87. Here then we see some evidence for threshold effects: Inflation seems to have significant adverse consequences only after it exceeds some critical level.

Table 4 repeats the same regression procedure for Taiwan, but lack of daily data prevents us from constructing a volatility of returns measure. Here we see a similar pattern to that for Korea. During the period of low inflation (1983–88), the effect of inflation on the growth rate of real stock market activity is insignificant and similarly for the real returns on equity. However, in the period of high inflation, both the growth rate of real equity market activity and the real returns on equity were adversely effected by inflation in a statistically significant way. Nominal equity returns are negatively associated with inflation, with a t-value of about 1.6. Here we see further evidence that inflation may be detrimental only after it exceeds some threshold level.

Figure 9 displays Taiwan's value of stock market transactions to GDP ratio, as well as its inflation rate. Clearly, this measure does not suggest that inflation has been detrimental to the level of equity market activity.

Table 5 shows simple correlations of the financial variables with the inflation rate for each of the countries and subperiods. These results are quite consistent with the regression results. On the whole, this empirical evidence seems to support our model's predictions. We have even seen evidence that inflation's adverse consequences may only be observed if the rate of inflation is sufficiently high.

Table 5 Comparisons Across Countries: Simple Correlations of Market and Macrovariables with the Inflation Rate (INF)

	Country/Period					
Variable	U.S. 1958–93	Chile 1981–91	Korea* 1982–87	Korea† 1988–94	Taiwan‡ 1983–88	Taiwan§ 1989–93
RV	−.06	.02	−.12	−.19	.12	−.20
RR	−.25	−.15	−.06	−.20	.12	−.23
V	.05	—	.09	.12	—	—
GIP	.02	.10	−.14	.03	−.12	−.24
RRAT	−.70	−.67	−.95	−.91	−.99	−.99

*Average monthly inflation rate 0.27 percent. † Average monthly inflation rate 0.54 percent. ‡ Average monthly inflation rate 0.07 percent. §Average monthly inflation rate 0.33 percent.

CONCLUSIONS

Both our theoretical analysis and our empirical evidence indicate that higher rates of inflation tend to reduce the real rates of return received by savers in a variety of markets.[32] When credit is rationed, this reduction in returns worsens informational frictions that interfere with the operation of the financial system. Once inflation exceeds a certain critical rate, a potential consequence is that the financial system provides less investment capital, resulting in reduced capital formation and long-run levels of real activity. Such forces need not operate at low rates of inflation, providing an explanation of why the consequences of higher inflation seem to be so much more severe once inflation exceeds some threshold level.

In addition, high enough rates of inflation force endogenously arising economic volatility to be observed. Thus, as we observe, high inflation induces inflation variability and variability in rates of return on all savings instruments. Theory predicts that this volatility should be transmitted to real activity as well.

Obviously, these results have been obtained in the context of a highly stylized and simplified model of the financial system. How general are they? We believe they are quite general. Boyd and Smith (1998) produce a model of a financial system that is subject to a costly state verification problem, one where investors provide some internal financing of their own investment projects. Again, two monetary steady-state equilibria exist and both can potentially be approached. Thus development trap phenomena arise. In the steady state with higher levels of real activity, higher inflation interferes with the provision of internal finance, thereby exacerbating the costly state verification problem. As a result, greater inflation reduces the long-run level of real activity, the level of financial market activity, and real returns to savers. Moreover, as is the case here, high enough rates of inflation force endogenously generated economic volatility to emerge. And, interestingly, Boyd and Smith (1998) obtain a result that is not available here: Once inflation exceeds a critical level, it is possible that only the low-activity steady state can be approached. Inflation rates exceeding this level can then force the kinds of crises discussed by Bruno and Easterly (1995). Related results are obtained by Schreft and Smith (1998 and 1994) in models where financial market frictions take the form of limited communication, as in Townsend (1987) and Champ, Smith, and Williamson (1996).

A shortcoming of all of the models mentioned—including ours—is that they do not give rise to distinct and/or interesting roles for debt and equity markets. Empirical evidence suggests that both kinds of markets are adversely affected by high inflation.[33] This is a natural topic for future investigation.

NOTES

1. This explanation has been articulated in a number of recent papers. See, for example, Azariadis and Smith (1996), Boyd and Smith (1998), and Schreft and Smith (1998 and 1994).
2. The same phenomena we report here occur in the presence of a costly state verification problem (Boyd and Smith forthcoming), or in a model where spatial separation and limited communication affect the financial system (Schreft and Smith 1998 and 1997).
3. In particular, in the absence of financial market frictions, our model reduces to one in which higher rates of inflation (easier monetary policy) stimulates real output growth. This occurs in a variety of monetary growth models: see Mundell (1965); Tobin (1965); Diamond (1965); or especially Azariadis (1993) (for an exposition); Sidrauski (1967); and Shell, Sidrauski, and Stiglitz (1969).
4. See Azariadis and Drazen (1990) for one of the original theoretical expositions of development traps.
5. If $\rho \geq 0$, our analysis is a special case of that in Azanadis and Smith (forthcoming). We therefore restrict attention here to $\rho < 0$. The assumption that $\rho < 0$ holds implies that the elasticity of substitution between capital and labor is less than unity. Empirical evidence supports such a supposition.
6. It is easy to verify that this assumption implies no real loss of generality.
7. This assumption implies that all capital investment must be externally financed, as will soon be apparent. This provides the link between financial market conditions and capital formation that is at the heart of our analysis.
8. Risk neutrality implies that there are no potential gains from the use of lotteries in the presence of private information.
9. The hallmark of models of credit rationing based on adverse selection or moral hazard is that different agents have different probabilities of loan repayment and hence regard the interest rate dimensions of a loan contract differently. See, for instance, Stiglitz and Weiss (1981) or Bencivenga and Smith (1993). Ours is the simplest possible version of such a scenario: Type 2 agents repay loans with probability one, while type 1 agents default with the same probability. Matters are somewhat different in models of credit rationing based on a costly state verification problem in financial markets. See, for instance, Williamson (1986 and 1987) and Labadie (1995). We will discuss such models briefly in the conclusion.
10. For models of informational frictions that do generate debt and equity claims, see Boot and Thakor (1993), Dewatripont and Tirole (1994), Chang (1986), or Boyd and Smith (1995a and b).
11. For a canonical adverse selection model, see Rothschild and Stiglitz (1976).
12. It is easy to verify that nondissembling type 1 agents will not wish to borrow if $R_{t+1} \geq \max(r_{t+1}, x)$. This condition will hold in equilibrium.
13. See Rothschild and Stiglitz (1976), or in this specific context, Azariadis and Smith (1996).
14. See equation 6.
15. An additional requirement of equilibrium is that intermediaries perceive no incentive to "pool" dissembling type 1 agents with type 2 agents and to charge an interest rate that compensates for the defaults by dissembling type 1 agents. Azariadis and Smith (forthcoming) show that there is no such incentive if $f'(k_{t+1}) \leq r_{t+1}/(1 - \lambda)$ holds for all t.
16. See, for example, Diamond (1965), Tirole (1985), or Azariadis (1993, chapter 26.2).
17. See, for instance, Diamond (1965), Tirole (1985), or Azariadis (1993, chapter 26.2).
18. In this analysis, inflation is inversely related to the return on real balances and hence to the return on savings. However, the intuition underlying our results is not dependent on real balances

earning the same real return as other savings instruments. Higher inflation will also reduce the return on savings in economies where nominal interest rate ceilings bind or where binding reserve requirements subject intermediaries to inflationary taxation. Binding interest rate ceilings and reserve requirements are very common in developing countries and are hardly unknown in the United States. Finally, our empirical results do support the notion that higher inflation does reduce the real returns received by investors (see the section on some empirical evidence).

19. Clearly $1/x > (1 - \lambda)/\lambda x$ can hold only if $\lambda > 0.5$. Equation a4 obviously implies this.
20. Equation 32 holds if and only if equation A19 holds, as established in the Appendix section on the existence of steady-state equilibria. Thus A19 gives a primitive condition under which Case 1 obtains.
21. The Appendix section that covers Result 6 proves that there are at most two intersections and that there are exactly two intersections in this particular case.
22. The existence of two saddles is possible because dynamical equilibria follow different laws of motion depending on whether a Walrasian regime or a regime of credit rationing pertains.
23. Except, possibly, for their initial capital stocks.
24. Strictly speaking, in any steady state with credit rationing, it is necessary that intermediaries perceive no arbitrage opportunities associated with "pooling" type 2 and dissembling type 1 agents (see footnote 11). The Appendix establishes that intermediaries perceive no such incentive, for any value of $\sigma \in [\underline{\sigma}, \bar{\sigma}]$, so long as $-\rho/x \leq \sigma^{-2}$ and $\underline{\sigma}^2 \geq (1 - \lambda)\,\hat{\sigma}^2$ both are satisfied.
25. For both outcomes to be consistent with positive levels of real balances, it is necessary that $(1 - \lambda)/x\lambda < \sigma^{-}$ holds. The Appendix establishes that $(1 - \lambda)/x\lambda < \sigma^{-}$ holds if either equations A27 or A28 and A29 are satisfied.
26. However, increases in σ can still result in a reduction in the steady-state capital stock if they induce transitions from the Walrasian to the credit rationing regime. The current analysis provides no guidance as to when such transitions might or might not occur.
27. Obviously we are assuming here that $\sigma_c > \bar{\sigma}$.
28. As we have seen, this is true along either branch of the steady-state equilibrium correspondence if $\sigma < \bar{\sigma}$. The statement in the text does require some qualification, though. In particular, as noted above, if higher inflation causes the economy to shift from the Walrasian to the credit rationed equilibrium for $\sigma < \bar{\sigma}$. then an increase in the inflation rate can cause long-run output to fall.
29. Data sources are listed in the Appendix.
30. We also ran the regressions reported without removing the sample means. This led to no differences in results.
31. Boudoukh and Richardson (1993), using a much longer time series, also find that higher rates of inflation have reduced real stock market returns in the United States and in the United Kingdom.
32. Further evidence on this point appears in Boyd, Levine, and Smith (1995).
33. See Boyd, Levine, and Smith (1995).

REFERENCES

Azariadis, C. *Intertemporal Macroeconomics.* Oxford: Basil Blackwell, 1993.

Azariadis, C. and Drazen, A. Threshold Externalities in Economic Development. *Quarterly Journal of Economics* (May 1990), pp. 501–26.

Azariadis, C. and Smith, B. D. Private Information, Money and Growth: Indeterminacy, Fluctuations, and the Mundell-Tobin Effect. *Journal of Economic Growth* vol. 1, no. 3 (Sept. 1996) 309–32).

Bencivenga, V. R., and Smith, B. D. Some Consequences of Credit Rationing in an Endogenous Growth Model. *Journal of Economic Dynamics and Control* (January–March 1993), pp. 97–122.

Boot, A. W. A., and Thakor, A. V. Security Design. *Journal of Finance* (September 1993), pp. 1349–78.

Boudoukh, J., and Richardson, M. Stock Returns and Inflation: A Long-Horizon Perspective. *American Economic Review* (December 1993), pp. 1346–55.
Boyd, J. H., Levine R., and Smith, B. D. Inflation and Financial Market Performance. Manuscript. Federal Reserve Bank of Minneapolis, 1995.
Boyd, J. H. and Smith, B. D. Capital Market Imperfections in a Monetary Growth Model. Economic Theory vol. 11, no. 2 (March 1998) 241–73.
Boyd, J. H. and Smith, B. D. The Evolution of Debt and Equity Markets in Economic Development. Manuscript. Federal Reserve Bank of Minneapolis, 1995a.
Boyd, J. H. and Smith, B. D. The Use of Debt and Equity in Optimal Financial Contracts. Manuscript. Federal Reserve Bank of Minneapolis, 1995b.
Bruno, M., and Easterly, W. Inflation Crises and Long-Run Growth. Manuscript. The World Bank, 1995.
Bullard, J., and Keating, J. The Long-Run Relationship Between Inflation and Output in Postwar Economics. *Journal of Monetary Economics* vol. 36, no. 3 (December 1995) 477–96.
Champ, B. Smith, B. D., and Williamson, S. D. Currency Elasticity and Banking Panics: Theory and Evidence. *Canadian Journal of Economics* vol. 29, no. 4 (November 1996) 828–64.
Chang, C. Capital Structure as Optimal Contracts. Manuscript. University of Minnesota, 1986.
Dewatripont, M., and Tirole, J. A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence. *Quarterly Journal of Economics* (November 1994), pp. 1027–54.
Diamond, P. A. National Debt in a Neoclassical Growth Model. *American Economic Review* (December 1965), pp. 1126–50.
Friedman, M. *Money Mischief: Episodes in Monetary History.* Harcourt-Brace-Jovanovich, 1992.
Labadie, P. Financial Intermediation and Monetary Policy in a General Equilibrium Banking Model. *Journal of Money, Credit and Banking* (November 1995), pp. 1290–1315.
Mundell, R. Growth, Stability, and Inflationary Finance. *Journal of Political Economy* (April 1965), pp. 97–109.
Rothschild, M., and Stiglitz, J. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics* (November 1976), pp. 629–650.
Schreft, S. L., and Smith, B. D. The Effects of Open Market Operations in a Model of Intermediation and Growth. *Review of Economic Studies* vol. 65, no. 3 (July 1998) 519–50.
Schreft, S. L. and Smith, B. D. Money, Banking, and Capital Formation. *Journal of Economic Theory* vol. 73, no. 1 (March 1997) 157–82.
Shell, K., Sidrouski, M. and Stiglitz, J. Capital Gains, Income, and Saving. *Review of Economic Studies* (January 1969), pp. 15–26.
Sidrauski, M. Inflation and Economic Growth. *Journal of Political Economy* (December 1967), pp. 796–810.
Stiglitz, J., and Weiss, A. Credit Rationing in Markets with Imperfect Information. *American Economic Review* (June 1981), pp. 393–410.
Tirole, J. Asset Bubbles and Overlapping Generations. *Econometrica* (November 1985), pp. 1499–1528.
Tobin, J. Money and Economic Growth. *Econometrica* (October 1965), pp. 671–84.
Townsend, R. M. Economic Organization with Limited Communication. *American Economic Review* (December 1987), pp. 954–71.
Williamson, S. D. Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing. *Journal of Monetary Economics* (September 1986), pp. 159–79.
Williamson, S. D. Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing. *Quarterly Journal of Economics* (February 1987), pp. 135–45.

This Page Intentionally Left Blank

38

Is Noninflationary Growth an Oxymoron?

David Altig
Vice President and Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Terry Fitzgerald and Peter Rupert
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Just before the Federal Open Market Committee's (FOMC) May 20 meeting, popular opinion about the near-term future of U.S. monetary policy was summarized by John O. Wilson, chief economist at BankAmerica Corp.:

> Mr. Wilson views the economy as continuing to expand too fast for the Fed's comfort and anticipates that a series of central bank moves will be needed to bring the economy back onto what economists call the sustainable non-inflationary growth path.[1]

The FOMC did not choose to alter the average level of the federal funds rate at its May meeting. A typical interpretation of this decision appeared in the May 21, 1997, *Los Angeles Times:*

> The decision by the . . . Federal Open Market Committee was designed to provide time for analysts to determine whether the economy is slowing down on its own . . . or will require additional reining in.[2]

These observations underlie one of the most widely held and persistent beliefs about the "theory" of inflation; that is, inflationary pressures will inevitably result from high levels of economic activity, defined as real GDP growth that exceeds some "natural," or normal, rate. The obvious consequence of such a belief—duly expressed in the quotations above—is that if the Fed desires to contain inflation, it must also contain economic growth.

This is indeed a predicament for a central bank that by its own pronouncements desires to conduct monetary policy to maximize the well-being of the average citizen. There is, of course, a distinction between a policy aimed at stabilizing output growth near its long-term trend and one designed to "fight growth" more generally. But the distinction is a subtle one, and the casual observer might be forgiven for not understanding why the goal of long-term economic growth appears to require periodic policy actions that seem aimed at slowing growth.

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (May 1, 1997).

This confusion is unnecessary and unproductive, because much of the popular commentary about monetary policy, inflation, and the pace of real economic activity is based on a none-too-accurate portrayal of economic theory and evidence. Economic growth is not the enemy of low inflation, and expanding employment and income do not, in and of themselves, threaten the Federal Reserve's legitimate role in protecting the purchasing power of money.

The contrary perception is, at least in part, due to a failure to communicate (for which those of us in the business of central banking are not blameless). In particular, the long-established and widely held theory of money, prices, and income does not suggest an obvious linkage between high levels of economic activity and high rates of inflation (or, more specifically, between accelerating inflation and growth in excess of "potential").[3] Just the opposite, in fact: Higher GDP growth should put *downward,* not upward, pressure on prices.

This *Economic Commentary* reviews the theoretical and empirical case for disinflationary economic growth. The basic story line is as follows: Rising prices follow from nominal money supply growth in "excess" of its demand.[4] More rapid GDP growth, however, implies an increase in the growth of money demand. Thus, everything else being equal, an uptick in GDP growth should lead to disinflation, not rising inflation.

The tricky step between theory and reality, of course, is that all else is rarely equal. Inflation and above-trend growth have tended to coincide in the past. But it is important to recognize that this can arise because growth is sometimes associated with other changes that exert upward pressure on prices, not because growth per se is inflationary. This message has been lost as the correlation between "excessive" output growth and changes in the inflation rate has become enshrined in the "Phillips curve" (discussed below). However, the stability of this relationship and the statistical regularities that underlie it are as much apparent as real. Appreciating this goes a long way toward explaining why the U.S. economy can safely buck the conventional wisdom and experience substantial noninflationary economic growth.

SOME SIMPLE THEORY

At the most basic level, the average price level—let's call it P—is the total units of money required to purchase one unit of a hypothetical, representative real good or service.[5] Holding the growth rate of money fixed, a positive productivity shock that raises production in the economy increases the private sector's desire to hold monetary assets. This requires the purchasing power of money in terms of goods and services to rise (that is, disinflation results) in order to maintain equality between demand and supply. Conversely, when there is an increase in the supply of money that does not directly affect money demand, the purchasing power of money will fall (that is, inflation occurs).

This is the essence of the theory of inflation: When the (nominal) money supply grows faster than the demand for (real) money balances, P grows, which is to say inflation occurs.[6] Thus, given the growth rate of money (which is ultimately controlled by the central bank), the rate of inflation is dependent on the growth rate of money demand.[7]

What, then, determines money demand? According to accepted economic theory, part of the answer is income, which for practical purposes can be measured by real GDP. Because income is related to spending, and money is held precisely because of its usefulness in facilitating transactions, higher income (GDP) translates into higher money demand (all else equal).

Thus, the simple theory of money, growth, and inflation yields the following syllogism:

1. The price level rises less rapidly (inflation falls) when the demand for money rises more rapidly than its supply.
2. Money demand rises when GDP rises, all else equal.
3. Thus, holding the growth rate of money fixed, inflation falls when GDP rises.

Inflation that persists when output is growing at its long-run average rate is thus attributable to monetary growth in excess of its demand, which, as an empirical matter, also increases at about the long-run average rate of GDP growth. Temporary accelerations of output growth beyond the normal rate will therefore cause inflation to deviate from its trend. However, holding all else constant, prices in this circumstance should grow *more slowly* than normal, not *more quickly,* as is often asserted.

IS EVERYONE CRAZY?

If theory speaks so clearly on the relationship between growth and inflation, why do so many people think that rapidly rising GDP is inflationary? Part of the answer can be found by expanding on the simple theory developed thus far. In addition to income, the theory on the determination of money demand identifies a second key variable: "the" nominal interest rate.

The nominal interest rate determines the opportunity cost of holding monetary assets. The higher the interest rate, the greater is the loss from holding wealth in the form of money instead of alternative, higher-yielding nonmonetary assets.[8] Thus, an increase in market interest rates will tend to reduce the demand for money which, all else equal, will put upward pressure on prices.

There is one more piece to the puzzle. If, at a time of expanding output, the demand for goods and services grows even faster—as might happen if businesses and consumers expect times to be even better in the future—interest rates will rise. Holding monetary policy (money growth) constant, inflation will tend to increase (at least in the short run) if the negative impact on money demand from rising interest rates dominates the positive influence of more rapid GDP growth. Rising prices in this event are not the result of growth per se, but rather of demand-driven interest rate pressures that are correlated with expanding economic activity, which in turn reduces the demand for money relative to its supply.[9]

Two related and important lessons are suggested by this discussion. First, the "fact" that a high level of economic activity causes inflation is not a fact at all. To the extent that price pressures and accelerations of short-run growth are positively correlated, this relationship results from the tendency for goods and services demand and market interest rates to accelerate along with output, and for money demand to decline as a consequence.

Second, the "inevitability" of inflationary pressures when GDP growth rises substantially above trend is critically dependent on the stability of these historical correlations. In other words, the prediction that growth "causes" inflation can rest securely only on the presumption that the impulse for growth in the final demand for goods and services will always outpace that for supply in periods of rapidly expanding GDP.

This scenario, however, suggests a different perspective than the one offered by the conventional wisdom. Although it may be appropriate to "tighten" monetary policy in periods of high demand, this need not be construed as an attempt to rein in output growth. An equally plausible interpretation is that the intent of such a policy is to slow money growth to match the realities of the changing demand for monetary assets.

THE PHILLIPS CURVE: A RELIABLE RULE OF THUMB?

"So what?" might be a reasonable response to the discussion above. As long as there is a stable and predictable relationship between changes in the inflation rate and GDP growth in excess of its long-run average, theoretical niceties are just that: Nice stories that, although intellectually interesting, have little practical importance for the appropriate conduct of monetary policy. As long as above-normal growth ultimately yields higher inflation, the policy implication—restrain money growth—is the same whether you surround the observation with a simple story or a complicated one. And, the argument goes, the case that "above-normal" GDP growth is inevitably associated with inflationary pressures is strongly supported by two well-known empirical propositions known as Okun's law and the Phillips curve.

Okun's law, named after the late economist Arthur Okun, is a rule-of-thumb relationship between output and unemployment. In its simple form, it is no more than a statement about the negative correlation between output growth and changes in the unemployment rate.[10] From Okun's law, one might divine the relationship between inflation and output growth via the so-called Phillips curve. The Phillips curve is yet another statistical rule of thumb that posits a negative relationship between changes in inflation and changes in the unemployment rate. Because output rises as the unemployment rate falls (from Okun's law), the Phillips-curve relationship suggests a predictable (positive) connection between changes in GDP growth and changes in price-level growth.

Although the high-inflation, high-unemployment experience of the 1970s had caused older representations of the Phillips curve to fall into some disrepute, the incorporation of inflation expectations and subsequent statistical refinements have resulted in its resurrection as a widely used tool for thinking about policy. It is common now to hear the Phillips-curve and Okun's-law relationships referred to as among the most reliable in macroeconomics. Because they form the foundation for arguing that overly robust GDP growth creates inflationary pressures, it is clear that this opinion is widely held.

We argue that the implicit message of modern versions of the Phillips curve—"too rapid growth causes accelerating inflation"—deserves further scrutiny. Figure 1 compares

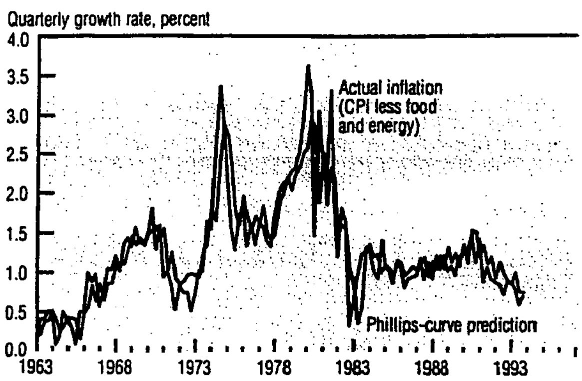


Figure 1 Actual and predicted inflation.

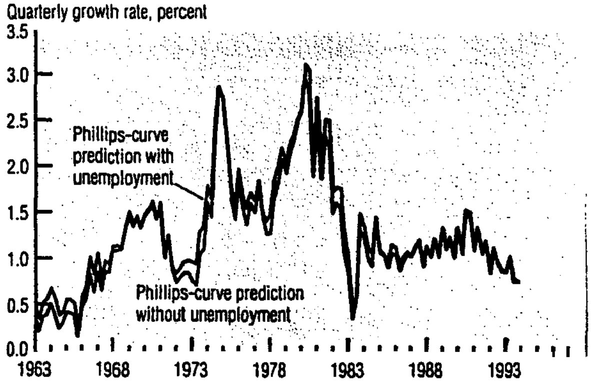


Figure 2 Predicted inflation with and without unemployment. **Sources:** U.S. Department of Labor, Bureau of Labor Statistic and authors' calculations.

actual quarterly inflation rates from 1963 to the present with rates predicted by one variant of the Phillips curve (based on Jeff Fuhrer's "The Phillips Curve Is Alive and Well," which appeared in the March/April 1995 edition of the Federal Reserve Bank of Boston's *New England Economic Review*). This particular model was chosen because it represents a particularly careful, thoughtful, and presumably successful variant of the Phillips curve.

As Figure 1 shows, the model appears to conform quite well to the actual inflationary experience of the U.S. economy over the past 30 years. The fact that it was estimated for this entire period is one of its particularly important features, because the most common criticism of the Phillips curve is its reputed instability as a forecasting tool. However, another important feature of the model is little appreciated: The success of this version of the Phillips curve appears, at least in recent years, to result in large part from the inclusion of very long lags of the inflation rate. Figure 2 shows inflation predictions with and without unemployment included in the specification. Over the 1963–93 period, unemployment rate changes—which through Okun's law relate inflation to output growth—do add to the model's predictive power. Since the late 1980s, however, the predictive value of changes in the unemployment rate is virtually zero. (Estimates are calculated through 1993:IVQ, reflecting the last available observation of the unemployment series before the survey redesign.)

IS IT TIME TO RETHINK THE CONVENTIONAL WISDOM?

In light of our earlier discussion, it is not particularly surprising that a rule of thumb relating changes in GDP growth relative to some notion of potential (sometimes called an "output gap") and changes in the inflation rate might, at least periodically, fail to capture the dynamics of price-level growth. The statistical relationship between output gaps and accelerating inflation is several steps removed from the direct determinant of price-level pressures, which is the relationship between the growth rates of money demand and supply. The notion that growth causes inflation—even growth in excess of normal levels—never was complete because it critically omits the "money part" of the story, and accepted theories of money demand and price-level determination clearly predict that rising GDP should cause the inflation rate to fall rather than rise.

This is not to say that the popular view of growth and inflation is utterly without foundation. However, the case for a positive connection between expanding GDP and inflationary pressures was always contingent on the presumption that demand pressures inevitably arise as a normal characteristic of the rapid expansion phases of a business cycle. The operationalization of this presumption has traditionally come from reportedly reliable and stable relationships between changes in inflation and measures of real activity. But the reliability and stability of these relationships are sufficiently suspect to draw into question their usefulness in thinking about policy today.

The recent economic environment of rapid growth and nonaccelerating inflation has left many people puzzled. But such a scenario is clearly possible from a theoretical standpoint: If accelerating inflation and presumed output gaps went together in the past, that is certainly no guarantee they must do so now or in the future. Furthermore, the simple statistical framework underlying the conclusion that an acceleration of price-level growth must follow from an acceleration of output growth "beyond capacity" is not as compelling as is often assumed.

It is an opportune time to reevaluate the language of monetary policy discussions. As with the inflationary episode in the 1970s, conventional rules of thumb have been hard-pressed to account for recent events. Perhaps the information revolution brought on by rapid advances in computer technology has broken down many of the traditional macroeconomic regularities that have informed our thinking about economic policy, resulting in an absence of money demand pressures that once may have accompanied output growth above levels considered normal. (Perhaps the answer is as simple as a significant change in the "normal" rate of GDP expansion.)

In any event, it is incumbent upon economists and policymakers alike to strive to communicate a deeper understanding of how various shocks to our economy affect output, unemployment, and inflation. Rules of thumb that equate rapid output growth with accelerating inflation do more than create bad advertising for monetary policies aimed at pursuing price stability. They enshrine as theory statistical connections that are, at best, indirectly connected to the ultimate determinant of price-level growth, which is to say the demand and supply of money. As such, they retard a more informed public discussion of monetary policy and make the job of the policymaker that much more difficult.

NOTES

1. Gordon Matthews, "Brace Yourself: 10 Out of 10 Economists Expect Fed Hike," *American Banker,* May 19, 1997.
2. Art Pine, "Wary Fed Decides against Interest Rate Hike for Now," *Los Angeles Times,* May 21, 1997.
3. "Potential" GDP growth is typically taken to be synonymous with "long-run average" GDP growth. Economists often refer to this as the "steady-state" rate.
4. In equilibrium, supply equals demand. More specifically, we are describing a condition in which prices rise precisely because money would be in excess supply if they didn't.
5. More detailed accounts of the simple, and thoroughly standard, theory discussed in this section can be found in almost any introductory economics textbook. See, for example, Alan Stockman, *Introduction to Economics,* Fort Worth: Dryden Press, 1996, chapter 27.
6. A simple example clarifies the distinction between nominal and real variables. Suppose that the money supply consists solely of dollar bills. The nominal supply of money would then just be the number of dollar bills in circulation. The real money supply would be the nominal stock ex-

pressed in terms of "purchasing power": How many units of goods and services can be purchased with the stock of money? For example, suppose that the stock of money, *M,* is $5 million, and the price level, *P,* is 2. Because the price level is the number of units of money required to purchase one unit of output, the real stock of money (in units of output) is 5/2 = 2.5.

7. This statement—which implicitly invokes the economist's standard "all-else-equal" clause—is not meant to minimize the difficulties inherent in controlling the money supply.
8. To be a bit more precise, opportunity cost is typically measured as the difference between the return on short-term Treasury securities and a measure of the return on a particular monetary aggregate, such as M2. For a recent discussion of the operational relationship between money and opportunity cost, see John B. Carlson and Benjamin D. Keen, "M2 Growth in 1995: A Return to Normalcy?" Federal Reserve Bank of Cleveland, *Economic Commentary,* December 1995.
9. There is another possible source for rising interest rates: rising expectations of inflation. The role of inflation expectations can significantly complicate the simple theory presented here and make things difficult indeed for monetary policymakers.
10. For a more complete discussion of Okun's law, see David Altig, Terry Fitzgerald, and Peter Rupert, "Okun's Law Revisited: Should We Worry about Low Unemployment?" Federal Reserve Bank of Cleveland, *Economic Commentary,* May 15, 1997.

This Page Intentionally Left Blank

39

Conducting Monetary Policy with Inflation Targets

George A. Kahn
Vice President and Economist, Federal Reserve Bank of Kansas City, Kansas City, Missouri

Klara Parrish
Assistant Economist, Federal Reserve Bank of Kansas City, Kansas City, Missouri

Since the early 1990s, a number of central banks have adopted numerical inflation targets as a guide for monetary policy. The targets are intended to help central banks achieve and maintain price stability by specifying an explicit goal for monetary policy based on a given time path for a particular measure of inflation. In some cases the targets are expressed as a range for inflation over time, while in other cases they are expressed as a path for the inflation rate itself. The measure of inflation that is targeted varies but is typically a broad measure of prices, such as a consumer or retail price index.

At a conceptual level, adopting inflation targets may necessitate fundamental changes in the way monetary policy responds to economic conditions. For example, inflation targeting requires a highly forward looking monetary policy. Given lags in the effects of monetary policy on inflation, central banks seeking to achieve a target for inflation need to forecast inflation and adjust policy in response to projected deviations of inflation from target. But central banks without an explicit inflation target may also be forward looking and equally focused on a long-run goal of price stability. Thus, at a practical level, adopting inflation targets may only formalize a strategy for policy that was already more or less in place. If so, targets might improve the transparency, accountability, and credibility of monetary policy but have little or no impact on the way policy instruments are adjusted to incoming information about the economy.

This article examines how central banks have changed their policy procedures after adopting explicit inflation targets. The first section summarizes the key features that characterize and motivate most inflation targeting regimes. The second section documents the procedural changes that a number of central banks have taken to implement inflation targeting. The third section examines empirical evidence to see if and how inflation targets have changed the way monetary policy reacts to economic information. The article concludes that, while inflation targets have perhaps improved the transparency, accountability,

---

Reprinted from: *Federal Reserve Bank of Kansas City Economic Review,* v. 83, n. 3 (Third Quarter 1998) 5–32.

and credibility of monetary policy, it is difficult to discern any significant and systematic changes in the way policymakers adjust policy instruments to incoming information after adopting inflation targets.

RATIONALE FOR AND KEY FEATURES OF INFLATION TARGETING

Central banks have adopted inflation targets as a strategy for achieving, and then maintaining, price stability. Inflation targeting regimes share several common features. This section describes the conceptual rationale for most inflation targeting regimes, as well as their common features.[1]

Rationale

Inflation targets can be set by the government, jointly agreed upon by the central bank and the government, or set by the central bank itself. The ultimate rationale for targets is to help the central bank achieve a desired long-run level of inflation, usually a measured rate of inflation consistent with "price stability." The inflation rate deemed consistent with price stability varies from country to country but generally falls within a range of 0 to 3 percent annually as measured by a broad index of consumer prices.[2] Inflation targets are designed to help the central bank achieve long-run price stability in three principal ways: by providing a nominal anchor for monetary policy, by improving the transparency and accountability of monetary policy, and by enhancing the central bank's inflation-fighting credibility.

Providing a Nominal Anchor

One rationale for inflation targets is that they supply a nominal anchor for monetary policy. Without such an anchor, policy actions can drift under the influence of short-run economic disturbances and, in the process, become inconsistent with long-run goals. With a nominal anchor, policy is bound to a long-run goal—such as price stability—that ties down inflation expectations but retains the slack needed to respond to short-run disturbances. Traditional nominal anchors for monetary policy have included monetary aggregates and exchange rates.

During the 1980s and 1990s, a number of countries abandoned these more traditional anchors. One reason was that the relationship of monetary aggregates to economic activity broke down in many countries, leaving those central banks that targeted monetary aggregates relying more on discretion and looking at a wide range of information for guidance. With most of these central banks using a very short-term nominal interest rate as the instrument of monetary policy, some analysts became concerned that, without an explicit target, monetary policy could develop an inflationary bias. For example, if policymakers were slow to react to rising inflation expectations, short-run real interest rates would fall, leading to an increasingly accommodative policy at a time when policy might need to be tightened.

Other countries that used exchange rate targets abandoned them when exchange rates became misaligned. The problem with using exchange rates as targets is that monetary policy must be directed at keeping the exchange rate within its target range, sometimes at the expense of promoting favorable domestic macroeconomic performance. If exchange rate targets are consistent with favorable macroeconomic performance, they can work well as an anchor for policy. When exchange rates become misaligned, however, a central bank may find itself defending the foreign exchange value of its currency at the cost of achieving goals for the domestic economy. When this happens, speculators may attack the currency, leading possibly to a realignment of exchange rate targets or, at the extreme, their demise as a guide to monetary policy. An example of this phenomenon occurred in the United Kingdom during the crisis in Europe's exchange rate mechanism (ERM) in 1992, when the UK left the ERM and established inflation targets as its anchor for monetary policy.

Improving Transparency and Accountability

Another rationale for inflation targets is that they can improve the transparency of monetary policy and the accountability of monetary policymakers. Inflation targets are highly transparent because they convey to the public a precise, readily understood goal for monetary policy. For example, an inflation target under which a central bank commits to "keep increases in the consumer price index between 1 and 3 percent annually from now until the end of the year 2001" gives the public a clear signal of both near-term and longer term plans. A central bank with such a target provides a clearer signal than a bank that simply commits to "achieving price stability in the long run," without specifying a numerical definition of price stability or a time frame for achieving it. Of course, the more precise a target, the easier it is to tell whether a target is hit or missed. And when targets are missed, policymakers have to explain why. Advocates of inflation targeting argue that explaining target misses increases transparency. Critics contend that inflation targets might give policymakers too strong an incentive to hit the target at the expense of adverse short-run fluctuations in output and employment.[3]

Along with increased transparency, inflation targets enhance accountability.[4] They do this by making it easier to judge whether policy is on track. An explicit numerical target for a specific measure of inflation is either hit or missed. When the target is missed, policymakers can be called on to explain why the target was missed. In many cases, the target will be missed because of special circumstances that are entirely justifiable. For example, if oil prices rise sharply and unexpectedly, policymakers might be unable to prevent a temporary increase in the overall inflation rate. And, to the extent the increase in inflation stemming from a one-time increase in oil prices was expected to be temporary, the appropriate monetary policy response might be to do nothing, accepting temporarily higher inflation.[5]

If the target is missed because of monetary policy mismanagement—admittedly a difficult thing to prove given the technical nature of monetary policy and the wide range of views about the effects of monetary policy on the economy—the government can hold the central bank accountable. Theoretically and in the extreme, the government could dismiss the chief monetary policymaker(s) or restructure the central bank. More realistically, the government can ask the central bank to improve its performance under threat of a range of sanctions.

Enhancing Credibility

One reason transparency and accountability are important is that they potentially enhance central bank credibility. That is, they help the public understand the goal of monetary policy and the commitment of the central bank to the goal. Credibility is important in central banking because it feeds into the public's formulation of expectations about future inflation. If consumers and businesses believe the central bank is committed to achieving price stability, they will accept lower nominal wage increases, incorporate lower inflation and inflation risk premiums into asset prices, and be more willing to make long-term commitments based on economic fundamentals instead of inflation expectations. This "credibility effect" can help reduce the output loss that typically accompanies disinflationary monetary policies.[6]

Features

Many inflation targeting regimes share common features designed to help the central bank achieve a more transparent, accountable, and credible monetary policy.

Reliance on Forecasts.

Because monetary policy actions affect the economy with significant lags, policymakers must rely on inflation forecasts to help them aim for an inflation target. Specifically, monetary policy actions generally affect output and employment with lags of six months or longer and affect inflation with lags of 18 months or more. As a result, policymakers must take action based on forecasts of inflation one to two years into the future. For example, if under the current setting of monetary policy instruments, inflation is projected to rise above target one year from today, policymakers might need to take action *now* to tighten the current stance of monetary policy. Waiting to see inflation rise before tightening policy may result in missing the inflation target.

Policymakers can, and do, use a variety of methods to forecast inflation. They can look at private forecasts, use information from financial markets, and make projections based on various econometric models of the economy. Whatever the approach, a necessary condition for the successful use of inflation targets is that the central bank has some capability of forecasting inflation based on the current stance of monetary policy. And, except in unusual circumstances, the central bank has to be willing to take timely actions to change the stance of policy when an unchanged stance would lead to a target miss.

Use of Inflation Reports

In helping to achieve transparency, most central banks that target inflation regularly issue an inflation report. The publication of these inflation reports is one of the key innovations of inflation targeting regimes. The purpose of the report is to explain what the central bank's inflation target is, describe how inflation has behaved relative to its target, and indicate where inflation may be headed in the future. Toward this end, some inflation targeting central banks actually publish their forecast for inflation, as well as a discussion of the risks surrounding that forecast. In addition, the central bank may use the inflation report to explain why a target may have been missed and what actions, if any, might be necessary to bring inflation back to its target.

Allowance for Flexibility

A final common feature of inflation targeting regimes is built-in flexibility. Given the difficulty of forecasting inflation and the likelihood that many economic shocks will have only temporary effects on inflation, all inflation targeting regimes allow the central bank to sometimes miss its target. When such misses occur, however, the central bank is expected to explain why. Acceptable explanations include transitory inflation shocks beyond the control of the central bank, such as changes in oil prices, natural disasters, and the first-round effect of excise tax changes. In addition, when the economy is weak and unemployment rising, most inflation targeting regimes allow inflation to temporarily exceed its target. To the extent rising unemployment is associated with future declines in inflation, an easing of monetary policy may very well be consistent with the long-run target for inflation even though it may exacerbate the near-term overshooting of the inflation target.

Why Doesn't Every Central Bank Target Inflation?

Given the favorable rationale and features of inflation targets, why don't all central banks target inflation? There are three main reasons. First, a number of central banks have effectively managed monetary policy without the use of inflation targets. For example, in the United States the Federal Reserve has managed to lower inflation over the last several years while fostering sustained economic growth without explicit numerical inflation targets. Second, some governments have mandated that the central bank achieve multiple goals. Again, in the United States, Congress has required the Federal Reserve "to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." Some analysts argue that inflation targets place too great an emphasis on the long-run inflation objective without providing explicit short-run objectives for output and employment variability (Cecchetti, 1998). And third, inflation targets require the support of the government and, in some cases, are given to the central bank by the government—usually the finance minister. Unless a political consensus has emerged that inflation targets are useful, governments are unlikely to impose them and central banks are less likely to adopt them on their own.

PROCEDURAL CHANGES FROM INFLATION TARGETING: CASE STUDIES

Central banks in nine countries currently conduct monetary policy with explicit inflation targets (Table 1). Following the lead of the Reserve Bank of New Zealand in 1990, these banks generally target an inflation rate below 3 percent. The Central Bank of Chile and the Bank of Israel target higher inflation rates—along with exchange rates—as they seek to bring inflation down from relatively high levels. While some banks specify a particular date for reaching their inflation targets, others do not give a time frame—either for when the target must be reached or for how long it will be valid.[7] Most countries target a broad price measure, such as an all-items consumer or retail price index, but allow temporary departures or "exemptions" from the targets. Some countries, such as the United Kingdom and New Zealand, have incorporated certain exemptions into the price measure itself. In roughly half of the countries, the government establishes the target, often in consultation with the central bank. In the other half, the central bank sets the target itself. Almost all banks publish an inflation report, but only a few currently publish an inflation forecast.

Despite the variety of institutional procedures used in implementing inflation targets, most inflation targeting regimes can be characterized along a few key dimensions. This section examines, in detail, how four of the banks using explicit inflation targets have changed monetary policy procedures to accommodate inflation targets. The banks—the Reserve Bank of New Zealand, the Bank of Canada, the Bank of England, and the Swedish Riksbank—were among the first to explicitly target inflation and represent a range of procedures and experiences. For each bank, this section examines the price index used as a target, some of the caveats or exemptions employed, the inflation forecasting procedures, and the different approaches used to ensure transparency and accountability.

New Zealand

Prior to 1984, the Reserve Bank of New Zealand acted on behalf of the government in implementing the government's daily monetary policy decisions (Brash 1996). To do so, the

Table 1 Summary of Inflation Targeting Frameworks

	New Zealand	Canada	United Kingdom	Sweden	Finland	Australia	Spain	Israel*	Chile*
Date first issued	March 1990	February 1991	October 1992	January 1993	February 1993	Approx. April 1993**	Summer 1994**	December 1991***	Approx. 1990****
Current target	0%–3%	1%–3% with "mid point" 2%	2.5%	2% ± 1%	2%	2%–3% ("thick point")	Less than 3%	7%–10%	4.5%
Time frame	5 years (to 2003)	through end-2001	1997 onward	1995 onward	1996 onward	On average over the cycle	By late 1997, less than 2% thereafter	1 year	1 year
Inflation measure	CPIX (CPI excluding credit services)	CPI (Underlying inflation used operationally)	RPIX (retail price index excl. mortgage interest payments)	CPI	Underlying CPI	Underlying CPI	CPI	CPI	CPI
Target announcement	Defined in Policy Target Agreement (PTA) between the Minister of Finance and the Governor of the central bank	Joint agreement between the Minister of Finance and the Governor of the central bank	Chancellor of the Exchequer	Governing Board of the Bank of Sweden (Sveriges Riksbank), which is an authority of the parliament	Bank of Finland	Reserve Bank of Australia	Bank of Spain	Minister of Finance in consultation with the Prime Minister and the Governor of the central bank	Central Bank of Chile
Inflation report	Since March 1990. Quarterly today, formerly semi-annually	Semi-annual, since May 1995	Quarterly, since February 1993	Since October 1993. Quarterly today, formerly three times per year	No	Semi-annual, since May 1997	Semi-annual, since March 1995	Since March 1998	Annual, every September
Inflation forecasts published?	Yes	No	Yes	Yes	No	No	No	No	No

* Israel and Chile also target the exchange rate.
** The Reserve Bank of Australia dates the introduction of inflation targets to approximately April 1993 and the Bank of Spain to summer 1994. However, Bernanke, Laubach, Mishkin, and Posen argue that Australia did not introduce targets until September 1994 and Spain until November 1994.
*** *Financial Times*, December 18, 1990.
**** Since 1990, the Central Bank of Chile has been required by law to announce each September an inflation rate to be reached the following year. By the mid-1990s, these "targets" had gained credibility.
Source: Debelle; Bernanke, Laubach, Mishkin, and Posen; Reserve Bank of New Zealand's Policy Targets Agreement (December 1997); Bank of Canada (May 1998); Bank of England (August 1997); Sveriges Riksbank (June 1998); Bank of Israel; and Banco Central de Chile.

bank used "a web of regulations and direct controls" until a new government, which came into power in July 1984, introduced substantial economic reforms (Brash 1993). The reforms included assigning the Reserve Bank of New Zealand the goal of reducing inflation and granting the central bank more independence in its actions. The Reserve Bank of New Zealand Act 1989, which took effect in 1990, formalized the goal of lower inflation and required that the primary function of the central bank be to achieve and maintain price stability. The Act stipulated that the Minister of Finance and the Governor of the Reserve

Bank of New Zealand establish an economic target for monetary policy through a *Policy Targets Agreement* (*PTA*), a formal agreement between the Minister and the Governor. A new *PTA* is issued whenever economic circumstances demand renegotiation of the previous target, but no later than the year before a new five-year term of the Governor. All five *PTAs* issued so far (the first one in March 1990) have called for the economic target to be an inflation target (Fischer 1995).

Price Index and Caveats

The most recent *PTA* of December 1997 states that the inflation target should be measured in 12-month changes in the "All Groups Consumers Price Index excluding Credit Services" (CPIX), and that the CPIX should be between 0 and 3 percent.[8] The *PTA* mentions "unusual events" that can lead to CPIX inflation moving temporarily outside the targeted range. These events include exceptional movements in the prices of commodities traded in world markets, changes in indirect taxes, significant government policy changes that affect prices directly, or natural disasters (Reserve Bank of New Zealand, December 1997). Before December 1997, the *PTAs* stated an inflation target for the "All Groups Consumers Price Index" (CPI), but the bank focused on *underlying* inflation in explaining monetary policy and measuring performance (Bernanke, Laubach, Mishkin, and Posen; Reserve Bank of New Zealand). The bank calculated underlying inflation by excluding from the CPI large movements in components that reflected interest rate changes, government policy initiatives, and oil price changes (Reserve Bank of New Zealand, June 1997 and September 1997).

Inflation Forecast

The Reserve Bank of New Zealand uses a system of models complemented by judgment to prepare its economic projections and inflation forecast. The system consists of a comprehensive "core" macroeconomic model, partial "satellite" models that analyze specific components of the economomy in more detail, and "indicator" models which use statistical techniques to make short-term projections based on a range of current economic indicators. The forecast is conditioned on inflation staying at the midpoint of the inflation target range, and its goal is to pinpoint the path for monetary conditions (exchange rate and interest rates) needed to maintain the target (Reserve Bank of New Zealand 1997).

Transparency

The Reserve Bank of New Zealand Act 1989 requires the bank to produce a statement at least every six months. The statement must explain the bank's plan to implement monetary policy in a way consistent with the *PTA's* objective and comment on the bank's inflation forecast and on various leading indicators (Fischer 1995, Reserve Bank of New Zealand 1996). To fulfill these requirements, the bank publishes a report called the *Monetary Policy Statement* every June and December. In addition, the bank issues a brochure called *Economic Projections* every March and September, the Governor gives numerous speeches, and the bank maintains a comprehensive, easy-to-understand Internet site. The *Monetary Policy Statements* and *Economic Projections* review recent monetary policy, give forecasts of many economic variables, including inflation, and assess monetary conditions and risks to the forecasts (Reserve Bank of New Zealand, June 1997; Reserve Bank of New Zealand 1998). Because of a highly transparent policy, the bank has often been able to achieve adjustments to monetary conditions by communicating its desired path of policy, without actually intervening directly in the market (McCallum).

Accountability

According to the Reserve Bank Act, the Governor is held accountable for the outcome of monetary policy. The Governor can be dismissed if either the Minister of Finance or the bank's Board of Directors believes performance has been inadequate. Inadequate performance in monetary policy can be measured by comparing the *PTA,* which forms a contract between the government and the Governor, with the inflation outcome. To explain monetary policy actions and their compliance with the *PTA* goals, the Governor is required to produce a policy statement at least once every six months (Fischer 1995, Reserve Bank of New Zealand 1996).

Canada

Since the mid-1980s, the Bank of Canada has focused on attaining price stability as the underlying objective of monetary policy (Crow). Prior to introducing explicit inflation targets, the bank used various monetary and credit aggregates as guides for monetary policy, but it had neither an explicit target nor a time path for achieving a long-term goal (Freedman 1995a). In February 1991, the Governor of the Bank of Canada and the Minister of Finance jointly announced a series of formal targets for reducing inflation. The goal was to lower inflation to the midpoint of an assigned range—between 1 and 3 percent—by the end of 1995 (Bank of Canada, November 1997).

Price Index and Caveats

The Bank of Canada and the Minister of Finance identify the inflation targets in terms of the 12-month rate of change in the overall consumer price index. The target is announced as a range whose midpoint is the ultimate target level for the specified time horizon. The original goal to keep inflation close to the midpoint of a range between 1 and 3 percent has been extended twice: In December 1993 to last until the end of 1998, and in February 1998 to last until the end of 2001 (Bank of Canada, May 1998). For the short run, the bank focuses on "underlying" inflation, which excludes volatile food and energy prices as well as large first-round changes in indirect taxes from the overall consumer price index (Freedman 1995a; Bank of Canada, November 1997). In addition, large increases in oil prices and "unexpected developments," such as natural disasters, are treated as caveats or exemptions. After such shocks, inflation needs to return to the target level with no compensation for the resulting change in the price level (Ammer and Freeman, Almeida and Goodhart).

Inflation Forecast

While the Bank of Canada does not publish a quantitative inflation forecast (it only publishes a nontechnical, qualitative "outlook" for inflation), it still uses forecasts in conducting monetary policy. According to Freedman and Longworth, the Bank of Canada uses a structural forecasting model of the economy as the key initial input into the policy process. The bank supplements the model by using its judgment, especially for the first few quarters. The first step in this process is for the staff to make a projection based on an econometric model. This projection is conditional on inflation being on target and states what path monetary conditions (short-term interest rates and the exchange rate) must take for inflation to be within the target range in six to eight quarters. The bank's staff prepares a base scenario and some alternative scenarios, which the management uses as a starting-point for policy decisions. Since the staff projections are based on models—which naturally simplify reality and therefore cannot account for all potentially important market events—the management adjusts the projections at the end of the forecasting process to make them more realistic (Freedman 1995b).[9]

Transparency

Inflation targets were introduced in Canada to make the bank's actions "more readily understandable to financial market participants and to the general public" (Freedman 1995a). To communicate the targets and the policy actions, the bank publishes an easy-to-read *Monetary Policy Report* twice a year. The report explains the goals of the policy, discusses recent developments in inflation, reviews the policy actions taken to keep inflation on target, and describes the inflation outlook. The outlook covers aggregate demand and supply conditions, temporary influences on inflation, inflation expectations, and monetary indicators. While the report describes the bank's inflation projections qualitatively, it does not give a quantitative inflation forecast, nor does it give a specific time frame for the projections (Bank of Canada, November 1997). In addition to the *Monetary Policy Report,* the bank periodically issues press releases about changes in policy and has an outreach program informing the public about monetary policy. Furthermore, the Governor meets regularly with the Minister of Finance and appears before the committees of the House and the Senate (Mishkin and Posen, Almeida and Goodhart).

Accountability

Even though inflation targets are specified in a contract between the Bank of Canada and the Canadian government, no formal arrangements are in place for holding the Bank of Canada accountable if targets are missed. Indeed, the bank is required neither to explain its policies nor to publish an outlook for monetary policy, and the agreement between the bank and the government is subject to cancellation in the event of a political crisis (Fischer 1993). However, by observing the bank's actions, the public can monitor the bank's performance. Thus, the Bank of Canada is informally accountable to the public (Mishkin and Posen).

United Kingdom

The primary goals of monetary policy in the United Kingdom since the early 1980s have been to achieve price stability and to improve the credibility of monetary authorities. In the early 1980s, the authorities used monetary aggregates as targets. In 1986, they switched to informal exchange rate targets. From 1990 to 1992, the UK participated in the Exchange Rate Mechanism (ERM) of the European Monetary System (EMS), pegging the exchange rate of the pound to other European exchange rates. In this way, the UK "imported" price stability by linking its monetary policy to the policies of other European countries (mainly Germany). After suspending sterling from the ERM, the United Kingdom needed a new anchor for its monetary policy. In October 1992, the Chancellor of the Exchequer announced the initial inflation target (Bowen, McCallum).

At the time, the Bank of England had little independence from the Treasury. In May 1997, however, the bank gained more autonomy when the Chancellor announced that the government was giving the bank "operational responsibility for setting interest rates to meet the Government's inflation target." In addition, the Chancellor announced that the bank's Monetary Policy Committee (MPC) was to make operational decisions (Bank of England, August 1997).

Price Index and Caveats

The Chancellor expresses the inflation target in terms of 12-month changes in the retail price index excluding mortgage interest payments. The current target rate of inflation is set at 2.5 percent. Recognizing that shocks and disturbances can lead to a deviation from the target, the rate of inflation is allowed to stray from the target by up to 1 percent before the bank must justify the deviation (Bank of England, August 1997).[10] In addition, effects of indirect taxes, subsidies, and interest costs are treated as caveats or exemptions, allowing inflation to move temporarily away from the target (Almeida and Goodhart).

Inflation Forecast

The Bank of England bases its forecasts on several models and key information variables. The bank uses atheoretical single-equation techniques for its short-term projections (three months) and a structural model for its medium-term projections (approximately two years). To complement the structural model, the bank's staff has turned to using information from a wider set of quantitative models (including detailed sectoral models) and qualitative information (such as surveys), encompassing a variety of real and monetary information (Haldane).

Britton, Fisher, and Whitley describe the forecasting process since the MPC took charge of monetary policy as a "series of meetings between the MPC and the Bank staff." The focus of the first meeting is to identify the key assumptions of the forecasts. The staff then uses the assumptions to prepare a central projection and a risk distribution. In following meetings, the MPC reviews the forecasts, and the staff and MPC make adjustments to the assumptions based on perceived risks and new data. This process continues until the MPC agrees on a final forecast and risk distribution, which it then publishes in its *Inflation Report*. The *Inflation Report* presents the forecast and the risk scenarios in what the bank calls a "fan chart." The fan chart shows the projected path for inflation as a narrow confidence band, with the MPC's subjective assessment of risks as shaded distributions around the main band. The distributions "fan out" as the time horizon moves further into the future and are often asymmetric, suggesting that forecast risk can lean more in one direction than another (Britton, Fisher, and Whitley, 1998).

Transparency

The Bank of England strives to achieve a transparent monetary policy by communicating policy actions in a variety of publications and by publishing its inflation forecasts.[11] Accordingly, the bank has introduced a variety of new publications since adopting inflation targets. The most important new publication is the fairly technical *Inflation Report*. It states the goals of monetary policy, examines an array of recent monetary and real variables critical to inflationary developments, explains recent monetary policy actions, and publishes the quantitative inflation forecast of the MPC. In addition, the MPC meets on a monthly basis and publishes minutes from its meetings two weeks after the following meeting. The minutes also appear as an annex to the *Inflation Report*.[12] When the bank changes interest rates, it also issues a press notice informing the public. In addition, the Governor and other members of the MPC appear before the Treasury on a regular basis and give more public speeches than before (Almeida and Goodhart).

Accountability

The Bank of England is expected to explain deviations of inflation from the inflation target. If inflation strays more than 1 percent from target, the Governor of the bank must send an open letter on behalf of the MPC to the Chancellor of the Exchequer. In this letter, the

Governor must explain the divergence, describe the policy actions being taken to deal with it, and state how long before inflation is expected to return to target (Bank of England 1997). In this way, the Bank of England is accountable directly to the Treasury.

Sweden

The main goal of the Swedish central bank (the Sveriges Riksbank) since it gave up the gold standard in 1931 has been to "stabilize domestic purchasing power" (Bäckström 1998a). Through the postwar period until 1992, the Riksbank operated under a fixed exchange rate regime. From 1991 to 1992, the bank pegged the Swedish krona unilaterally to the ECU—the European currency unit—based on a basket of currencies from members of the European Union. When Sweden was forced to abandon the fixed exchange rate regime in November 1992, the Riksbank needed a new anchor for monetary policy. In January 1993, the Governing Board of the Riksbank introduced explicit inflation targets as the new anchor. The targeted level of inflation was to be reached by 1995 and maintained thereafter (Andersson and Berg, 1995; Padoa-Schioppa, 1992).

Price Index and Caveats

The Riksbank expresses the inflation target in 12-month changes of the "headline" consumer price index (CPI).[13] The bank gives a central value with tolerance bands for the target rate of inflation. The current target rate is 2 percent, with a tolerance interval of plus or minus one percentage point—the same as the bank had established at the announcement of its inflation targeting regime in 1993 (Andersson and Berg, 1995; Sveriges Riksbank, March 1998). The bank does not have a list of automatic caveats. When a shock occurs, the Riksbank analyzes whether the shock will have a temporary or permanent effect on inflation, and changes monetary policy actions only if it anticipates a permanent effect. As a result, the bank allows inflation to move outside the target temporarily in response to transitory shocks such as the first-round effects of changes in taxes and subsidies. Other potential caveats include changes in mortgage interest costs that are due to monetary policy adjustments and large external shocks, such as oil price shocks (Bäckström 1998b).

Inflation Forecast

The Riksbank uses a wide range of real and financial indicators to monitor inflationary pressures and to forecast inflation. To assess immediate inflationary pressures, the bank monitors indicators such as various price indexes, the output gap, capacity utilization, and the unemployment rate.[14] For a medium-term time horizon (up to one year), the bank analyzes variables such as wages, import prices, changes in the exchange rate, and raw material and intermediate goods prices. For one-year to two-year forecasts, the bank uses monetary aggregates, interest rates, and surveys of inflation expectations as indicators. The bank evaluates the influence of these indicators on future inflation and develops an inflation forecast, assuming the repo rate is unchanged.[15] The bank then publishes the quantitative inflation forecast for the following two years and discusses alternative inflation paths and potentially asymmetric risk scenarios (Andersson and Berg, 1995; Sveriges Riksbank, March 1998).

Transparency

The Riksbank publishes a quarterly *Inflation Report.*[16] The report explains monetary policy, analyzes recent economic developments, discusses determinants of inflation and their projections, derives a quantified inflation forecast from them, and draws conclusions for monetary policy from the forecast. The report also contains a discussion of alternative paths for inflation under various additional scenarios (Sveriges Riksbank, March 1998). Further means to increase transparency include public hearings before the Finance Committee of the parliament, reviews of central bank actions in speeches by the Governor and the staff (Andersson and Berg, 1995; Almeida and Goodhart), and an Internet site providing information about monetary policy actions.

Accountability

While there is no contract, agreement, or directive that binds the Riksbank to pursue the inflation target, the bank is accountable for achieving the target through the political process. The Sveriges Riksbank is an authority of the parliament, not a government agency like many other central banks, and seven of the eight members of its Governing Board are appointed by the parliament. Because the parliament monitors the bank's actions and the public elects the parliament, the bank is ultimately accountable to the public.

CHANGES IN THE SHORT-RUN CONDUCT OF POLICY: EMPIRICAL EVIDENCE

While inflation targets have clearly changed the institutional framework and many of the procedures of monetary policymaking, it is less clear that inflation targets have led to discernable changes in the way policy instruments are adjusted in response to economic information. On the surface, changes in the short-run conduct of policy seem plausible because of the magnitude of the procedural changes and the fact that inflation has tended to be lower and in some cases more stable after the introduction of inflation targets. However, over the last several years, inflation has also been lower and more stable in many countries without inflation targets.

This section analyzes fluctuations in short-term interest rates for evidence of significant and systematic changes in the way monetary policy is conducted after the introduction of inflation targets. As in the previous section, the analysis focuses on the four countries with the most experience using inflation targets. In addition, however, the experience of the United States—a country that has not introduced explicit numerical targets for inflation—is examined as a "control" case for weighing evidence on the operational significance of inflation targets.

The Behavior of Inflation

In the sample of four countries that adopted inflation targets in the early 1990s, inflation has declined and, in some cases, become more stable (Figure 1). Inflation targets were introduced in these countries when inflation was in the 4 to 8 percent range—substantially higher than the inflation rate most central banks considered consistent with price stability. Moreover, in New Zealand, Canada, and Sweden, inflation was rising at the time targets were introduced. In the 10 to 12 years prior to the introduction of targets, inflation had reached as high as 20 percent in New Zealand and the United Kingdom and 12–15 percent in Canada and Sweden.

All of the inflation targeting countries in the sample have successfully brought inflation down to a level that is near or within target. In New Zealand, except for a temporary increase in 1995–96, consumer price inflation has remained below 2 percent. In Canada, inflation fell rapidly after the introduction of targets and, in fact, frequently undershot the tar-

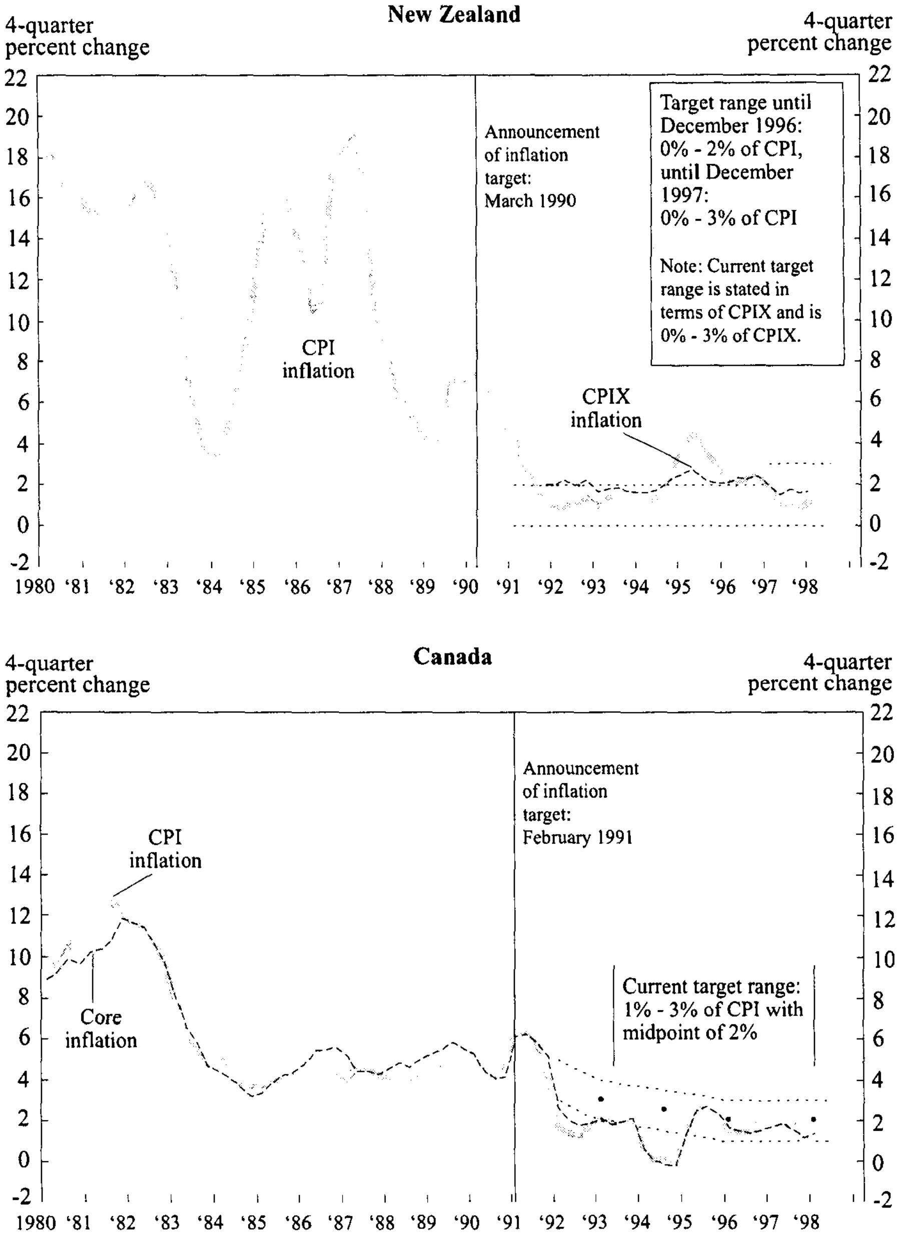
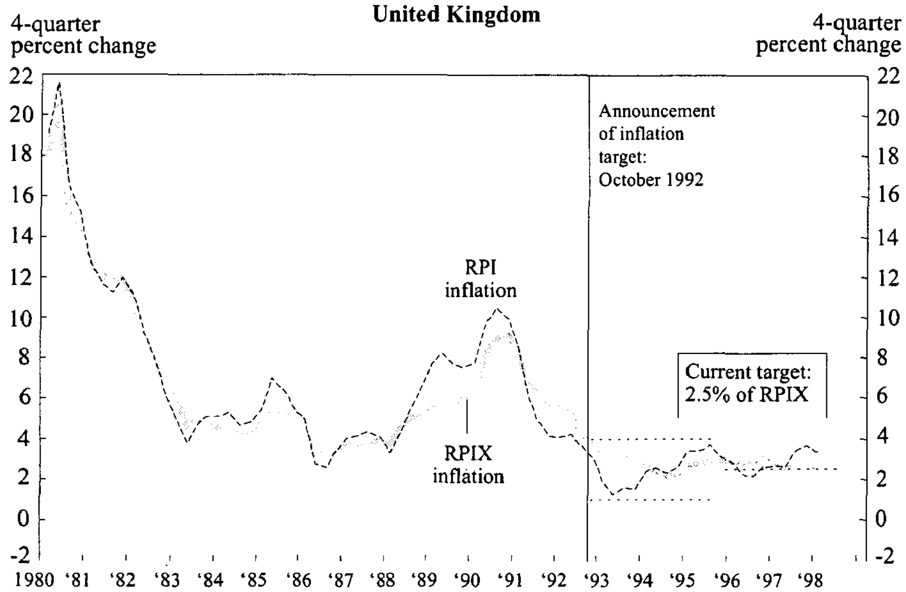


Figure 1 Inflation and inflation targets. Source: See Table 3 for inflation data sources. Target data are from Reserve Bank of New Zealand, Bank of Canada, Bank of England, and Riksbank.

get range in the first few years of inflation targeting. Since 1995, however, consumer price inflation has remained in a range of 1 to 3 percent. In the United Kingdom, retail price inflation fell relatively quickly after the introduction of targets, from around 4 percent to a level in 1997 that was very close to its current target of 2.5 percent. Finally, in Sweden consumer price inflation has fallen to a rate within or below the target of 2 percent plus or minus 1 percent.



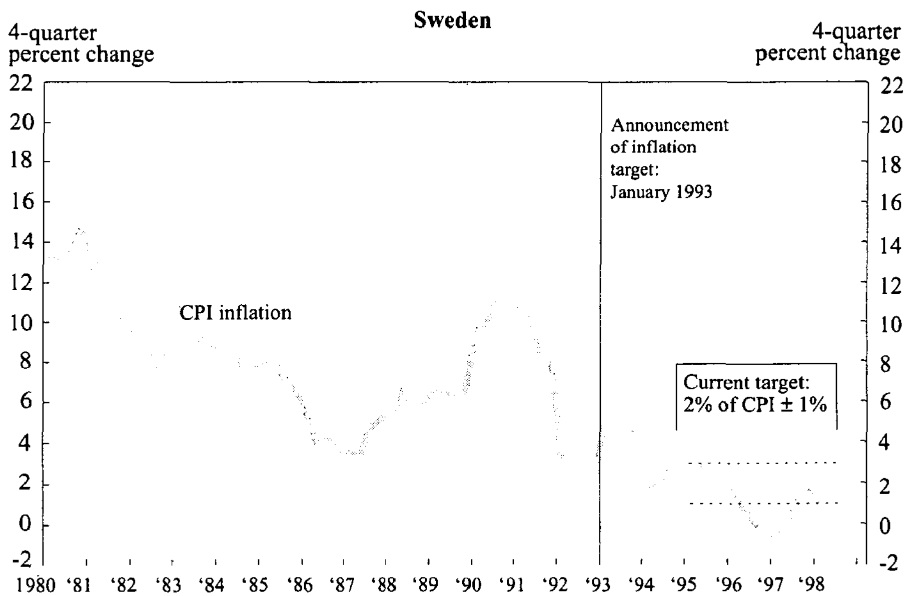


Figure 1 Continued.

The behavior of inflation has clearly improved after the introduction of inflation targets. But the behavior of inflation has also improved since the early 1990s in a number of countries that have not introduced explicit numerical inflation targets. In the United States, for example, consumer price inflation fell from over 5 percent in 1990 to around 2 percent in 1997 (Figure 2). Thus, while inflation targeting may help some central banks lower inflation, the adoption of inflation targets has not been necessary for other central banks to lower inflation.

Behavior of Short-Term Interest Rates

Can the improved behavior of inflation in countries that target inflation be attributable to changes in the way monetary policy responds to incoming information? One approach to answering this question is to examine the behavior of short-term interest rates that are used as policy levers by central banks. Most central banks in recent years—whether they explicitly target inflation or not—have come to use short-term interest rates as the primary instrument of monetary policy (Sellon and Weiner). Through their control over reserves in the commercial banking system, central banks seek to achieve levels of short-term interest rates that are consistent with the goals of monetary policy.

Because of the critical role of short-term interest rates in monetary policy, fundamental changes in the conduct of policy should affect how these interest rates move over time. If inflation targets imply a change in the central bank's inflation fighting commitment, they should cause systematic changes in the behavior of short-term interest rates. Alternatively, if the introduction of inflation targets is mainly a communication device designed to increase transparency and accountability—and not representative of fundamental changes in the central bank's inflation fighting resolve—the introduction of inflation targets may have little impact on the behavior of short-term interest rates.

Means and Standard Deviations

Nominal short-term interest rates have, on average, been lower and less volatile after the introduction of inflation targets than before their introduction (Table 2). Mean "official" rates—rates that central banks focus on in their day-to-day provision of reserves to the banking system—declined by anywhere from 622 basis points in New Zealand to 284 basis points in Sweden.[17] Because nominal interest rates incorporate an inflation premium, the decline in their mean primarily reflects lower inflation, not necessarily a shift in monetary policy. The standard deviation of official rates also declined in all countries, with the

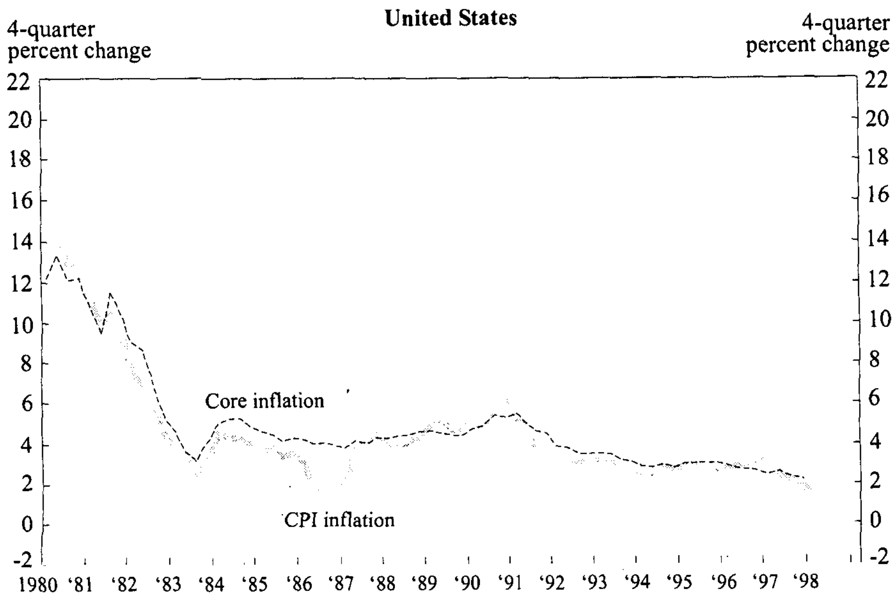


Figure 2 Inflation and inflation targets. Source: See Table 3 for inflation data sources.

Table 2 "Official" Rates Before and After Inflation Targets (Summary Statistics)

	New Zealand	Canada	United Kingdom	Sweden	United States*
Mean					
Nominal rate:					
Before	15.73	10.86	11.38	9.51	8.58
After	9.51	5.52	6.05	6.67	4.77
Real rate:					
Before	4.72	4.30	2.76	1.75	2.52
After	7.45	4.07	3.20	5.01	2.12
Standard deviation					
Nominal rate:					
Before	4.45	2.78	2.67	3.46	3.28
After	2.43	1.87	.85	1.96	1.11
Real rate:					
Before	19.40	4.76	8.69	8.45	4.22
After	4.43	3.05	4.92	5.37	2.15
Sample period					
Before	1979:2–1990:2	1977:10–1991:1	1975:2–1992:9	1971:1–1992:12	1971:1–1990:12
After	1990:3–1998:3	1991:2–1998:2	1992:10–1998:5	1993:1–1998:4	1991:1–1998:5

* In the absence of a regime shift to inflation targets, the sample is arbitrarily split at 1990:12 in the United States.
Note: The "official" rate is the discount rate for New Zealand, the day-to-day call money rate for Canada, the two-day London local authority call money rate for the united Kingdom, the 3-month Treasury bill rate for Sweden, and the federal funds rate for the United States. The real rate was calculated as the nominal rate minus the annualized monthly change in the consumer price index over the same period. Because monthly CPI data were not available for New Zealand, a monthly rate was calculated by subtracting a quarterly CPI inflation series from the monthly average nominal rate.
Sources: IMF (discount rate of New Zealand), Federal Reserve Board of Governors ("official" rates for Canada, the United Kingdom, Sweden, and the United States), New Zealand Department of Statistics (New Zealand's CPI), Statistics Canada (Canadian CPI), British Central Statistical Office (British Retail Price Index), Swedish Official Statistics Office (Swedish CPI), and U.S. Department of Labor (U.S. CPI).

United Kingdom showing the sharpest decline and Sweden the least. Again, the decline in standard deviation likely reflected more stable inflation and therefore, perhaps, a decline in the inflation risk premium. Thus, while nominal rates indicate that inflation expectations changed over the two periods, they are not necessarily informative about whether the short-run conduct of monetary policy changed. In addition, nominal rates in the United States followed the same trends with no shift to inflation targets.

Real short-term rates provide a better measure of the stance of monetary policy than nominal rates because they are adjusted for changes in inflation expectations. Over the long run, real rates tend toward an equilibrium level determined by nonmonetary factors. Because inflation expectations adjust slowly, monetary policy has some control over short-term real rates through its ability to affect the level of short-term nominal rates. When real rates are high relative to their long-run equilibrium, monetary policy is restrictive and, other things equal, inflation tends to fall. When real rates are low relative to their long-run equilibrium, monetary policy is accommodative and, other things equal, inflation tends to rise.

Real "official" rates have tended to be higher after the introduction of inflation targets than before their introduction (Table 2). Of the four countries in the inflation targeting sample, only Canada has seen a decline in real rates after introducing inflation targets, but the decline was from an already high level. In contrast, in the United States real short-term interest rates have fallen in the 1990s relative to their average in the previous 20-year period. Thus, the introduction of inflation targets seems to be associated with an increase in real rates—possibly suggesting tighter monetary policy has been associated with inflation targets. Alternatively, real rates may have been higher in the inflation targeting period because, for most countries, the period has been characterized by cyclical expansion.[18] At the same time, real rates have been less volatile in the 1990s in both the inflation targeting countries and in the United States. The decline in the standard deviation of the real official rate could reflect a more muted response of monetary policy to economic shocks or simply a more stable economic environment. In any event, the period of inflation targeting has been universally associated with relatively high and stable real rates.

Policy Reaction Functions

To the extent high real interest rates are associated with tight monetary policy, they indicate a desire by central banks to lower inflation. They do not necessarily indicate that inflation targets have led to a change in the response of policy to near-term economic conditions. One way to estimate how policymakers respond to incoming economic information is to estimate a "policy reaction function." A policy reaction function is a relationship between the level of the nominal official rate and key economic indicators that policymakers use as information variables in adjusting policy. Possible information variables include recent data on inflation, unemployment, and exchange rates. Policymakers may use variables such as these in an implicit reaction function because they care about them directly or because the variables are reliably related to other variables they care about—such as future inflation.

While no central bank uses an explicit policy reaction function in setting the short-run stance of policy, analysts have estimated such functions as a way to summarize the decision making process of monetary policymakers. Estimating a policy reaction function before and after a fundamental shift in the monetary policy regime can provide insight about how the regime shift affected the short-run conduct of policy. If, for example, the introduction of inflation targets affected short-run policy decisions, one would expect a change in the nature of an estimated reaction function.

Estimated reaction functions for the four inflation targeting countries show some indications of changes in the short-run conduct of monetary policy resulting from the introduction of inflation targets (Table 3). These reaction functions were estimated as a linear regression of the nominal "official" rate on a constant, seasonal dummies (not reported), and lagged values of the official rate, the inflation rate (as measured by the operationally relevant price index), the unemployment rate, and the trade-weighted exchange rate. The same number of lags was included for each explanatory variable, but varied from country to country based on goodness of fit criteria. The regressions were estimated using monthly data for all countries except New Zealand, where quarterly data were used because of the unavailability of monthly statistics on inflation.

For all of the inflation targeting countries except New Zealand, at least one of the variables that was statistically significant in explaining the official rate before the introduction of inflation targets lost its significance after the introduction of targets. In Canada, lagged unemployment and the lagged exchange rate lost explanatory power. In the UK, lagged inflation and the lagged exchange rate lost explanatory power.[19] And in Sweden, the lagged exchange rate lost explanatory power.[20] In contrast, in the United States, where the sample period was arbitrarily split in January 1991, the significance of coefficients in the policy reaction function remained fairly similar across the two periods.

Table 3 "Official" Rate Equations

		New Zealand		Canada		United Kingdom		Sweden		United States	
		Before	After	Before	After	Before	After	Before	After	Before	After
Explanatory variables											
Constant	Coefficient	5.075	8.170**	2.100***	.712	.922***	2.153***	.348	.279	.382	.627**
	Standard error	3.247	3.107	.605	.658	.335	.691	.294	.597	.329	.273
Lagged official rate	Sum of coefficients	.839***	.325	.912***	.955***	.983***	.782***	.970***	.974***	.973***	.940***
	Standard error	.104	.291	.031	.031	.019	.068	.016	.021	.020	.025
	F-statistic	17.113***	5.495***	464.777***	463.393***	1494.078***	65.797***	3513.525***	2056.632***	1253.486***	761.478***
Lagged inflation	Sum of coefficients	–.224	.786**	.000	–.030	–.018***	–.020	.003	.004	.005	–.026
	Standard error	.186	.327	.031	.024	.009	.033	.008	.010	.020	.026
	F-statistic	2.055	2.403	.000	.786	1.916	.515	.123	.161	.067	.521
Lagged unemployment	Sum of coefficients	–.035	–.453**	–:105**	–.027	–.053**	–.092***	–.031	–.029	–.014	–.062*
	Standard error	.562	.153	.042	.063	.022	.031	.071	.070	.036	.034
	F-statistic	5.949***	2.550*	3.296**	.106	5.893***	4.879**	.187	.171	6.003***	3.052*
Lagged exchanged rate	Sum of coefficients	–.158***	–.062*	.009**	–.004	–.012***	.002	.009**	.001	–.003	.000
	Standard error	.047	.034	.003	.003	.003	.003	.004	.002	.003	.001
	F-statistic	4.810***	1.657	3.294**	1.269	15.408**	.371	6.035**	.457	1.253	.345
Summary Statistics											
Adjusted R-squared		.905	.931	.917	.941	.943	.861	.932	.974	.948	.982
Standard error of estimate		1.401	.666	.793	.463	.644	.313	.905	.311	.746	.148
Number of lags		4		2		2		1		2	
Breakpoint test	F-statistic	2.207**		1.067		2.032***		.233		.401	
	Significance	(.022)		(.387)		(.007)		(.999)		(.991)	
Sample period		1979:2–1989:4	1990:1–1997:4	1977:10–1991:1	1991:2–1998:2	1975:2–1992:9	1992:10–1998:5	1971:1–1992:12	1993:1–1998:4	1971:1–1990:12†	1991:1–†1998:5

*Significant at 10 percent level.
** Significant at 5 percent level.
*** Significant at 1 percent level.
† Sample split arbitrarily at 1990:12.

Note: Observations are monthly for all countries except New Zealand. For New Zealand, the regressions were run with quarterly data. For each country, all explanatory variables were constrained to have the same number of lags. The optimal lag length was determined using the Akaike criterion and the Schwarz criterion on the sample before the introduction of inflation targets. All regressions included seasonal dummy variables.

Sources: Federal Reserve Board of Governors, IMF, Statistics Canada, UK Central Statistical Office, UK Department of Employment, Swedish Official Statistics Office, Swedish Central Bank, New Zealand Department of Statistics, Reserve Bank of New Zealand, Reuters, and the U.S. Department of Labor.

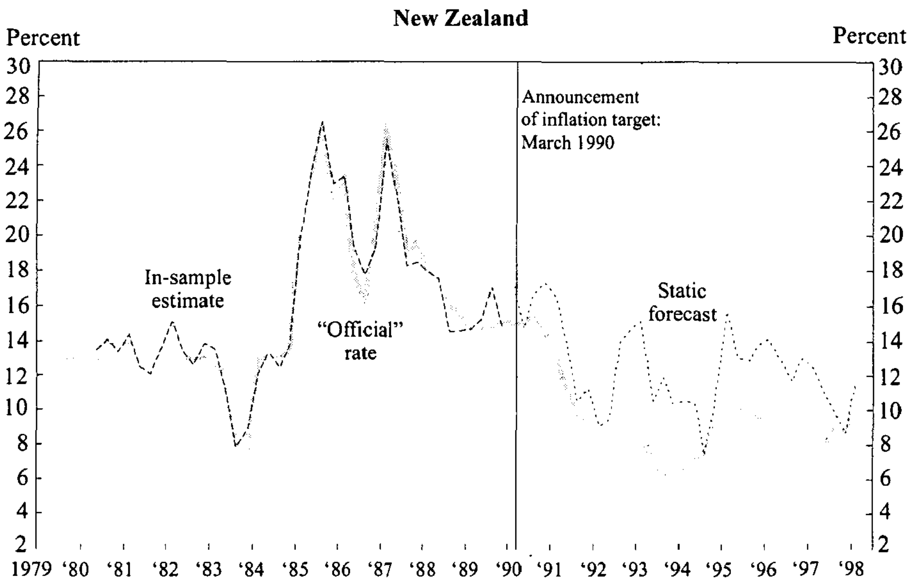
While the overall deterioration in the significance of explanatory variables in the policy reaction functions of the inflation targeting countries may indicate changes in the conduct of policy, the deterioration may also reflect the smaller sample size in the post-inflation targeting period. In addition, some coefficients in the reaction function were likely less precisely estimated because of less variation in the nominal official rate and the explanatory variables in the inflation targeting period than in the earlier period.

Another approach to evaluating whether the policy reaction function changed fundamentally in response to the introduction of inflation targets is to test the equation's overall statistical stability (Table 3, "breakpoint test"). The hypothesis that all of the coefficients jointly are the same in the two sample periods is rejected at the 5 percent significance level in New Zealand and at the 1 percent level in the United Kingdom—indicating the possibility of a structural shift in the policy reaction function in those two countries.[21] The strong rejection of stability in the UK likely reflects the changing role of the exchange rate in the policy reaction function after the UK broke away from the ERM. After leaving the ERM, monetary policy in the UK was freed up to pursue domestic goals such as employment growth and low inflation as opposed to maintaining the foreign exchange value of the pound against the mark.[22] It is by no means clear whether the instability in the UK policy reaction function is attributable to a shift in focus of monetary policy away from exchange rate stability or to the accompanying shift to inflation targets.

In contrast, the hypothesis of no structural shift could not be rejected for Canada and Sweden, indicating no change in the policy reaction function. Similarly, in the United States, where the sample was split arbitrarily and inflation targets were not introduced, there is no evidence of a structural shift in the policy reaction function.

To get a sense of the economic importance of possible structural changes—whether statistically significant or not—the policy reaction functions estimated for the pre-inflation targeting period were simulated over the post-inflation targeting period. A "static" forecast of the equation—in which actual values of the lagged official rate and other explanatory variables were fed into the policy reaction function estimated for the preinflation targeting regime—shows a relatively good performance (Figure 3). Although in New Zealand, Canada, and the United Kingdom predicted official rates were consistently above the actual rate for prolonged periods of time, the equations in all countries except New Zealand picked up most of the turning points in the actual official rates. In New Zealand, the predicted official rate is much more volatile than the actual rate—suggesting the possibility of a reduced sensitivity of official rates to economic information in the post-inflation targeting period.

A more challenging test of forecast performance comes from a "dynamic" simulation of the policy reaction function estimated from the pre-inflation targeting regime (Figure 4). In the dynamic simulation, forecasts of the official rate from the estimated reaction function are plugged into the right-hand-side instead of actual rates. The results clearly show a persistent overestimation of the level of official rates in New Zealand, Canada, and the United Kingdom based on the policy reaction function from the pre-inflation targeting regime.[23] In the UK, the overprediction is the result of the sharp decline in the value of the pound at the beginning of the post-inflation targeting regime. While the "before" policy reaction function predicts official rates would rise in response to such an exchange rate depreciation (in accordance with a policy that targeted the exchange rate), this reaction did not occur in the period after the introduction of inflation targets. In Canada, the overprediction is also due to exchange rates, but for the opposite reason. While the Canadian dollar appreciated sharply after the introduction of inflation targets, the "before" reaction function predicted official rates would rise in response.[24]



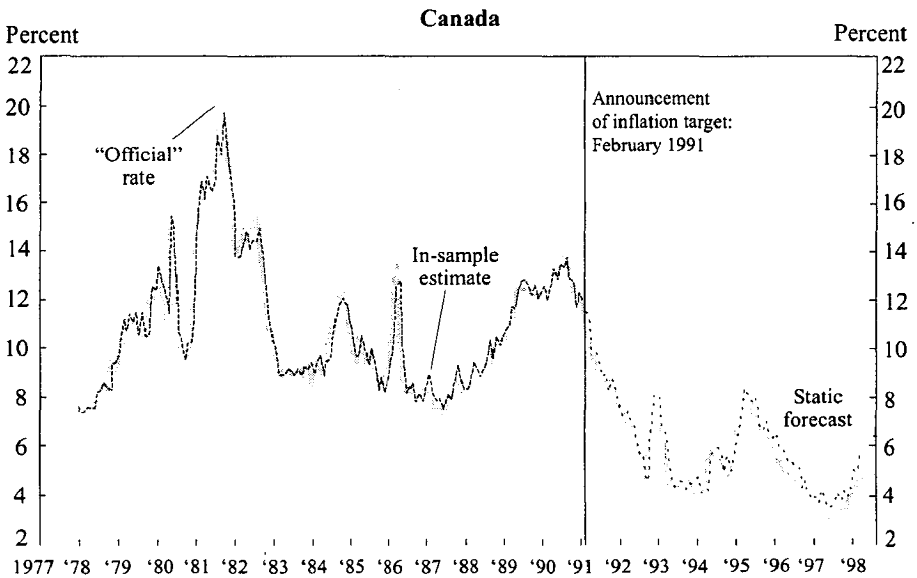
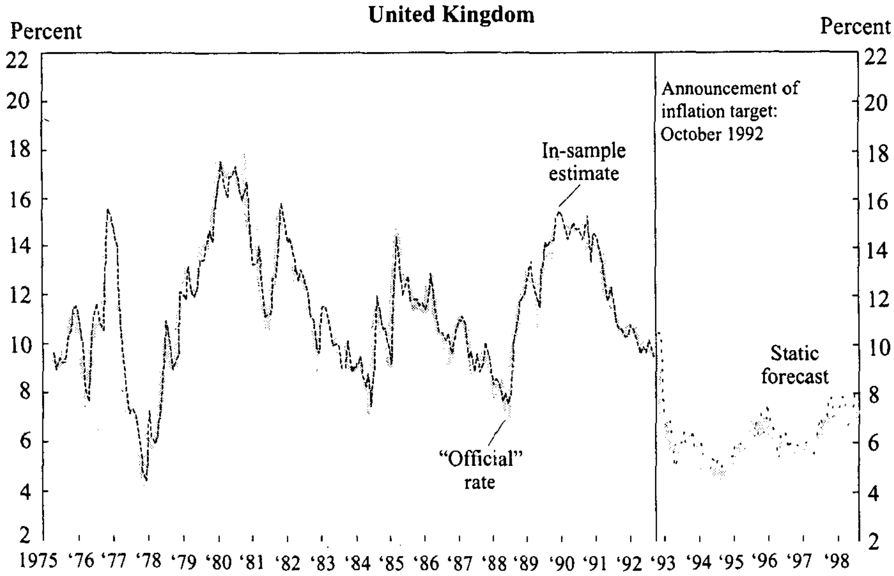


Figure 3 Projected and actual "official" rates (static forecasts). Note: Observations are monthly for Canada, the United Kingdom, and Sweden, and quarterly for New Zealand. Source: See Table 2 for actual official rate. Projections based on authors' calculations using "before" regression from Table 3.



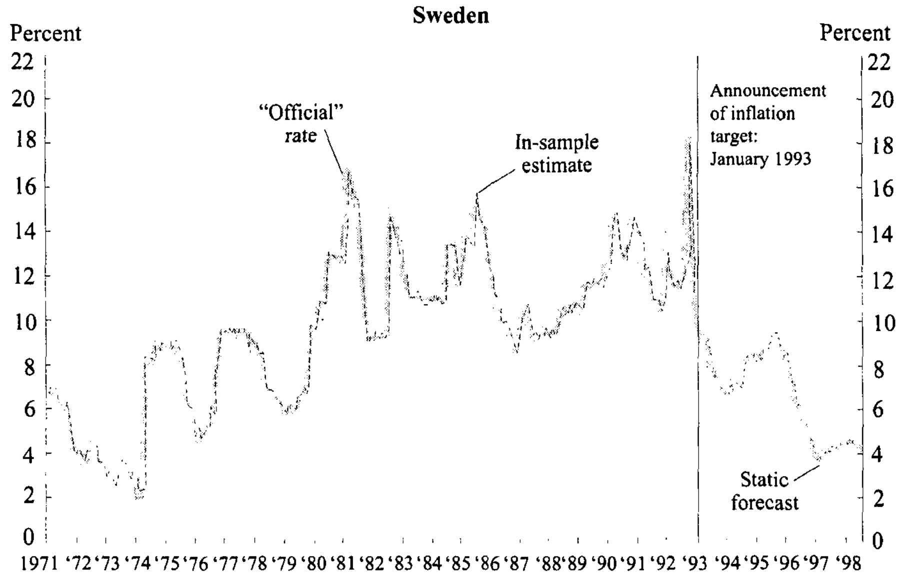
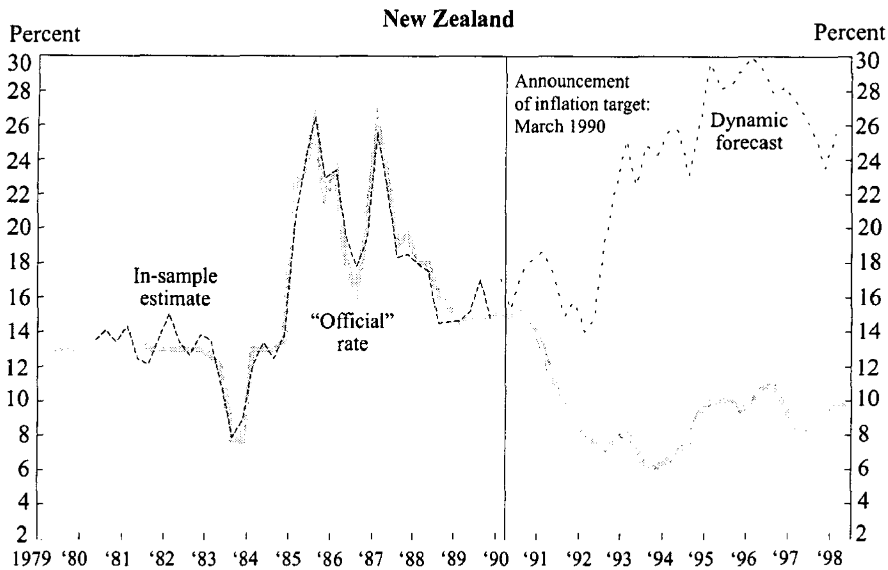


Figure 3 Continued.



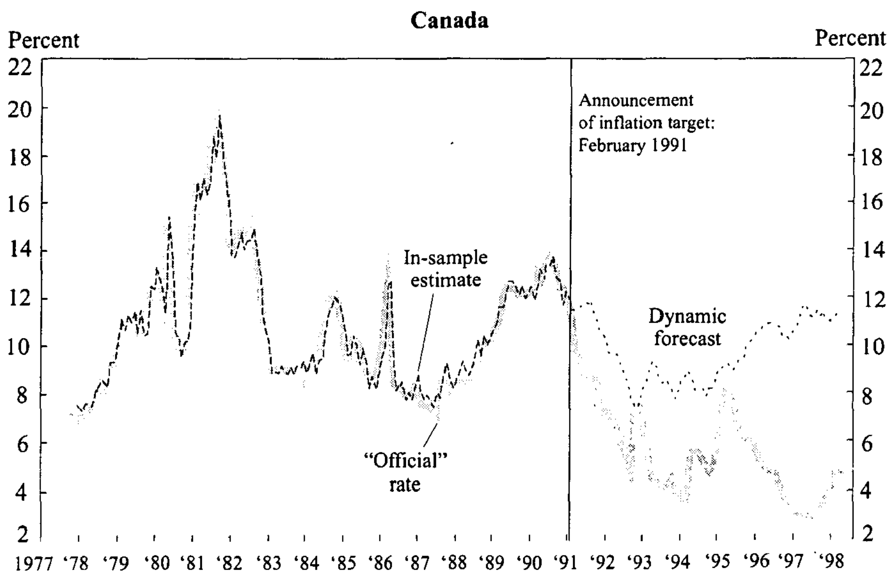
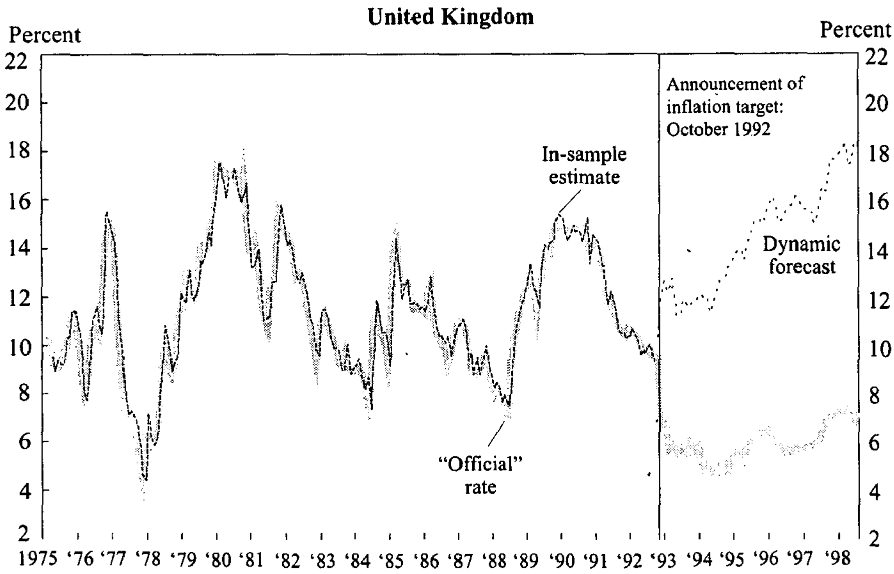


Figure 4 Projected and actual "official" rates (dynamic forecasts). Note: Observations are monthly for Canada, the United Kingdom, and Sweden, and quarterly for New Zealand. Source: See Table 2 for actual official rate. Projections based on authors' calculations using "before" regression from Table 3.



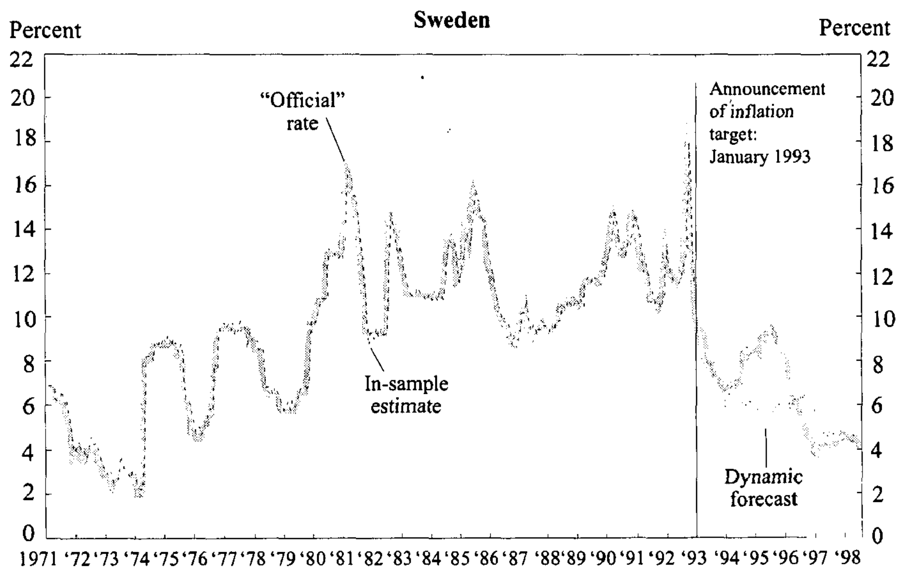


Figure 4 Continued.

To put these results in perspective, the same experiment was carried out for the United States (Figure 5). Although the United States did not adopt explicit inflation targets, a dynamic simulation of the policy reaction function estimated over the period from 1971 to 1990 shows the same characteristic overprediction of official rates after 1990 as did the simulations for New Zealand, Canada, and the United Kingdom. The implication is that

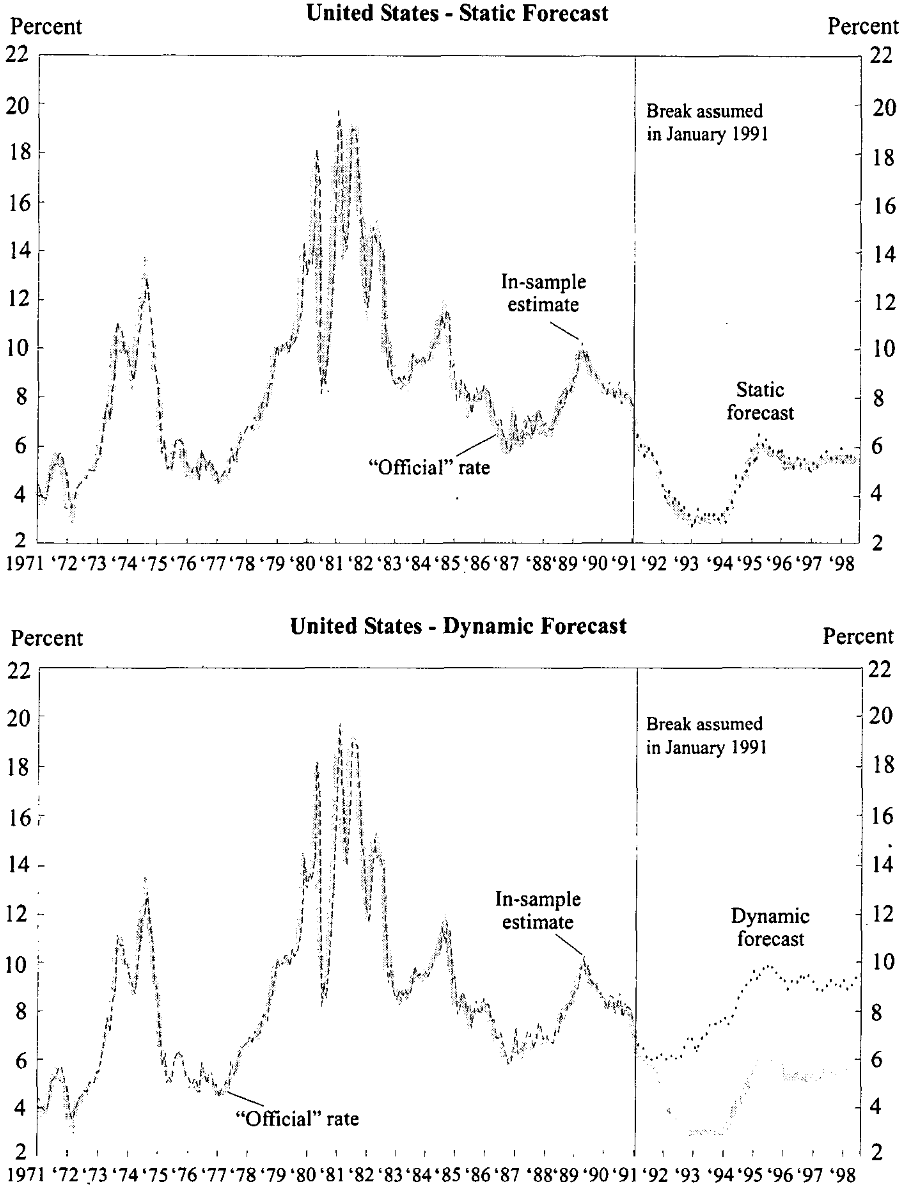


Figure 5 Projected and actual "official" rates. Note: Observations are monthly. Source: See Table 2 for actual official rate. Projections based on authors' calculations using "before" regression from Table 3.

policy reaction functions may be a poor way to characterize policy or that they shift in response to a variety of circumstances instead of or in addition to the establishment of inflation targets.

CONCLUSION

The introduction of explicit numerical targets for inflation has led to a number of highly visible changes in the procedures central banks follow in conducting monetary policy. These procedural changes have arguably increased the transparency and accountability of monetary policy. Through this channel, policymakers have possibly improved their inflation fighting credibility.

Evidence is mixed, however, as to whether inflation targets have led to changes in the way central banks conduct monetary policy in the short run. The lack of firm evidence may be due to the relative scarcity of data from inflation targeting regimes, which have been in place for less than a decade. Alternatively, inflation targets may simply formalize a monetary policy strategy that, in many cases, was already implicitly in place.

NOTES

1. For a comprehensive discussion of the rationale for and features of inflation targets, see Bernanke, Laubach, Mishkin, and Posen.
2. See Kahn for a summary of views about how central bankers should define price stability for monetary policy purposes.
3. Proponents counter that inflation targets merely keep discretionary policy actions consistent with long-run goals and therefore do not prevent policymakers from countering short-run disturbances (Bernanke and Mishkin, 1997).
4. This does not imply that central banks that do not have inflation targets are not accountable. For example, in the United States, the Federal Reserve Board of Governors is required to submit a report on the economy and the conduct of monetary policy twice a year to Congress. In addition, the Chairman of the Federal Reserve Board of Governors is called to testify on the report before the Senate Committee on Banking, Housing, and Urban Affairs, and the House Committee on Banking and Financial Services.
5. For a discussion of the effects of oil prices on inflation see Kahn and Hampton.
6. Evidence on whether inflation targets have increased credibility and therefore reduced the cost of disinflation is inconclusive (Johnson).
7. Some countries, such as Sweden, set an initial target that remains in place today.
8. Credit services represent the consumer cost of repaying debt and therefore fluctuate with interest rates.
9. The bank monitors a variety of indicators to help its policy decisions. The main focus is on estimates of excess demand or supply in goods and labor markets. Other variables, such as the growth rates of monetary aggregates, credit, total spending, and wage settlements, are used as additional guides for policy decisions (Freedman 1995a).
10. The Chancellor has changed the interpretation of targets three times: In October 1992 (range between 1 percent and 4 percent), June 1995 (at or below 2.5 percent, acknowledging that shocks can make inflation move between 1 percent and 4 percent), and June 1997 (2.5 percent, acknowledging that shocks can make inflation move plus or minus 1 percent). The targets are valid from the announcement date onward.

11. The bank typically publishes a short-term forecast and an approximately two-year-ahead forecast. The forecasting horizon has been extended some recently.
12. Prior to the bank's operational independence in May 1997, the Governor met with the Chancellor on a monthly basis.
13. The bank uses several measures of underlying inflation in its analysis. They are used as indicators of inflationary pressures, not as official targets. The "headline" CPI is regarded as the most transparent and unambiguous measure of inflation, which gives it a more objective appeal than other measures, and makes it more suitable for a target (Andersson and Berg; Sveriges Riksbank, March 1998).
14. The output gap measures the divergence of actual production from what is considered its potential level.
15. The Riksbank uses the repo rate (rate at which it agrees to repurchase securities), the lending rate, and the deposit rate for monetary policy purposes (Sveriges Riksbank 1997).
16. The Bank introduced the inflation report initially in October 1993 under the title "Inflation and Inflation Expectations in Sweden," which was published three rather than four times a year (Almeida and Goodhart, Andersson and Berg).
17. For Canada and the UK, the interest rate used as the "official" rate is the rate monitored or controlled by the central bank. For Sweden, the 3-month interest rate was used as a proxy for an official rate because data on official rates were not available. In New Zealand, the discount rate—which moved closely with the 3-month rate over the part of the sample that the 3-month rate was available—was used as a proxy for the official rate because it was the only short-term rate available for the entire sample.
18. In addition, to the extent inflation rose in the period before the introduction of inflation targets, real rates may have been kept too low.
19. This result might be explained by a forward looking monetary policy focused on inflation targets if the unemployment rate helped predict future inflation but other variables did not. Surprisingly, in all countries except New Zealand, the coefficients on lagged inflation were insignificant (and sometimes negative) in the inflation targeting period.
20. Because in Sweden there appears to be a break in the unemployment series in the early 1990s, the regressions were also run using industrial production in place of unemployment. However, in these alternative regressions, industrial production also came in insignificant in both the before and after periods.
21. In New Zealand, the breakpoint also corresponds to the granting of operational independence to the Bank of New Zealand.
22. The UK targeted exchange rates during only part of the sample period before the introduction of inflation targets. Prior to the use of exchange rate targets, the Bank of England targeted various monetary aggregates.
23. When industrial production is substituted for unemployment in the Swedish "before" regression, the Swedish official rate is also persistently over-predicted.
24. It is not clear why the policy reaction function indicates that policy was tightened in Canada in the before period in response to an exchange rate appreciation.

REFERENCES

Almeida, A., and Goodhart, C. A. E. 1998. "Does the Adoption of Inflation Targets Affect Central Bank Behavior?" *Banco Nazionale del Lavoro Quarterly Review* vol. 51, no. 204 (supplement March 1998) 19–107.

Ammer, J., and Freeman, R. T. 1994. Inflation Targeting in the 1990s: The Experiences of New Zealand, Canada, and the United Kingdom. Board of Governors of the Federal Reserve System. *International Finance Discussion Papers,* no. 473, June.

Andersson, K., and Berg, C. 1995. The Inflation Target in Sweden. In A. G. Haldane, ed., *Targeting Inflation—A Conference of Central Banks on the Use of Inflation Targets Organised by the Bank of England.* Bank of England, pp. 207–25.

Bäckström, U. 1998a. Five Years with the Price Stability Target. Speech at the Swedish Economics Association, available on the Internet at *http://www.riksbank.se/eng*. February.

Bäckström, U. 1998b. Inflation and the Interest Rate. Speech at the Stockholm Chamber of Commerce and Veckans Affärer, available on the Internet at *http://www.riksbank.se/eng*. January.

Bank for International Settlements. 1996. 66th Annual Report. Basle: Bank for International Settlements, June.

Bank of Canada. Various issues. *Monetary Policy Report.*

Bank of England. Various issues. *Inflation Report.*

Bank of Israel. 1998. *Inflation Report.* Available on the Internet at *http://www.bankisrael.gov.il.* March.

Bernanke, B. S., Laubach, T. Mishkin, F. S. and Posen, A. S. Forthcoming. *Inflation Targeting—Lessons from the International Experience.* Princeton University Press: Princeton.

Bernanke, B. S., and Frederic S. Mishkin. 1997. Inflation Targeting: A New Framework for Monetary Policy? *Journal of Economic Perspectives,* vol. 11, no. 2, Spring, pp. 97–116.

Bowen, A. 1995. British Experience with Inflation Targetry. In Leiderman and L. E. O. Svensson, eds., *Inflation Targets.* London: Centre for Economic Policy Research, pp. 53–68.

Brash, D. T. 1996. New Zealand's Remarkable Reforms. Speech at the Institute of Economic Affairs, London, available on the Internet at *http://www.rbnz.govt.nz.* June.

Brash, D. T. 1993. Reconciling Central Bank Independence with Political Accountability—The New Zealand Experience. Speech at the European Policy Forum, London, available on the Internet at *http://www.rbnz.govt.nz.* June.

Britton, E., Fisher, P. and Whitley, J. 1998. The Inflation Report Projections: Understanding the Fan Chart. *Bank of England Quarterly Bulletin,* February, pp. 30–37.

Cecchetti, S. G. 1998. Policy Rules and Targets: Framing the Central Banker's Problem. Federal Reserve Bank of New York, *Economic Policy Review,* vol. 4, no. 2, June, pp. 1–14.

Crow, J. W. 1993. Monetary Policy Under a Floating Rate Regime: The Canadian Experience. *Bank of Canada Review,* Summer, pp. 37–47.

Debelle, G. 1997. Inflation Targeting in Practice. International Monetary Fund, *IMF Working Paper WP/97/35,* March.

Duguay, P., and Poloz, S. 1994. The Role of Economic Projections in Canadian Monetary Policy Formulation. *Canadian Public Policy,* vol. 20, no. 2, pp. 189–99.

Fischer, A. M. 1995. New Zealand's Experience with Inflation Targets. In L. Leiderman and L. E. O. Svensson, eds., *Inflation Targets.* London: Centre for Economic Policy Research, pp. 32–52.

Fischer, A. M. 1993. Inflation Targeting: The New Zealand and Canadian Cases. *CATO Journal,* vol. 13, no. 1, Spring/Summer, pp. 1–27.

Freedman, C. 1995a. The Canadian Experience with Targets for Reducing and Controlling Inflation. In L. Leiderman and L. E. O. Svensson, eds., *Inflation Targets.* London: Centre for Economic Policy Research, pp. 19–31.

Freedman, C. 1995b. The Role of Monetary Conditions and the Monetary Conditions Index in the Conduct of Policy. *Bank of Canada Review,* August, pp. 53–59.

Freedman, C., and Longworth, D. 1995. The Role of Staff Economic Projections in Conducting Canadian Monetary Policy. In A. G. Haldane, ed., *Targeting Inflation—A Conference of Central Banks on the Use of Inflation Targets Organised by the Bank of England.* Bank of England, pp. 101–12.

Freeman, R. T., and Willis, J. L. 1995. Targeting Inflation in the 1990s: Recent Challenges. Board of Governors of the Federal Reserve System, *International Finance Discussion Papers* no. 525, September.

Haldane, A. G. 1997. Some Issues in Inflation Targeting. *Bank of England Working Paper Series,* no. 74, December.

Johnson, D. 1998. The Credibility of Monetary Policy: International Evidence Based on Surveys of Expected Inflation. Bank of Canada, *Price Stability, Inflation Targets, and Monetary Policy—Proceedings of a Conference Held by the Bank of Canada,* Ottawa, Ontario, pp. 361–95.

Kahn, G. A. 1996. Symposium Summary. Federal Reserve Bank of Kansas City, *Achieving Price Stability—A Symposium Sponsored by the Federal Reserve Bank of Kansas City,* Kansas City, Mo., pp. ix–xix.

Kahn, G. A., and Hampton, R., Jr. 1990. Possible Monetary Policy Responses to the Iraqi Oil Shock. Federal Reserve Bank of Kansas City, *Economic Review,* November/December, pp. 19–32.

King, M. 1998. The Inflation Target Five Years On. London School of Economics, Financial Markets Group, *Special Paper,* no. 99, February.

King, M. 1994. Monetary Policy Instruments: The UK Experience. *Bank of England Quarterly Bulletin,* August, pp. 268–76.

McCallum, B. T. 1996. Inflation Targeting in Canada, New Zealand, Sweden, the United Kingdom, and in General. *National Bureau of Economic Research Working Paper 5579,* May.

Mishkin, F. S., and Posen, A. S. 1997. Inflation Targeting: Lessons from Four Countries. Federal Reserve Bank of New York, *Economic Policy Review,* August, pp. 9–110.

Padoa-Schioppa, T. 1992. Problems of the Exchange Rate Mechanism of the European Monetary System. In P. Newman, M. Milgate, and J. Eatwell, eds., *The New Palgrave Dictionary of Money and Finance.* New York: The Stockton Press, pp. 822–27.

Reserve Bank of New Zealand. 1998. Economic Projections. Available on the Internet at *http://www.rbnz.govt.nz.* March.

Reserve Bank of New Zealand. 1997. The Forecasting and Policy System: An Introduction. *Reserve Bank Bulletin,* vol. 60, no. 3, September, available on the Internet at *http://www.rbnz.govt.nz.*

Reserve Bank of New Zealand. 1996. Briefing on the Reserve Bank of New Zealand. Available on the Internet at *http://www.rbnz.govt.nz.*

Reserve Bank of New Zealand. Various issues. *Monetary Policy Statement.* Available on the Internet at *http://www.rbnz.govt.nz.*

Reserve Bank of New Zealand. Various issues. *Policy Targets Agreement.* Available on the Internet at *http://www.rbnz.govt.nz.*

Sellon, G. H., and Weiner, S. E. 1996. Monetary Policy Without Reserve Requirements: Analytical Issues. Federal Reserve Bank of Kansas City, *Economic Review,* Fourth Quarter, pp. 5–24.

Siklos, P. L. 1996. Charting a Future for the Bank of Canada: Inflation Targets and the Balance Between Autonomy and Accountability. In D. Laidler, ed., *Where We Go from Here—Inflation Targets in Canada's Monetary Policy Regime,* Policy Study 29. Toronto: C. D. Howe Institute, pp. 101–84.

Svensson, L. E. O. 1996. Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets. *National Bureau of Economic Research Working Paper 5797,* October.

Sveriges Riksbank. 1997. The Discount Rate. Press release, available on the Internet at *http://www.riks.bank.se/eng.* October.

Sveriges Riksbank. Various issues. *Inflation Report,* available on the Internet at *http://www.riksbank.se/eng.*

Yates, A. 1995. On the Design of Inflation Targets. In A. G. Haldane, ed., *Targeting Inflation—A Conference of Central Banks on the Use of Inflation Targets Organised by the Bank of England,* Bank of England, pp. 135–69.

40

The Shadow of the Great Depression and the Inflation of the 1970s

J. Bradford DeLong

Associate Professor, University of California, Berkeley, California

The inflation of the 1970s was a time when uncertainty about prices made every business decision a speculation on monetary policy. During that decade, the annual U.S. inflation rate rose in the 5–10% range, compared to a 0–3% range typical of peacetime America.

If you asked a representative sample of economists why there is inflation, many would refer you to Kydland and Prescott (1977). They argue that higher than desirable inflation springs from the inability of central bankers to commit themselves to low-inflation policies. Central banks benefit when workers, firms, and investors have confidence that inflation will be low; but they also benefit—production is higher and unemployment is lower—when inflation turns out to be a little bit higher than expected. Thus workers, firms, and investors lack complete confidence in declarations that policy is to keep inflation low.

But this explains why inflation might be higher than desirable *in general.* It does not explain why inflation would be higher than desirable for one particular decade—the 1970s—while remaining low in the surrounding decades of the 1950s, the 1960s, the 1980s, and the 1990s. Thus, it is necessary to look elsewhere, and looking elsewhere leads to the conclusion that the inflation of the 1970s had three causes, depending on how deeply one wishes to look at the underlying situation. (These issues and more detail on the history of the period are in DeLong 1997.)

At the first surface level, the United States had a burst of inflation in the 1970s because until the 1980s no influential policymakers—until Paul Volcker became Chairman of the Federal Reserve—placed a sufficiently high priority on stopping inflation; other goals took precedence. At the second deeper level, the United States had a burst of inflation in the 1970s because economic policymakers during the 1960s dealt their successors a bad hand. And bad cards coupled with bad luck made inflation in the 1970s worse than anyone expected it might be. At the third level, the most profound cause of the inflation of the 1970s was the shadow cast by the Great Depression—an event that made it impossible for a while to believe that the business cycle was a *fluctuation around,* rather than a *shortfall below,* potential output and employment. Only after the experiences of the 1970s were

Reprinted from: *FRBSF Economic Letter,* No. 98-14 (May 1, 1998).

policymakers persuaded that the minimum sustainable rate of unemployment attainable by macroeconomic policy was relatively high, and that the costs—at least the political costs—of even moderately high one-digit inflation were high as well.

INFLATION REDUCTION: A LOW PRIORITY

By the end of the 1960s, the United States had finished its experiment to see if it was possible to push unemployment below 4% without accelerating inflation. The answer was "no." Nonfarm nominal wage growth had fluctuated around or below 4% per year between the end of the Korean War and the mid-1960s, and by 1968 it was more than 6% and rising.

Could U.S. monetary policy have responded to rising wage inflation at the end of the 1960s and ended the inflation of the 1970s before it started? At a technical level, yes. For example, Germany managed it: Germany's inflation peaked at the very beginning of the 1970s, at which point the Bundesbank made inflation reduction its first priority; by the mid-1970s, its cyclical peak was lower than the 1971 peak, and the early 1980s peak is invisible.

But in the United States there were obstacles. One was Richard Nixon, who was wary of policies to reduce inflation at the potential cost of higher unemployment. Nixon attributed his defeat in 1960 to Eisenhower's unwillingness to stimulate employment at the risk of increasing inflation (Nixon 1962). Thus, during his presidency he authorized his labor secretary to declare that his administration would "control inflation without a rise of unemployment" (Stein 1984).

A second obstacle was that Federal Reserve Chair Arthur Burns was skeptical about the value of using monetary policy to control inflation in the postwar era. In Burns (1960), he noted that "During the postwar recessions the average level of prices in wholesale and consumer markets has declined little or not at all," and he cited several factors—including a sharp rise in long-term interest rates and an "abnormally low" yield on stocks relative to bonds—that appeared to be "symptoms of a continuation of inflationary expectations or pressures." Before World War II such inflationary expectations and pressures would have been erased by a severe recession and by the pressure put on workers' wages and manufacturers' prices by falling aggregate demand. But Burns could not see how such pressures to moderate wage increases could be generated in the postwar world in which workers and firms rationally expected recessions to be short and shallow.

Third, there was no political support anywhere else for a policy of reducing inflation first. Congressional Republicans did not call for more aggressive fights against inflation. And Democrats complained that the Nixon administration policy was too contractionary.

Indeed, only with the acceleration of inflation at the end of the 1970s did political sentiment begin to shift. By the time Paul Volcker became Chairman of the Federal Reserve, inflation reduction—rather than full employment—had become the highest priority. The Volcker-led Fed quickly signaled its intention to place first priority on controlling inflation by shifting its operating procedures to place a greater emphasis on money supply targets.

BAD LUCK DURING THE 1970S

On top of the unwillingness to give inflation reduction high priority came bad luck. Alan Blinder (1982) has argued that double-digit inflation in the 1970s had a single cause: supply shocks that sharply increased the nominal prices of a few categories of goods. But those in the monetarist tradition never found this explanation convincing. As Milton Friedman (1975) said: "The special conditions that drove up the price of oil and food required purchasers to spend more on them, leaving them less to spend on other items. Did that not force other prices to go down, or to rise less rapidly than otherwise? Why should the *average* level of prices be affected significantly by changes in the price of some things relative to others?"

Ball and Mankiw (1995) believe the missing link is to be found in menu-cost models of aggregate supply. Firms and workers adjust their prices and wages to large economic changes, but ignore small ones because it is not worthwhile to figure out how to respond to them. Thus, a concentrated sharp increase in the nominal prices of a few commodities produces a much larger effect on the average level of prices and inflation than a more diffused increase in the nominal prices of money commodities.

But U.S. inflation in the 1970s was not solely the result of supply shocks. Taking the rate of wage inflation at the start of the 1970s and adding the effects of the 1970s productivity slowdown on the difference between wage and price inflation gives a baseline price inflation rate of 7% per year at the end of the 1970s without including any persistent effect of supply shocks. The contribution of bad luck during the 1970s is less than the contribution of the situation at the start of the decade.

THE SHADOW OF THE GREAT DEPRESSION

It is not enough to explain the inflation of the 1970s to say that the U.S. economy had bad luck during the 1970s, that the political consensus to support a policy of inflation reduction did not exist until the very end of the 1970s, and that economic policymakers in the 1960s dealt their successors a bad hand. The questions remain: Why did the political consensus to reduce inflation not exist until the end of the 1970s? And why did makers of economic policy during the 1960s watch with little concern as inflation crept upward, and as expectations of rising rates of price inflation became embedded in labor contracts and firm operating procedures?

The source of these attitudes and frames of mind is, in a strong sense, the most profound cause of the inflation of the 1970s. And that source is the shadow cast by the Great Depression. The extraordinarily high unemployment of the Great Depression made it very difficult to believe that the business cycle was a fluctuation around a growing level of potential output and sustainable employment. Instead, the Great Depression created a strong presumption that business cycles were shortfalls below potential output and sustainable employment. In the context of the Great Depression, it seemed absurd to take the average level of capacity utilization or unemployment achieved as a measure of the economy's sustainable productive potential.

In the shadow of the Great Depression, economists thought of fiscal and monetary policy in terms of "closing the gap" between actual and potential output, not in terms of moderating fluctuations around some long-run trend. But how were makers of fiscal and monetary policy to know when they had "closed the gap" and attained the maximum sustainable level of employment and capacity utilization? Neither economic theory nor economic history gave guidance, so there was a strong tendency to rely on hope and optimism.

Thus, in the early 1960s, economists as prominent as Paul Samuelson and Robert Solow (1960) could write of a 4% "interim" target for the unemployment rate that should be achieved. Moreover, they saw no reason that the "final" target for unemployment might not be even lower. The hope that unemployment could be pushed lower than the U.S. economy had ever achieved before without producing unacceptable consequences for the rate of inflation was reinforced by A.W. Phillips's late 1950s study of inflation and unemployment in the United Kingdom. This "Phillips curve" appeared to show a stable, unchanging relationship between inflation and unemployment.

The temptation to believe that Phillips's results were robust and could be generalized was very strong. The average level of capacity utilization and the unemployment rate were not guides to the long-run sustainable level, so there was an overwhelming temptation to "close the gap"—to see if unemployment could not be pushed even lower and capacity utilization higher. Politicians and their advisors had very strong incentives to believe that the Phillips curve taught that unemployment below 4% could be sustained indefinitely. Only the experience of the end of the 1960s and 1970s—when attempts to sustain low unemployment generated strongly accelerating inflation, as Milton Friedman (1968) had predicted it would—could change people's minds.

Thus there is a strong sense that something like the inflation of the 1970s was nearly inevitable. Had macroeconomic policy been less stimulative in the 1960s, and had inflation been lower at the end of that decade, there *still* would have been calls for increasing efforts to reduce unemployment in the 1970s.

Only after the experiences of the 1970s were policymakers persuaded that the flaws and frictions in American labor markets made it unwise to try to use stimulative macroeconomic policies to push the unemployment rate down to a very low level and to hold it there.

REFERENCES

Ball, L., and Mankiw, N. G. 1995. Relative Price Changes as Aggregate Supply Shocks. *Quarterly Journal of Economics* 110, pp. 161–193.

Blinder, A. 1982. The Anatomy of Double-Digit Inflation in the 1970s. In *Inflation: Causes and Effects,* ed. R.E. Hall, pp. 261–282. Chicago: University of Chicago Press.

Burns, A. 1960. Progress towards Economic Stability. *American Economic Review* 50, pp. 1–19.

DeLong, J. B. 1997. America's Peacetime Inflation: The 1970s. In *Reducing Inflation: Motivation and Strategy,* eds. C. Romer and D. Romer. Chicago: University of Chicago Press.

Friedman, M. 1975. Perspectives on Inflation. *Newsweek* (June 24) p. 73.

Friedman, M. 1968. The Role of Monetary Policy. *American Economic Review* 58(1) pp. 1–17.

Kydland, F., and Prescott, E. 1977. Rules Rather Than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy* 87 (June) pp. 473–492.

Nixon, R. 1962. *Six Crises.* New York: Doubleday.

Samuelson, P., and Solow, R. 1960. Analytical Foundations of Anti-Inflation Policy. *American Economic Review* 50 (May) pp. 185–197.

Stein, H. 1984. *Presidential Economics.* New York: Simon and Schuster.

41

Federal Reserve Credibility and Inflation Scares

Chan G. Huh
Economist, Federal Reserve Bank of San Francisco, San Francisco, California

Kevin J. Lansing
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio, and Federal Reserve Bank of San Francisco, San Francisco, California

After two decades of rising inflation during the 1960s and 1970s, the Federal Reserve under Chairman Paul Volcker undertook a deliberate disinflationary policy that was successful in reducing the U.S. inflation rate from well over 10 percent in 1980 to around 3 percent by 1985. The cost of this victory, however, was an extremely severe recession: the civilian unemployment rate peaked at about 11 percent in 1982—the highest level observed in the U.S. economy since the Great Depression.

It is widely recognized that an important factor governing the cost of disinflationary policies is the degree of central bank credibility.[1] Credibility is important because it influences the public's expectations about future inflation.[2] These expectations, in turn, affect the *current* state of the economy because they are incorporated into wages via forward-looking labor contracts and into the level of long-term nominal interest rates, which govern borrowing behavior. When the central bank enjoys a high degree of credibility, rational agents will quickly lower their inflation expectations in response to an announced policy to reduce the prevailing rate of inflation. This shift in expectations helps to lower current inflation, leading to a faster and less costly disinflation episode. In contrast, when central bank credibility is low, agents' expectations respond only gradually as they become convinced of the central bank's commitment to reducing inflation. In such an environment, nominal wages and long-term interest rates adjust slowly to the new inflation regime, contributing to a misallocation of resources and a more costly transition to low inflation.

The above reasoning suggests that low credibility on the part of the Federal Reserve may help to explain the severity of the recession induced by the Volcker disinflation. Indeed, it seems likely that the Federal Reserve's commitment to reducing inflation was viewed with considerable skepticism in 1980. Two previous attempts to reduce inflation begun in April 1974 and August 1978 had proven unsuccessful.[3] Contributing to this skepticism in the early stages of the disinflation were large and erratic fluctuations of monetary

Reprinted from: *Federal Reserve Bank of San Francisco Economic Review* (No. 2, 1998) 1–16.

aggregates, which were frequently outside their target ranges.[4] Moreover, U.S. fiscal policy during the early 1980s was characterized by large and growing federal budget deficits which, if projected forward, might have been seen to imply the need for future monetization of the debt to maintain solvency of the government's intertemporal budget constraint.[5]

In this chapter, we develop a simple, quantitative model of the U.S. economy to demonstrate how imperfect credibility on the part of the Federal Reserve may give rise to an episode known as an "inflation scare." Following Goodfriend (1993), we define an inflation scare as a significant increase in the long-term nominal interest rate that takes place in the absence of any aggressive tightening by the Fed that would serve to push up short-term rates. Hence, during an inflation scare, the increase in the long rate is driven primarily by an upward shift in agents' expectations about future inflation. In our model of an inflation scare, an adverse movement in the inflation rate triggers a deterioration in the public's beliefs about the Federal Reserve's commitment to maintaining low inflation in the future. This leads to a sudden increase in the long-term nominal interest rate, even while the short-term rate can actually be falling. We find that simulations from our model capture some observed patterns of U.S. interest rates in the 1980s.

The framework for our analysis is a version of the rational expectations macroeconomic model developed by Fuhrer and Moore (1995a,b). This model is quite tractable and has the advantage of being able to reproduce the dynamic correlations among U.S. inflation, short-term nominal interest rates, and deviations of real output from trend. The model consists of an aggregate demand equation, a nominal wage contracting equation (that embeds a version of an expectations-augmented Phillips curve), a Fed reaction function that defines monetary policy, and a term structure equation. A simple version of Okun's law relates the unemployment rate to the deviation of real output from trend.

We consider an experiment where the economy is initially in a regime of high and variable inflation and the Fed announces a program to reduce both the mean and variance of the inflation rate. The announced program (which is immediately implemented) involves a change to the parameters of the reaction function. Specifically, the inflation target is lowered and more weight is placed on minimizing the variance of inflation versus stabilizing output. We formalize the notion of credibility as the public's subjective probabilistic belief that the reaction function parameters have in fact been changed. The true parameters are assumed to be unobservable due to the presence of exogenous stochastic shocks that enter the reaction function. These policy shocks, together with stochastic disturbances to other parts of the economy, give rise to a distribution of observed inflation rates around any given inflation target.

Under full credibility, the economy is assumed to be populated by agents who, upon hearing the Fed's announcement, assign a probability of one to the event that the reaction function has changed. These agents continue to assign a probability of one regardless of the time path of inflation that is subsequently observed. In contrast, partial credibility implies that agents update their prior assessment of the true reaction function in a (quasi) Bayesian way on the basis of the Fed's success or failure in reducing inflation over time. Our setup is similar to one used by Meyer and Webster (1982) in which agents' expectations are constructed as a probability weighted average of the expectations that would prevail under an "old" and "new" policy rule.[6]

The behavior of the long-term nominal interest rate in the model is governed by the pure expectations hypothesis, that is, the long-term rate is a weighted average of current and expected future short-term rates. If the short rate rises as a result of tighter monetary policy, the implications for the long rate are theoretically ambiguous. In particular, upward pressure stemming from the increase in the current short rate may be offset by downward pressure from the anticipation of lower short rates in the future, due to lower expected inflation. Thus, by affecting the level of expected inflation, the degree of Fed credibility can exert a strong influence on the long-term nominal interest rate.

Using reaction function parameters estimated over the two sample periods 1965:1 to 1979:4 and 1980:1 to 1996:4 we trace out the economy's dynamic transition path for the two specifications of credibility described above. The speed at which agents adjust their inflation expectations in response to the change in monetary policy depends crucially on the Fed's credibility: expectations adjust quickly with full credibility and slowly with partial credibility.

Under both specifications of credibility, we find that the inflation rate exhibits damped oscillations as the economy transitions to the new stationary equilibrium. Following the change in Fed policy, the inflation rate undergoes an initial drop, but ends up overshooting the new target level. The inflation rate then starts to increase as it approaches the new target from below. When the Fed does not have full credibility, agents interpret this interval of rising inflation as evidence that monetary policy has not in fact changed and therefore will continue to tolerate an environment of high and variable inflation. Consequently, agents' expectations of future inflation are revised upward, and the long-term nominal interest rate experiences a sudden increase. In this way, our model generates an endogenous inflation scare.

Numerical simulations of our model produce a 2 percentage point jump in the long-term nominal interest rate that begins about 24 quarters after the change in Fed policy. A similar pattern can be observed in the U.S. data about 29 quarters after the start of the Volcker disinflation.[7] Specifically, from 1986:4 to 1987:4, the yield on a 10-year Treasury bond increased sharply, despite only a small increase in the 3-month Treasury bill rate. Over this same period, the inflation rate (based on the GDP deflator) was rising. This pattern suggests that the increase in the U.S. long rate was driven by an upward shift in the public's expectations about future inflation, thus conforming with our definition of an inflation scare. Given this interpretation of the data, the 1987 scare episode illustrates the long memory of the public in recalling the high and variable inflation of the 1970s, and serves as an important reminder of the fragility of Federal Reserve credibility.

Although Goodfriend (1993) identifies three other inflation scare episodes in U.S. data that occurred much closer to the start of the Volcker disinflation,[8] we choose to emphasize the 1987 scare for two reasons: First, the magnitude and timing of the 1987 episode is reasonably close to the inflation scare that we are able to generate using the model, and second, the episode stands out readily in a plot of quarterly U.S. data. Interestingly, the 1987 scare occurred shortly after U.S. inflation "bottomed out" and again started to rise. This feature of the data resembles the dynamic overshooting behavior of inflation in our model. The point of the exercise, however, is simply to illustrate the mechanics by which an inflation scare may occur—not to identify any one episode as being more significant than the others.

The remainder of the paper is organized as follows. The first section describes the model and our specification of Federal Reserve credibility. The second section presents our parameter estimates and examines their sensitivity to different sample periods. The third section presents our simulation results. The last section concludes.

THE MODEL

The model is a version of the one developed by Fuhrer and Moore (1995a,b). This framework has the advantage of being able to reproduce the pattern of dynamic correlations exhibited by an unconstrained vector autoregression system involving U.S. inflation, short-term nominal interest rates, and deviations of real output from trend. In the model, agents' expectations are rational and take into account the nature of the monetary policy regime, as summarized by the parameters of the Fed reaction function. However, since the other parts of the economy are specified as reduced-form equations, the model is susceptible to Lucas's (1976) econometric policy critique. Our estimation procedure attempts to gauge the quantitative importance of the Lucas critique for our results by examining the stability of the model's reduced form parameters across different sample periods. The equations that describe the model are as follows:

Aggregate Demand/I-S Curve

$$\tilde{y}_t = a_1\tilde{y}_{t-1} + a_2\tilde{y}_{t-2} + a_\rho(\rho_{t-1} - \bar{\rho}) + \varepsilon_{yt}, \tag{1}$$

where $\tilde{y}_t$ is the so-called "output gap" defined as the deviation of log per-capita real output from trend and ρ_{t-1} is the lagged value of the ex ante long-term real interest rate. The error term $\varepsilon_{yt} \sim N(0,\sigma^2_{\varepsilon_y})$ captures random fluctuations in aggregate demand. We assume that the steady-state value of $\tilde{y}_t$ is zero, which implies that $\bar{\rho}$ is the steady-state real interest rate.

Wage Contracting Specification/Short-Run Phillips Curve

$$\pi_t = \frac{1}{2}(\pi_{t-1} + E_t\pi_{t+1}) + \frac{\gamma}{2}(\tilde{y}_t + \tilde{y}_{t-1}) + \varepsilon_{\pi t}, \tag{2}$$

where π_t is the inflation rate defined as the log-difference of the price level, E_t is the expectation operator conditional on information available at time t, and $\varepsilon_{\pi t} \sim N(0, \sigma^2_{\varepsilon\pi})$ is an error term. Fuhrer and Moore (1995a) show that (2) can be derived from a two-period model of staggered nominal wage contracts, where the real value of the contract price negotiated at time t is a simple average of the real contract price negotiated at $t - 1$ and the real contract price that agents expect to negotiate at $t + 1$, adjusted for the level of aggregate demand. The forward-looking nature of wage contracts creates an environment where current inflation depends on expected inflation. The error term represents a stochastic disturbance that affects labor supply decisions.[9]

Equation (2) can also be interpreted as a version of an expectations-augmented Phillips curve.[10] Evidence of a short-term Phillips curve trade-off can be found in the positive correlation between inflation and the real output gap in postwar U.S. data, and the corresponding negative correlation between inflation and the unemployment rate.[11] The steady-state version of (2) implies that there is no long-run trade-off between inflation and real output.

Federal Reserve Reaction Function

$$r_t = r_{t-1} + \alpha_\pi(\pi_t - \bar{\pi}) + \alpha_y\tilde{y}_t + \varepsilon_{rt}, \tag{3}$$

where r_t is the short-term nominal interest rate, $\bar{\pi}$ is the inflation target, and $\varepsilon_{rt} \sim N(0, \sigma^2_{\varepsilon r})$ is an exogenous stochastic shock that is not directly observed by the public. The policy rule implies that the Fed strives to smooth short-term interest rates, but responds to deviations of inflation from target and to deviations of output from trend. The strength of the interest rate response to these deviations is governed by the parameters α_π and α_y.[12] We interpret ε_{rt} as capturing random, nonsystematic factors that arise from the political process or the interaction of policymakers with different preferences, different target rates of inflation, etc. Alternatively, we could interpret ε_{rt} as reflecting operational or institutional features that preclude perfect control of r_t.[13] The presence of the unobservable shock term is crucial for our credibility analysis because it prevents agents from being able to quickly learn the true values of $\bar{\pi}$, α_π, and α_y from a sequence of four observations on r_t, π_t, and $\tilde{y}_t$. Equation (3) implies that the steady-state inflation rate is $\bar{\pi}$.

Real Term Structure

$$\rho_t - D\,(E_t\,\rho_{t+1} - \rho_t) = r_t - E_t\,\pi_{t+1}, \tag{4}$$

where D is the duration of a real consol that is used here to approximate a finite maturity long-term bond. Equation (4) is an arbitrage condition that equates the expected real holding-period return on a long-term bond (interest plus capital gains) with the expected real yield on a short-term Treasury security. In steady-state, (4) implies the Fisher relationship: $\bar{r} = \bar{\rho} + \bar{\pi}$. By repeatedly iterating (4) forward and solving the resulting series of equations for ρ_t, we obtain the following expression:

$$\rho_t = \frac{1}{1+D}\,E_t \sum_{i=0}^{\infty} \left(\frac{D}{1+D}\right)^i (r_{t+i} - \pi_{t+1+i}), \tag{5}$$

which shows that the ex ante long-term real rate is a weighted average of current and expected future short-term real rates.[14]

Nominal Term Structure

$$R_t - D\,(E_t R_{t+1} - R_t) = r_t, \tag{6}$$

$$R_t = \frac{1}{1+D}\,E_t \sum_{i=0}^{\infty} \left(\frac{D}{1+D}\right)^i r_{t+i}, \tag{7}$$

where R_t is the nominal yield on the long-term bond. The above equations are the nominal counterparts of (4) and (5). In steady-state, equation (6) implies $\bar{R} = \bar{r}$.

Okun's Law

$$u_t = (1 - b_1)\,\bar{u} + b_1 u_{t-1} + b_2 \tilde{y}_t + b_3 \tilde{y}_{t-1} + b_4 \tilde{y}_{t-2} + \varepsilon_{ut}, \tag{8}$$

where u_t is the unemployment rate, $\bar{u}$ is the corresponding steady-state, and $\varepsilon_{ut} \sim N\,(0, \sigma^2_{\varepsilon u})$ is an error term.[15]

Credibility

In modeling the role of credibility during the Volcker disinflation, we abstract from the Fed's adoption of a new operating procedure for targeting nonborrowed reserves from October 1979 to October 1982. Studies by Cook (1989) and Goodfriend (1993) indicate that the majority of federal funds rate movements during this period were the result of deliberate, judgmental policy actions by the Fed, and not automatic responses to deviations of the money stock from its short-run target.[16] Moreover, it has been suggested that the Fed's emphasis on monetary aggregates during this period was simply a device that allowed it to disclaim responsibility for pushing up short-term nominal interest rates to levels that would otherwise have been politically infeasible. Based on the above reasoning, we interpret the Fed's statement on October 6, 1979, as an announcement of a change in the parameters of the reaction function.[17]

We consider an experiment where the economy is initially in a regime of high and variable inflation and the Fed announces a program to reduce both the mean and variance of the inflation rate. The announced program (which is immediately implemented) involves a change to the parameters of the reaction function (3). Specifically, the inflation target $\overline{\pi}$ is lowered, the parameter α_π is increased, and the parameter α_y is decreased. This constitutes a regime shift that is consistent with the empirical evidence of a statistical break in U.S. inflation occurring around October 1979.[18] The increase in α_π relative to α_y implies a decision on the part of the Fed to place more emphasis on minimizing the variance of inflation and less emphasis on stabilizing output.[19] It is important to recognize that we have simply posited the Fed's decision to change monetary policy, since our model abstracts from any economic benefits of low and stable inflation. Moreover, we do not attempt to explain how the Fed allowed inflation to become too high and variable in the first place.[20]

We define credibility as the public's subjective probabilistic belief that the announced policy change has in fact occurred. To formalize this idea, we endow agents with the knowledge of two possible reaction functions and the corresponding equilibrium distributions of π_t that arise under each. The two reaction functions are defined by the parameter combinations $\{\overline{\pi}^H, \alpha_\pi^H, \alpha_y^H\}$ and $\{\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L\}$, where $\overline{\pi}^L < \overline{\pi}^H$, $\alpha_\pi^L > \alpha_\pi^H$, and $\alpha_y^L < \alpha_y^H$. In a stationary equilibrium, the linearity of the model, together with the assumptions that ε_{yt}, $\varepsilon_{\pi t}$, and ε_{rt} are i.i.d. normal implies

$$\pi_t \sim N(\overline{\pi}, \sigma_\pi^2), \tag{9}$$

where the mean of the inflation distribution is the steady-state and the variance σ_π^2 depends on the variances of the stochastic shocks.

We assume that the economy is initially in a stationary equilibrium with the reaction function parameters $\{\overline{\pi}^H, \alpha_\pi^H, \alpha_y^H\}$. These parameters give rise to the distribution $\pi_t \sim N(\overline{\pi}^H, \sigma_{\pi H}^2)$. At $t = t^*$ the Fed adopts the new reaction function parameters $\{\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L\}$ and announces this action to the public. The unobservable error term ε_{rt} in (3) prevents the public from being able to verify the Fed's announcement from a sequence of four observations of r_t, π_t, and $\tilde{y}_t$. Hence, the public's beliefs regarding the reaction function parameters are used to form expectations while the true parameter values are used in (3) to compute the period-by-period values of r_t. Learning takes place (as described below), and the economy eventually converges to a new stationary equilibrium with $\pi_t \sim N(\overline{\pi}^L, \sigma_{\pi L}^2)$, where $\sigma_{\pi L}^2 < \sigma_{\pi H}^2$. In other words, the change in Fed policy ultimately brings about an inflation distribution with a lower mean and a lower variance.

We consider two specifications of credibility, labeled "full" and "partial." Full credibility implies that agents assign the probability $p_t = 1$ to the parameter combination $\{\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L\}$ for all $t \geq t^*$. Under partial credibility, agents assign a "prior" probability to the parameter combination $\{\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L\}$ at the time of the Fed's announcement. This prior is a free parameter that is influenced by the Fed's past track record in maintaining control over inflation. Agents compute a sequence of posterior probabilities $\{p_t\}_{t=t^*}^{\infty}$ by updating their prior in a (quasi) Bayesian way on the basis of observed realizations of the inflation rate and knowledge of the two (long-run) distributions of inflation centered a $\overline{\pi}^H$ and $\overline{\pi}^L$. The degree of Fed credibility is indexed by p_t.

We make the simplifying assumption that agents do not take into account the evolving nature of the inflation distribution during the transition to the new stationary equilibrium. Furthermore, we follow Meyer and Webster (1982), Baxter (1989), and Fuhrer and Hooker (1993), in assuming that the Fed's policy action is a once-and-for-all change. Thus, agents do not consider the possibility of any future regime shifts when forming their expectations.[21]

Under partial credibility, the public's beliefs regarding the reaction function parameters for $t \geq t^*$ evolve according to a version of Bayes' rule:

$$p_t = \frac{p_{t-1}\Pr(\pi_t \leq \pi_{t-1} \mid \overline{\pi}^L,\alpha_\pi^L,\alpha_y^L)}{p_{t-1}\Pr(\pi_t \leq \pi_{t-1} \mid \overline{\pi}^L,\alpha_\pi^L,\alpha_y^L) + (1-p_{t-1})\Pr(\pi_t \leq \pi_{t-1} \mid \overline{\pi}^H,\alpha_\pi^H,\alpha_y^H)}, \tag{10}$$

with p_{t^*-1} given. The posterior probability $p_t \equiv \Pr(\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L \mid \pi_t \leq \pi_{t-1})$ is computed by combining the prior probability $p_{t-1} \equiv \Pr(\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L)$ with in-sample information. Specifically, the prior is weighted by $\Pr(\pi_t \leq \pi_{t-1} \mid \overline{\pi}^L, \alpha_\pi^L, \alpha_y^L)$, which represents the probability that inflation in period t will be lower than inflation observed in period $t-1$, conditional on the parameter combination $\{\overline{\pi}^L, \alpha_\pi^L, \alpha_y^L\}$. The relevant probability weights in (10) are given by

$$\Pr(\pi_t \leq \pi_{t-1} \mid \overline{\pi}^L,\alpha_\pi^L,\alpha_y^L) = \int_{-\infty}^{\pi_{t-1}} \ell(z)dz, \tag{11}$$

$$\Pr(\pi_t \leq \pi_{t-1} \mid \overline{\pi}^H,\alpha_\pi^H,\alpha_y^H) = \int_{-\infty}^{\pi_{t-1}} h(z)dz, \tag{12}$$

where $\ell(z)$ and $h(z)$ are the normal density functions that describe the stationary inflation distributions centered at $\overline{\pi}^L$ and $\overline{\pi}^H$, respectively.

Three features of the above specification warrant comment. First, the integrals in (11) and (12) are computed using the observation of π_{t-1}, *not* π_t. This is done to preserve the model's linearity in π_t. In particular, since p_t is used to construct the expectation $E_t\pi_{t+1}$ (as described below), the specification $p_t = p(\pi_t)$ would imply that (2) is nonlinear in the current period inflation rate. Maintaining linearity in π_t is desirable because it greatly simplifies the model solution procedure.[22]

Second, (11) and (12) imply that probability inferences are made using observations of a single economic variable (inflation), and that the relevant data sample includes only the most recent inflation rate, not the whole history of inflation rates $\{\pi_{t-i}\}_{i=1}^{t-t^*}$ observed since the announcement.[23] While our setup maintains tractability, it introduces some nonrationality into agents' forecasts to the extent that they ignore the potentially valuable information contained in the history of joint observations on inflation, interest rates, and the real output gap.[24]

Third, equation (10) differs from the standard classification formula for computing the conditional probability that a given observation comes from one of two populations with known densities.[25] In our model, the standard formula would take the form

$$p_t = \frac{p_{t-1}\,\ell(\pi_{t-1})}{p_{t-1}\,\ell(\pi_{t-1}) + (1-p_{t-1})h(\pi_{t-1})}, \tag{10'}$$

which says that p_t depends on the relative *heights* of the two density functions evaluated at π_{t-1}. In contrast, equation (10) says that p_t depends on the relative *areas* of the two density functions to the left of π_{t-1}. In quantitative simulations, we find that (10) quickens the pace of learning in comparison to (10′) and thus leads to more a realistic transition time between steady states. This occurs because (10) introduces an implicit bias into agents' inferences such that p_t is higher than that implied by (10′) for any given value of p_{t-1}. For the parameter values we consider, both specifications exhibit the desirable property that the credibility index p_t declines monotonically as inflation rises, for any given p_{t-1}.[26]

After computing the posterior probability, agents' expectations are formed as a weighted average of the rational forecasts that would prevail under each of the two possible reaction functions:

$$E_t\pi_{t+1} = p_t E_t[\pi_{t+1} \mid \overline{\pi}^L, \alpha_\pi^L, \alpha_y^L] + (1 - p_t)E_t[\pi_{t+1} \mid \overline{\pi}^H, \alpha_\pi^H, \alpha_y^H], \tag{13}$$

$$E_t\rho_{t+1} = p_t E_t[\rho_{t+1} \mid \overline{\pi}^L, \alpha_\pi^L, \alpha_y^L] + (1 - p_t)E_t[\rho_{t+1} \mid \overline{\pi}^H, \alpha_\pi^H, \alpha_y^H], \tag{14}$$

$$E_tR_{t+1} = p_t E_t[R_{t+1} \mid \overline{\pi}^L, \alpha_\pi^L, \alpha_y^L] + (1 - p_t)E_t[R_{t+1} \mid \overline{\pi}^H, \alpha_\pi^H, \alpha_y^H], \tag{15}$$

where p_t is given by (10). Since p_t is a function of past inflation, the model with rational expectations and partial credibility will now exhibit some of the backward looking characteristics of a model with adaptive expectations.[27]

ESTIMATION AND CALIBRATION

For the purpose of estimating parameters, we adopt a baseline model specification that incorporates full credibility. The resulting parameter set is then used for both credibility specifications in order to maintain comparability in the simulations. The data used in the estimation procedure are summarized in Table 1.

The model's reduced-form parameters are assumed to be "structural" in the sense that they are invariant to changes in the monetary policy reaction function (3). We attempt to gauge the reasonableness of this assumption by examining the sensitivity of the parameter estimates to different sample periods. Following Fuhrer (1996), we do not estimate the duration parameter but instead calibrate it to the value $D = 28$. This coincides with the sample average duration (in quarters) of a 10-year constant maturity Treasury bond. Equations (1) through (4) form a simultaneous system that we estimate using full-information maximum likelihood.[28] The estimation results are summarized in Table 2.

Despite small differences in our model specification and data, estimates from the full sample (1965:1 to 1996:4) are very much in line with those obtained by Fuhrer and Moore (1995b). With the exception of a_ρ and γ, the parameter estimates are all statistically signif-

Table 1 Quarterly Data, 1965:1 to 1996:4

Variable	Definition
$\tilde{y}_t$	Deviation of log per capita real GDP from its linear trend.
π_t	Log-difference of GDP implicit price deflator.
r_t	Yield on 3-month Treasury bill.
R_t	Yield on 10-year constant-maturity Treasury bond.
u_t	Nonfarm civilian unemployment rate.

Table 2 Maximum Likelihood Parameter Estimates

	1965:1 to 1996:4		1965:1 to 1979:4		1980:1 to 1996:4	
Parameter	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
a_1	1.23	0.09	0.94	4.97	1.24	0.10
a_2	−0.26	0.08	0.10	4.62	−0.31	0.09
a_ρ	−0.20	0.12	−0.57	2.17	−0.05	0.05
$\bar{\rho}$	0.02	0.01	0.02	0.36	0.00	0.04
γ	0.01	0.01	0.04	0.47	0.01	0.01
α_π	0.06	0.03	0.07	1.04	0.10	0.05
α_y	0.08	0.03	0.07	1.05	0.05	0.06
$\bar{\pi}$	0.05	0.01	0.04	0.45	0.05	0.01

icant. In contrast, the estimates from the first subsample (1965:1 to 1979:4) are highly imprecise, most likely due to the strong upward trends in U.S. inflation and nominal interest rates over this period. Estimates from the second subsample (1980:1 to 1996:4) are much closer to the full-sample results. Evidence of subsample instability seems to be concentrated mostly in the I-S curve parameters a_1, a_2, and a_ρ. Notice, however, that all subsample point estimates lie within one standard error of each other. We interpret these results to be reasonably supportive of the hypothesis that the reduced-form parameters a_1, a_2, a_ρ, $\bar{\rho}$ and γ do not vary across monetary policy regimes.

A comparison of the subsample point estimates of α_π and α_y suggests that the Fed has placed more emphasis on targeting inflation and less emphasis on stabilizing output in the period after 1980. For the simulations, we choose $\alpha_\pi^H = \alpha_y^H = 0.07$ for the high inflation regime and $\alpha_\pi^L = 0.10$ and $\alpha_y^L = 0.05$ for the low inflation regime. To complete the specification of the reaction function, we require values for $\bar{\pi}^H$ and $\bar{\pi}^L$. We choose $\bar{\pi}^H = 0.06$ to coincide with the sample mean from 1965:1 to 1979:4. Thus, we assume that the U.S. inflation rate prior to October 1979 can be characterized by a stationary distribution centered at 6 percent. While this assumption is undoubtedly false, it serves to illustrate the effects of partial credibility on the disinflation episode. Since $\bar{\pi}^L$ is intended to represent the new steady-state after the disinflation has been completed, we choose $\bar{\pi}^L = 0.03$ to coincide with the sample mean from 1985:1 to 1996:4. In computing this average, we omit the period of rapidly falling inflation from 1980:1 to 1984:4 because this can be interpreted as the transition to the new steady state.[29] For the other model parameters, we adopt the full-sample estimates in Table 2.

Our disinflation simulations abstract from stochastic shocks because these have the potential to obscure differences between the dynamic propagation mechanisms of the two credibility specifications.[30] We assume, however, that agents make decisions *as if* stochastic shocks were present. This assumption is necessary for a meaningful analysis of credibility because without stochastic shocks, agents can always learn the true values of $\bar{\pi}$, α_π, and α_y within four periods. To compute the integrals in (11) and (12), we simply calibrate the standard deviations of the two long-run inflation distributions centered at $\bar{\pi}^H$ and $\bar{\pi}^L$. For the high inflation regime, we choose $\sigma_{\pi H} = 0.023$ to coincide with the sample standard deviation from 1965:1 to 1979:4. For the low inflation regime, we choose $\sigma_{\pi L} = 0.011$ to coincide with the sample standard deviation from 1985:1 to 1996:4. In computing this statistic, we once again exclude the transition period from 1980:1 to 1984:4.

For the steady-state unemployment rate, we choose $\bar{u} = 0.06$ to coincide with the average over the full sample. Given $\bar{u}$, we estimate the parameters of Okun's law (8) using ordinary least squares to obtain $b_1 = 0.96$, $b_2 = -0.30$, $b_3 = 0.10$, and $b_4 = 0.18$, which are all statistically significant.

Our solution procedure can be briefly summarized as follows. Given a set of parameters, we solve the full-information version of the model for each of the two reaction functions described by $\{\bar{\pi}^H, \alpha_\pi^H, \alpha_y^H\}$ and $\{\bar{\pi}^L, \alpha_\pi^L, \alpha_y^L\}$. In each case, the solution consists of a set of time-invariant linear decision rules for π_t, ρ_t, and R_t, defined in terms of the "state" vector $s_t = \{\tilde{y}_{t-1}, \tilde{y}_{t-2}, \pi_{t-1}, \rho_{t-1}, r_{t-1}\}$. The decision rules for $\tilde{y}_t$ and r_t are simply given by (1) and (3), respectively. For each reaction function, we use the decision rules to construct linear expressions for the conditional expectations $E_t[\pi_{t+1} \mid \bar{\pi}^i, \alpha_\pi^i, \alpha_y^i]$, $E_t[\rho_{t+1} \mid \bar{\pi}^i, \alpha_\pi^i, \alpha_y^i]$, and $E_t[R_{t+1} \mid \bar{\pi}^i, \alpha_\pi^i, \alpha_y^i]$, $i = L, H$. Next, we form the unconditional expectations $E_t\pi_{t+1}$, $E_t\rho_{t+1}$, and E_tR_{t+1} using the current value of p_t (which does not depend on π_t) and (13) through (15). Finally, the unconditional expectations are substituted into (2), (4), and (6) which, together with (1) and (3), form a system of five linear equations in the five unknowns $\tilde{y}_t$, π_t, ρ_t, r_t, and R_t.

Under full credibility, it is straightforward to show that the model possesses a unique, stable equilibrium for the parameter values we employ.[31] Under partial credibility, agents use observations of an endogenous variable (inflation) to form expectations that are crucial for determining the period-by-period values of that same variable. The presence of this dynamic feedback effect between the trajectory of inflation and the inputs to the learning process may create an environment where learning goes astray. In particular, there is no way to guarantee that the model will converge to a new steady state with $\bar{\pi} = \bar{\pi}^L$.[32] However, for the parameter values we employ, we find that convergence is achieved in the quantitative simulations.[33]

QUANTITATIVE RESULTS

Deterministic Disinflation Simulations

In our experiments with the model, we find that a very low prior p_{t^*-1} is needed for the model to generate an endogenous inflation scare. Therefore, in our specification with partial credibility, we set the initial prior to 0.001 percent. This choice also reflects our view (noted earlier) that the Federal Reserve had very little credibility at the start of the Volcker disinflation.[34]

The evolution of credibility is shown in Figure 1A. With full credibility, p_t jumps immediately to 100 percent on the strength of the Fed's announcement at $t^* = 0$. With partial credibility, p_t increases slowly over time as agents observe that π_t is falling (see Figure 1B). This feature of the model is consistent with the findings of Hardouvelis and Barnhart (1989) who show that an empirical proxy for Fed credibility increased only gradually in the period following October 1979. Moreover, they find that credibility is statistically linked to the rate of inflation.[35]

Credibility approaches the value $p_t = 100$ percent approximately 16 quarters after the change in Fed policy. Once full credibility is reached, Bayes' rule (10) implies that $p_t = 100$ percent will be sustained forever. However, as long as $p_t < 100$ percent by even a single decimal point, the economy will be susceptible to an inflation scare. In the simulation, credibility peaks at a value of 99.97742 percent and then begins to deteriorate

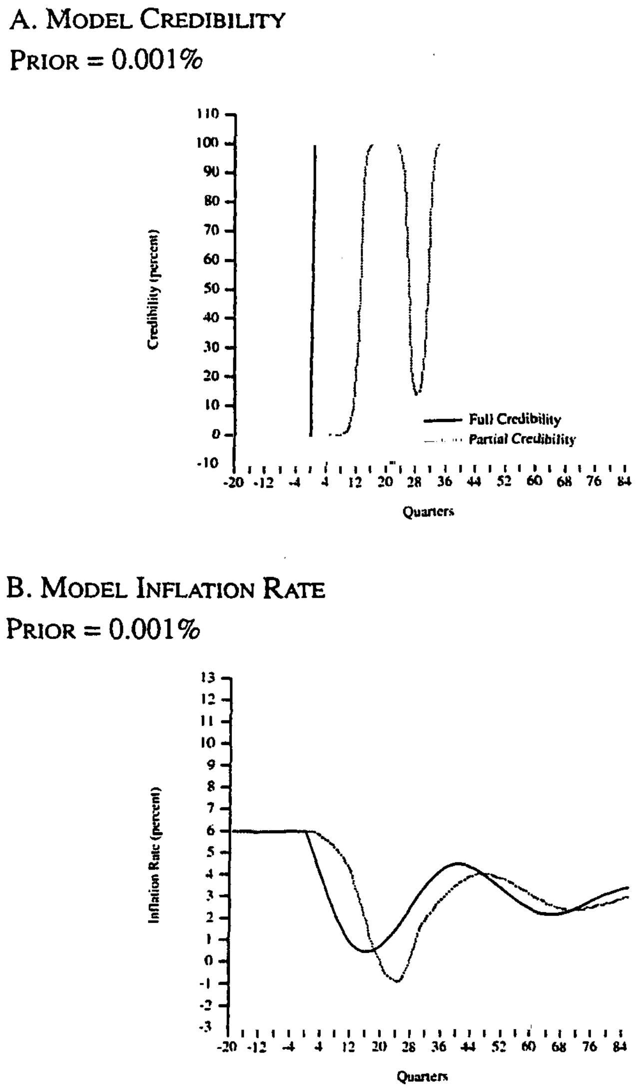


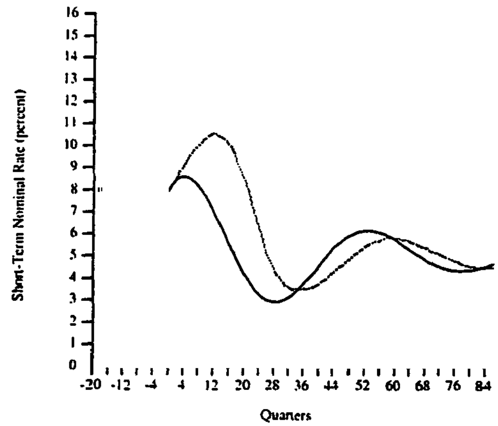
Figure 1 Credibility and inflation.

rapidly. This loss of credibility is triggered by the period of rising inflation (observed in Figure 1B) that results from the dynamic overshooting characteristics of the model.[36]

Figure 1B shows that disinflation proceeds more slowly under partial credibility. The intuition for this result follows directly from equation (2). With partial credibility, the sluggish behavior of $E_t\pi_{t+1}$ delays the response of current inflation π_t to the policy change. This, in turn, delays the accumulation of credibility, which feeds back to inflation expectations.[37]

Figure 2A shows that both credibility specifications imply an initial monetary contraction, as evidenced by an increase in the short-term nominal interest rate r_t.[38] With partial credibility, the Fed undertakes a greater degree of monetary tightening, as measured by

A. MODEL SHORT-TERM NOMINAL INTEREST RATE
PRIOR = 0.001%



B. MODEL LONG-TERM NOMINAL INTEREST RATE
PRIOR = 0.001%

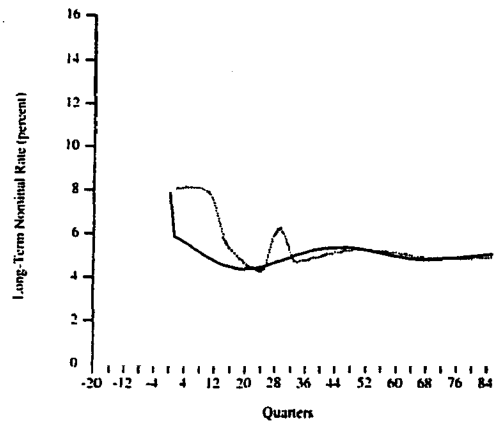


Figure 2 Nominal interest rates.

the peak level of r_t. This is due to the form of the reaction function (3) that makes r_t an increasing function of the distance $\pi_t - \overline{\pi}^L$. Since π_t falls more slowly under partial credibility, the level of r_t implied by (3) is higher. Moreover, the sluggish adjustment of $E_t\pi_{t+1}$ means that a higher level of inflation is built into expectations of *future* short rates. These two effects combine to raise the level of the current long rate R_t in comparison to the model with full credibility. Figure 2B shows that, under partial credibility, the inertia built into agents' inflation forecasts is sufficient to cause R_t to increase slightly in response to the tighter monetary policy. In contrast, full credibility generates an immediate *fall* in R_t as agents quickly lower their inflation expectations. Empirical studies generally indicate that tighter monetary policy leads to an increase in long-term nominal interest rates.[39]

PRIOR = 0.001%

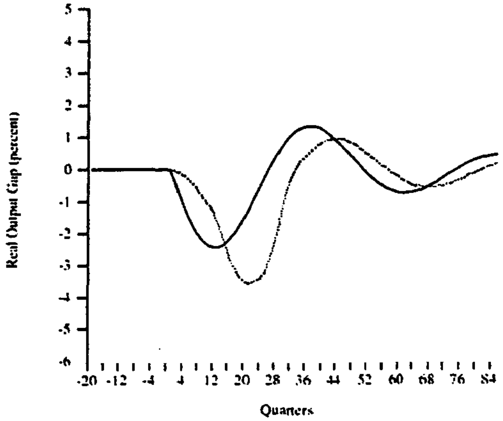


Figure 3 Model real output gap.

The key feature of Figure 2B is the inflation scare that occurs about 24 quarters after the change in Fed policy. The scare produces a 2 percentage point jump in the long-term rate R_t that coincides with the interval of deteriorating credibility and rising inflation described above. Notice that the jump in R_t takes place in the absence of any aggressive tightening by the Fed. In fact, Figure 2A shows that the short-term rate r_t is actually *falling* during the inflation scare. Equation (13) implies that a decrease in p_t will cause expectations of future inflation to be revised upward. This forecast of higher inflation implies higher future values of r_t which, in turn, are incorporated into R_t via the term structure equation (7). In this way, the model generates an endogenous inflation scare.

Figures 3 and 4 show that the Fed's tighter monetary policy leads to a prolonged recession: real output declines relative to trend, and the unemployment rate goes up. Notice

PRIOR = 0.001%

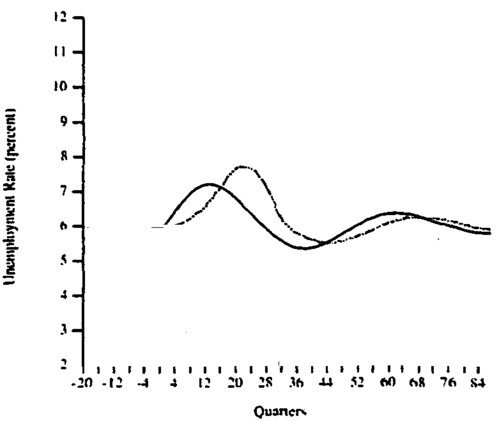


Figure 4 Model unemployment rate.

that the recession is considerably more severe in the case of partial credibility. This result helps to provide some insight into the high unemployment rates observed during the Volcker disinflation which, as we argued earlier, was initiated when the Fed's credibility was very low.

The time paths of the model variables in Figures 3 and 4 illustrate a potentially important stabilization property of full credibility. In particular, stabilization of the model is aided by the elimination of the backward-looking dynamics associated with the learning process. This result is consistent with the findings of Fuhrer (1997), who shows that a stronger *forward-looking* component in the contracting equation (2) helps to stabilize the model.[40]

Comparison with Volcker Disinflation

Figures 5 and 6 compare the evolution of U.S. macroeconomic variables during the Volcker disinflation with the corresponding variables in our model. The vertical line in the U.S. figures marks the start of the Volcker disinflation in October 1979. The model captures many of the qualitative features of the Volcker disinflation. Notice that the U.S. variables appear to exhibit some low frequency, damped oscillations that resemble the dynamic overshooting characteristics of the model variables. It should be noted, however, that the 16-year period following October 1979 may include some additional monetary policy actions that are not present in the model. For example, Taylor (1993) shows that the time path of the federal funds rate since 1987 is well-described by a policy rule with an inflation target of 2 percent (see footnote 12). In addition, Romer and Romer (1994) find evidence that the Federal Reserve made a deliberate decision to reduce inflation in December 1988.

In Figures 5A–C, we highlight the classic pattern of an inflation scare that can be observed in U.S. data about 29 quarters after the start of the Volcker disinflation. Specifically, from 1986:4 to 1987:4, the yield on a 10-year Treasury bond increased sharply from 7.3 percent to 9.1 percent (Figure 5C), despite only a small increase in the 3-month Treasury bill rate from 5.3 to 6.0 percent (Figure 5B). Over this same period, the inflation rate increased from 2.9 to 3.9 percent (Figure 5A). This pattern fits our definition of an inflation scare, suggesting that the increase in the U.S. long rate was driven by an upward shift in the public's expectations of future inflation. Notice that the 1987 scare episode occurred shortly after U.S. inflation "bottomed out" and again started to rise. Interestingly, this feature of the data resembles the dynamic overshooting behavior of inflation in the model (Figure 6A). Given our interpretation of the data, the 1987 scare episode illustrates the long memory of the public in recalling the high and variable inflation of the 1970s, and serves as an important reminder of the fragility of Federal Reserve credibility.[41]

As noted earlier in the introduction, Goodfriend (1993) identifies three other inflation scare episodes in U.S. data that occur much closer to the start of the Volcker disinflation. Our model does not capture these episodes because the dynamic overshooting behavior of the inflation rate (which triggers the inflation scare) takes a long time to evolve. We note, however, that our simulations abstract from stochastic shocks which may have played a role in triggering these earlier episodes.

Another feature of the U.S. data that we do not capture is the dramatic increase in the long-term nominal interest rate in the period following October 1979 (Figure 5C). In Huh and Lansing (1998), we show that a version of this model that combines adaptive expectations with partial credibility can exhibit more sluggish adjustment in inflation expectations. As a result, we find that R_t can rise significantly in response to a tightening of monetary policy.

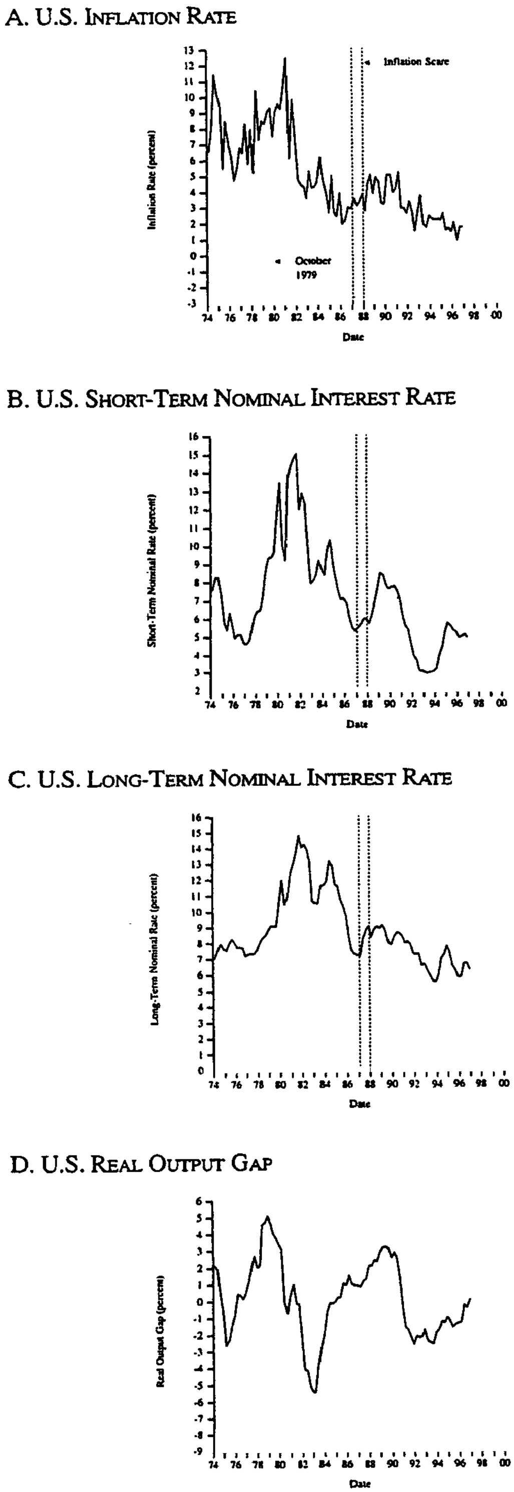
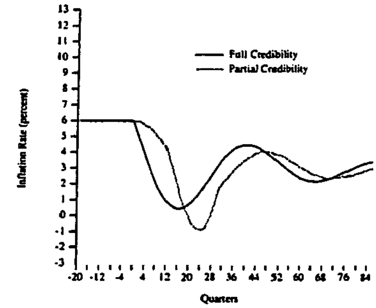
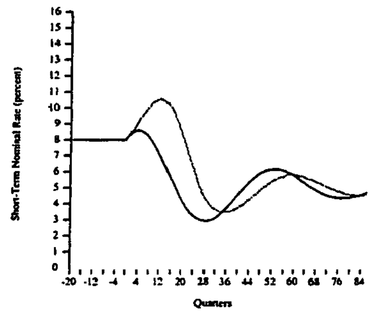


Figure 5 U.S. data.

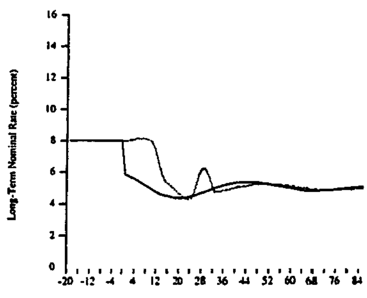
A. Model Inflation Rate: Prior = 0.001%



B. Model Short-Term Nominal Interest Rate: Prior = 0.001%



C. Model Long-Term Nominal Interest Rate: Prior = 0.001%



D. Model Real Output Gap: Prior = 0.001%

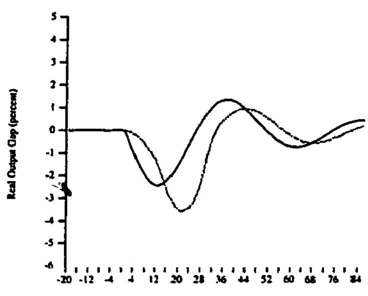


Figure 6 Model simulations.

CONCLUSION

This developed a simple, quantitative model of the U.S. economy to show how an inflation scare may occur when the Federal Reserve lacks full credibility. Our simulation exercise was reasonably successful in capturing the magnitude and timing of the 1987 U.S. inflation scare episode that produced a sharp increase in the 10-year Treasury bond yield. Our model also captures many of the qualitative features of the Volcker disinflation of the early 1980s.

The potential for an inflation scare will continue to exist so long as the public believes that the U.S. economy may someday return to an environment of high and variable inflation. One way of addressing this problem is through legislation designed to enhance credibility by requiring the Fed to pursue some notion of "price stability" as its primary or sole objective. An arrangement such as this was put in place for the central bank of New Zealand in 1989.[42]

NOTES

1. See, for example, Sargent (1982, 1983), Taylor (1982), and Fischer (1986).
2. This idea is the basis for many game theoretic models of credibility in monetary policy. See, for example, Barro and Gordon (1983), Backus and Driffill (1985a,b), Barro (1986), and Cukierman and Meltzer (1986). For an excellent survey of this literature, see Blackburn and Christensen (1989).
3. See Shapiro (1994) for an analysis of the relative success of Federal Reserve attempts to reduce inflation following seven postwar dates marking the start of an explicit disinflationary policy, as identified by Romer and Romer (1989, 1994).
4. For details on monetary policy in the early 1980s, see Friedman (1984). Blanchard (1984). Hetzel (1986), and Goodfriend (1993, 1997).
5. The crucial importance of the fiscal regime in determining the credibility of disinflationary policies is emphasized by Sargent (1982, 1983, 1986). For applications of this idea, see Flood and Garber (1980) and Ruge-Murcia (1995).
6. Other research that applies Bayesian learning to models of monetary policy includes Taylor (1975), Flood and Garber (1980), Backus and Driffill (1985a,b), Barro (1986), Lewis (1989), Baxter (1989), Bertocchi and Spagat (1993), Gagnon (1997), and Andolfatto and Gomme (1997). For related models with least squares learning, see Friedman (1979), Fuhrer and Hooker (1993), and Sargent (1998).
7. We take the starting date of the Volcker disinflation to be October 6, 1979, which coincides with Fed's announcement of a new operating procedure for targeting nonborrowed reserves. This starting date is consistent with the findings of Romer and Romer (1989), who use evidence from the minutes of Federal Open Market Committee meetings to identify October 1979 as a date when the Federal Reserve decided to undertake an explicit disinflationary policy.
8. The approximate dates of these episodes are: (1) December 1979 to February 1980, (2) December 1980 to October 1981, and (3) May 1983 to June 1984.
9. We do not explicitly link the supply shock $\varepsilon_{\pi t}$ to the real price of oil. Fuhrer and Moore (1995a, footnote 15) report that oil prices are uncorrelated with the residuals of their contracting equation, suggesting that their omission does not affect the model's performance. See Bernanke, Gertler, and Watson (1997) for an empirical study of the potential links between oil prices and monetary policy.
10. See Roberts (1997).
11. King and Watson (1994) document the robust negative correlation between inflation and unemployment at business cycle frequencies.

12. The policy rule is similar to one proposed by Taylor (1993), which takes the form: $r_t = (\bar{\rho} + \pi_t) + \alpha_\pi (\pi_t - \bar{\pi}) + \alpha_y \bar{y}_t$, where $\bar{\rho}$ is the steady-state real interest rate. The Taylor rule uses $\bar{\rho} = 0.02$, $\alpha_x = \alpha_y = 0.5$, and $\bar{\pi} = 0.02$. See Taylor (1998) and Judd and Rudebusch (1998) for historical analyses of how policy rules of this form fit U.S. interest rate data.
13. Cukierman and Meltzer (1986) develop a model in which the central bank intentionally adopts an imprecise monetary control process in order to obscure its preferences, and thereby exploit a more favorable output-inflation trade-off.
14. In going from (4) to (5) we have applied the law of iterated mathematical expectations.
15. Since $\bar{u}$ is independent of π_t, it can be interpreted as the "Natural Rate of Unemployment."
16. It is straightforward to append a money demand equation that determines how much money the Fed must supply in order to achieve the value of r_t given by (3). This would have no effect on the model's dynamics.
17. Evidence that the public perceived the statement in this way can be found in published newspaper reports of the time. See, for example, "Fed Takes Strong Steps to Restrain Inflation, Shifts Monetary Tactic," *The Wall Street Journal,* October 8, 1979, p. 1.
18. See, for example, Walsh (1988).
19. See Svensson (1997) and Ball (1997) for analyses of "efficient" monetary policy rules that minimize a discounted weighted-sum of the variances of inflation and output.
20. See Sargent (1998) for a model that seeks to endogenize the rise and fall of U.S. inflation.
21. See Gagnon (1997) for a univariate model of inflation that relaxes both of the foregoing assumptions.
22. Our solution procedure is described in Section II.
23. The history of inflation *does* influence credibility, however, because it is incorporated into agents' prior beliefs, which are summarized by P_{t-1} in (10).
24. See Ruge-Murcia (1995) for a model where credibility is inferred using joint observations on fiscal and monetary variables.
25. See Anderson (1958), Chapter 6.
26. This property will obtain when the ratios $\left(\int_{-\infty}^{\pi} \ell(z)dz\right) / \left(\int_{-\infty}^{\pi} h(z)dz\right)$ and $(\ell(\pi))/(h(\pi))$ are monotonically decreasing in π.
27. A similar effect obtains in the models of Fisher (1986), Ireland (1995), King (1996), Bomfim and Rudebusch (1997), and Bomfim, et al. (1997). In these models, credibility is determined by a backward-looking, linear updating rule. In contrast, Ball (1995) models credibility using a purely time-dependent probability measure.
28. We use the Matlab programs developed by Fuhrer and Moore (1995b), as modified to reflect the differences in our model specification and data.
29. The values $\bar{\pi}^H = 0.06$ and $\bar{\pi}^L = 0.03$ are very close to those used by Fuhrer (1996, figure IIb) to help reconcile the pure expectations theory of the term structure with U.S. nominal interest rate data.
30. For studies that explore disinflation dynamics in models subject to stochastic shocks, see Meyer and Webster (1982), Orphanides, et al. (1997), and Bomfim and Rudebusch (1997).
31. The steady states associated with the two reaction functions both exhibit the well-known saddle point property.
32. In contrast, Taylor (1975), Meyer and Webster (1982), Baxter (1989), and Andolfatto and Gomme (1997), among others, consider Bayesian learning models in which agents' expectations do not affect the evolution of the variables they form expectations about. Hence, convergence follows from standard results on the asymptotic properties of estimators.
33. Marcet and Sargent (1989) develop an analytical framework for proving the convergence of "self-referential" models in which the evolution of an endogenous variable is governed by an adaptive learning process.
34. A similar view is put forth by Mankiw (1994), who shows that forecasts made by the Council of Economic Advisers in January 1981 predicted a gradual and moderate decline in the inflation rate, in contrast to the rapid and pronounced disinflation that actually occurred under Fed Chairman Volcker.

35. The Hardouvelis-Barnhart measure of credibility is inversely proportional to the response of commodity prices (such as gold and silver) to unanticipated changes in the M1 money stock.
36. For the parameter values we employ, the model's dynamical system exhibits complex eigenvalues which give rise to damped oscillatory behavior.
37. In the words of Fed Chairman Volcker. "Inflation feeds in part on itself, so part of the job of returning to a more stable and more productive economy must be to break the grip of inflationary expectations." See Volcker (1979), pp. 888–889.
38. Since r_t rises and $\bar{y}_t$ falls, a traditional Keynesian money demand equation with a predetermined price level would imply a contraction of the nominal money stock.
39. See Akhtar (1995) for a survey of the enormous empirical literature on this subject.
40. For a related discussion, see Taylor (1980, section IV).
41. See Gagnon (1996) for some cross-country evidence that inflation expectations exhibit a "long memory" of past inflation.
42. See Romer and Romer (1997) for a discussion regarding the merits of legislated rules and other institutional arrangements for the conduct of monetary policy.

REFERENCES

Akhtar, M. A. 1995. Monetary Policy and Long-Term Nominal Interest Rates: A Survey of the Empirical Literature. *Contemporary Economic Policy* 13, pp. 110–130.

Anderson, T. W. 1958. *An Introduction to Multivariate Statistical Analysis.* New York: John Wiley & Sons Inc.

Andolfatto, D., and Gomme, P. 1997. Monetary Policy Regimes and Beliefs. Federal Reserve Bank of Minneapolis Discussion Paper 118.

Backus, D., and Driffill, J. 1985a. Inflation and Reputation. *American Economic Review* 75, pp. 530–538.

Backus, D., and Driffill, J. 1985b. Rational Expectations and Policy Credibility Following a Change in Regime. *Review of Economic Studies* 52, pp. 211–221.

Ball, L. 1995. Disinflation with Imperfect Credibility. *Journal of Monetary Economics* 35, pp. 5–23.

Ball, L. 1997. Efficient Rules for Monetary Policy. National Bureau of Economic Research Working Paper 5952.

Barro, R. J. 1986. Reputation in a Model of Monetary Policy with Incomplete Information. *Journal of Monetary Economics* 17, pp. 3–20.

Barro, R. J. and Gordon, D. B. 1983. Rules, Discretion, and Reputation in a Model of Monetary Policy. *Journal of Monetary Economics* 12, pp. 101–121.

Baxter, M. 1989. Rational Response to Unprecedented Policies: The 1979 Change in Federal Reserve Operating Procedures. *Carnegie-Rochester Conference Series on Public Policy* 31, pp. 247–296.

Bernanke, B. S., Gertler, M. and Watson, M. 1997. Systematic Monetary Policy and the Effects of Oil Price Shocks. *Brookings Papers on Economic Activity* 1, pp. 91–157.

Bertocchi, G., and Spagat, M. 1993. Learning, Experimentation, and Monetary Policy. *Journal of Monetary Economics* 32, pp. 169–183.

Blackburn, K., and Christensen, M. 1989. Monetary Policy and Credibility: Theories and Evidence. *Journal of Economic Literature* 27, pp. 1–45.

Blanchard, O. J. 1984. The Lucas Critique and the Volcker Deflation. *American Economic Association Papers and Proceedings* 74, pp. 211–215.

Bomfim, A. N., and Rudebusch, G. D. 1997. Opportunistic and Deliberate Disinflation Under Imperfect Credibility. Federal Reserve Bank of San Francisco Working Paper 97–07.

Bomfim, A. N., Tetlow, R., von zur Muehlen, P., and Williams, J. 1997. Expectations, Learning, and the Costs of Disinflation: Experiments Using the FRB/US Model. Federal Reserve Board, Finance and Economics Discussion Series Paper 1997–42.

Cook, T. 1989. Determinants of the Federal Funds Rate: 1979–1982. Federal Reserve Bank of Richmond, *Economic Quarterly* 75(1), pp. 3–19.

Cukierman, A., and Meltzer, A. H. 1986. A Theory of Ambiguity, Credibility, and Inflation Under Asymmetric Information. *Econometrica* 54, pp. 1099–1128.

Fisher, S. 1986. Contracts, Credibility, and Disinflation. In *Indexing, Inflation, and Economic Policy,* ed. S. Fisher. Cambridge, MA: MIT Press.

Flood, R. P., and Garber, P. M. 1980. An Economic Theory of Monetary Reform. *Journal of Political Economy* 88, pp. 24–58.

Friedman, B. M. 1979. Optimal Expectations and the Extreme Information Assumptions of Rational Expectations Models. *Journal of Monetary Economics* 5, pp. 23–41.

Friedman, M. 1984. Lessons from the 1979–82 Monetary Policy Experiment. *American Economic Association, Papers and Proceedings* 74, pp. 397–400.

Fuhrer, J. C. 1996. Monetary Policy Shifts and Long-Term Interest Rates. *Quarterly Journal of Economics* 111, pp. 1183–1209.

Fuhrer, J. C. 1997. The (Un)Importance of Forward-Looking Behavior in Price Specifications. *Journal of Money Credit and Banking* 29, pp. 338–350.

Fuhrer, J. C., and Hooker, M. A. 1993. Learning About Monetary Regime Shifts in an Overlapping Wage Contract Model. *Journal of Economic Dynamics and Control* 17, pp. 531–553.

Fuhrer, J. C., and Moore, G. 1995a. Inflation Persistence. *Quarterly Journal of Economics* 110, pp. 127–160.

Fuhrer, J. C., and Moore, G. 1995b. Monetary Policy Trade-offs and the Correlation between Nominal Interest Rates and Real Output. *American Economic Review* 85, pp. 219–239.

Gagnon, J. H. 1996. Long Memory in Inflation Expectations: Evidence from International Financial Markets. Federal Reserve Board, International Finance Discussion Paper 538.

Gagnon, J. H. 1997. Inflation Regimes and Inflation Expectations. Federal Reserve Board, International Finance Discussion Paper 581.

Goodfriend, M. 1993. Interest Rate Policy and the Inflation Scare Problem: 1979–1992. Federal Reserve Bank of Richmond *Economic Quarterly* 79(1), pp. 1–24.

Goodfriend, M. 1997. Monetary Policy Comes of Age: A 20th Century Odyssey. Federal Reserve Bank of Richmond *Economic Quarterly* 83(1), pp. 1–22.

Hardouvelis, G. A., and Barnhart, S.W. 1989. The Evolution of Federal Reserve Credibility: 1978–1984. *Review of Economics and Statistics* 71, pp. 385–393.

Hetzel, R. L. 1986. Monetary Policy in the Early 1980s. Federal Reserve Bank of Richmond *Economic Review* 72(2), pp. 20–32.

Huh, C. G., and Lansing, K. J. 1998. Expectations, Credibility, and Disinflation in a Small Macroeconomic Model. Federal Reserve Bank of San Francisco Working Paper 98–01.

Ireland, P. N. 1995. Optimal Disinflationary Paths. *Journal of Economic Dynamics and Control* 19, pp. 1429–1448.

Judd, J. P., and Rudebusch, G. D. 1998. Taylor's Rule and the Fed: 1970–1997. Federal Reserve Bank of San Francisco *Economic Review,* vol. 0, no. 3, pp. 3–16.

King, M. 1996. How Should Central Banks Reduce Inflation?—Conceptual Issues. In *Achieving Price Stability,* symposium sponsored by the Federal Reserve Bank of Kansas City.

King, R. G., and Watson, M. W. 1994. The Post-War U.S. Phillips Curve: A Revisionist Econometric History. *Carnegie-Rochester Conference Series on Public Policy* 41, pp. 157–219.

Lewis, K. K. 1989. Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange. *American Economic Review* 79, pp. 621–636.

Lucas, R. E., Jr. 1976. Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy* 1, pp. 19–46.

Mankiw, N. G. 1994. Discussion of "Optimal Monetary Policy and the Sacrifice Ratio," in *Goals, Guidelines, and Constraints Facing Monetary Policymakers,* proceedings of a conference sponsored by the Federal Reserve Bank of Boston.

Marcet, A., and Sargent, T. J. 1989. Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models. *Journal of Economic Theory* 48, pp. 337–368.

Meyer, L. H., and Webster, C., Jr. 1982. Monetary Policy and Rational Expectations: A Comparison of Least Squares and Bayesian Learning. *Carnegie-Rochester Conference Series on Public Policy* 17, pp. 67–98.

Orphanides, A., Small, D. H., Wieland, V., and Wilcox, D. W. 1997. A Quantitative Exploration of the Opportunistic Approach to Disinflation. Federal Reserve Board, Finance and Economics Discussion Series Paper 1997–36.

Roberts, J. 1997. Is Inflation Sticky? *Journal of Monetary Economics* 39, pp. 173–196.

Romer, C. D., and Romer, D. H. 1989. Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. In *NBER Macroeconomics Annual* 4, eds. O.J. Blanchard and S. Fischer. Cambridge, MA: MIT Press.

Romer, C. D., and Romer, D. H. 1994. Monetary Policy Matters. *Journal of Monetary Economics* 34, pp. 75–88.

Romer, C. D., and Romer, D. H. 1997. Institutions for Monetary Stability. In *Reducing Inflation: Motivation and Strategy,* eds. C. Romer and D. Romer. Chicago: University of Chicago Press.

Ruge-Murcia, F. J. 1995. Credibility and Changes in Policy Regime. *Journal of Political Economy* 103, pp. 176–208.

Sargent, T. J. 1982. The Ends of Four Big Inflations. In *Inflation: Causes and Effects,* ed. R.E. Hall. Chicago: University of Chicago Press.

Sargent, T. J. 1983. Stopping Moderate Inflations: The Methods of Poincaré and Thatcher. In *Inflation, Debt, and Indexation,* eds. R. Dornbusch and M. H. Simonsen. Cambridge, MA: MIT Press.

Sargent, T. J. 1986. Reaganomics and Credibility. In *Rational Expectations and Inflation.* New York: Harper Collins Publishers.

Sargent, T. J. 1999. *The Conquest of American Inflation.* Princeton, NJ: Princeton University Press.

Shapiro, M. D. 1994. Federal Reserve Policy: Cause and Effect. In *Monetary Policy,* ed. N.G. Mankiw. Chicago: University of Chicago Press.

Svensson, L..E. O. 1997. Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets. *European Economic Review* 41, pp. 1111–1146.

Taylor, J. B. 1975. Monetary Policy During a Transition to Rational Expectations. *Journal of Political Economy* 83, pp. 1009–1021.

Taylor, J. B. 1980. Aggregate Dynamics and Staggered Contracts. *Journal of Political Economy* 88, pp. 1–23.

Taylor, J. B. 1982. Establishing Credibility: A Rational Expectations Viewpoint. *American Economic Association, Papers and Proceedings* 74, pp. 207–210.

Taylor, J. B. 1993. Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, pp. 195–214.

Taylor, J. B., 1998. An Historical Analysis of Monetary Policy Rules. Stanford University, unpublished manuscript.

Volcker, P. A. 1979. "Statement" before the Joint Economic Committee, *Federal Reserve Bulletin* 65 (November) pp. 888–890.

Walsh, C. E. 1988. Testing for Real Effects of Monetary Policy Regime Shifts. *Journal of Money Credit and Banking* 20, pp. 393–401.

This Page Intentionally Left Blank

42

Inflation, Asset Markets, and Economic Stabilization
Lessons from Asia

Lynn Elaine Browne
Senior Vice President and Director of Research, Federal Reserve Bank of Boston, Boston, Massachusetts

Rebecca Hellerstein
Research Associate, Federal Reserve Bank of Boston, Boston, Massachusetts

Jane Sneddon Little
Assistant Vice President and Economist, Federal Reserve Bank of Boston, Boston, Massachusetts

In the 1980s, a new convention emerged in the economics profession—that central banks' primary, or even single, responsibility should be controlling consumer price inflation. Such a view, in various forms, has been around for a long time.[1] But the idea gained broader acceptance as the dramatic rise in world inflation during the 1970s, coupled with a broad deterioration in economic performance, seemed to highlight the pernicious effects of rapidly rising prices, just when various theoreticians were arguing that monetary policy was prone to an inflationary bias. Moreover, while aggressive efforts to combat inflation by the Volcker Fed and central bankers in Japan and Germany led to recession in the early 1980s, the output costs of disinflation were not as severe as many had feared and the subsequent recovery and expansion in both the United States and Japan reinforced the notion that if only consumer price inflation could be brought under control, economic growth would take care of itself.

By the 1990s, this view was gaining credibility within policy circles, and various countries mandated that their central banks make inflation their primary, if not their sole, focus—although these mandates usually contain an escape clause for an economic shock that affects employment and growth severely.[2] New Zealand and Canada were two leaders in this regard. Meeting a low inflation target also became a major criterion for joining the European Monetary Union, and ensuring price stability is now the main goal of the European Central Bank. Moreover, while the concept of inflation targeting originated in the industrial world, governments in many developing countries now adopt annual inflation

---

Reprinted from: *New England Economic Review*, Federal Reserve Bank of Boston (Sept./Oct. 1998) 3–32.

targets, and achieving low inflation is widely perceived as a sign of "success" under IMF surveillance and in international capital markets.[3]

Here in the United States, this orthodoxy has never gained official status, although legislation advocating an inflation target was proposed.[4] Rather, the U.S. policy goal remains promoting stable long-term growth, and the policy approach remains eclectic. (Of course, even in countries with an inflation target, like Britain and Germany, policymakers admittedly watch and react to a variety of indicators as practicality requires.) Nonetheless, the strong performance of the United States during the mid 1990s, with inflation declining even as the unemployment rate fell below most estimates of full employment, seemed to reinforce the case for a focus on price stability.

That Japan had experienced a severe recession early in this decade that was not preceded by a significant rise in inflation was not generally seen as a challenge to this view. Although the country had experienced an asset price "bubble" to which monetary policy mistakes may have contributed, Japan's ongoing problems were widely laid at the door of a directed model of industrial development that had outlived its usefulness; an unwillingness to deal decisively with its banks' bad loan problems; and more recently, an overly stringent fiscal policy as the economy struggled to recover.

Now, however, world policymakers are dealing with a currency and financial crisis in East Asia that has produced serious recessions in South Korea, Indonesia, Thailand, and Malaysia; threatens their neighbors; and is adversely affecting trading partners and other nations seemingly far removed.[5] Yet the Asian countries had not previously experienced any pronounced acceleration in consumer price inflation; nor had they suffered the deterioration in their fiscal position or other economic "fundamentals" most commonly viewed as forerunners of financial crisis.[6]

The recent problems in East Asia, as well as the earlier one in Japan, raise the question of whether such a concentrated focus on consumer price inflation has become tunnel vision, dulling the sensitivity of policymakers and market participants to other signs of overheating. This article argues that the focus on consumer price inflation may indeed have been too narrow. Drawing upon the crises in Japan and other Asian countries, with reference to comparable episodes in the United States, it suggests that a preoccupation with inflation may have lulled policymakers and investors, both domestic and foreign, into ignoring useful signals from stock, real estate, and currency markets and from emerging imbalances in the real economy. Whether such imbalances would have been better addressed by monetary policy or by improved disclosure, supervisory intervention, or tax policy, a broader perspective might have identified problems in Asia at an earlier stage, before they assumed such crippling proportions. As it was, widespread recognition of Asia's relatively good inflation and fiscal performance may have provided unwarranted comfort.

The first section of this article discusses the "normal" relationship between inflation and economic activity seen during much of the postwar period in the United States, while the second section contrasts this familiar background with the recent experience of Japan and the "Asian Tigers" currently in crisis. The third section explores developments in Asian asset markets in more detail. The discussion focuses on Japan and South Korea, the largest of the crisis-ridden countries, with Japan illustrating the perils of rapid increases in asset prices and South Korea the more subtle dangers of excessive investment. The text also highlights the role of the shift from directed to (partially) liberalized financial markets, both in contributing to these excesses and in exposing them to the market's unforgiving judgment. Section four examines distortions in the real economy and suggests that very high levels of investment spending and an imbalance between employment growth in "nontradables" industries like construction and "tradables" industries like manufacturing may signal future problems even in the absence of consumer price inflation. Section five describes how the global context can help to reconcile divergent trends in asset and consumer goods markets. The sixth section concludes by suggesting that policymakers may want to look out for signs of overheating emanating from asset markets and from emerging imbalances in the real economy, even when consumer prices are well behaved. Signs that high levels of debt may be financing increasingly optimistic investments warrant particular concern. This final section also stresses the vulnerabilities that newly liberalized financial markets may introduce and the importance of measures that encourage the private sector to price risk more accurately and force it to bear the costs of international financial crises more fully. Overall, it advocates an eclectic approach to assessing economic performance rather than a focus on any single indicator.

INFLATION, SPECULATION, AND THE BUSINESS CYCLE

The rationale for central banks' focus on inflation[7] has several dimensions. Not least is the strong public aversion to inflation found in many countries. To some degree, this aversion arises because high rates of inflation are frequently associated with political instability. Extreme inflation, in particular, often occurs in the context of social upheaval and political turmoil. The German hyperinflation of the 1920s is a classic instance and is generally thought to explain why the German public remains adamant that its policymakers not let inflation rise above very modest rates.

A government that allows inflation to rise to very high levels is usually ineffective in many respects; so determining whether high rates of inflation are a cause of poor economic performance or a symptom of other problems that impinge on economic activity can be difficult. Nevertheless, numerous researchers have documented a link between high rates of inflation and poor economic outcomes. The picture is less clear at more moderate rates of inflation (Bruno 1995; Ball 1994) but plausible arguments can be advanced that even relatively low rates of inflation distort economic decisionmaking. Thus, monetary policymakers commonly justify their concern with restraining inflation in terms of the beneficial impact on economic output over the long term. They also point to the role of low inflation expectations and enhanced central bank credibility in strengthening the policymakers' hand.

But for many, the most compelling argument for vigilance against inflation is the economic pain of the subsequent disinflation. In the United States and most other countries throughout the post–World War II era, rising inflation has preceded most recessions. As can be seen in Figure 1, quickening inflation in the United States in the late 1960s was followed by the recession of 1970–71; the much sharper rise in the early 1970s was followed by the deep recession of 1973–75, and the still steeper acceleration in inflation in the second half of the decade was followed by two recessions in the early 1980s, the second of which was very severe. Inflation picked up more gradually in the late 1980s, and the relatively mild recession of 1990 to 1991 ensued. Germany and the United Kingdom show similar patterns, as can be seen in Appendix Figures 1a and 1b.

To say that the acceleration in inflation "caused" the recessions would not be accurate. Rather, rising inflation led to tighter monetary policy, and the ensuing downturn brought about a slowing in inflation. In most cases, an unexpected shock also played an important role in bringing on the recession—in the United States, a major auto strike in 1970,

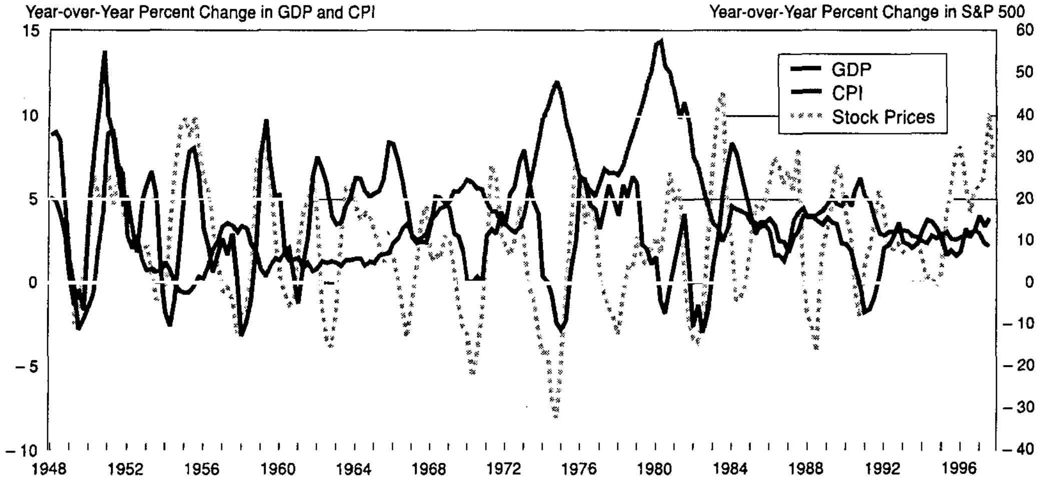


Figure 1 U.S. CPI, GDP, and stock prices. Source: U.S. Bureau of Economic Analysis.

the confrontation with Iraq in 1990, and most dramatic of all, the spikes in oil prices in 1973 to 1974 and again in 1979. The last involved oil shortages that not only drove inflation to yet higher levels but also disrupted economic activity, contributing directly to the downturn.[8]

The lesson many have taken from this experience is that central banks should pay close heed to inflation and act promptly to forestall increases. By so doing, they can avoid the need for more aggressive action later and spare the economy the severe fluctuations suffered in the past. Even those who might otherwise have tolerated somewhat higher rates of inflation or been tempted to exploit a possible short-term trade-off between inflation and unemployment seem to have accepted the argument that resisting inflation in the present helps avert economic disruptions in the future.

The present-day focus on inflation contrasts with much of the past thinking about business cycles, which gave considerably greater emphasis to the role of investment and speculation in the prices of stocks, real estate, and other assets.[9] Thus, Irving Fisher, in the wake of the Great Depression, placed a good deal of the blame on an excess of debt. Such excesses commonly followed an earlier period when technological, regulatory, or financial innovations improved the environment for investment. The initial period of rational growth in investment then spiraled into a "euphoric" phase of overinvestment and overborrowing. Businesses and investors, in turn, sought to grow their way out of their heavy debt burdens through speculation, often worsening their financial positions in the process. Eventually, some event, perhaps a rise in interest rates or an increase in credit rationing, would lead to a retrenchment in which financial intermediaries called in loans, forcing borrowers with insufficient liquidity to go bankrupt. The subsequent deflation, in both asset and goods prices, further aggravated the circumstances of debtors and their creditors and intensified the downturn.

Figure 2 shows the sharp run-up in stock prices that characterized the years leading to the Great Depression in the United States. Stock prices rose throughout the 1920s while, by contrast, consumer and wholesale goods prices drifted down slightly. The bull market gained momentum as the decade wore on, soaring in the year prior to October 1929. The rapid rise in stock prices was followed by an even more precipitous decline; within three years, stock prices had fallen below the level of 10 years earlier. Goods prices also fell.

Unfortunately, speculative bubbles are clearly evident only after the fact. While a marked deviation from past experience and widespread investor focus on capital gains without regard to the strength of underlying earnings may suggest to the cautious observer that a speculative bubble is under way, during the bubble eminent scholars and financiers usually emerge with plausible explanations for why circumstances have changed and why higher asset prices are justified. Fisher himself defended the valuations of financial assets in October 1929![10]

Speculation in stocks does not appear to have been a major factor in the recessions of the post-World War II era, however. Stock prices have been quite cyclical, typically soaring at the start of recoveries, rising more moderately as expansions age, and then falling in the year immediately preceding a recession (Figure 1). Thereafter, stock prices have tended to recover quite rapidly, suggesting that valuations were not unreasonably optimistic. The most striking exceptions to this pattern occurred in the 1970s. Stock prices languished throughout this high-inflation period; prices only began to rise on a consistent basis as inflation came down in the 1980s.

The mid 1980s also represent a modest departure from the normal pattern. Following the recessions of the early 1980s and a huge dollar appreciation, inflation fell sharply from prior levels, and U.S. stock prices soared. However, earnings growth, especially for manufacturers, sputtered; thus, price-to-earnings ratios reached a 25-year high early in 1987. With overall economic activity accelerating and the dollar now retracing its previous climb, inflation threatened to pick up, leading the Fed to raise interest rates in the spring. When weak trade data and a further weakening of the dollar threatened additional rate increases, stock prices plunged over 20 percent in the fourth quarter. The Fed provided liquidity, and the economic expansion continued. Earnings growth improved, and by mid 1989, stock

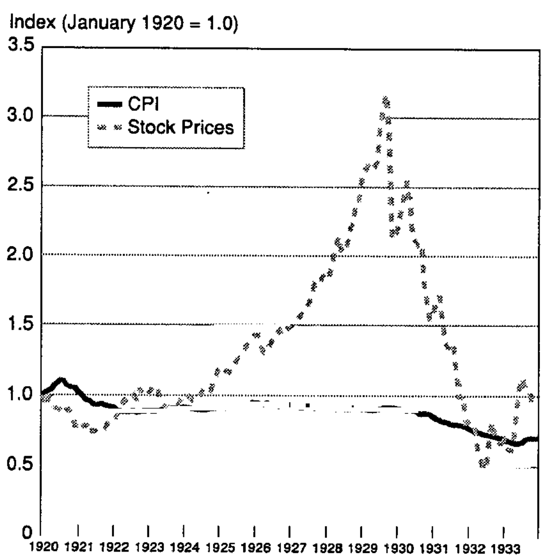


Figure 2 U.S. consumer price level and stock price index (S&P 500 Index) monthly data, 1920–1933. Source: NBER Macrohistory Database.

prices had returned to their prior level with price-earnings ratios in a more normal range. Whether or not the earlier run-up had a speculative element, it was short-circuited; and neither the run-up nor the subsequent decline in stock prices had serious consequences for overall economic activity.

Real estate is another potential vehicle for speculation, but real estate prices in the United States have followed a rather different pattern than have prices of stocks. For most of the post-World War II era, real estate values have risen faster than inflation. In contrast to stocks, real land prices rose especially fast during the high inflation years of the 1970s.[11] As inflation subsided, land prices followed an erratic path, with collapsing land and real estate values implicated in several important regional financial crises. For instance, the price of farm land, which doubled in value in the second half of the 1970s, fell by 30 percent in the mid 1980s, creating severe difficulties for farmers in the Midwest and the Farm Credit banks that served them. Prices did not regain previous peaks for another 10 years.

Speculation in real estate also played an important role in the savings and loan debacle in the Southwest in the middle of the 1980s and in New England's banking crisis at the end of the decade. In both cases, and consistent with Fisher's theories, an earlier period of financial deregulation and prosperity led to increased competition in financial services and increased investment in real estate. Eventually, rising prices encouraged speculation, leading to yet higher prices, and to construction outstripping demand. When economic growth finally began to slow, in response to declining oil prices in the Southwest and a series of shocks in New England, real estate values collapsed. Highly leveraged developers were unable to meet their obligations to the local banks that had provided the bulk of the financing for these projects. Many institutions failed. While the 1990 recession was relatively mild for the country as a whole, New England experienced a 10 percent decline in employment, in large part because of the real estate and banking debacle. The Southwest also experienced a prolonged period of sluggish growth. (See Browne 1992.)

In sum, while speculation in asset markets has at times contributed to economic fluctuations in the United States, particularly in the 1980s and early in the century, most of the recessions of the past 50 years have been preceded by rising inflationary pressures. The U.S. postwar pattern has come to be viewed as the norm, leading some central bankers and many investors to discount signals from asset markets and to look primarily to consumer price inflation for evidence of overheating. (See Box 1.)

Box 1 Asset Prices and Central Banks

Arguments that central banks should pay attention to unusual increases in equity, real estate, and other asset prices per se (that is, beyond considering the impact of such swings on consumer prices through wealth and credit channel effects) commonly elicit the question: How can the central bank know better than the market the appropriate value of assets?

But the benefit of paying attention to unusually rapid or prolonged increases in asset prices does not hinge on the central bank being a more astute judge of fundamental asset values than the market—any more than the central bank is asked to judge the appropriate *level* of consumer prices. Rather, the need for the central bank to consider the implications of unusual *increases* in asset prices stems from its systemic responsibilities, a focus that may well produce assessments that differ from those appropriate for private market participants. In addition, asset price inflation may be a symptom of an overheated economy. Since consumer prices are only a subset of all transactions prices, under some circumstances, excess liquidity may find other outlets. When that outlet is rapid asset price inflation, the resulting perverse incentives may create economic imbalances that may damage the real economy. Lastly, unusually rapid and prolonged increases in asset prices pose many of the same issues for individual well-being as does consumer price inflation.

Gains in asset prices that do not seem solidly supported by earnings prospects may warrant central bank scrutiny because of the potential for a subsequent destabilizing price decline (although central banks may be able to limit the spillovers by providing liquidity, as the Fed did in 1987). For example, rapid increases in stock and real estate prices are often supported by rapid increases in bank and other forms of credit. Thus, unusual surges in asset prices may provide a warning that lenders' portfolios and underwriting standards deserve more intense scrutiny. When asset prices are rising, the most successful investors are those who are most concentrated in the assets with the greatest gains. Competition to do business with these investors can lead lenders to become dependent on these same assets and to reduce their lending standards, since successful customers can readily take their business elsewhere if they feel loan conditions are onerous.

But unfortunately, from time to time, these risky concentrations culminate in large declines in asset prices that disrupt banking and other financial markets and contribute to serious economic downturns. To the degree that central banks see themselves as having some responsibility for financial and economic stability, they will want to respond to these developments and have a legitimate interest in forestalling them. After all, if central bankers feel compelled to offset the systemic/real impacts of a significant decline in asset prices, they may validate investor mistakes. And if investors assume that the central bank will act to offset the fallout from a major market correction, an element of "moral hazard" may creep into the equation. The appropriate preemptive response in these circumstances is likely to vary from case to case.

The main point here is that individual investors care only about their own risk exposure, but the central bank must consider the possibility that rapidly rising asset prices will have negative repercussions for the economy as a whole. Thus, the central bank's evaluation of the risks linked to various asset price paths must include the public costs and will differ from the appropriate calculation for private market participants. But admittedly, if removing the punch bowl seems hard when the concern is accelerating consumer prices, the difficulty will be even greater in the case of overheated asset prices.

Further, many of the arguments in favor of central banks' focus on consumer price inflation also apply to increases in asset prices. Central banks generally choose to target consumer rather than producer or other input prices because their ultimate concern lies in promoting consumer welfare, broadly measured, now and in the future. But fluctuations in asset prices, which affect the value of society's accumulated wealth, also matter to people's well-being, as much perhaps as changes in consumer prices. Indeed, as economic development proceeds and the ability to accumulate wealth spreads, as age spans lengthen and as individuals are increasingly asked to provide for their own retirements, avoiding large contractions in the value of wealth may gain greater importance. And just as consumer price inflation can distort economic decisions and make planning for the future difficult, so too can volatile asset prices. All told, the unfortunate consequences of "excessive" asset price inflation have a lot in common with the consequences of "excessive" consumer price inflation.

GROWTH AND INFLATION IN JAPAN, SOUTHEAST ASIA, AND SOUTH KOREA: THE RECENT HISTORY

This section traces the chronology of inflation and real GDP growth in troubled Asia since the early 1970s. As this review demonstrates, contrary to the postwar "norm," a marked pickup in inflation did not herald recent downturns in Asia.

Japan

Financial and economic difficulties in Japan through most of the 1990s and recent developments in South Korea and other East Asian countries caught much of the world by surprise. During the 1980s, Japan's economic performance was the object of world envy. It enjoyed the fastest growth of all the G-7 industrial countries,[12] with growth especially strong late in the decade, despite a yen appreciation and the maturity of the expansion. Inflation also remained muted throughout the 1980s, despite Japan's poor record in this regard in the previous decade. As Table 1 shows, from 1970 to 1979, Japan's annual inflation averaged 9 percent, slightly above the norm for the G-7; by contrast, the Japanese CPI rose less than 3 percent a year on average in the 1980s. While inflation did edge up at the end of the decade, from roughly 0 in 1987 to 3.3 percent in 1991, such a rate was still very low by historical standards and below rates in most of the G-7.[13] Japan's econ-

Table 1 Annual Change Percent in Real GDP and Consumer Prices

	Real GDP			Consumer prices		
	1970–1979	1980–1989	1990–1997	1970–1979	1980–1989	1990–1997
G-7 Countries[a]	3.6	2.8	2.0	8.3	5.5	3.0
United States	3.5	2.8	2.2	7.2	5.5	3.3
Japan	4.6	3.8	2.1	9.3	2.5	1.5
South Korea	8.8	7.9	7.5	15.1	6.2	6.1

[a] The G-7 include the United States, Japan, Germany, France, Italy, the United Kingdom, and Canada.
Source: *OECD Economic Outlook*.

omy slowed sharply in 1992 and has languished ever since, with growth in real GDP averaging only 1 percent per year. With fiscal stringency in 1996 and 1997 and the crisis elsewhere in Asia, Japan is now in the midst of a sharp downturn. Underlying inflation is roughly zero.

South Korea

South Korea was commonly thought to be following the path traced by Japan, before the latter's recent problems. As Table 1 shows, Korea enjoyed rapid growth in real GDP, averaging 8 to 9 percent, from 1970 through 1997. Over the same years, inflation was higher in Korea than Japan and fairly volatile, starting as high as 15 percent in the 1970s and falling to 3 percent for part of the 1980s. From 1990 to 1997, South Korea's inflation averaged 6 percent—well below the standard for developing countries and just slightly above the OECD average; it fell below 5 percent in 1995 and 1996 and below 4 percent in early 1997.

South Korea's strong economic performance was acknowledged in the fall of 1996 when it was invited to join the Organisation for Economic Cooperation and Development (OECD), a group of developed countries. OECD assessments of the Korean economy around that time indicated that difficulties in the export sector and reduced government spending would result in a slightly slower economic growth and some deceleration in inflation in 1997. Korea was also seen to face some risk from increased exposure to world capital markets and from the impact of business failures on the banking system. Nothing suggested a decline of the magnitude that now seems likely, however.

Southeast Asia

The economic situation in the other Asian Tigers during the 1990s was roughly similar to South Korea's. Although circumstances varied across countries, some generalizations apply. As can be seen in Table 2, in all of these countries, growth was strong, inflation was reasonably low (despite a rapid expansion of bank credit), and fiscal balances were generally prudent. In other words, most traditional measures of "fundamental" economic health looked good.

To be sure, current account balances were deteriorating. But many observers interpreted this trend as a market response to the investment opportunities available in rapidly developing countries. And foreign investors seemed eager to supply these countries with large amounts of capital, although much of it was short term and denominated in dollars. As Table 2 shows, Bank for International Settlements (BIS) data revealed high ratios of loans from foreign banks relative to GDP in most of these countries in the mid 1990s.

As analysts around the world look for clues as to what went wrong in Asia, attention is turning to financial asset markets.[14] This current focus on asset markets has two strands—one emphasizes the overinvestment that drove asset values to unrealistic levels; the other emphasizes "directed" or "relationship" lending that channeled resources to unproductive ends.[15] Both blame lax banking supervision and inadequate disclosure requirements for allowing untenable positions to build. The next section looks at Asian asset markets in more detail.

Table 2 Economic Indicators, Selected Asian Countries

	Real GDP (percent change)				Consumer prices (percent change)				Government balance[a] (percent of GDP)				Current account (percent of GDP)				Foreign bank claims[b] (percent of GDP)		
	1994	1995	1996	1997	1994	1995	1996	1997	1994	1995	1996	1997	1994	1995	1996	1997	End '95	End '96	End '97
All Developing Countries	6.8	6.0	6.6	5.8	51.4	22.7	13.7	8.5	–2.8	–2.6	–2.4	–2.2							
Indonesia	7.5	8.2	8.0	5.0	8.5	9.4	7.9	6.6	2.0	3.0	2.3	1.4	–1.7	–3.4	–3.5	–2.2	24.8	25.6	39.1
Malaysia	9.2	9.5	8.6	7.8	3.7	3.5	3.5	2.7	2.3	.9	1.5	3.3	–6.3	–8.5	–5.2	–3.3	21.8	26.1	28.0
Philippines	4.4	4.8	5.7	5.1	9.1	8.1	8.4	5.0	1.1	.5	.3	.2	–4.6	–4.4	–4.3	–4.8	11.1	16.2	20.1
Thailand	8.6	8.8	5.5	–.4	5.1	5.8	5.8	5.6	2.8	3.2	.7	–.6	–5.6	–8.0	–7.9	–3.0	55.4	53.7	49.5
South Korea	8.6	8.9	7.1	5.5	6.3	4.5	4.9	4.5	.6	.5	.0	–.1	–1.2	–2.0	–4.9	–1.8	18.3	22.5	22.2
Taiwan	6.5	6.0	5.7	6.8	4.1	3.7	3.1	.9	–5.7	–7.4	–8.0	–7.6	2.6	1.9	3.8	4.5	8.5	8.4	7.5
Hong Kong	5.4	3.9	5.0	5.3	8.2	8.6	6.0	5.7	1.7	–1.3	1.5	5.2	1.6	–3.2	–.7	–6.0	365.8	290.2	287.4
Singapore	10.5	8.7	6.9	7.8	3.1	1.7	1.4	2.0	12.5	11.5	10.7	10.8	17.0	16.9	15.0	13.7	331.4	290.2	273.0

[a] Central government fiscal balances.

[b] Claims of BIS-reporting banks. Large ratios for Hong Kong and Singapore reflect their role as major financial centers.

Source: Data for all developing countries from the IMF; individual countries' data for real GDP, consumer prices, government balances, and current account from Standard & Poor's DRI, and foreign bank claims data from the Bank for International Settlements (BIS).

ASSET VALUES AND INVESTMENT IN ASIA

Japan

In Japan, the decade of the 1980s was characterized by rapidly rising asset prices, including, after 1985, a strongly appreciating yen. By contrast, consumer prices increased very modestly, despite a declining unemployment rate and an accommodative monetary policy as measured by the growth in M2 + CDs and in bank credit. As Appendix Table 1a shows, these aggregates grew faster than nominal GDP by substantial margins in the second half of the 1980s, as interest rates fell to levels that were among the lowest in the OECD.

Between 1981 and 1989, the value of the Nikkei, the main index of stock prices in Japan, increased fourfold. Price-earnings ratios jumped from 20 in the early 1980s to more than 30 a few years later, with essentially unchanged earnings (see Figure 3). Earnings then picked up, and P/E ratios soared, surpassing 60 in 1987, as low domestic interest rates helped to support high-flying equity prices. Over the next two years, accelerating earnings further buoyed the Nikkei but let P/E ratios moderate. Tighter monetary policy starting in 1989 and flattening earnings in 1990 led to a sharp decline in stock prices. A further drop in earnings in the 1992 economic slowdown sent the Nikkei plunging. Earnings have remained depressed ever since, at less than half the level in the peak year of 1989 and below the levels of the early 1980s.[16]

The fluctuations in Japan's stock market have been mirrored by developments in real estate. Real estate values soared in the late 1980s, particularly in Japan's largest cities, only to plunge in the early 1990s.[17] Land prices also rose and then fell in Japan's smaller cities, but the swings have been more modest.

These changes in asset values were linked to shifts in real economic activity. A surge in business investment accompanied the bullish stock market as soaring equity prices and

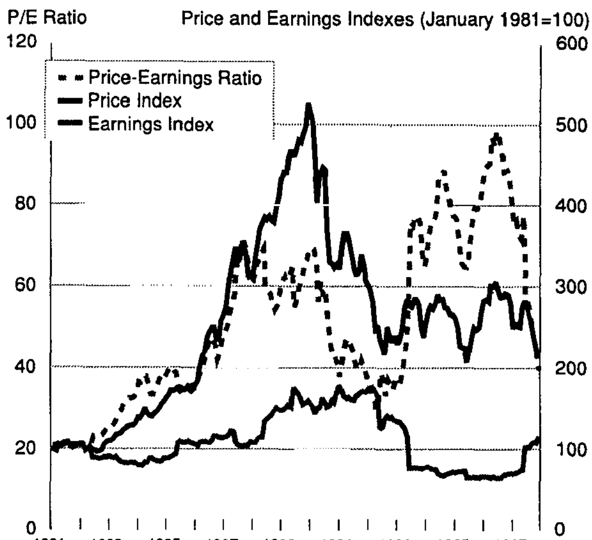


Figure 3 Price-earnings ratio and price and earnings indexes for Japanese stock market; January 1981 to December 1997 (Arithmetic Average of Stocks in First Section of Tokyo Stock Exchange). Note: Earnings index was derived from the P/E series. Source: Nikkei Macroeconomic Statistics.

an accommodative monetary policy made the cost of capital relatively cheap. Business investment then collapsed in the early 1990s. Residential investment also grew very rapidly for a brief period, roughly coinciding with the sharpest escalation in land prices. However, in contrast to business investment, the increase was short-lived. (Table 3 highlights the timing of these changes.)

With the benefit of hindsight, during the "endaka" or "bubble" years, Japan fell victim to damaging speculation in asset prices. As Kähköonen (1995) points out, almost all empirical studies attempting to explain Japanese asset price movements since the mid '80s find evidence of a speculative bubble, with fundamentals unable to explain the sharp rise and subsequent fall of these prices. Kähkönen's own empirical work leads him to conclude, in the case of stock prices, that cyclical improvements in corporate profits should not have affected a rational investor's valuation of equities and, thus, that "the possibility of a bubble cannot be excluded." In the case of land prices, he finds that fundamentals, like real GDP growth and monetary policy, and changed expectations for the growth in rents (admittedly hard to measure), leave "the bulk" of the increase in land prices in the late 1980s "unexplained."

Whether investment was also "excessive" is more problematic. As noted above, except for a brief spurt, residential investment was not extraordinarily high by past standards, and the Japanese are certainly underhoused in comparison with U.S. residents.[18] In contrast, however, Japan's rate of business investment in the late 1980s probably should have been recognized as unsustainable. At the end of the decade, investment was increasing two to three times as fast as consumption (or consumption plus exports). Such a rate seems rapid, given the maturity of the Japanese expansion and the advanced level of development that Japan had achieved.

Moreover, between 1985 and 1988, the yen almost doubled in value vis-à-vis the dollar and other East Asian currencies tied to the dollar. More broadly, by 1988 the yen was

Table 3 Asset Price Changes (Percent) and Real GDP Growth in Japan

	Real GDP	Nikkei 225	Earnings[a]	Land Prices: Large Cities	Land Prices: Small Cities	Real Effective Exchange Rate
1985	4.3	18.8	+	8.4	2.4	-.5
1986	2.9	30.4	=	21.3	2.7	17.4
1987	4.1	41.5	+	35.1	9.4	2.9
1988	6.3	16.6	+	19.2	5.6	5.1
1989	4.8	26.0	++	27.4	8.9	-6.8
1990	5.2	-13.4	–	20.3	16.0	-10.8
1991	3.8	-17.6	+	-6.9	3.3	4.7
1992	1.0	-25.2	–	-19.4	-3.9	1.8
1993	.3	5.0	––	-15.1	-5.0	14.1
1994	.7	4.3	–	-11.0	-3.7	3.9
1995	1.4	-12.9	=	-13.8	-3.6	.8
1996	4.1	21.5	–	-9.7	-4.5	-14.7
1997	.9	-12.9	++	-5.8	-3.6	-5.0

[a] Calculated from stock price index and price-earnings ratios.
Source: Real GDP data from FAME database of the Board of Governors of the Federal Reserve System; Nikkei 225 price index and price-earnings ratios from Haver Analytics; land price data from the Japan Real Estate Institute; and real effective exchange rate data from J.P. Morgan.

almost 30 percent above its average for the 1980–85 period on an inflation-adjusted, trade-weighted basis. Thus, even as Japan launched a major investment program, the sharp appreciation of the yen was undermining the competitiveness of its products in world markets and leading to a decline in its share of world exports in the early 1990s. While these investments may have produced efficiencies that helped prevent an even greater loss of market share, they also contributed to excess capacity. Even now, by some estimates, about 30 percent of the productive capacity in Japan's housing, retailing, construction, and major appliance industries is unneeded; in autos the figure is roughly 20 percent (Fuji Research Institute 1998).

As in the U.S. real estate crises, financial liberalization also played a role in Japan's saga. Starting in 1980, Japan began phasing in measures to liberalize its financial markets so that, for the first time, nonfinancial firms had alternatives to domestic bank credit. As firms gained readier access to foreign bond and equity markets, the increased competition led to a fall in the cost of domestic financing, encouraging investment.[19] But increased competition also meant that Japanese banks faced narrower margins, particularly after 1985 when interest rates on most types of deposits were deregulated. The banks responded by expanding loans to small firms in real estate and construction and in other sectors with limited access to nonbank capital. Although these borrowers were seen as riskier than the banks' traditional customers, they were also potentially more profitable. And indeed, as real estate and stock prices rose, this strategy appeared successful. Demand for financing increased, but so too did the value of borrowers' collateral. Moreover, since the banks themselves were permitted to own stocks and real estate and to count unrealized capital gains as part of their capital base, the rise in asset prices directly increased this base, supporting greater lending.

Subsequent declines in land and stock values have put great stress on Japan's banking system and on the real economy. The net worth of many businesses has collapsed. Bankruptcies have soared. Banks have found themselves with increasing numbers of nonperforming real estate loans just as the other components of their capital base were becoming more precarious. The growth in bank credit, which had been very rapid in the 1980s, has been negligible in the 1990s; recently, outstanding bank credit has actually been falling. Japanese borrowers, particularly small and medium-sized firms, are facing a severe credit crunch.

Asset Values and Investment in Southeast Asia

Identifying asset price bubbles, always difficult, is especially so in developing countries, where data are limited and investment opportunities are both large and risky. Nonetheless, such data as do exist suggest that, during the years preceding the recent crisis, rapid increases in asset prices and optimistic levels of investment occurred in much of Southeast Asia—most clearly in real estate markets. Indeed, in its recent analysis of the Asian crisis, the International Monetary Fund (IMF) includes among the most important triggers the excessive investment in property and commercial real estate (and in certain industrial sectors, particularly in South Korea and Thailand), along with the resulting rise in asset prices. And again, while consumer price inflation remained moderate, bank credit grew substantially faster than nominal GDP as capital inflows into these pegged-currency countries soared (Appendix Table 1b).

Stock markets in Southeast Asia did not follow a common pattern in the years just before the recent crisis (Figure 4). Between 1995 and mid 1997, equity prices rose briskly

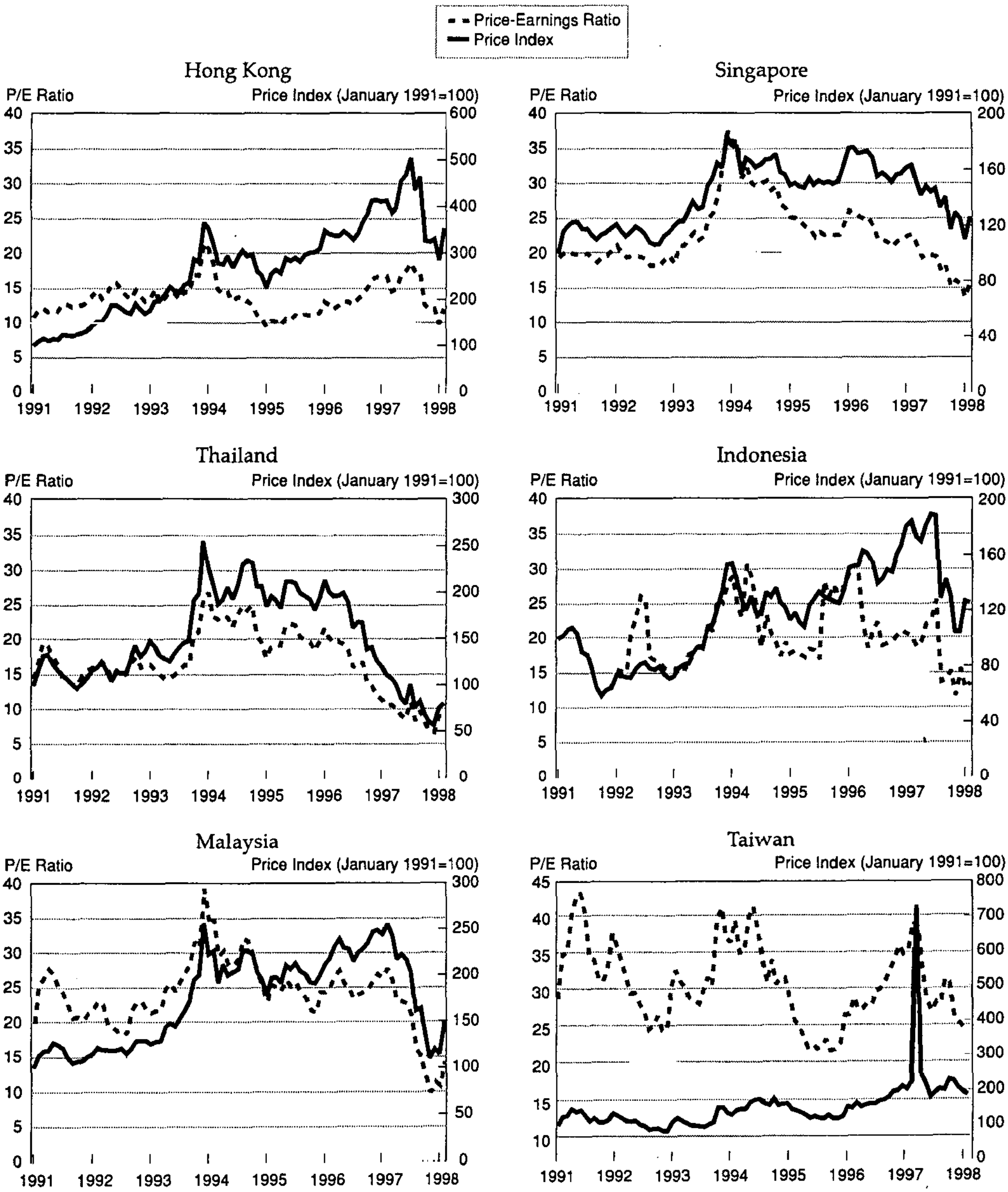


Figure 4 Price-earnings ratio and price index for selected Asian stock markets; January 1991 to February 1998. Source: Standard & Poor's DRI.

in Hong Kong and Indonesia and modestly in Taiwan and Malaysia; but stock prices exhibited no clear trend in Singapore and the Philippines and declined sharply in Thailand starting in 1996. While P/E ratios were elevated by U.S. standards—especially, as the BIS points out, given the comparatively high domestic interest rates prevailing in these countries and the much greater volatility of these markets,[20] P/Es were far below those reached in Japan and generally below these countries' previous peaks. And earnings, with the notable exception of South Korea, were rising. In other words, while P/Es may have been high given the risks faced by investors in developing countries, for the most part they were not obviously out of line.

In real estate markets, by contrast, the situation was more clearly ominous (Figure 5). In Singapore, industrial property prices rose over 250 percent between 1990 and late 1996, while residential property prices rose 300 percent. Construction outran demand,

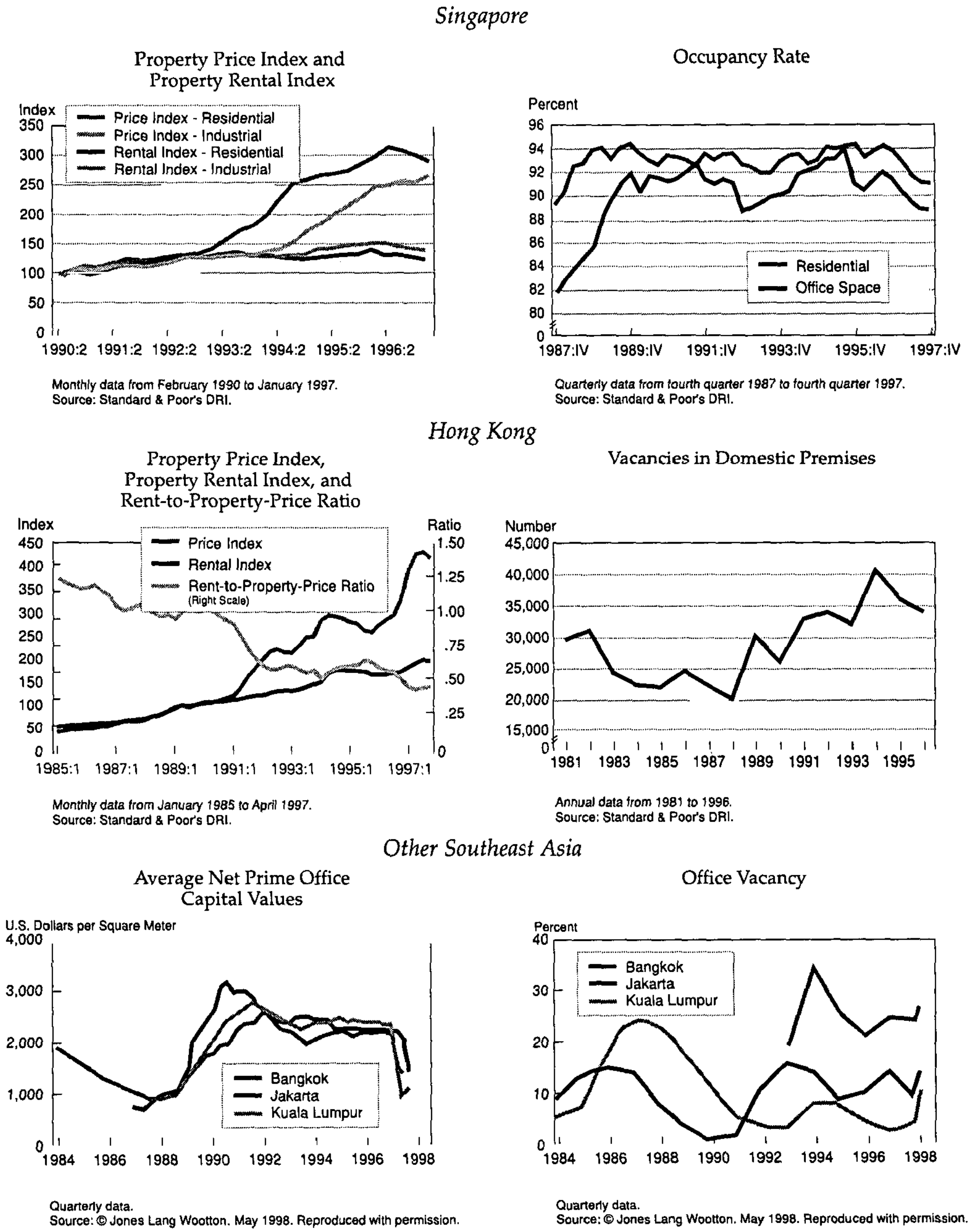


Figure 5 Real estate markets in Singapore, Hong Kong, and other Southeast Asian Countries.

and occupancy rates began to fall in 1995. In Hong Kong, property prices quadrupled between 1990 and late 1996. Rents rose much less rapidly and the rent-to-property price ratio (the equivalent of earnings-to-price (E/P) ratios for equity markets) fell steadily. In Thailand, where real estate data are limited, the collapse of a property company in early 1997 drew the world's attention to the thousands of empty condo units ringing Bangkok. This bankruptcy initiated a reassessment of the Tiger economies and precipitated the present crisis.

Since mid 1997, property prices have fallen in much of the region. In particular, office property prices have fallen 30 to 35 percent in Hong Kong and Bangkok and 10 percent in Singapore and Kuala Lumpur since March 1997 as office vacancy rates have risen in all four cities. The real estate collapse has put severe pressures on banking systems in the region.

Asset Values and Investment in South Korea

Signals that something might be amiss were especially subtle in South Korea, particularly as information on property values is limited. The Korean stock market shadowed the Japanese market during the 1980s, soaring almost eightfold in value (see Figures 3 and 6). However, at its 1980s peak, the average P/E ratio in Korea was just 15, suggesting that the rise in the market, spectacular as it was, may not have been fueled by expectations of future price increases but by the high earnings growth of a rapidly developing economy. In particular, Korea received a very strong boost from its external sector in the 1980s. Export growth was extremely rapid: Korea's current account swung from a deficit amounting to 8 percent of GDP in 1980 to a comparable surplus in 1988.

In 1989, however, Korea suffered a reversal as its relative unit labor costs jumped sharply and its competitive position took a turn for the worse. Exports (in won) declined, while imports grew strongly. The deterioration in exports was broad-based, involving sales to most developed countries and cutting across many product categories. The stock market plummeted as both earnings and P/Es fell. But the effect on overall growth was delayed,

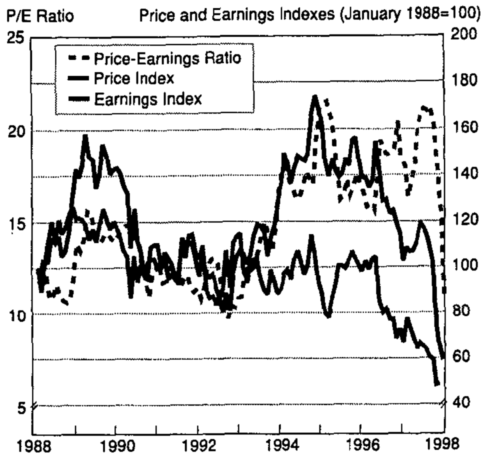


Figure 6 Price-earnings ratio and price and earnings indexes for Korean stock market; January 1988 to December 1997 (Arithmetic Average of All Stocks Listed). Source: Standard & Poor's DRI.

because the government started a massive construction program that doubled the share of GDP devoted to housing investment over the course of two years.

By 1994, the South Korean stock market had regained its earlier peak. Earnings, in contrast, remained well below their previous high. Thus, unlike the 1980s, the market's gains in the 1990s were based primarily on higher stock valuations, as P/Es fluctuated between 16 and 21 over the next three years. While low by Japanese standards, P/Es in this range exceed the U.S. experience in all but a few years. Expectations of strong future earnings might justify such values. But Korean corporations faced serious challenges.

Although total exports resumed growing in the 1990s, important categories of exports, including apparel, footwear, and some consumer electronics, did not. Korea faced intensified competition on several fronts. China was proving an increasingly virile global competitor, and from mid 1995, the appreciation of the U.S. dollar, and thus the dollar-linked won, vis-à-vis the yen placed Korea at a disadvantage versus Japan in some export markets. Mexico's late 1994 peso devaluation also posed a challenge, as did Taiwan's 1996 entry into semiconductor production, since semiconductors rank among Korea's most important exports. By 1996, thus, South Korea's current account deficit had reached 5 percent of GDP and was exerting a significant drag on the Korean economy. All in all, Korea's economic outlook was turning increasingly precarious. Yet the high levels of capital spending that had become a key component of GDP growth continued.

South Korea is often seen as following the Japanese model of economic development, in that saving and investment rates were very high and the government intervened actively in the economy to encourage the development of favored industries. During the 1980s, constant-dollar private nonresidential investment grew an average of almost 12 percent per year, compared to consumption and GDP growth of 8 percent and 9 percent respectively. In the 1990s, growth in nonresidential investment averaged 9 percent per year, versus 7 percent in consumption and GDP.

Such rapid growth in investment had boosted investment's share of GDP to over 35 percent by the mid 1990s. Rates of return to capital had fallen quite precipitously, however. According to the OECD, capital's share of income in Korea had dropped from over 50 percent in the 1970s, to 45 percent in the first half of the 1980s, to just over 30 percent in the mid 1990s. Given Korea's low level of development 20 years ago, it is not surprising that the return to capital was considerably higher then. But the rapid decline in capital's share of income implied that investment could not continue to grow faster than the rest of the economy without driving rates of return down to very low levels. With the external sector struggling, however, Korea needed strong investment to achieve the rates of GDP growth it had come to expect.

As in Japan, financial liberalization also played a role in Korea—with similar results. Interest rate deregulation was phased in, starting in late 1991.[21] Deposit rates rose, as lending rates fell, pinching bank earnings. In response, South Korean banks looked for new lending opportunities, particularly in real estate and construction.[22] Korea also began to liberalize its capital account. Historically, Korea had financed most of its investment domestically. In 1993, however, Korea began to relax restrictions on capital inflows, particularly on short-term bank debt. Limits on foreign ownership of Korean stock were eased, and Korean companies were permitted to borrow from foreign banks indirectly through authorized Korean banks. Thus, they were able to take advantage of the very low interest rates available on short-term dollar-denominated interbank loans. As foreign borrowing soared, Korea's highly leveraged firms and weak banking system grew more exposed to the discipline of international capital markets just as the viability of Korean assets turned increasingly suspect.[23] Previously, the ultimate judge of the value of Korea's investment projects had been the South Korean government, which could overlook disappointing returns in favor of social goals. Financial liberalization imposed the judgment of international capital markets—eventually.[24]

DISTORTIONS IN THE REAL ECONOMY—A CROSS-COUNTRY VIEW

That Japan, South Korea, and other East Asian countries engaged in excessive investment spending is obviously much clearer with hindsight than it was in the years before the crises. Even now, it must be recognized that financial crises create self-fulfilling prophecies in which projects that would have been viable in the absence of crisis fare no better than those that could never have generated a competitive return. (Conversely, loans based on inflated net worth and real estate collateral do not look weak until asset and currency prices collapse.) Nevertheless, the levels of investment spending in some of these countries were quite extraordinary.

Figures 7a through 7k show investment spending relative to GDP in Japan and seven other Asian countries and compare these levels with investment spending in the OECD. In these figures investment includes public and private, nonresidential and residential; data are nominal. As can be seen, in the 1990s, investment averaged some 40 percent of GDP in Thailand and Malaysia, while in South Korea and Singapore investment spending surpassed 35 percent of GDP. In Japan and Hong Kong investment's share of GDP was a more modest 30 percent. In contrast, investment spending in the OECD countries averaged about 22 percent of GDP in the 1960s and 1970s, and 20 percent or less more recently.

Admittedly, most of the Asian countries have grown much faster than the OECD nations. Indeed, higher investment is generally regarded as one of the reasons for their rapid growth; and presumably a faster-growing and less developed economy can make effective use of a higher level of investment spending. But South Korea, Hong Kong, and Singapore are now considered "advanced" states.[25] And in Korea, Thailand, Malaysia, and Hong Kong, investment levels in the 1990s were high by the standards of their own past, as well as in comparison with investment shares in developed countries.[26]

In Japan, the figures are less startling. While the share of GDP devoted to investment rose in the late 1980s, it remained below the levels achieved in the 1960s and early 1970s. Japan's growth was much faster back then, however, and Japan was at an earlier stage of economic development, with more opportunity to "catch up." A more relevant comparison may be between Japan in the 1980s and the OECD countries in the 1960s. Growth rates were comparable; but, the OECD nations devoted less than 25 percent of GDP to capital formation in the 1960s.

Data from the Penn World Table, presented in Appendix Figure 2, argue even more persuasively that levels of investment spending in the Asian countries in the late 1980s and early 1990s were very high. The Penn World Table data attempt to measure expenditures in "a common set of prices in a common currency so that *real* quantity comparisons can be made, both between countries and over time" (Summers and Heston 1991, p. 327). The rationale behind the Penn World Table is that a given category of goods can cost more in real terms in some countries than in others. In general, investment goods—both producer durables and construction—are relatively more expensive in poor countries than in rich ones. Consequently, the same fraction of GDP devoted to investment buys much less "real"

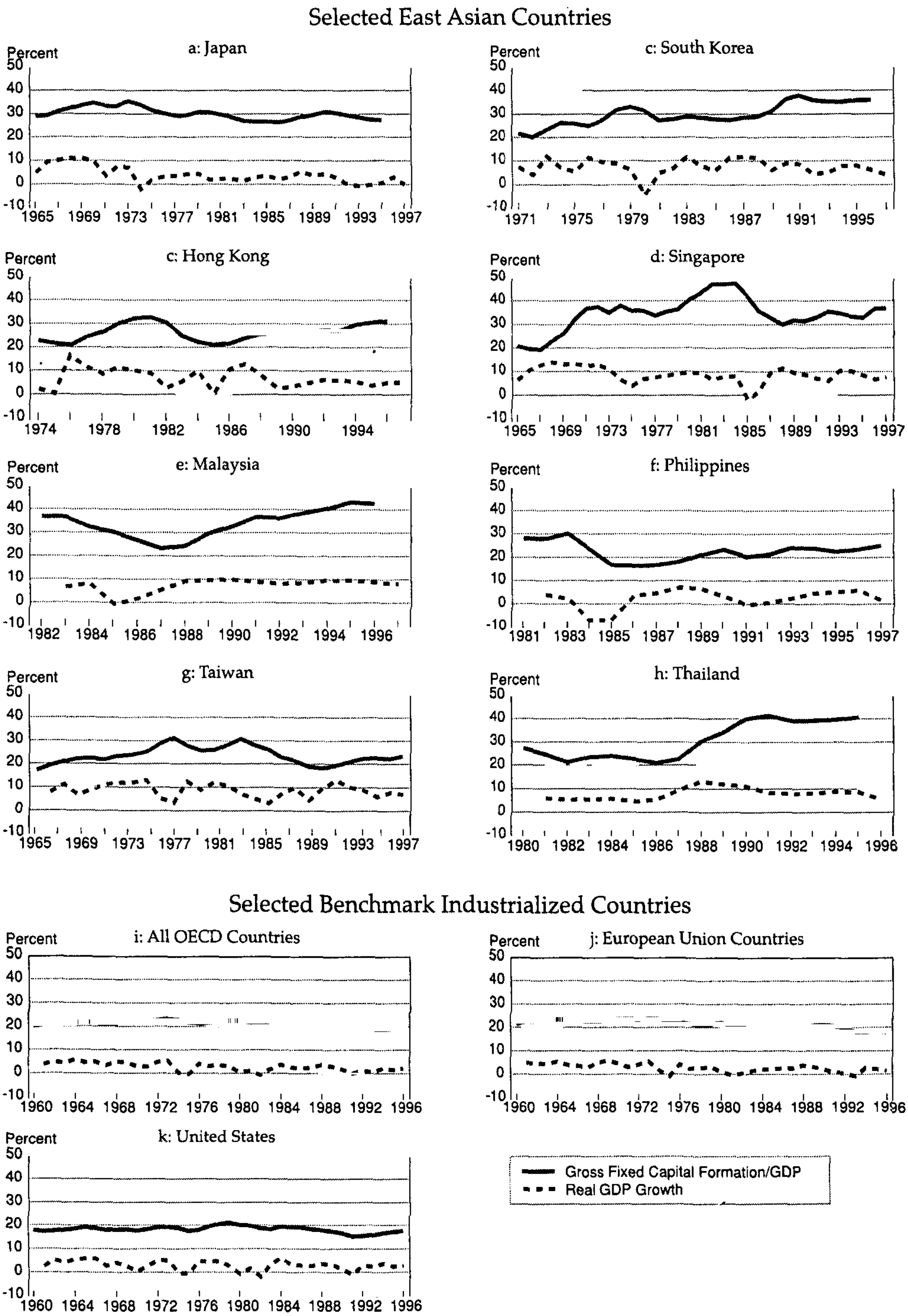


Figure 7 Investment/GDP and growth in real GDP. Source: Standard & Poor's DRI; OECD.

investment in low-income than in high-income countries. The difference is most pronounced in the very poorest countries and largely disappears as countries' output per capita rises above one-third of the U.S. level.

Thus, the very high shares of nominal GDP devoted to investment by South Korea, Thailand, and Malaysia back in the 1960s and 1970s actually provided relatively little in the way of real investment. However, by the late 1980s and early 1990s, reflecting these countries' gains in development and the associated declines in the cost of investment goods, the ratio of investment to GDP was soaring, far surpassing earlier levels and the levels in the developed world. The Penn World Table data do not extend beyond 1992, unfortunately, but based on the nominal figures for GDP and investment, the upward trajectory must have continued.

Real estate has played a particularly important role in the current crisis and in Japan's earlier problems. Here, the U.S. real estate and banking crises in Texas in the mid 1980s and New England in the late 1980s are instructive. In neither U.S. case were signs of impending problems obvious. While rising property prices in New England could have been seen as a sign of overheating, California had experienced an earlier escalation without serious consequences, while more elastic construction in Texas damped down price increases. Nor was the emergence of excess capacity immediately evident in higher vacancy rates. (Similarly, in Japan, the number of bankruptcies and default rates fell to very low levels during the "bubble" years between 1986 and 1989; they soared, along with the value of liabilities involved in bankruptcies, in 1990 and 1991. And corporate profits, which had been rising until 1990, did not turn down until 1991. (See Bank of Japan 1998a, pp. 22–32 and 1998b, chart 37.)

However, an examination of employment patterns in manufacturing and other traded goods industries relative to those in construction would have revealed that the expansion in construction lacked solid underpinnings, as both regions were encountering difficulties in industries that historically had provided the primary impetus to their growth (see Browne 1992). In effect, the real estate developers were building offices for themselves and the service industries supported by their growth. Figures 8 and 9 compare changes in traded goods

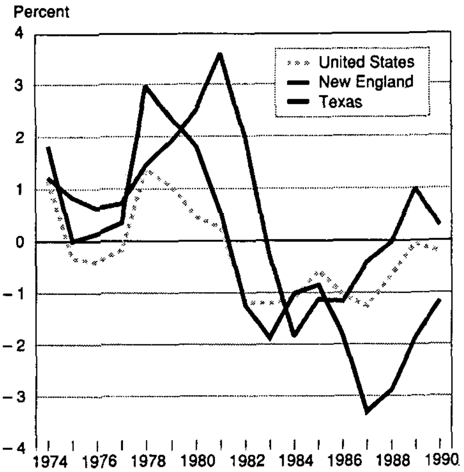


Figure 8 Divergence between changes in export employment and changes in construction employment. Note: Calculations described in the text. Source: U.S. Bureau of Economic Analysis.

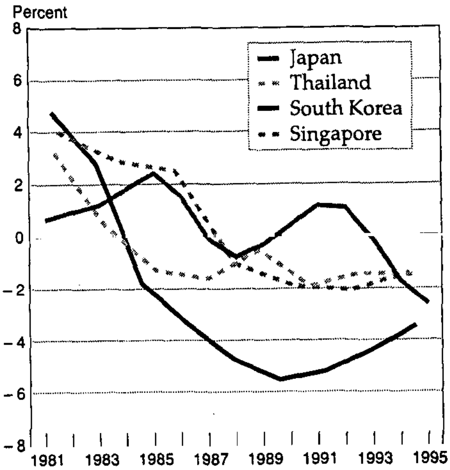


Figure 9 Divergence between changes in export employment and changes in construction employment (East Asian Countries). Source: OECD; Standard & Poor's DRI.

employment to changes in construction employment, over three-year periods, for New England, Texas and selected Asian countries.[27] (Box 2 presents a similar comparison for three Scandinavian countries that had serious financial crises in the early 1990s after a period of rapid increases in stock and real estate prices.) In most cases, employment trends in traded goods industries and in construction diverged sharply in the years before the crisis. In Asia, the divergence is especially pronounced in South Korea; Thailand and Singapore also show clear gaps. In Japan, by contrast, the divergence in the late 1980s was more modest. The notable recent divergence reflects weakness in manufacturing employment concurrent with new public works projects intended to stimulate the economy.

Box 2 Asset Prices, Consumer Prices, and Banking Crises in Selected Scandinavian Economies

In the early 1990s Norway, Finland, and Sweden experienced severe banking crises that shared a number of features with other financial crises, like the regional banking problems in the United States and the more recent problems in East Asia. The resulting stresses in Scandinavia were sufficiently severe that a number of the largest banks required substantial public funding to continue operating.

As in Asia, Scandinavian consumer prices stayed relatively flat by historical standards in the pre-crisis period, despite rapid growth in bank credit; strong real effective exchange rates and increased competition in export markets probably contributed to the subdued behavior of consumer prices. By contrast, rapid increases in real estate and equity prices first stimulated bank lending but later, as asset prices turned down, posed a severe challenge to the banking sector's viability.

Financial market deregulation also played a role in each of these countries' banking crises.[a] Before the early 1980s, all three countries had maintained a tradition of strict bank regulation. According to the head of financial research at the Norwegian Central Bank, Sigbjorn Berg, these rules were in place "not for prudential reasons, but as important components of their monetary policy." Interest rates and lending volumes were regulated "to control macroeconomic impulses from credit markets" (Berg 1993, p. 442). But following financial liberalization the countries experienced rapid credit expansion and robust economic growth.[b] Banks sought higher returns, and households and corporations began to borrow aggressively; thus, bank loans became increasingly risky and an ample supply of credit flowed into the real estate market; real estate prices rose.[c] Meanwhile, overvalued real effective exchange rates, the collapse of exports to the former Soviet Union, and weakness in the world market for forestry products all contributed to a deterioration in these countries' current account.[d]

As the figure shows, employment growth in the nontradables sectors (real estate and construction) diverged from that in tradables sectors starting in Finland in 1983 and lasting through the decade. In Norway and Sweden, an initial divergence occurred in 1983 and reappeared in a more pronounced form in 1987 and 1988 in Norway and in 1991 and 1992 in Sweden. When the real estate market turned down abruptly in the early 1990s, banks faced substantial losses from nonperforming real estate loans. The devaluation of the markka, the krone, and the krona during the 1992 ERM crisis left domestic firms, many of which had high levels of foreign-currency-denominated debt, in difficult circumstances that exacerbated the recession that followed.[e]

---

[a] Problems in Norway preceded those in Finland and Sweden by roughly two years.

[b] Norway liberalized its housing market starting in 1982, and abolished volume controls in 1984 and interest rate ceilings in 1985. As of 1990, all foreign currency transactions were liberalized, following the abolition of limits on capital flows from foreign private financial institutions in the late 1980s. Land price inflation began almost immediately after deregulation was phased in, and lasted from 1985 to 1989. Finland and Sweden followed a similar path of financial deregulation in the early 1980s. Land price inflation picked up there in 1985, lasting through 1987 in Finland and through 1991 in Sweden.

[c] Berg also noted that although faced with mounting evidence of severe problems in bank balance sheets, "Supervisors had (had) few problems under previous regulations and had been lulled into the belief that banking is an inherently stable industry with no great need for supervision . . . The supervisors and economic policymakers did not do much to prevent the banking crisis from developing. One can even argue that they in some cases contributed to make things worse" (Berg 1993, p. 442).

[d] In response to these pressures, the Finnish markka, the Norwegian krone, and the Swedish krona were initially devalued in 1991. All three currencies were unpegged from the ECU in 1992 during the ERM crisis, and the real effective exchange rate for each depreciated substantially.

[e] GDP slowed in 1989 as Norway went into recession. In 1990, GDP grew 0 percent in Finland and fell in 1991 by 7 percent. In Sweden a quickly deepening recession began in 1990.

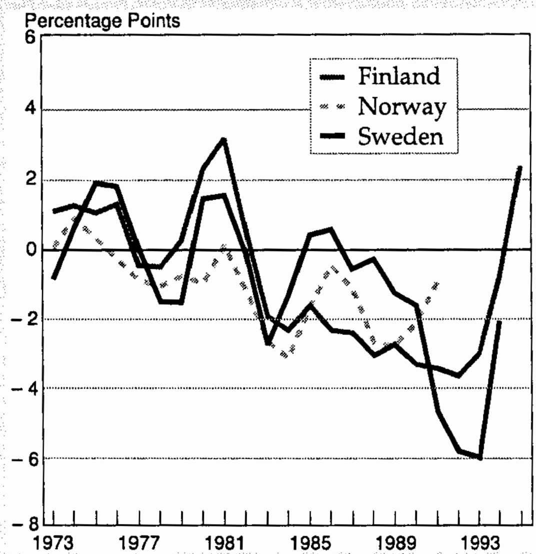


Figure B-1 Divergence between export employment and construction employment (selected Scandinavian countries). Source: OECD.

In Scandinavia in the 1980s, as in the United States in the 1980s in the East Asian economies in the 1990s, deregulation appears to have unleashed competitive forces that contributed importantly to the crisis and seemingly caught the authorities off guard. The cost of their myopia was high. The Norwegian and Swedish governments spent public funds equivalent to about 4 percent of their GDPs to bail out failing banks, while the cost for the Finnish government was a stunning 8 percent (IMF 1998, p. 118).

Rebecca Hellerstein and Anna Sokolinski

Admittedly, the effect of high levels of construction on real estate and financial markets depends in part on what is built. Public works projects create little or no capacity overhang in the private sector, although completion of large projects will still pose adjustment challenges, and some public projects may not be economically productive. Also, timing is unpredictable; as Finland's experience illustrates (see Box 2), employment in traded goods industries and construction may diverge for a considerable period before a correction. Nonetheless, a feature of many recent crises in both developed and developing countries has been a bout of construction activity running ahead of sustainable demand.

In sum, the East Asian countries now in crisis, as well as Japan before them, appear to have fallen victim to problems that showed up first or early in asset markets. In some cases, notably Japan, huge fluctuations in asset prices left a crippled banking system and an overhang of excess capacity in some sectors. In others, notably South Korea, asset markets were less clearly involved while the high level of investment spending (and, recently, associated foreign debt) seems more problematic. In Thailand, excessive investment in both financial assets and industrial capacity appears to have contributed to the crisis. But in no case did consumer price indexes indicate overheating.

CONSUMER PRICES AND ASSET MARKETS IN A GLOBAL CONTEXT

The global context is key to understanding how Asian asset markets could become overheated or distorted without an acceleration in consumer price inflation. In Japan, a soaring exchange rate seems to be a critical part of the story. In contrast, effective exchange rates in most of the Asian Tigers did not increase markedly until the second half of 1995, when the yen began to fall in value.[28] Nonetheless, producers of tradable goods already faced intense competition both at home and abroad. The good behavior of tradable goods' prices most likely helped restrain domestic inflation, even as a surge in foreign capital inflows in response to financial liberalization encouraged investment.

In theory, an appreciating exchange rate can damp down consumer price inflation in an open economy, even as it shifts economic activity toward nontradables industries. A rising exchange rate lowers the cost of imported products and puts competitive pressure on domestic producers of tradable goods. At a given overall rate of growth, nominal interest rates will tend to be lower, favoring construction and real estate. If speculative tendencies are present, they are likely to be reinforced by a rising exchange rate. When speculative and momentum trading is significant, the process can feed on itself, as expectations of rising stock and property values will tend to attract foreign as well as domestic investors, further bidding up the currency and adding currency gains to foreign investors' returns in equity and real estate markets. Late in the 1980s, for example, foreign inflows to Japan's equity markets soared, bringing that country's foreign liabilities to 50 percent of GDP, up from less than 20 percent at the start of the decade (Kähkönen 1995).

Producers of traded goods are likely to respond to the combination of intense competitive pressure, on the one hand, and the low cost of capital, on the other, by undertaking investments intended to boost productivity. While these investments may indeed improve productivity, and further damp down output prices, they may also lead to overcapacity in some industrial sectors.[29] Nor are rising asset prices likely to feed back into consumer price measures immediately or fully;[30] equity values and commercial real estate do not enter into such measures, and, while shelter is an important consumer expenditure, its price is typically calculated according to a rental equivalence method with only a loose link to housing prices. (See Appendix Table 2.)

The point is not that a rising exchange rate will cause a boom in asset prices or lead to excessive levels of investment; but rather that, if conditions are favorable, it is likely to reinforce such tendencies while simultaneously helping to keep consumer prices relatively subdued. This scenario appears to have occurred in Japan in the late 1980s. In reprise, Japan's very strong export performance earlier in the decade provided a powerful impetus to the economy and generated great optimism about the future. Stock prices were bid up, then real estate values. Borrowing rose. Inflated real estate prices magnified the value of many borrowers' collateral, while higher equity prices fed directly into bank capital, supporting an expansion in lending.

Meanwhile, the yen was appreciating; between 1985 and 1988, the yen's effective value rose more than 50 percent. Along with the drop in world oil prices and a modest decline in unit labor costs (in yen), this appreciation seemingly helped to damp down consumer price inflation in the face of an accommodative monetary policy and declining unemployment. Export and import prices both fell sharply in the second half of the 1980s, and prices for (nontraded) consumer services rose faster than prices for (tradable) consumer goods throughout the period; indeed, consumer goods prices actually fell in the three years from 1986 to 1988. Possibly, thus, policymakers and investors were lulled into believing, as some observers suggested at the time, that the extraordinary P/Es and rapidly rising property values were measures of Japan's competitive strength and promising future rather than indicators of an overheated economy. In fact, the strong yen was eroding the competitive position of Japan's traded goods industries; and the combination of disappointing earnings and an eventual tightening of monetary policy subsequently led to the stock market collapse and ultimately to the financial crisis from which Japan has still not recovered.

In South Korea and most of the other Asian Tigers, rising real exchange rates were not an issue until after mid 1995 (Appendix Figure 3). Nevertheless, tradable goods producers in these countries were under severe competitive pressure. Demand from the industrial countries, especially Europe, was sluggish, while the emergence of China as a viable competitor in export markets, the devaluation of the Mexican peso, Taiwan's move into semiconductors, and increases in their own productive capacity had intensified competition. Tradable goods prices were flat through much of the 1990s, helping to damp down consumer price inflation. The situation became untenable when, in response to financial liberalization, these countries attracted huge capital inflows from abroad (Table 4).

These inflows supported rapid expansion in bank credit and increased investment relative to GDP. In several countries, construction activity accelerated and property prices rose sharply. Meanwhile, strong rates of overall growth bid up compensation costs, further weakening traded goods producers, who could not afford to pass higher labor costs on to their customers. (Korean unit labor costs, in won, rose 6 percent a year in 1995 and 1996, for instance.) Thus, problems in the external sector were both offset and obscured by strong investment spending.

That South Korea and other East Asian countries were so attractive to foreign investors was not entirely a reflection of their own attributes. Laudable reductions in fiscal deficits and inflation in many industrial countries and sluggish growth in Japan and much of Europe, despite generally accommodative monetary policies, had brought interest rates in the industrial world down to their lowest levels in many years. When investors sought higher returns, they were drawn to the East Asian countries, most of which had just started to open their financial markets to foreigners and seemingly promised fast growth, low inflation, sensible government budgets, and "predictable" exchange rates. Banks in Canada, France, Germany, and the United Kingdom, all of which had large declines in interest rates or sharp improvements in their fiscal deficits, increased their loans to the East Asian countries rapidly.[31]

Table 4 Five Asian Economies: Net Foreign Borrowing (in Billions of Dollars)

	1992	1993	1994	1995	1996	1997
Private Flows, net	24.7	44.8	37.9	79.2	97.1	–11.9
Equity	9.5	21.1	12.1	15.4	18.7	2.1
Private Creditors	15.2	23.8	25.8	63.8	78.4	–14.0
Commercial Banks	10.2	7.5	23.4	49.9	55.7	–26.9
Nonbanks	5.0	16.3	2.4	13.8	22.7	12.9

The five Asian economies are South Korea, Indonesia, Malaysia, Thailand, and the Philippines.
Source: Institute of International Finance, Inc.

Most of the lending was short-term, maturing within less than one year, and denominated in unhedged dollars. In South Korea, two-thirds of the lending was bank to bank.

By lending to the Korean banks, which then lent to the domestic firms, foreign banks could take advantage of the low-risk weights assigned to interbank loans in determining capital needs. Foreigners may have presumed some form of government guarantee, given the close relationships between Asian banks and their governments. And with most loans maturing in less than a year and no signs of rising inflation, foreign lenders probably thought they were protected against a broad-based deterioration in economic fundamentals and had little fear of a tightening in monetary policy. What many private investors and public officials seemingly failed to appreciate, however, was the degree to which these traditional indicators of economic soundness were being distorted by the combination of intense pressure on tradable goods producers and massive investment spending.

Now that Asian asset prices and currencies have collapsed, borrowers' collateral values have plummeted even as the burden of their debt obligations has soared. Bank capital has evaporated, and financing for even routine working capital is scarce, particularly as alternatives to bank finance remain relatively limited in most of these countries.

CONCLUSION

The recent financial crisis in Asia has prompted considerable debate over many of the policy prescriptions that economists have been giving central banks and other government agencies in recent years.[32] It will probably take some time before any consensus emerges on the lessons to be drawn from this unfortunate episode.

Among the questions that should be considered is whether the recent emphasis on inflation has had some unpredicted and undesirable consequences. In particular, have academic arguments in favor of inflation targeting been interpreted by some policymakers and market participants to mean that as long as inflation is low, there is little cause for concern or reason for action? For instance, did low inflation rates help delay corrective action in Asia by reassuring policymakers and regulators in those countries that all was well? And did investors assume that they could scrutinize investments less closely than they might otherwise, construing low inflation in these countries as evidence that the economy was in balance or that, in any event, the monetary authorities were unlikely to tighten policy any time soon?[33]

This article has presented some evidence suggesting that consumer prices may provide an incomplete picture of the pressures on the economy and that other indicators, notably rising asset prices, extraordinary levels of investment spending, and rapid growth in construction relative to traded goods industries, may also signal distortions that threaten future growth. To date, discussions of the causes of the Asian crisis have centered, with good reason, on the role of inadequate transparency and lax supervision. But in the absence of good financial data and strong supervisory systems, a more critical examination of Asian investment trends might have alerted policymakers and investors to the growing instabilities. Looking ahead, improved regulation and increased disclosure are, of course, essential. But including the behavior of asset markets among the indicators receiving careful consideration and asking more probing questions about the sustainability of real investment patterns may be a useful supplement to these widely proposed financial system reforms—particularly since bad loans appear to be a lagging indicator of problems stemming from overinvestment in asset markets.

The East Asian crisis may also provide some guidance for future financial liberalizations, while, incidentally, resolving a recent debate among economists. According to the standard (Mundell-Fleming) open economy model, countries cannot simultaneously have pegged exchange rates, liberalized capital accounts, and substantial autonomy in setting monetary policy. But several authors have claimed that Malaysia, Indonesia, South Korea, and Thailand all retained considerable policy independence, despite having pegged their exchange rates and opened their capital markets to international investors (Woo and Hirahama 1996; Frankel 1993; Fischer and Reisen 1993). These authors argued that the East Asian economies achieved this feat because policymakers were able to exploit their regulatory authority over domestic institutions to influence their lending activities. The Asian sequence—international deregulation before domestic deregulation—violated the generally preferred order (McKinnon 1993). But given recent developments in these same countries, the conventional prescription appears to have been right; government efforts to guide the lending activities of domestic financial institutions tend to produce serious distortions under most circumstances but particularly in an economy with a fixed exchange rate and open capital accounts.

From this perspective, the high, probably excessive, levels of investment spending in the East Asian countries may be seen not simply as a consequence of opaque, inadequate financial standards and poor financial supervision, but also as the result of government policies, both implicit and explicit, to sustain the pegged exchange rates and the high growth rates to which their populations had become accustomed, despite competitive pressures on traded goods industries. Policymakers failed to recognize that, as they began to open their economies to foreign capital, they could no longer remain the ultimate judges of those investments. They had substituted the markets' values for their own.

At the same time, investors with insufficient data to form sound judgments about private sector loans had made unfounded assumptions about government guarantees. In effect, they assumed that the old way of doing business still operated in these countries. Then, belatedly recognizing the impossibility of governments' meeting such large foreign currency liabilities, these investors looked at their loans with a more jaundiced eye and fled.

In the future, thus, policymakers, in borrowing *and* lending countries, may want to revisit their supervisory and tax systems to ensure that they are not creating incentives that are potentially destabilizing. This review may be particularly important in countries where the financial system is undergoing major change, for, as the Asian experience illustrates, moving from a highly regulated to a more market-oriented system can ignite forces that inexperienced lenders, borrowers, and regulators are ill equipped to handle. Since interbank lending was such an important component of the huge capital flows to Southeast Asia, high priority should go to raising the cost of interbank transactions to reflect their true degree of risk more fully. Possible measures include imposing modest reserve requirements on short-term interbank borrowing or increasing the risk weight for interbank loans in risk-based capital standards. Developing international bankruptcy procedures for nations might also help to shift the cost of international financial crises from the public to the private sector.

As the financial crises in Asia and elsewhere illustrate, unsustainable increases in asset prices and excessive investment spending can seriously damage economies, without a simultaneous pickup in consumer prices raising an alarm. Accordingly, policymakers may want to consider broadening their present focus on consumer price inflation to include developments in asset markets and other indicators of imbalances in the real economy. In some cases, the appropriate tools for preventing asset market imbalances may be supervisory or fiscal measures rather than monetary policy actions. However, monetary authorities still have a responsibility to use their expertise and stature to draw attention to the dangers of speculation and excessive investment. They also have a responsibility to avoid validating supervisory shortcomings and investment mistakes.

Appendix Table 1A Money and Credit Aggregates in Japan (in Annual Percent Change)

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
M2 + CDs	8.4	8.7	10.4	11.2	9.9	11.7	3.6	.6	1.1	2.1	3.2	3.3	3.1
Bank Credit	10.2	10.0	10.8	11.3	11.5	11.2	6.3	3.6	.0	–.6	1.1	1.6	.4
Nominal GDP	6.6	4.7	4.3	6.9	7.0	7.5	6.6	2.8	.9	.8	.8	3.4	1.5

Source: OECD, Board of Governors of the Federal Reserve System.

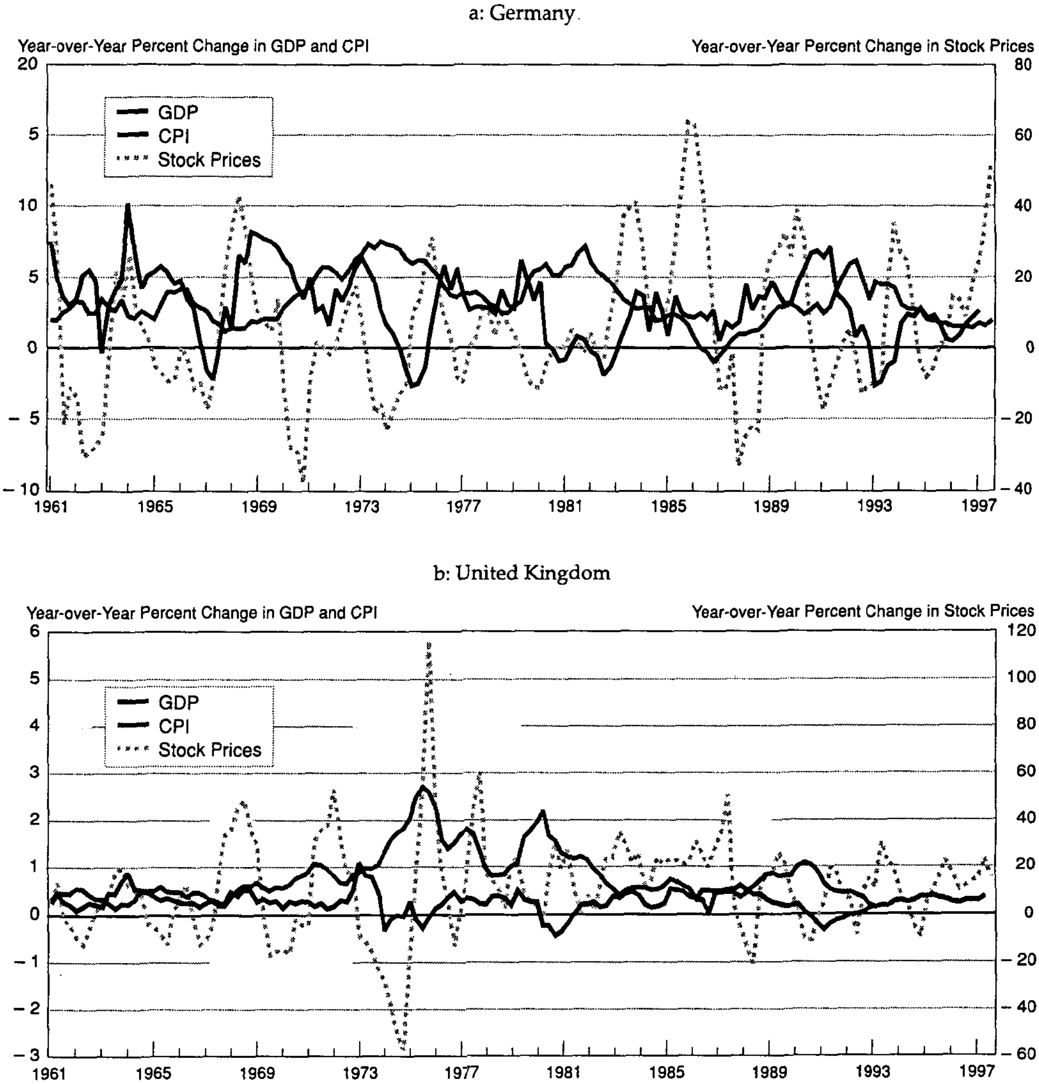
Appendix Table 1B Money and Credit Aggregates in East Asian Countries (in Annual Percent Change)

	1993	1994	1995	1996	1997
Hong Kong					
M3	13.9	12.3	15.1	11.1	12.5
Bank Credit	21.0	25.0	8.6	18.0	18.8
Nominal GDP	15.2	12.7	6.5	10.7	11.9
Indonesia					
M2	19.8	21.2	24.8	28.2	25.8
Bank Credit	21.0	22.9	21.7	22.7	n.a.
Nominal GDP	16.8	15.9	18.9	17.2	17.2
South Korea					
M3	21.5	21.9	20.1	18.9	16.1
Bank Credit	12.7	18.4	14.7	19.4	23.3
Nominal GDP	11.1	14.5	15.0	10.8	8.0
Malaysia					
M3	19.5	21.1	15.1	23.6	19.7
Bank Credit	7.8	10.3	26.5	28.0	26.3
Nominal GDP	11.2	15.2	15.0	14.2	10.9
Philippines					
M3	13.9	24.9	32.6	19.7	20.5
Bank Credit	28.5	19.0	31.3	40.3	30.8
Nominal GDP	9.1	14.8	12.6	15.2	11.6
Singapore					
M3	7.3	13.9	13.9	9.6	10.5
Bank Credit	12.0	12.8	17.4	17.3	19.5
Nominal GDP	16.4	15.1	11.2	8.3	9.4
Taiwan					
M2	16.1	16.2	11.6	9.2	8.3
Bank Credit	15.1	20.5	10.2	6.1	17.2
Nominal GDP	10.1	8.6	8.1	8.5	8.8
Thailand					
M2	16.1	13.0	17.3	16.6	13.6
Bank Credit	22.7	28.9	23.1	14.0	32.3
Nominal GDP	12.0	14.5	15.4	9.8	5.0

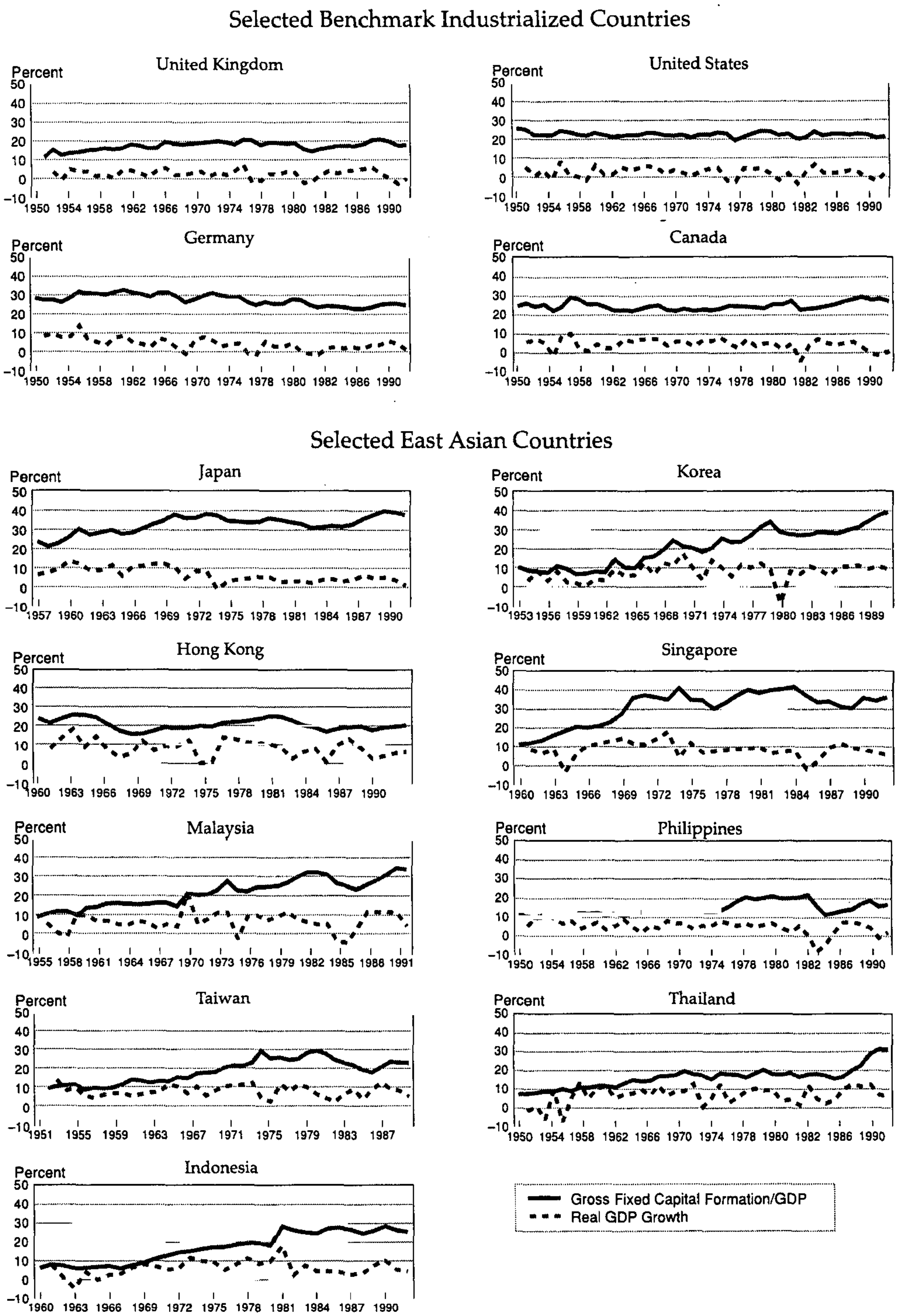
Source: IMF, Standard & Poor's DRI.

Appendix Table 2 Consumer Price Indexes: How Do They Measure Asset Prices?

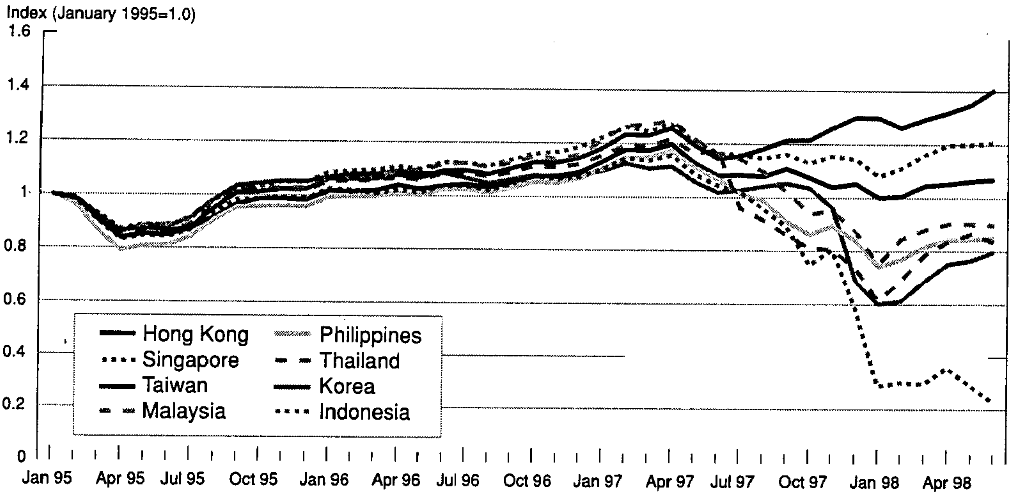
Basic Concept	The CPI has become an almost universally accepted indicator of inflation as well as a barometer preferred by many policymakers. The CPI covers household *consumption* expenditures; excluded are household expenditures of other kinds, in particular those representing investment, savings or transfers. Asset expenditures excluded by definition from the index include purchases of a dwelling and stocks. Thus, the CPI is designed rather narrowly to measure changes over time in the average retail prices of a fixed basket of goods and services that represent the consumption habits of households. Across countries, house prices are the only asset weighted prominently in the calculations, generally through owners' equivalent rent calculations derived from prices in the rental market.
United States	Home ownership: represented by owners' equivalent rent, defined as the cost of renting housing services equivalent to those provided by owner-occupied housing and calculated based on changes in the rent of rental units and by household insurance exclusive of the house structure. Rent: Derived from survey of 40,000 tenants. Weighting: As of December 1993, rent equals 27 percent of entire index.
Japan	Home ownership: Rental equivalence approach used to calculate the housing cost of owner-occupied dwellings. Rent: Monthly survey of a group of tenants in both the private and the public sectors. Weighting: As of 1995, rent equals 17 percent of entire index.
Finland	Home ownership: Represented by repair and maintenance costs, insurance premiums, interest on housing loans, depreciation, water charges, and the like. It does not include an imputation of housing benefits gained by owner-occupiers. Rent: A quarterly mail survey of 24,000 tenants is conducted to obtain data on changes in rents. Weighting: As of 1990, rent, repairs and maintenance, and home ownership costs equal 17 percent of index.
Norway	Home ownership: Represented by mortgage interest, repair and maintenance, insurance, water charges, and the like. No price survey of owner-occupied dwellings is conducted; thus, price changes are assumed to reflect rent changes. Rent: Data on rent for private houses are gathered quarterly for 1500 homes. Weighting: As of 1995, rent equals 16 percent of entire index.
South Korea	Housing expenditure: Calculated using a rental equivalence approach. Rent: Data collected as part of the "Family Income and Expenditure Survey." Weighting: CPI weights are rebased every five years: for 1990–1994, rent was weighted at 11.9 percent of total expenditure; since 1995, it has been weighted at 12.8 percent.



Appendix Figure 1 Classic cases: CPI, GDP, and stock prices. Source: OECD.



Appendix Figure 2 Investment/GDP and growth in real GDP. Source: Penn World Table Mark 5.6 (NBER Penn Table Database.)



Appendix Figure 3 East Asian exchange rates vis-à-vis the yen. Source: Board of Governors of the Federal Reserve System; Standard and Poor's DRI.

NOTES

1. This focus on consumer price inflation stems from the belief that monetary policy can have no effect on real variables over the long run. It is closely akin to the monetarism that Milton Friedman and his colleagues forcefully and eloquently espoused in the 1960s and 1970s in that both offer a relatively straightforward, rules-based approach to policy. While monetarism's simplicity had great appeal, innovation in the financial services industry during the 1980s (for example, the development of money market and sweep accounts and the resulting blurring of distinctions between bank and nonbank financial institutions) increased the variability of velocity and made the monetarist approach impractical for policy purposes. Kindleberger (1978) links monetarism, in turn, to the Currency School of the early 1800s. Thus, the current focus on the consumer price index draws on a long legacy.
2. For example, Stanley Fischer (1996) has suggested that "long-run price stability should be the primary goal of the central bank, with the promotion of full employment and growth being permitted to the extent that they do not conflict with the primary goal." And Goodhart and Viñals (1994) have documented that price stability has become the central bank's primary objective throughout Europe, Canada, and Australia/New Zealand. But, of course, many economists disagree with this prescription, recommending instead that central banks look at both prices and output or at nominal GDP. Moreover, as noted in the text, most statutes that require the monetary authorities to follow an inflation target contain an escape clause in the event that major supply shocks have a severe impact on employment and growth.
3. For example, in 1995 Michael Bruno, then chief economist of the World Bank, wrote, "Very low inflation is again becoming the norm, not only in the industrial world but also in developing regions." After discussing stabilization strategies, like fiscal retrenchment, exchange rate pegs, currency boards, and wage freezes, he argues that "getting inflation down to single digits is important . . . for long-term growth reasons" and cautions that "the upward bias of inflationary persistence argues for keeping the inflation genie tightly in the bottle" (Bruno 1995). In the opening of its May 1997 *World Economic Outlook,* the International Monetary Fund approvingly sees "few signs of the tensions and imbalances that usually foreshadow significant downturns in the business cycle: global inflation remains subdued, and commitments

to reasonable price stability are perhaps stronger than at any other time in the postwar era . . ." (The same overview warns of potential dangers posed by fragile banking systems exposed to large foreign exchange risk by large and possibly unsustainable capital inflows.) Later, the IMF writes, "In Chile, the most *successful* economy in (Latin America), inflation fell to a 36-year low of 6 1/2 percent . . ." (italics added). Rudiger Dornbusch has also described Chile's approach as "exemplary" as it brought annual inflation down from 30 percent in the mid 1980s to 7 percent in the mid 1990s—in part because "the central bank has refused to overreach and squeeze inflation down to the fashionable 2 percent of the industrialized countries." He contrasts Chile with Mexico where he sees "exaggerated emphasis put on inflation, exaggerated urgency to get to 2 percent, dangerous imperviousness to overvaluation." He concludes, "The right message is that inflation must come down and that there is never room for complacency; that is not the same as inflation reduction first, growth later" (Dornbusch 1996).

4. *The Economic Growth and Price Stability Act of 1997* was sponsored by Florida Senator Connie Mack. The bill proposed that an explicit numerical definition of price stability be established, and that the promotion of long-term price stability, so defined, should be the sole mandate of the Federal Reserve System.
5. As of mid 1998, Hong Kong, Malaysia, Indonesia, Thailand, South Korea, and Japan report that their economies have been shrinking since the beginning of the year, Japan by a stunning annual rate of 5.3 percent in the first quarter. Those East Asian economies that managed to maintain some momentum, Taiwan and Singapore, have nonetheless slowed as well. The Philippines economy now has a negative growth rate.
6. Many investors and policymakers—both monetary and supervisory—apparently ignored other, less standard signs of trouble, like soaring ratios of foreign-currency debt to GDP and questionable levels of investment in certain asset markets and industrial sectors. The fact that these countries did not display the symptoms—rapid inflation and large fiscal deficits and "excessive" consumption relative to GDP—seen in the Latin American economies on the brink of their crises was seemingly reassuring.
7. This article uses the term "inflation" to connote consumer price inflation. Price inflation in asset markets will always be referred to as "asset-price inflation."
8. Like any major supply shock, a sharp shift in oil prices pushes employment and price trends in opposite directions, creating a dilemma for monetary policy. But in the 1970s, U.S. inflation probably contributed to the oil exporters' decision to raise prices. Although the precipitating event for the first increase was the Arab-Israeli War, oil is priced in dollars and rising U.S. inflation in the early 1970s had contributed to a nominal depreciation of the dollar and a decline in the value of oil revenues in world markets. The second increase also followed an extended period of rising U.S. inflation and nominal dollar depreciation.
9. In this article, asset prices refer to prices of land, real estate, financial assets such as stocks and bonds, and, on some occasions, the price of foreign currency in terms of domestic currency. Purchases of such assets are often motivated by hopes of capital gains.
10. *The New York Times* of October 22, 1929 reported "Fisher Says Prices of Stocks Are Low: Quotations Have Not Caught Up with Real Values as Yet, He Declares," page 24, col. 1. See also Carosso (1970), pages 300–302.
11. Rising real land prices in the 1970s may reflect the fact that the baby boom cohort began forming households at that time; however, for much of the period, prices for farm properties were rising faster than prices for nonfarm, noncorporate land. The boom in prices for farm land may have been linked to the unusually sharp and temporary rise in prices for foodstuffs in the early years of the decade.
12. The G-7 consists of the United States, Japan, Germany, France, Italy, the United Kingdom, and Canada.
13. The modest pickup in Japanese inflation from 1989 to 1991 reflects several developments, including some one-time factors: the Japanese introduced a consumption tax in 1989; the al-

ready low unemployment rate fell even lower, from 2.8 percent in 1987 to 2.1 percent in 1990; the yen weakened a bit in 1989 and 1990 before resuming its rise in 1991; and bad weather conditions in 1990 and 1991 aggravated the impact of the 1990 oil price increase sparked by the Gulf War. (Japan is highly dependent on imported oil.) (Government of Japan 1998.)

14. For example, chapter four of the IMF's May 1998 *World Economic Outlook* examines indicators of vulnerability to financial crises in 50 advanced and emerging market countries from 1975 to 1997. It finds that "typically, in the lead-up to a currency crisis, the economy is overheated: inflation is relatively high, the real exchange rate is appreciated, the current account deficit has widened, domestic credit had been growing at a rapid pace, and asset prices have often been inflated" (IMF May 1998, page 96).
15. Paul Krugman, Nouriel Roubini, and other commentators in the fundamentals camp argue that the crisis reflected excessive investment fueled, first, by international speculation that drove regional asset values to unrealistic levels, and second, by an East Asian variant of crony capitalism that directed investment to unproductive ends (Krugman 1998). Krugman has further argued that East Asia's spectacular economic growth since the 1960s was based on an accelerated use of the inputs of labor and capital rather than on the absorption of new technology. So, by the 1990s, the pace of the region's growth was likely to slow as diminishing returns set in (Krugman 1994).

    Jeffrey Sachs and Steven Radelet espouse a contrary view, arguing that IMF errors caused a crisis that was, at base, avoidable. After the initial devaluation of the baht in mid 1997, the IMF effectively yelled "fire" by announcing that severe shortcomings in East Asian financial markets would require fundamental restructuring; thus alerted, investors began to run for the exits. But the fundamentals of East Asian economies were sound, in this view, and any vulnerabilities in such sectors as finance or real estate would have been manageable in the absence of the IMF spooking private investors (Radelet and Sachs 1998a, b).
16. Because earnings have been so weak, in recent years P/E ratios have at times surpassed levels seen in the heady days of the late 1980s, even though stock prices are down substantially.
17. Commercial land prices in the six largest cities quadrupled between 1985 and 1989; residential and industrial land prices rose only slightly less rapidly. Prices peaked in 1991 and plummeted in 1992. They then continued to fall steadily. As of late 1997, commercial land prices had almost completely retraced their earlier run-up and were back to levels of 12 years earlier. As for land prices in the smaller cities, the index of commercial land values for 200-plus cities was down "only" 40 percent from its peak, as compared to the 75 percent decline suffered in the largest cities (Japan Real Estate Institute 1998).
18. Remarkably, however, the fraction of GDP devoted to residential investment in Japan surpasses that in the United States, despite the U.S. penchant for large and comparatively luxurious dwellings and the faster U.S. population growth.
19. Starting with a revision of the *Foreign Exchange Control Law* in 1980, Japanese nonbank firms gained increased access to foreign banks and foreign and domestic bond and equity markets, thereby reducing their dependence on the Japanese banks. Until then, financial regulations had ensured that the Japanese banks were the primary beneficiaries of Japan's high savings rate. But the increased competition spurred by financial deregulation broke the banks' cartel. (See Weinstein and Yafey 1998.)
20. See Bank for International Settlements, *67th Annual Report,* page 105.
21. Interest rate deregulation followed the following schedule: November 1991, deposits maturing in three or more years and corporate bonds maturing in two years; November 1993, loans (except government loans) and deposits maturing in two or more years; December 1994, deposits maturing in one or more years; in July 1995, deposits maturing in more than six months and less than one year; and in November 1995, demand deposits of three or more months (W.A. Park 1996).

22. The percentage of securities in banks' assets rose from 19.6 percent in 1990 to 32.4 percent in 1995. Trust accounts, where risky real estate investments could be held without regulatory oversight, also grew as a share of bank assets, from 17.9 percent in 1990 to 34.5 percent in 1995 (W.A. Park 1996).
23. As of 1994, the only restriction on such borrowing was that over half of the loans must have an initial maturity of more than three years. In May of 1995, in an effort to stem domestic investment growth financed by foreign-currency borrowing through domestic banks, the government tightened some of the access regulations, but with little impact on overall borrowing levels (*OECD Economic Survey, Korea,* 1995–1996, p. 136).
24. While improved accounting and disclosure standards are essential to preventing *future* crises, the transition to new standards can be difficult when the slate is not clean. As better information has become available in Asia, it has often alerted investors to the poor quality of credits already extended.
25. The IMF has recently recategorized Hong Kong, South Korea, Singapore, and Taiwan (as well as Israel) as "advanced" countries because their per capita incomes and industrial structures put them on a par with the members of the OECD.
26. Singapore, in contrast, had achieved even higher levels of investment spending in the early 1980s, with almost 50 percent of GDP devoted to capital formation. As this earlier episode ended with a sharp economic downturn and a scaling back of investment's share, it does not provide much support for the sustainability of such spending.
27. Divergence is calculated as $[(X_t - X_{t-3}) - (C_t - C_{t-3})]/E_t$ where X_t is export employment, C_t is employment in construction and E_t is total employment, all in year t. Browne's original analysis looked at the growth in employment in both construction and real estate relative to the growth in export employment. Figure 8 shows only construction employment in order to be comparable with the Asian data, which generally include real estate in another employment category.
28. Effective exchange rates are inflation-adjusted averages of exchange rates for a country's important trading partners.
29. Businesses may also respond to an appreciating exchange rate by increasing foreign direct investment overseas. If so, the resulting shift in labor demand will dampen wage costs and, indirectly, consumer prices in the appreciating country.
30. In time, through wealth and credit channel effects, rising asset prices are likely to lead to somewhat higher prices for the goods and services measured by the consumer price index.
31. European integration has spurred considerable restructuring and increased competition in Europe's financial services sector. These forces may also have encouraged the surge in European lending to emerging markets.
32. Topics of debate include preferred exchange rate regimes, the need for restrictions on short-term capital inflows, the proper approach to financial deregulation and liberalization, and the best response to a combined currency/banking crisis.
33. Mishkin and Posen (1997) explore some of these issues in four countries that have adopted numerical inflation targets for monetary policy, New Zealand, Canada, the United Kingdom, and Germany.

REFERENCES

Archer, D. J. 1997. The New Zealand Approach to Rules and Discretion in Monetary Policy. *Journal of Monetary Economics,* vol. 39 (June), pp. 3–15.

Ball, L. 1994. What Determines the Sacrifice Ratio? In N. G. Mankiw, ed., *Monetary Policy: NBER Studies in Business Cycles,* vol. 29, pp. 155–82. Chicago: University of Chicago Press.

Bank for International Settlements. 1996/1997. *67th Annual Report.*

Bank for International Settlements. 1997 and 1998. *The Maturity, Sectoral, and Nationality Distribution of International Bank Lending.* January and July.

Bank of Japan. 1998a. Utilization of Financial Institutions' Self-Assessment in Enhancing Credit Risk Management, *Quarterly Bulletin,* vol. 6, no. 1 (February), pp. 22–32.

Bank of Japan. 1998b. *Monthly Report of Recent Economic and Financial Developments,* June 26.

Berg, S. A. 1993. The Banking Crises in the Scandinavian Countries. *Proceedings of the Annual Conference on Bank Structure and Competition,* Federal Reserve Bank of Chicago, pp. 441–49.

Bernanke, B. S. and Gertler, M. 1995. Inside the Black Box: The Credit Channel of Monetary Policy Transmission. *The Journal of Economic Perspectives,* vol. 9, no. 4 (Fall), pp. 27–48.

Bernanke, B. S. and Mishkin, F. S. 1997. Inflation Targeting: A New Framework for Monetary Policy? *The Journal of Economic Perspectives,* vol. 11, no. 2 (Spring), pp. 97–116.

Browne, L. E. 1992. Why New England Went the Way of Texas Rather Than California. *New England Economic Review,* January/February, pp. 23–41.

Browne, L. E. and Hellerstein, R. 1997. Are We Investing Too Little? *New England Economic Review,* November/December, pp. 29–50.

Bruno, M. 1995. Does Inflation Really Lower Growth? *Finance & Development,* September, pp. 35–38.

Carosso, V. P. 1970. *Investment Banking in America: A History.* Cambridge, MA: Harvard University Press.

Chinn, M. 1998. Before the Fall: Were East Asian Currencies Overvalued? NBER Working Paper No. 6491, April.

Cline, W. and Barnes, K. 1998. Spreads and Risks in Emerging Markets Lending. Institute of International Finance Research Paper No. 97-1.

Corsetti, G., Pesenti, P. and Roubini, N. 1998. What caused the Asian currency and financial crisis? Mimeo, New York University.

Dornbusch, R. 1996. Commentary: How Should Central Banks Reduce Inflation?—Conceptual Issues. In *Achieving Price Stability: A Symposium Sponsored by The Federal Reserve Bank of Kansas City,* pp. 93–103. Jackson Hole, Wyoming, August.

Fischer, B. and Reisen, H. 1993. *Liberalising capital flows in developing countries: Pitfalls, prerequisites, and perspectives.* Paris: Development Studies Centre, OECD.

Fischer, S. 1996. Why Are Central Banks Pursuing Long-Run Price Stability? In *Achieving Price Stability: A Symposium Sponsored by The Federal Reserve Bank of Kansas City,* pp. 7–34. Jackson Hole, Wyoming, August.

Fischer, I. 1932. *Booms and Depressions.* New York: Adelphi.

Fischer, I. 1933. The Debt-Deflation Theory of Great Depressions. *Econometrica,* vol. 1, pp. 337–57.

Frankel, J. 1993. Sterilization of money inflows: Difficult (Calvo) or easy (Reisen)? Working Paper No. 93-024. Center for International and Development Economics Research, University of California, Berkeley, October.

Friedman, M. and Schwartz, A. J. 1963. *A Monetary History of the United States, 1867–1960.* New York: National Bureau of Economic Research.

Fuji Research Institute Corp. 1998. Japan's Economic Outlook. June.

Goodhart, C. A. E. and Viñals, J. 1994. Strategy and Tactics of Monetary Policy: Examples from Europe and the Antipodes. In *Goals, Guidelines, and Constraints Facing Monetary Policymakers,* J. C. Fuhrer, ed., pp. 139–87. Federal Reserve Bank of Boston, Conference Series No. 38, June.

Government of Japan. 1998. Consumer Price Index: January 1997. Internet document (www.stat.go.jp/158c.htm), January 30.

Institute of International Finance. 1998. Capital Flows to Emerging Market Economies. April 30.

International Monetary Fund. 1997. *World Economic Outlook.* Washington, DC: IMF, May, October and December.

International Monetary Fund. 1998. *World Economic Outlook.* Washington, DC: IMF, May.

Ito, T. and Iwaisako, T. 1996. Explaining Asset Bubbles in Japan. Bank of Japan *Monetary and Economic Studies,* vol. 14, no. 1 (July), pp. 143–93.

Ito, T. and Krueger, A. O., eds. 1996. *Financial Deregulation and Integration in East Asia.* NBER-East Asia Seminar on Economics, Vol. 5. Chicago, IL: University of Chicago Press.

Jones Lang Wooton. 1998. *Asia Pacific Property Digest. The Market as at April 1998.* Market Analysis Series. May. Hong Kong.

Japan Real Estate Institute. 1998. Urbanland Price Index. March.

Kähkönen, J. 1995. "Japan's Capital Flows" and "Movements in Asset Prices Since the Mid-1980s." In U. Baumgartner and G. Meredith, eds., *Saving Behavior and the Asset Price "Bubble" in Japan,* pp. 16–35 and 51–62. IMF Occasional Paper No. 124, April.

Kindleberger, C. P. 1978. *Manias, Panics, and Crashes, A History of Financial Crises.* New York: Basic Books.

Kregel, J. A. East Asia Is Not Mexico: The Difference between Balance of Payments Crises and Debt Deflations. Jerome Levy Economics Institute, Working Paper No. 235.

Krugman, P. 1994. The Myth of Asia's Miracle. *Foreign Affairs,* November/December, pp. 62–78.

Krugman, P. 1998. What happened to Asia? Mimeo, MIT.

Little, J. 1997. Anatomy of a Currency Crisis. Federal Reserve Bank of Boston *Regional Review,* Fall, pp. 8–13.

Lucas, R. 1981. Econometric Policy Evaluation; A Critique. *Studies in Business Cycle Theory,* pp. 66–89.

McKinnon, R. I. 1993. *The order of economic liberalization: Financial control in the transition to a market economy.* 2nd edition. Baltimore, MD: Johns Hopkins University Press.

Miranda, K. 1995. Does Japan Save Too Much? In U. Baumgartner and G. Meredith, eds., *Saving Behavior and the Asset Price "Bubble" in Japan,* pp. 4–15. IMF Occasional Paper No. 124, April.

Mishkin, F. S. and Posen, A. S. 1997. Inflation Targeting: Lessons from Four Countries. NBER Working Paper No. 6126, August.

Obstfeld, M. 1994. The logic of currency crises. *Cahiers Economiques et Monetaires,* vol. 43, pp. 189–213.

OECD Economic Outlook. Various years.

OECD Economic Surveys. Finland. Various years.

OECD Economic Surveys. Japan. Various years.

OECD Economic Surveys. Korea. Various years.

OECD Economic Surveys. Norway. Various years.

Park, W. A. 1996. Financial Liberalization: The Korean Experience. In T. Ito and A. O. Krueger, eds., *Financial Deregulation and Integration in East Asia,* pp. 247–76. NBER-East Asia Seminar on Economics, Vol. 5. Chicago, IL: University of Chicago Press.

Park, Y. C. 1996. East Asian Liberalization, Bubbles, and the Challenges from China. *Brookings Papers on Economic Activity,* No. 2, pp. 357–71.

Peek, J. and Rosengren, E. S. 1997. Collateral Damage: Effects of the Japanese Real Estate Collapse on Credit Availability and Real Activity in the United States. Federal Reserve Bank of Boston Working Paper No. 97-5, September.

Radelet, S. and Sachs, J. 1998a. The Onset of the East Asian Financial Crisis. Harvard Institute for International Development, mimeo, February.

Radelet, S., and Sachs, J. 1998b. The East Asian Financial Crisis: Diagnosis, Remedies, Prospects. Paper presented at the Brookings Panel, Washington, D.C., March 26–27.

Sachs, J., Tornell, A., and Velasco, A. 1996. Financial Crises in Emerging Markets: The Lessons of 1995. *Brookings Papers on Economic Activity,* No. 1, pp. 147–217.

Shigemi, Y. 1995. Asset Inflation in Selected Countries. Bank of Japan *Monetary and Economic Studies,* vol. 13, no. 2, pp. 89–130.

Smets, F. 1997. Financial Asset Prices and Monetary Policy: Theory and Evidence. Bank for International Settlements Working Paper No. 147, September.

Summers, R. and Heston, A. 1991. The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988. *Quarterly Journal of Economics,* vol. 106, no. 2 (May), pp. 327–68.

Suzuki, Y., ed. 1987. *The Japanese Financial System.* New York: Oxford University Press.

Weinstein, D. E. and Yafey, Y. 1998. On the Costs of a Bank-Centered Financial System: Evidence from the Changing Main Bank Relations in Japan. *Journal of Finance,* vol. 53, April, pp. 635–72.

Woo, W. T. and Hirahama, K. 1996. Monetary Autonomy in the Presence of Capital Flows: And Never the Twain Shall Meet, Except in East Asia? In T. Ito and A. O. Krueger, eds., *Financial Deregulation and Integration in East Asia,* pp. 307–29. NBER-East Asia Seminar on Economics, Vol. 5. Chicago, IL: University of Chicago Press.

Young, A. 1995. The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience. *Quarterly Journal of Economics,* CX (August), pp. 641–80.

43
Globalization and U.S. Inflation

Geoffrey M. B. Tootell
Vice President and Economist, Federal Reserve Bank of Boston, Boston, Massachusetts

Where has all the inflation gone for what is getting to be a long time passing? Estimates of the Phillips curve suggest that the low level of unemployment over the last few years should have produced a fairly significant acceleration in prices, yet inflation has continued to decline. Some, like Robert Gordon (1997) and Staiger, Stock, and Watson (1997), take this occurrence as evidence that the non-accelerating-inflation rate of unemployment, the NAIRU, has declined. Others argue that special factors, such as recent movements of employee health coverage to health maintenance organizations, have temporarily masked the increase in inflation. Yet, perhaps the most widely cited explanation for the surprisingly good inflation performance of late concerns the increasing sensitivity of the U.S. economy to foreign economic conditions; specifically, since capacity utilization abroad has been slack in recent years, U.S. inflation has remained mild. This study uses a variety of approaches to examine whether U.S. inflation depends on foreign, rather than domestic, capacity constraints. It is shown that foreign capacity plays little, if any, role in the determination of U.S. inflation independent of any role it might play in U.S. capacity utilization.

Monetary policymakers must understand the determinants of inflation in order to attain their inflation goal. If foreign capacity constraints help determine domestic price increases, then U.S. policymakers should modify their concerns about domestic capacity utilization. Going beyond, or falling below, some domestic measure of full capacity would not be sufficient to increase, or decrease, inflation, and the concept of a domestic NAIRU might not necessarily be an important or even meaningful policy goal. The rationale for the importance of foreign capacity to U.S. inflation is fairly clear-cut. If, for example, domestic demand exceeded domestic capacity, while foreign capacity remained underutilized, either the excess domestic demand for goods would be absorbed by imports or profits would be squeezed as labor markets tighten and costs rise. Little pressure on prices would result, as attempts by domestic producers to raise prices would immediately decrease demand for domestic goods and relieve pressure on costs; all this without monetary policy action. Similarly, if the U.S. economy were below full employment, while the rest of the world was beyond it, U.S. inflation would tend to rise.[1] In such a world, monetary policymakers would be less likely to achieve both their inflation and their output targets, since they would no longer control a key determinant of U.S. inflation.

Reprinted from: *New England Economic Review, Federal Reserve Bank of Boston* (July/August 1998) 21–33.

This article begins with a simple description of how global capacity could affect domestic inflation. Several measures of foreign excess capacity are then constructed in order to test the relationship empirically. As "excess capacity" is not directly measurable, several proxies are calculated. First, Phillips curves are estimated for our major trading partners and their NAIRUs calculated, in order to produce gauges of unemployment rate gaps. Additionally, deviations from trend GDP in those countries are examined. Neither of these measures of excess foreign capacity was found to have a direct effect on U.S. inflation.

Concerns about the precision of these estimates of foreign capacity utilization motivate another approach. Several empirical relationships should hold if foreign capacity affects U.S. inflation: Specifically, foreign capacity should affect U.S. import prices, and U.S. import prices should affect U.S. inflation. There is little evidence in support of either of these two links. In total, the results indicate that the Phillips curve, relating some measure of U.S. capacity utilization to U.S. inflation, is alive, if ailing a bit, even as the world gets more integrated.

CHANNELS THROUGH WHICH FOREIGN CAPACITY MIGHT AFFECT U.S. INFLATION

The posited link between global capacity utilization and U.S. inflation is straightforward. Foreign capacity utilization is assumed to affect the prices of foreign goods. Foreign goods prices, then, help determine U.S. import prices. And, finally, U.S. import prices may affect U.S. inflation through a variety of channels, both direct and indirect.

The direct effect of import prices on U.S. inflation is simple arithmetic. The inflation measure examined throughout this article is changes in the Consumer Price Index less food and energy (core CPI) as it captures the goods most directly related to the public's welfare.[2] Many foreign goods prices are included in the CPI. Thus, CPI inflation,

$$\pi_i^{CPI} = \alpha^* \pi_i^M + (1 - \alpha)^* \pi_i^{NM}, \tag{1}$$

is a weighted average of inflation in domestically produced goods prices, π^{NM}, and inflation in imported goods prices, π^M, where α is the share of imported goods in the CPI. Equation 1 illustrates that any change in the dollar price inflation of foreign goods at the consumer level will be directly captured by the domestic inflation measure. Note that this direct effect makes the CPI more likely than the GDP deflator to exhibit any effects from foreign capacity utilization.[3]

The inflation of foreign goods prices could also have several indirect effects on U.S. goods price inflation. First, and perhaps most important, price movements of foreign goods should affect the prices of U.S. goods with which they directly compete. In fact, much of the recent debate attributes the surprisingly low current inflation to foreign competition. It has been argued that U.S. producers cannot raise their prices even when cost pressures begin to appear, because doing so when foreign prices remain moderate would seriously diminish their market share. Thus, the smaller the U.S. producers are in this traded goods market, or the more substitutable the goods, the greater the discipline imposed by foreign prices. This greater discipline could manifest itself in lower wage demands by workers in these domestic firms, in a profit squeeze, or in a decline in U.S. production as a response to rising imports; any of these alternatives would offset or relieve the pressure of U.S. capacity constraints on U.S. prices. Second, movements in foreign goods prices can indirectly affect U.S. goods prices if those foreign goods are important inputs to U.S. goods production. Fluctuations in the prices of foreign goods used as inputs, and particularly permanent changes in these prices, alter the costs, and thus the prices, of U.S.-produced goods and services.

Without these two major indirect channels, global capacity constraints would affect only a small subcomponent of U.S. inflation, since imports represent a small share of total U.S. output. If only the direct effect were operative, inflation for U.S. produced goods and services would still essentially be determined by domestic capacity constraints. Thus, a finding that import price inflation affects the U.S. inflation rate need not imply that foreign goods are acting as a constraint on the price inflation of U.S.-produced goods or that foreign capacity determines the rate of U.S. inflation. For foreign capacity utilization to be a significant determinant of U.S. inflation, import prices must affect the price inflation of domestically produced goods, π^{NM}.

Nevertheless, for several reasons, one might expect foreign capacity utilization to have little effect on U.S. inflation. Movements in the dollar value of the production costs faced by foreign producers, resulting from a change either in their home currency costs or in the value of the dollar, may have little effect on the prices of imported goods in the United States. Several margins—the manufacturer's, the importer's, the retailer's—may be squeezed or expanded by such price changes. For a variety of reasons foreign producers may let profits bear the brunt of any change in the dollar value of production costs. If, for example, the dollar appreciates—reducing the dollar value of foreign costs—foreign firms might allow profits to increase rather than cut U.S. prices, either because they are price followers in the U.S. market or because they perceive the appreciation as temporary. Changes in costs resulting from excess capacity abroad may also be offset by movements in the dollar. Finally, even if changes in foreign prices or the value of the dollar did affect import prices, this change in import prices may have little indirect effect on U.S. goods prices. U.S. and foreign goods may be imperfect substitutes, so that domestic firms do not find that foreign prices greatly affect the demand for their products.

In fact, there is abundant evidence that changes in exchange rates, and thus changes in the dollar costs faced by foreign producers, have only a modest effect on the dollar prices foreign producers charge in the United States. The debate over the degree of "pass-through" has received considerable attention since the mid 1980s when, as shown in Figure 1, dollar import prices failed to fluctuate nearly as much as the value of the dollar itself. The experiences in the mid 1980s resulted in several theoretical explanations for only a partial pass-through.[4] This literature emphasized that large foreign players here, firms that face a downward-sloping demand curve, will pass through only part of a change in the dollar value of their production costs to their U.S. prices. Of course, if foreign producers are small players in the U.S. market, they will tend to take dollar prices as given, with very little pass-through. This theoretical work has also received broad empirical support at the micro level.[5] The industry-level data suggest that foreign producers are likely to change their U.S. prices significantly only if their U.S. competitors do so also, and U.S. competitors alter their prices when strains in capacity in the United States raise their costs. These results imply that import prices should be correlated with capacity utilization in the United States, not capacity constraints abroad.

Since theory is agnostic about whether global capacity determines U.S. inflation, it is left to empirical analysis to answer the question. Accordingly, the next few sections examine the data. Foreign capacity since the early 1970s is calculated, first by estimating the NAIRU in each of our six major trading partners and, alternatively, by taking a simple trend of real GDP in these countries. It is then shown that deviations from full employment or trend GDP in these trading partners have had no apparent effect on U.S. inflation. Examining the

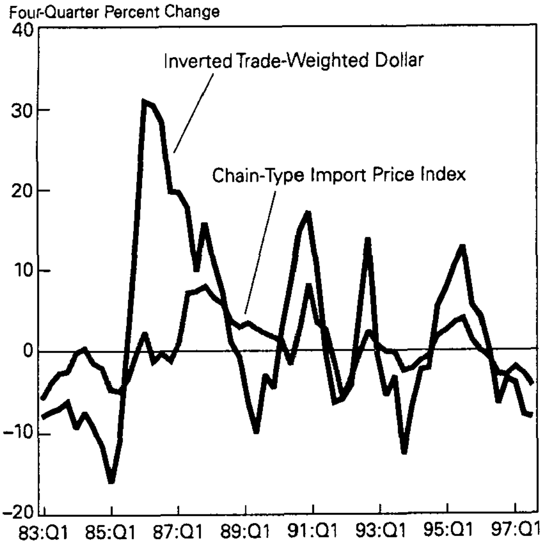


Figure 1 Trade-weighted dollar and import prices. Source: Trade-weighted dollar data, from G.5 release, Federal Reserve Board. Chain-type import price index, from U.S. Bureau of Labor Statistics.

various possible sources of the break in the transmission mechanism could help sidestep the problems with measuring foreign capacity. Either foreign capacity could have little effect on foreign prices, foreign prices could have little correlation with U.S. import prices, or U.S. import prices could have no effect on U.S. inflation. The Phillips curve regressions for the foreign countries in the next section show that the link between foreign capacity utilization and foreign inflation is fairly strong. The subsequent section reveals that the link is weak between foreign inflation and inflation in U.S. import prices. But perhaps more important, U.S. import price inflation fails to have a significant effect on U.S. CPI inflation, and this last result is not dependent on any particular measure of foreign capacity utilization.

MEASURING FOREIGN CAPACITY

In theory, measuring excess capacity is relatively straightforward. Full capacity is defined as the level of GDP or unemployment at which all of the economy's available resources are utilized voluntarily at stable wages and prices (or stable rates of wage and price inflation). Calculating the output level when all inputs are so utilized, and comparing that level to the actual output level of the economy at any given time, is one way of discerning whether the economy is running at, below, or above full capacity. However, actually ascertaining the amounts of labor and capital available to the economy at full employment and determining their productivities is problematic. Since it is almost impossible to measure capacity in this way, other approaches are usually taken.

One commonly used alternative calculates potential output as an average of past actual values, once allowances are made for stable growth. It is argued that, over the long run, prices and wages adjust to ensure that the economy is at full employment. The economy may occasionally operate above or below its capacity, but on average it should run near its potential level. In reality, output may differ from its full-employment level for a significant length of time, so this method of calculating potential output could produce a serious over- or underestimation of capacity.[6] Shifts in potential output, or output growth, also confound this approach; including output levels from the period before any such shift will bias the estimate of potential away from its true value. These failings aside, this is one approach used in this study.

Inflation is an important indicator of whether the economy is above or below full capacity, one that the previous approach ignores. If the economy is operating beyond (below) capacity, inflation will tend to rise (fall). Estimating a country's level of full employment by estimating its Phillips curve captures the information about capacity utilization contained in the inflation numbers.[7] Defining full capacity by its effect on inflation is also the relevant measure for the focus of this study, since foreign capacity affects U.S. inflation through its effect on foreign inflation. For this reason, estimates of foreign capacity derived fròm estimates of the Phillips curve for each of our six major trading partners are also used to examine the effect of foreign capacity utilization on U.S. inflation. As a by-product, estimating foreign Phillips curves tests whether foreign capacity affects foreign inflation, which is the first link in the chain between foreign capacity constraints and U.S. inflation.

Data limitations on foreign prices and unemployment rates make it impossible to estimate this relationship for all of our trading partners.[8] The most reliable data belong to our six major developed trading partners, Canada, France, Germany, Italy, Japan, and the United Kingdom, so capacity utilization in these countries is examined. Although these countries consistently account for roughly 50 percent of all U.S. imports throughout the sample period, it is likely that their capacity utilization would have a disproportionately important effect on U.S. inflation, since imports from these developed economies are more likely to compete directly with U.S. goods production. Thus, these imports are more likely to have both a direct and an indirect effect on U.S. inflation.

The estimates of a simple quarterly Phillips curve for each of our six major trading partners from 1971 to 1996 are presented in Table 1. The basic specification of the Phillips curve equation,

$$\pi_{tj} = \alpha + \sum \beta_j {*} UR_{t-i,j} + \sum \gamma_{t-i,j} {*} \pi_{t-i,j} \qquad (2)$$

for j = *Canada, France, Germany, Italy, Japan, U.K.* was allowed to vary slightly for the six different countries, although imposing the same specification across countries had no

Table 1 The Phillips Curve: Six Major U.S. Trading Partners

	Canada	France	Germany	Italy	Japan	United Kingdom
Constant	1.11*	.41*	.18	1.91*	.39	1.41*
Unemployment	–.13*	–.23	–.04*	–.85*	–.47*	–.71*
Lagged inflation	1.0	1.0	1.0	1.0	1.0	1.0
Estimated NAIRU	8.50*	6.91*	5.10*	9.50*	2.30*	7.30*
	(.35)	(1.33)	(1.17)	(.77)	(.50)	(.93)
Log likelihood	–68.0	–56.4	–55.6	–117.2	–133.3	–151.3

Note: The coefficients on 2 lags of the unemployment rate are summed. The sum of coefficients on 12 lags of inflation is constrained to 1. The sample period is from 1971 to 1996. Standard errors are in parentheses. *Significant at the 5 percent level. Source: Inflation and unemployment data for the six countries were obtained from the OECD, *Main Economic Indicators*.

effect on the results. In all cases the γ coefficients on the lagged domestic inflation realizations were constrained to sum to one. However, a nonlinear relationship between the unemployment rate and inflation was assumed for Italy, Japan, the United Kingdom, and France; in these countries a log linear specification appeared to fit the data better. Further, the speed of the response of inflation to the unemployment rate was allowed to vary by country; the number of quarterly lags of the unemployment rate ranged from two to three.

The second row of Table 1 reveals that each country's unemployment rate had a strong negative effect on that country's rate of inflation. In all but France, the unemployment rate was a significant determinant of foreign inflation, and even in France the unemployment rate was marginally significant.[9]

The next to the last row of Table 1 presents the estimated NAIRUs for each of these countries. They range from a low of 2.3 percent in Japan to a high of 9.5 percent in Italy. These estimates seem to coincide with conventional wisdom; Japan's unemployment rate rarely rose above 3 percent, and the relatively high estimated NAIRUs for France, Italy, Germany, and the United Kingdom are consistent with the diagnosis of Eurosclerosis in the 1980s. The estimate of Canada's NAIRU at 8.5 percent also seems reasonable, as the high unemployment rate that country has experienced over the last several years has reaped a relatively slow payoff in lower inflation.

More important, even for such parsimonious Phillips curves, the NAIRUs are estimated with relative precision.[10] This fact is important since the more precise the estimate of the NAIRU, the more reliable the estimate of foreign excess capacity. Based on the literature hypothesizing hysteresis of the NAIRU in Europe, such as Blanchard and Summers (1986) and Sachs (1986), one might expect the standard errors of the NAIRU estimates for the European countries to be quite large, given the shocks Europe experienced over the sample period. And, in fact, the estimates with the least certainty, France and Germany, are the two countries for which hysteresis in the NAIRU was most seriously discussed in the 1980s. Yet even the NAIRUs for these two countries are estimated fairly reliably. Furthermore, the estimates for Canada and Japan are the most precise, which is not surprising since little has been said about unstable NAIRUs in these countries.

Deviations of GDP from estimates of its potential level are used as an alternative measure of excess capacity in these countries. Potential output is calculated simply as the trend GDP over the previous 26 years. The results attained using deviations from potential GDP are robust to the inclusion of different trends over various parts of the sample, although the more parsimonious model is used in the estimation presented in the tables.

FOREIGN CAPACITY AND U.S. INFLATION

A standard Phillips curve for the United States,

$$\begin{aligned}\pi_t^{US} &= \alpha + \sum \beta_i * UR_{t-i} + \sum \gamma_i * \pi_{t-i}^{US} + \sum \Theta_i * \pi_{t-i}^{NOM} \\ &\quad + \sum \eta_i * \pi_{t-i}^{OIL} + \lambda * Nixon + \tau * Nixoff, \end{aligned} \tag{3}$$

variants of which are found in Gordon (1977, 1982, 1997), Motley (1990), Weiner (1993), Tootell (1994), and Fuhrer (1995), is used here to examine the effects of foreign capacity constraints on U.S. inflation. Foreign capacity utilization, $UR^*_t - UR^{*FE}$, should help determine non-oil import prices, π^{NOM},

$$\pi_t^{NOM} = \sum \beta_i^F * (UR^*_{t-1} - UR^{*FE}), \tag{4}$$

which could affect the U.S. rate of inflation both directly and indirectly.[11] Plugging equation 4 into equation 3,

$$\pi_t^{US} = \alpha + \sum \beta_i * UR_{t-i} + \sum \gamma_i * \pi_{t-i}^{US} + \sum \varphi_i * (UR^*_{t-1} - UR^{*FE}) + \sum \eta_i * \pi_{t-i}^{OIL} + \lambda * Nixon + \tau * Nixoff, \tag{5}$$

produces the base specification examined in Table 2.[12] As is usual, the coefficients on the lagged U.S. inflation variable are assumed to sum to one, to ensure inflation stability at a natural rate.[13]

The NAIRUs used to calculate the foreign excess capacity in equation (5) were derived from the results in Table 1. The measure of foreign excess capacity also requires aggregation of the six different excess capacities. The weights chosen to aggregate these excess capacities were the shares of U.S. imports from each country.[14]

These weights should be a good indicator of the pressure on U.S. import prices and U.S. inflation from a given country's capacity, since that pressure should depend on the country's relative importance in U.S. imports. Both U.S. and foreign capacity are included in the equation. If foreign capacity alone determined U.S. inflation, then measures of U.S. capacity utilization should be insignificant and measures of foreign capacity significant. If foreign excess capacity merely moderates U.S. inflation, particularly since U.S. capacity is an important part of global capacity, then both measures of capacity utilization should be important.

Finally, the effect of oil price movements on U.S. inflation is accounted for separately in these Phillips curve estimations, even though the equations examine the CPI excluding food and energy. The inflation in oil prices must be included because oil is an important input to production of so many goods; it should be included separately from other import prices because, to a large extent, changes in oil prices over this sample period were the result of oligopolistic price-setting behavior that was primarily a function of global politics, not rates of global capacity utilization. As a result, the effect of foreign capacity utilization on U.S. inflation would occur mostly through non-oil import prices.

The coefficients from the estimation of variants of equation (5) are presented in Table 2. The equation is estimated from 1973 to 1996 because the country-specific import weights required to calculate the foreign capacity utilization rate only go back that far.[15] The base specification includes 2 quarterly lags of the unemployment rate, 12 quarterly lags of the inflation rate, and 4 quarterly lags of oil import price inflation. The first column contains the coefficients from a standard Phillips curve—one estimated without foreign capacity. As expected, U.S. unemployment has a significant negative effect on U.S. inflation, and oil prices affect inflation in the expected direction. Since the sample period includes the end of the Nixon wage and price controls, an indicator variable capturing the effect of removing these controls is also included in the regression; its coefficient is correctly signed and of marginal significance.

The second column of Table 2 presents the coefficients from a Phillips curve estimation that adds two lags of the trade-weighted excess of each country's unemployment above its estimated NAIRU. If foreign capacity constraints significantly determine U.S. inflation, either directly or indirectly, the coefficient should be negative and significant. It is

Table 2 Foreign Capacity Utilization and U.S. Inflation

	Full sample: 1973:III–1996:II			Short sample: 1984:III–1996:II		
	Domestic capacity	Foreign capacity added	Nonlinear relation	Domestic capacity	Foreign capacity added	Nonlinear relation
Constant	.72*	.64*	.59	.40*	.28	.28
US unemployment	–.11*	–.10*	–.09*	–.06*	–.05*	–.05*
Lagged inflation	1	1	1	1	1	1
Oil import price inflation	.01*	.01*	.01*	.0001	.0003	.0001
Foreign output gap		–.05	–.30		–.08	–.56
Nixoff	.64*	.63*	.62*			
Observation	92	92	92	48	48	48
Log likelihood	–17.88	–16.46	–15.96	42.76	45.89	46.4

Note: The coefficients on two lags of unemployment rate are summed. The sum of coefficients on 12 lags of inflation is constrained to 1. Foreign Output Gap refers to the trade-weighted deviation of unemployment from the NAIRU in the six countries. *Significant at the 5 percent level. Source: Data for U.S. inflation, unemployment, and import prices are from the *Current Population* Survey, U.S. Bureau of Labor Statistics.

negative but far from significant.[16] Concerns about collinearity between the U.S. and foreign capacity measures can be assuaged, as the estimates of the other coefficients in the equation are little affected by the inclusion of the foreign unemployment gap. In fact, the estimated effect of U.S. unemployment on U.S. inflation remains essentially unchanged from column 1, as is its level of significance. It is not that we cannot tell the difference between the domestic and foreign capacity measures, but that the foreign measure does not appear to affect U.S. inflation.

The exact functional form of any possible relationship between foreign capacity and U.S. prices is not clear. For this reason, column 3 reestimates the Phillips curve allowing for a nonlinear connection between U.S. inflation and foreign capacity. Estimating a nonlinear specification does not increase the importance of foreign capacity. The size of the coefficient is larger in absolute terms because of the change in the functional form.[17] The insignificance of foreign capacity is unaffected by this change in the specification.[18] Other functional forms were tested, with the same result. There is little compelling evidence that this measure of foreign capacity utilization affects either U.S. inflation or the relationship between U.S. capacity and U.S. inflation.

It has been asserted that the importance of foreign capacity to U.S. inflation increased with the trade gap in the 1980s. If so, foreign capacity utilization may have become significantly more important in the mid-1980s. To ensure that the previous regressions are not masking the effect of foreign capacity on U.S. inflation by including the 1970s, the final three columns of Table 2 present the coefficients from reestimates of the first three columns over a shorter sample period, from 1984 on.[19] Foreign capacity remains insignificant.[20] The inclusion of the foreign capacity variable still has little effect on the other coefficient estimates. It appears that foreign capacity utilization has no effect on U.S. inflation, and that the relationship has not strengthened since the 1980s.

As mentioned earlier, one's confidence in the measure of foreign excess capacity in Table 2 depends on the reliability of the estimates of the foreign NAIRUs. To examine the robustness of the results to different measures of foreign capacity utilization,

Table 3 The U.S. Phillips Curve, with Trade-Weighted Output Gaps

	Full sample: 1973:III–1996:II		Short sample: 1984:III–1996:II	
	Domestic capacity	Foreign capacity added	Domestic capacity	Foreign capacity added
Constant	–.07*	–.07*	–.08	–.05
U.S. output gap	9.17*	8.07*	7.25*	4.1*
Lagged inflation	1.0	1.0	1.0	1.0
Oil import price inflation	.01*	.01*	0	0
Foreign output gap		–0.21		1.30
Nixoff	.90*	.92*		
Log likelihood	–20.97	–18.93	–16.23	–14.75

Note: Nixoff is a price control for the period of 1974:Q2–1975:Q1. U.S. Output GAP refers to the deviation of actual income from its potential, and foreign output gap is the trade-weighted deviation of actual income from potential income in the six countries. *Significant at the 5 percent level. Source: U.S. GDP data from the U.S. Bureau of Economic Analysis. Foreign GDP data from the Bank of International Settlements, Switzerland.

Table 3 repeats Table 2 using the deviation of GDP from its trend level for both the United States and, on a trade-weighted basis, the six foreign countries examined in this study. The same six countries are used in order to maintain consistency with the previous table.

The results are identical. The first column of Table 3 shows that estimates of a simple Phillips curve using the U.S. GDP gap are very similar to one using the unemployment rate. The positive and significant coefficient on the U.S. GDP gap indicates that GDP beyond trend significantly increases U.S. inflation, as expected. Oil prices remain significant in this reformulation of the basic Phillips curve. Including the trade-weighted deviation of foreign GDP from its trend, column 2, has little effect on the estimates in column 1; the coefficient estimate on the U.S. GDP gap is roughly the same, and the coefficient on foreign capacity utilization is small, insignificant, and incorrectly signed.[21] The last two columns of Table 3 show that the results, again, do not change if the shorter sample is used.[22] There appears to be little evidence that foreign excess capacity, no matter how it is measured, helps determine U.S. inflation.

THE EFFECT OF FOREIGN CAPACITY ON IMPORT PRICES

It is still possible that the results reveal more about these two measures of foreign capacity than about the effect of foreign capacity on U.S. inflation. For this reason, a different approach is also taken, one that attempts to sidestep the difficulties of quantifying foreign excess capacity. This section and the next examine the different stages of the transmission mechanism through which foreign capacity should affect U.S. inflation. The first stage of the transmission of foreign capacity to U.S. inflation is the effect of foreign capacity on inflation in foreign prices. The foreign Phillips curves in Table 1 show that foreign capacity utilization does affect foreign rates of inflation. Thus, the progression from foreign capacity utilization to U.S. inflation is being broken either at the link between foreign inflation and U.S. import prices or at the link between U.S. import prices and U.S. inflation. This section examines the first of the latter two links.

Table 4 Domestic and Foreign Capacity Utilization and U.S. Import Price Inflation

	Full sample 1974:I–1996:II		Short sample: 1984:I–1996:II	
	Unemployment	GDP GAP	Unemployment	GDP GAP
Constant	–.12	.29	–1.26	.18
U.S. output gap	.087	6.825	.21	3.929
Foreign output gap	–.29	–6.95	.078	–2.42
Non-oil import price inflation	.452*	.4849*	.5855*	.494*
Oil import price inflation	.045*	.0552*	.0058	–.01004
Observations	90	90	50	50
Log likelihood	–143.65	–146.06	–70.2	–65.01

Note: The coefficients on 2 lags of the unemployment rate are summed. The sum of coefficients on 12 lags of inflation is constrained to 1. *Significant at the 5 percent level.

Attempts to explain U.S. import prices are presented in Table 4. These regressions are less grounded in theory or history than the Phillips curve equation, so the results from several specifications were examined. The basic formulation of U.S. non-oil import price inflation,

$$\pi_t^{NOM} = \alpha + \sum \beta_i{*}GAPUS_{t-i} + \sum \theta_i{*}GAPF_{t-i} + \sum \gamma_i{*}\pi_{t-i}^{NOM} + \sum \eta_i{*}\pi_{t-i}^{OIL}, \quad (6)$$

includes the U.S. and the trade-weighted foreign capacity gaps, GAPUS and GAPF, measured either as deviations from trend GDP or as deviations from the estimated NAIRU, and π^{NOM}, non-oil U.S. import price inflation. Thus, non-oil import price inflation is explained using lags of itself, oil import price inflation, the trade-weighted foreign GDP or unemployment gaps, and the U.S. GDP gap or U.S. unemployment rate.[23]

The justification for including the foreign gap is clear; if foreign excess capacity is holding down foreign costs of production, it might restrain the price inflation of foreign goods to U.S. consumers and producers. Alternatively, the level of U.S. capacity might affect import prices, since foreign goods may be priced to market. The lagged oil prices are examined for the same reason as in the U.S. Phillips curve analysis: Inflation in oil prices tends to affect the prices of all goods because oil is such a ubiquitous input to production. The lags of non-oil import price inflation capture the inertia in the inflation rate of these prices, but their inclusion is far less justifiable than that of the lags of the dependent variable in the usual Phillips curve equation, which are usually assumed to capture inflation expectations. As a result, the coefficients on these lagged prices, the γs, are not constrained to sum to 1, particularly since foreign prices and wages will be determined by the expectations of foreign prices in general, rather than simply the prices of their exports to the United States.

The first column in Table 4 shows that neither foreign nor U.S. capacity adds independently to the explanation of import prices when both are included. The significance tends to be higher for the foreign measure of excess capacity. The second column reveals that the same is true when excess capacity is measured as a GDP gap. The final two columns highlight the fact that this effect did not strengthen over the most recent period. Only the coefficients on the lagged inflation terms are significant in these regressions.

As mentioned above, it is unclear which lagged inflation rate should be included as an explanatory variable for import price inflation. One might argue that, just as in the

Phillips curve estimation, foreign inflation in general should be included, since the determinant of the export prices foreigners charge U.S. customers will depend on the expectation of their own inflation rates; firms care about their real return relative to their domestic currency. In this case, lags of the trade-weighted inflation rate as a whole should be included as an explanatory variable. When lags of non-oil import price inflation are replaced with lags of the trade-weighted foreign price inflation, both U.S. capacity, when measured as a GDP gap, and foreign capacity, when measured as an unemployment gap, tend to have a significant effect on U.S. import prices. All in all, however, although there is some evidence that foreign capacity has an effect on U.S. import prices, it is not particularly robust to different specifications, samples, or the measure of capacity utilization used.

THE EFFECT OF IMPORT PRICES ON U.S. INFLATION

Examining the determinants of U.S. import price inflation sheds little light on the insignificance of foreign capacity for U.S. inflation. Thus, the last stage in the transmission of foreign capacity's effect on U.S. inflation is examined in this section. The importance of non-oil import price inflation to U.S. inflation can be directly estimated in the U.S. Phillips curve. If foreign capacity affects U.S. inflation, it must be the case that U.S. import price inflation influences U.S. inflation, either directly or both directly and indirectly. Furthermore, if import prices do significantly affect U.S. inflation, then the weakness of the link between foreign capacity utilization and U.S. import prices would appear to explain the failure of foreign capacity to affect U.S. inflation.

Table 5 presents the coefficients from a simple Phillips curve regression with and without various measures of import price inflation. The first column provides the estimated coefficients of the base specification when oil price inflation is included and non-oil import price inflation is omitted. The sample is extended back to the late 1960s, since the availability of country-specific import data is no longer a binding constraint.[24] The unemployment rate is significant and correctly signed, as is the sum of the coefficients on the oil import price inflation.

The second column of the table presents the coefficient estimates when both oil import price inflation and non-oil import price inflation are included.[25] Oil import price inflation still adds significantly to the explanation of U.S. inflation, while non-oil import

Table 5 Phillips Curve and Import Prices, 1968–1996

	Oil import prices	Oil & Non-oil import prices
Constant	.528*	.458*
U.S. unemployment	−0.876*	−0.84*
Lagged inflation	1.0	1.0
Oil import price inflation	.0105*	.0086*
Non-oil import price inflation		.047
Nixon	−.30*	−.35*
Nixoff	.61	.52
Log likelihood	−18.26	−15.0

*Significant at the 5 percent level.

price inflation does not.[26] These results suggest that any effect of import prices on U.S. inflation derives from their oil component. If, however, import prices were having a significant effect on U.S. inflation, either directly or indirectly, both oil and non-oil import prices should influence U.S. inflation. It appears that the only foreign price changes to significantly affect U.S. inflation were oil price changes.

The lack of a significant relationship between import price inflation and U.S. inflation is surprising.[27] Common sense, and equation 1, tell us that if import prices change radically enough, they must have some effect on U.S. inflation. Table 6 examines this hypothesis by distinguishing episodes when non-oil import price inflation changed significantly from those when the movements were more benign. Non-oil import price inflation is defined as significant if it is a standard deviation above or below its mean value. The first column shows that large changes in non-oil import prices do have a significant effect on U.S. inflation. Columns 2 and 3 examine whether this effect is symmetric: Do large decreases in import prices have the same effect as large increases in import prices? The last two columns of Table 6 reveal that all the significance of the threshold effects comes from large *increases* in non-oil import price inflation. This result is interesting because it is relevant for today's policy debate since import prices are declining, not increasing. There is no evidence that even large declines in import prices affect U.S. inflation.

As imports are a component of the CPI, one would expect that the coefficient on import prices would at least be equal to their share in the CPI, while coefficient estimates greater than that share would suggest pricing discipline on U.S. producers. Although imports are a relatively small share of the CPI, if those prices consistently moved with import prices, the coefficient may be small, but the effect would be statistically significant. A closer examination of the two price series may explain why it is not. The CPI measures the prices paid by consumers at the retail level. The import price index measures the prices paid by the importer at the docks. The importer charges a different price to the distributor, the distributor charges another price to the retailer, and finally the retailer charges still another price to the consumer. The profit squeeze or profit expansion caused by the differential between import prices and U.S. consumer prices may be borne by any of these links in the

Table 6 Phillips Curve and Import Prices: Threshold Values

	Large changes in non-oil import prices	Large movements up in non-oil import prices	Large movements down in non-oil import prices
Constant	.498*	.57*	.519*
Unemployment	–.087*	–.1011*	–.085*
Lagged inflation	1	1	1
Oil import price inflation	.0065*	.0065*	.0114*
Large changes in non-oil import prices	.0796*		
Large upward changes		.0967*	
Large downward changes			.0436
Nixon	–.32*	–.33*	–.31*
Nixoff	.397*	.305	.593*
Log likelihood	–11.28	–11.2	–15.3

*Significant at the 5 percent level.

chain. In other words, the pricing to market is not relevant just to the foreign producer's behavior, but it may also be important to the behavior of the domestic importers, distributors, and retailers. For this reason, the behavior of the dock price of imports may differ significantly from their retail price, especially in the short run.

CONCLUSION

It has been asserted repeatedly that the surprisingly good news on inflation, given the low level of the unemployment rate, is not only good luck but the direct result of increased global competition and excess global capacity. If this statement is true, then policymakers need not worry about the current low level of unemployment since, given the current level of excess global capacity, inflation will not rise. However, if this assertion is correct, several relationships should be clear in the data. Most obviously, global capacity should directly affect U.S. inflation in a traditional Phillips curve. This paper finds no evidence of such a relationship. Furthermore, if foreign capacity did affect U.S. inflation, U.S. import prices would depend on foreign capacity, and U.S. import prices would have a strong relationship to U.S. consumer prices. Neither relationship appears to occur. The results in this study suggest that anyone who believes in a world where we no longer need worry about domestic capacity constraints will eventually be rudely awakened by data that suggest otherwise.

NOTES

1. The response of the exchange rate to these different capacity utilization rates is usually left rather unclear in these arguments. Exchange rate movements should moderate any effects on domestic inflation. In fact, inflation rates routinely vary across countries, as do changes in these rates, with exchange rate movements helping to offset these differences.
2. Most of the conclusions in this paper do not depend on which price measure is used, however. The exceptions will be discussed as they arise.
3. The GDP deflator does not include imports directly. The effect of import prices on both U.S. input costs and U.S. export prices, which are included in the deflator, might produce some reaction of the deflator to foreign capacity utilization. However, this effect would be much less direct than the effect on CPI inflation.
4. Examples of this theoretical discussion can be found in Baldwin (1988) and Dixit (1989).
5. Empirical studies of pass-through can be found in Feenstra, Gagnon, and Knetter (1996), Gagnon and Knetter (1995), and Froot and Klemperer (1989), for example.
6. Many countries in Europe have operated well below full employment through most of the 1980s and 1990s. As a result, such estimates of potential output in these countries will tend to be biased down.
7. One drawback to a simple Phillips curve is that other forces besides each country's domestic capacity could affect inflation, such as oil price shocks or the sources investigated in this paper, import prices.
8. Data on unemployment and inflation rates for some of the currently important exporters to the United States, such as Mexico and China, are unreliable or unavailable for much of the sample. These countries were also much smaller exporters to the United States in the early part of the 1971–96 sample.
9. The data show that with more than 90 percent certainty the unemployment rate in France is an important determinant of its domestic inflation.

10. In fact, examining different specifications for the Phillips curve, such as estimates including oil price shocks, did not significantly alter the estimates of the NAIRUs, even though their inclusion often helped lower the standard errors.
11. Non-oil import prices could also depend on lagged prices and should depend on the exchange rate. Which prices to use—lagged import prices or lagged trade-weighted foreign prices, for example—is unclear. Both of these price measures were also included in this standard specification, along with the trade-weighted value of the dollar, with no effect on the results.
12. Note that any effect of import prices on U.S. capacity will be captured in the estimation through the unemployment rate. Also, Nixon and Nixoff are dummy variables capturing the quarters when wage and price controls were instituted and released.
13. A specification that examined relative price shocks, both import price inflation and oil import price inflation relative to U.S. inflation, was also tested. The results were identical.
14. The shares were based on total U.S. imports from these six countries, so they always summed to one. As a result, there is no trend in the foreign excess capacity variable due to any possible trend in the weights, although this share appeared to be relatively constant around 50 percent.
15. A longer sample was used when estimating the foreign NAIRUs since the import weights were not needed. The longer sample provides more information about each country's NAIRU.
16. The significance is unaffected by the number of lags of foreign capacity included in the equation. Since two lags of the trade-weighted foreign capacity variable are included, the log likelihood ratio testing whether we can reject that the coefficients on the foreign capacity are equal to zero is distributed as a chi-square with two degrees of freedom. The critical value for this ratio is 7.38. Its actual value is 2.8, providing little support that the coefficients are different from zero. In general, this result holds for other inflation measures. For the preferred specification of the chain GDP deflator, the total CPI, core PPI, and the deflator on personal consumption expenditures, foreign capacity has no statistically significant effect. Only for total PPI does foreign capacity appear significant. However, there is little evidence of a robust relationship between these foreign capacity measures and U.S. inflation.
17. In column 3, U.S. inflation is estimated as a linear function of the U.S. unemployment rate and a nonlinear function of the trade-weighted foreign unemployment gap; as a result, the coefficients are not comparable. The linear form is maintained for the U.S. unemployment rate, as it appears to fit the data better.
18. The critical value for the significance of the coefficient on foreign capacity is 7.38, and the log likelihood ratio remains far below that value, at 3.8.
19. The NAIRUs of our six major trading partners were reestimated over the shorter sample because of the concerns that the NAIRUs in Europe rose substantially in the shorter period. Only the estimate of the NAIRU in Italy rose significantly. The results are identical when the NAIRU estimates from the longer sample are used.
20. The critical value for accepting the importance of foreign capacity utilization in U.S. inflation in the shorter sample is 7.38. The actual value of the log likelihood ratio is 6.26, rejecting that foreign capacity plays an important role in the determination of U.S. inflation. The rejection is stronger if the sample begins in 1980.
21. The log likelihood ratio is distributed as a chi-square with two degrees of freedom. Its critical value is 7.38, while the ratio's actual value is 4.08.
22. The log likelihood ratio is again distributed as a chi-square with two degrees of freedom. Its critical value remains 7.38, while its actual value over this shorter sample is 2.96.
23. All the import price inflation regressions were also estimated including the trade-weighted exchange rate. These results are less relevant for this study since it is the total derivative of import prices and foreign capacity utilization that we are concerned about, not its partial derivative holding the change in the exchange rate constant. However, when the exchange rate was included in the import price inflation regressions reported in the text, the foreign capacity variable was more apt to reveal a significant effect on U.S. import prices, although the effect was not very robust to different samples or specifications.

24. The results are identical over the sample examined in Tables 2, 3, and 4, however.
25. The constraint on the coefficients on the lagged price variables in the Phillips curve is invalid if a subcomponent of the index is run independently. However, as will be discussed later, the import deflator prices do differ from the consumer prices, so the import prices are not exactly a subcomponent. Furthermore, examining import price inflation relative to U.S. inflation has no effect on the results.
26. Since four lags of all the import price inflation indexes were examined, the likelihood ratio is distributed as a chi-square with four degrees of freedom, the critical value of which is 11.1. The actual value for the log likelihood ratio for the test of the significance of the coefficients on oil import price inflation is 16.02. The value of the likelihood ratio testing whether non-oil import price inflation adds to the Phillips curve is 6.52, well below the critical value of 11.1.
27. The tests in Table 5 were performed with other measures of U.S. inflation. The results examining total CPI and the deflator for consumer expenditures were consistent with the results above; a specification with a statistically selected lag length of past inflation rates always rejects the importance of non-oil import price inflation in these other consumer price measures. The importance of the non-oil import prices in the PPI is consistently rejected over various specifications and samples. Over some specifications, there is some evidence that non-oil import prices influenced inflation of the GDP deflators, both the fixed-weight and the chain-weighted deflators. Since GDP inflation measures include no imported goods directly, the result might appear surprising. However, the close correlation between export and import prices could explain the result. Even with the GDP deflators, the importance of non-oil import prices is not robust to different specifications of the lags in the prices used or the sample selected. Other inflation measures do not offer strong evidence of a relationship between non-oil import prices and U.S. inflation.

REFERENCES

Baldwin, R. 1988. Hysteresis in Import Prices: The Beach-head Effect. *The American Economic Review,* vol. 78, September, pp. 773–85.

Blanchard, O. and Summers, L. 1986. Hysteresis and the European Unemployment Problem. In Stanley Fischer, ed., *NBER Macroeconomics Annual* 1986, pp. 15–78. Cambridge, MA: MIT Press.

Dixit, A. 1989. Hysteresis, Import Penetration, and the Exchange Rate Pass-Through. *Quarterly Journal of Economics,* vol. CIV, May, pp. 205–28.

Feenstra, R. C., Gagnon, J. E., and Knetter, M. M. 1996. Market Share and Exchange Rate Pass-Through in World Automobile Trade. *Journal of International Economics,* vol. 40, pp. 187–207.

Froot, K. A., and Klemperer, P. D. 1989. Exchange Rate Pass-Through When Market Share Matters. *The American Economic Review,* vol. 79, September, pp. 637–54.

Fuhrer, J. 1995. The Phillips Curve Is Alive and Well. Federal Reserve Bank of Boston, *New England Economic Review,* March/April, pp. 41–56.

Gagnon, J. E., and Knetter, M. M. 1995. Markup Adjustment and Exchange Rate Fluctuations: Evidence From Panel Data on Automobile Exports. *Journal of International Money and Finance,* vol. 14, no. 2, pp. 289–310.

Gordon, R. J. 1977. Can the Inflation of the 1970s Be Explained? *Brookings Papers on Economic Activity,* vol. 8, pp. 253–77.

Gordon, R. J. 1982. Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment. In Neil Baily, ed., *Workers, Jobs, and Inflation,* pp. 89–158. Washington, DC: The Brookings Institution.

Gordon, R. J. 1997. The Time-Varying NAIRU and Its Implications for Economic Policy. *Journal of Economic Perspectives,* vol. 11, no. 1, pp. 11–32.

Motley, B. 1990. Has There Been a Change in the Natural Rate of Unemployment? Federal Reserve Bank of San Francisco, *Economic Review,* Winter (no. 1), pp. 3–16.

Sachs, J. D. 1986. High Unemployment in Europe: Diagnosis and Policy Implications. NBER Working Paper no. 1830, February.

Staiger, D. J., Stock, J. H. and Watson, M. W. 1997. The NAIRU, Unemployment, and Monetary Policy. *Journal of Economic Perspectives,* vol. 11, no. 1, pp. 33–50.

Tootell, Geoffrey M. B. 1994. Restructuring, the NAIRU, and the Phillips Curve. Federal Reserve Bank of Boston, *New England Economic Review,* September/October, pp. 31–44.

Weiner, Stuart E. 1993. New Estimates of the Natural Rate of Unemployment. Federal Reserve Bank of Kansas City, *Economic Review,* vol. 78, no. 4, pp. 53–69.

44

On the Origin and Evolution of the Word Inflation

Michael F. Bryan
Assistant Vice President and Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

> Inflation is the process of making addition to currencies not based on a commensurate increase in the production of goods.
>
> —*Federal Reserve Bulletin* (1919)

> Most prominent among these inflationary forces were a drop in the exchange rate of the dollar, a considerable increase in labor costs, and severe weather.
>
> —*Federal Reserve Bulletin* (1978)

For many years, the word *inflation* was not a statement about prices but a condition of paper money—a specific description of a monetary policy. Today, *inflation* is synonymous with a rise in prices, and its connection to money is often overlooked.

Consider the opening quotations, taken from *Federal Reserve Bulletins* spanning a period of almost 60 years. The first defines inflation as a condition of the currency, while the latter makes no reference to money. Indeed, it would seem that in 1978, inflation was about many things other than excessive money growth.

This chapter considers the origin and uses of the word *inflation* and argues that its definition was a casualty in the theoretical battle over the connection between money growth and the general price level. What was once a word that described a monetary cause now describes a price outcome. This shift in meaning has complicated the position of anti-inflation advocates. As a condition of the money stock, an inflating currency has but one origin—the central bank—and one solution—a less expansive money growth rate. But as a condition of the price level, which may have originated from a variety of things (including a depreciating dollar, rising labor costs, bad weather, or a number of factors other than "too much money"), the solution to—and the prudence of—eliminating inflation is much less clear.

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (Oct. 15, 1997).

VALUE, MONEY, AND CURRENCY

> I smiled at myself at the sight of all this money. "Oh, drug," said I, aloud, "What art thou good for? Thou art not worth to me, no not the taking off the ground. One of these knives is worth all this heap."
>
> —Daniel Defoe (1719) *Robinson Crusoe*

The classical economists, by which I refer to the generation writing around the time that Adam Smith's *The Wealth of Nations* was published in 1776, were very exact in defining economic terms, because they were constructing a language on which an emerging science was being built. Among their first contributions was to make explicit the distinction between "real" and "nominal" prices.

A good's real price, or *value,* was defined as the effort required to produce it, while its nominal, or *money,* price was said to be its cost in money alone (fixed in terms of gold or some other precious metal). According to this view, the value of goods is anchored by the laws of nature—the effort of labor—but their nominal price fluctuates with the availability of the precious metal, and the laws of the sovereign, that define a nation's money.

> The real price of everything... is the toil and trouble of acquiring it. The same real price is always of the same value; but on account of the variations in the value of gold and silver, the same nominal price is sometimes of very different values.
>
> —Adam Smith (1776)[1]

Although the classical economists supposed that fluctuations in the money price of goods can have temporarily disruptive influences on the economy (such as producing capricious redistributions of wealth between parties bound by contracts with fixed money prices), in the end, these changes merely serve to alter the scale by which value is measured. They do not alter values or have any long-term consequences. The idea that changes in the quantity of money affect only the money price of goods, not their value, was championed by many of the early classical economists, most notably David Hume. The theory was more rigorously developed in the early twentieth century by American economist Irving Fisher, and has come to be known as the "quantity theory of money."

> If the history of commercial banking belongs to the Italians and of central banking to the British, that of paper money issued by a government belongs indubitably to the Americans.
>
> —John Kenneth Galbraith (1975)

To these early economists, the word *money* almost always referred to a metallic coin.[2] But the first generation of economists following Adam Smith in the nineteenth century was very interested in paper money, since this form of payment had become popular in the burgeoning American colonies.[3] The colonies offered a large variety of paper currencies, virtually all of which were conspicuous by their "overproduction" and their subsequent rapid loss of purchasing power.

The Continental Congress issued a paper note to help finance the American Revolution, and these "bills of credit" became a circulating medium.[4] In 1775, Congress issued $6 million of the new currency and urged the states to impose taxes for its ultimate redemption. The taxes were never raised, however, and larger continental issues were authorized. By the end of 1779, Congress had expanded the number of continental bills of credit more than 40-fold, and, to make matters worse, the states had issued their own paper monies in a similar magnitude.[5] In 1781, a dollar in paper was worth less than two cents in gold coin.[6]

By the early nineteenth century, economists were careful to distinguish among three sources of a change in the "cost" of goods—changes in *value,* referring to the real resource cost of a good, changes in *money prices,* caused largely by fluctuations in the metallic content of money, and *depreciation of the currency,* caused by a change in the quantity of currency relative to the metal that constitutes a nation's money. The latter distinction would become a focal point in American political economy.

INFLATION OF THE CURRENCY

The era between the mid-1830s and the Civil War—a period economists refer to as the "free banking era"—saw a proliferation of banks. Along with these institutions came "bank notes," a private paper currency redeemable for a specific amount of metal. That is, if the issuing bank had it. At times, banks did not have enough gold or silver to satisfy all of their claims. Bank notes, like the public notes that preceded them, also tended to depreciate. It is during this period that the word *inflation* begins to emerge in the literature, not in reference to something that happens to prices, but as something that happens to a paper currency.[7]

> The astonishing proportion between the amount of paper circulation representing money, and the amount of specie actually in the Banks, during the past few years, has been a matter of serious concern . . . [This] inflation of the currency makes prices rise.[8]
>
> —From the Bee (1855)[9]

During the Civil War, both the federal and the confederate governments issued paper currency to help finance expenditures. The federal government authorized the issue of $450 million of a paper money called "Greenbacks," and at war's end, President Johnson authorized the Treasury to repay these notes with gold. This reduced the outstanding Greenbacks by about 20 percent and had the predictable effect of propping up the "value" of Greenback dollars, or driving down the Greenback price of goods.

Restoring the purchasing power of Greenbacks worked in favor of creditors, since it meant they would be repaid in a currency that had greater purchasing power than would otherwise have been the case. But of course, what worked to the advantage of creditors worked to the disadvantage of debtors, who found their dollar liabilities rising in value. Debtor groups, predominantly farmers, advocated "inflating the Greenback" as a means of easing the debt burdens of borrowers and perhaps helping to redistribute income from the eastern to the western constituencies. In the election of 1868, the Democratic party endorsed the "Ohio Idea," which proposed that war debts be repaid with Greenbacks unless otherwise stipulated.[10] These predominantly western Democrats became known as "Inflationists."

Despite the election of Republican candidate Ulysses S. Grant to the presidency, Inflationist sentiment carried considerable influence in Congress. The movement was given further support by the Supreme Court decision of 1870, which reversed an earlier ruling and declared that the issuance of paper money as "legal tender" was constitutional. In 1874, Congress passed the "Inflation Bill," which provided for the additional issuance of $14 million in Greenbacks. President Grant vetoed the measure and resumed bond redemption in terms of coin.

The idea that the government can "create value" by issuing a paper money and merely stating that it is of value is in direct conflict with the quantity theory of money—and it was a subject of considerable scorn.

PRICE INFLATION

The term *inflation* was initially used to describe a change in the proportion of currency in circulation relative to the amount of precious metal that constituted a nation's money. By the late nineteenth century, however, the distinction between "currency" and "money" was becoming blurred.

> It has been rather the fashion with political economists to refuse the name Money to any medium of exchange which is not "a material recompense or equivalent." . . . For myself, I can see no valid objection to the scientific acceptance of the popular term, Paper Money. The presence of the word paper so far qualifies and explains the word money, as to show that a material recompense or equivalent is not meant.
>
> —Francis A. Walker (1883)

At the turn of the century, economists tended to refer to *any* circulating medium as money, and any change in the circulating medium relative to trade needs as an inflation of money. But this shift in meaning introduced another problem. Although it is easy to determine the amount of currency relative to the stock of a precious metal, how does one know when the amount of the circulating medium exceeds "trade needs"?

> Many current controversies about inflation are due not to conflicting ideas but to conflicting uses of the same word. When a nation has too much money, it is said to have inflation: that is about as near as we can come to an accepted definition of the term. As to what constitutes having too much money, there is not agreement . . . If we use the term inflation to denote any increase in the volume of money that is accompanied by a rise in the general price-level, we confine ourselves to a definite and logical use of the term, and one that directs attention at once to the practical monetary problem with which business is to-day chiefly concerned.
>
> —William Trufant Foster and Waddill Catchings (1923)

Economists appear to have reached a definitional crossroads during the first several decades of the twentieth century. Presumably, because they could be certain of the "excessiveness" of the circulating medium only by its effect on the price level, the notions of an inflated currency and prices became inextricably linked.

Consider the following quotations, from works published by the same economist at two different times during this period:

> . . . inflation occurs when, *at a given price level,* a country's circulating media—cash and deposit currency—increase relatively to trade needs. (Emphasis in original.)
>
> —Edwin Walter Kemmerer (1918)

> Inflation exists in a country whenever the supply of money and of [circulating] bank deposits . . . increases, relatively to the demand for media of exchange, in such a way *as to bring about a rise in the general price level.* (Emphasis added.)
>
> —Edwin Walter Kemmerer (1934)

In the earlier definition, inflation is something that happens to the circulating media at a given price level; in the later definition, an inflating currency is defined to exist when it produces a rise in the general price level, as suggested by the quantity theory. What originally described a monetary cause came to describe a price effect.

To the quantity theorists, this shift in emphasis may have had little direct consequence, since it is unlikely they could have seen an important distinction between these two ideas. Of course, increasing the quantity of currency relative to "trade needs" could have but one effect—to make prices rise. And a rising price level could have but one origin—an increase in the quantity of money relative to its demand!

Still, some economists attempted to maintain the distinction between a rise in the price level that originated from the "creation" of additional currency relative to trade, and one that resulted from a decrease in trade for a given supply of money. It was the former, not the latter, that caused problems for economies whose trade was conducted via paper money.

> Just as we can increase the size of a balloon either by pumping in more air, or by decreasing the outside pressures, . . . we can increase prices either by pumping more dollars into the monetary circulation, or by decreasing the pressure of the work that money has to perform. It seems best, however, not to extend the term inflation to cover failures to reduce the money in circulation when prices begin to rise. Such an extension of the use of the term would be at variance with its derivation, and would, moreover, leave the term less definitely applicable to the actual, current monetary problems of the world.
>
> —William Trufant Foster and Waddill Catchings (1923)[11]

Linking *inflation* to the price level proved to be another important turning point for the word. With the publication of John Maynard Keynes' *General Theory* in 1936, an assault on the quantity theory of money commenced, and it dominated macroeconomic thought for the next 40 years. By appealing to the belief that resources could be regularly and persistently underemployed—an idea given support by the worldwide depression of the time—Keynesian theory challenged the necessary connection between the quantity of money and the general price level. Moreover, it suggested that aggregate price increases could originate from factors other than money.

In addition to separating the price level from the money stock, the Keynesian revolution in economics appears to have separated the word *inflation* from a condition of money and redefined it as a description of prices. In this way, inflation became synonymous with *any* price increase. Indeed, Keynes spoke about different "types" of inflation, including income, profit, commodity, and capital inflation. Today, little distinction is made between a price increase and inflation, and we commonly hear reports of energy inflation, medical care inflation, and even wage inflation. Some go so far as to argue that the monetary definition forces the word to take on too specific a meaning:

> Even if we agree that an inflationary situation is to be taken to imply something about prices, precise definitions vary. . . . Part of the difficulty here is that definitions of the more popular variety such as "too much money chasing too few goods," not only purport to define inflation, but also imply something more about particular inflationary processes.
>
> —R. J. Ball (1964)

CONCLUSION

"That's a great deal to make one word mean," Alice said in a thoughtful tone. "When I make a word do a lot of work like that," said Humpty Dumpty, "I always pay it extra."

—Lewis Carroll (1872) *Through the Looking Glass*

Inflation, a term that first referred to a condition of the currency and later to a condition of money, is now commonly used to describe prices. This shift in meaning seems to have originated in an unfortunate—but perhaps inevitable—sequence of events. By referring to inflation as a condition of "too much money," economists were forced to struggle with the operational issue of "how much is too much?" The quantity theory offered a clear answer to that question: Too much money is an increase in the money stock that is accompanied by a rise in the general price level. In other words, an inflated money supply will reveal itself through its effect on the price level. When Keynesian economic theory challenged the direct link between money and the price level, inflation lost its association with money and came to be chiefly understood as a condition of prices.

Without being tied to the money supply, any price increase seems to have an equal claim to the word *inflation.* Indeed, today we regularly read reports of a seemingly endless variety of "inflations." When the word is used as a description of the price level, an anti-inflation policy can easily be characterized as being against any price increase, including higher wages! This is simply not the case. An anti-inflation strategy is concerned with a particular type of price increase—a rise in the general price level stemming from excessive money creation. When viewed in this light—the light provided by the word's original meaning—a zero-inflation objective for the central bank becomes a much more sensible goal.

NOTES

1. The idea that value is fixed by labor effort, called the "labor theory of value," is now generally discredited by economists. Still, we make clear distinctions between a good's real cost and its money cost.
2. Western economists of the time were certainly aware of paper money. Chinese notes called "chao" were known to have been used as early as the ninth century (they were also said to have depreciated rapidly in value).
3. A common lament in the New World was that paper money was necessary because of a lack of metallic coins.
4. Some historians note that the decision to issue continental currency was made in the conventions that occurred prior to the establishment of the Continental Congress.
5. See Charles J. Bullock, *Essays on the Monetary History of the United States,* New York: Macmillan, 1990, pp. 64–5.
6. The French also issued a paper money—"assignats"—around the time of their Revolution, with a similar result: They, too, rapidly lost their purchasing power. The French experience with paper money gave rise to the saying, "After the paper money machine comes the guillotine."
7. Bank notes were taxed out of existence by an act of Congress in 1865.
8. This is the earliest reference to *inflation* in the Federal Reserve Bank of Cleveland's library. *The Oxford English Dictionary* shows the earliest reference to be from D. D. Barnard (1838): "The property pledge can have no tendency whatever to prevent an inflation of the currency."
9. *Gold and the Currency: Specie Better than Small Bills.* Boston: Evans and Plumber, 1855.

10. However, "sound money man" Horatio Seymour, the reluctant Democratic candidate for the presidency in 1868, is said to have indicated that if elected, he would not support the plan.
11. Similar in spirit are the following:

> . . . we must distinguish between inflation and the rise in prices. The one is not necessarily synonymous with the other . . . An alteration in the general price level accordingly means a change in the relation between goods on the one hand and money on the other. Obviously, however, such a change in the relation may be ascribable, in its origin, to either of the two elements, the goods or the money.
>
> —Edwin R. A. Seligman (1921)

> Either the rise in prices might be due to the scarcity of goods or it might be due to the superabundance of money, but as a matter of actual historical fact it is, so far as I know, universally true . . . that it is the change in the money that makes the changes in the value of the money, and not changes in the goods.
>
> —Irving Fisher

This Page Intentionally Left Blank

45

Inflation, Growth, and Financial Intermediation

V. V. Chari
Professor of Economics, University of Minnesota, Minneapolis, Minnesota

Larry E. Jones
Professor of Managerial Economics, Northwestern University, Evanston, Illinois

Rodolfo E. Manuelli
Professor of Economics, University of Wisconsin, Madison, Wisconsin

The key role of government policies in the process of development has long been recognized. The recent availability of quality data has led to quantitative analyses of the effect such policies have on development. Most of the renewed research effort on this front, both theoretical and empirical, has emphasized the relationship between fiscal policy and the paths of development of countries.[1] Although there have been several empirical studies on the relationship between monetary policy and growth,[2] there has been very little theoretical work in this area.[3] We have two goals in this article. One is to summarize the recent empirical work on the growth effects of monetary policy instruments. The other is to compare the empirical findings with the implications of quantitative models in which monetary policy can affect growth rates. We ask, in particular, What is the relationship in the data between monetary policy instruments and the rate of growth of output? Are the predicted quantitative relationships from theoretical models consistent with the data?

Monetary policy plays a key role in determining inflation rates. In the next section, we summarize the empirical evidence on the relationship between inflation and growth in a cross-section of countries. This evidence suggests a systematic, quantitatively significant negative association between inflation and growth. Although the precise estimates vary from one study to another, evidence suggests that a 10 percentage point increase in the average inflation rate is associated with a decrease in the average growth rate of somewhere between 0.2 percentage points and 0.7 percentage points.

Some researchers are tempted to view this link as implying that if a country conducts monetary policy so as to lower its inflation rate by 10 percentage points, its growth rate will rise by anywhere from 0.2 percentage points to 0.7 percentage points. Obviously, the data alone cannot give us an answer to the policy question we care about. Therefore we explore the ability of various models with transactions demand for money to account for this asso-

Reprinted from: *Federal Reserve Bank of St. Louis Review* (May/June 1996) 9–29.

ciation. We use the growth rate of the money supply as our measure of the differences in monetary policies across countries. Although many models predict qualitatively that an increase in the long-run growth rate of the money supply decreases the long-run growth rate of output in the economy, we find that in these models, a change in the growth rate of the money supply has a quantitatively trivial effect on the growth rate of output. The reason is that in endogenous growth models, changes in output growth rates require changes in real rates of return to savings, and it turns out that changes in inflation rates have trivial effects on real rates of return and thus on output growth rates.

We go on, then, to broaden our notion of monetary policy to include financial regulations. We study environments in which a banking sector holds money to meet reserve requirements. We model banks as providing intermediated capital, which is an imperfect substitute for other forms of capital, and we consider three kinds of experiments.

In the first we hold reserve requirements fixed and examine the effects of changes in inflation rates on growth rates. Even though higher inflation rates distort the composition of capital between bank-intermediated capital and other forms of capital and thus reduce growth rates, the quantitative effects turn out to be small.

In the second kind of experiment, we simultaneously change money growth rates and reserve requirements in a way that is consistent with the association between these variables in the data. This avenue is promising because these variables are positively correlated, and changes in each of them have the desired effect on output growth rates. We find that monetary policy changes of this kind have a quantitative effect on growth rates that is consistent with the lower end of the estimates of the relationship between inflation rates and growth rates.

Our third experiment uses data on inflation rates and cash held by banks in each country to compute our model's implications for growth in that country. We regress growth on inflation using the data generated by our model and find that a reduction in inflation rates of 10 percentage points is associated with an increase in growth rates of as much as 0.08 percent.

Thus, although our models cannot reproduce the large association between inflation rates and growth rates found in the data, the policy implication is that reductions in inflation rates can indeed generate substantial increases in growth rates. We conclude by arguing that models which focus on the transactions demand for money alone cannot account for the sizable negative association between inflation and growth, while models that focus on the distortions caused by financial regulations can.

THE EVIDENCE ON INFLATION AND GROWTH

Numerous empirical studies analyze the relationship between the behavior of inflation and the rate of growth of economies around the world. Most of these studies are based on (some subset of) the Summers and Heston (1991) data sets and concentrate on the cross-sectional aspects of the data that look at the relationship between the average rate of growth of an economy over a long horizon (typically from 1960 to the date of the study) to the corresponding average rate of inflation over the same period and other variables. Some of the more recent empirical studies undertake similar investigations using the panel aspects of the data more fully.[4]

To summarize this literature, we begin with some simple facts about the data. According to Levine and Renelt (1992), those countries that grew faster than average had an average inflation rate of 12.34 percent per year over the period, while those countries that grew more slowly than average had an average inflation rate of 31.13 percent per year.[5] Similar results are reported in Easterly et al. (1994). Here fast growers are defined as those countries having a growth rate more than one standard deviation above the average (and averaging about 4 percent per year) and are found to have had an average inflation rate of 8.42 percent per year. In contrast, slow growers, defined as those countries having a growth rate more than one standard deviation below the average (and averaging about 0.2 percent per year), had an average inflation rate of 16.51 percent per year. Using the numbers from either Levine and Renelt (1992) or Easterly et al. (1994) to estimate an unconditional slope (which those studies do not do), we see that a 10 percentage point rise in the inflation rate is associated with a 5.2 percentage point fall in the growth rate. These groups of countries also differ in other systematic ways. For example, fast growers spent less on government consumption, had higher investment shares in gross domestic product (GDP), and had lower black-market premiums. However, this association between inflation and growth suggests that monetary policy differences are important determinants in the differential growth performances present in the data.[6]

In two recent studies, Fischer (1991 and 1993) analyzes the Summers and Heston (1991) data, using both cross-sectional and panel-regression approaches to control for the other systematic ways in which countries differ from one another. Fischer (1991) controls for the effects of variables such as initial income level, secondary school enrollment rate, and budget deficit size and finds that on average, an increase in a country's inflation rate of 10 percentage points is associated with a decrease in its growth rate of between 0.3 percentage points and 0.4 percentage points per year. Similar results are reported by Roubini and Sala-i-Martin (1992), who find that a 10 percentage point increase in the inflation rate is associated with a decrease in the growth rate of between 0.5 percentage points and 0.7 percentage points.[7] In his article in this issue of the Review, Barro, using a slightly different framework to control for the effect of initial conditions and other institutional factors, also finds a negative effect of inflation on growth that he estimates to be between 0.2 percentage points and 0.3 percentage points per 10 percentage point increase in inflation. He also finds the relationship to be nonlinear, although—contrary to the other studies—he estimates that the greater effect of inflation on growth comes from the experiences of countries in which inflation exceeds a rate of between 10 percent and 20 percent per year.

In summary, the standard regression model seems to suggest that a 10 percentage point increase in the inflation rate is associated with a decrease in the growth rate of between 0.3 percentage points and 0.7 percentage points.[8] Are these growth effects of higher inflation significant? As an illustration of the importance of these effects, note the difference in income levels between two countries that are otherwise similar but which have a 10 percentage point difference in annual inflation rates. Although these countries start in 1950 with the same levels of income, their income levels would differ by a factor of between 16 percentage points and 41 percentage points by the year 2000 (starting with the average growth rate of 1.92 percent per year as the base).[9]

MODELS OF GROWTH AND MONEY DEMAND

Two theoretical arguments in the literature concern the effect on output of changing the average level of inflation. One argument is based on what has become known as the Mundell-Tobin effect, in which more inflationary monetary policy enhances growth as investors move out of money and into growth-improving capital investment. The evidence we have summarized seems to contrast this argument sharply, at least as a quantitatively important alternative. The other argument is based on the study of exogenous growth models. In an early paper in this area, Sidrauski (1967) constructs a model in which a higher inflation rate has no effect on either the growth rate or the steady-state rate of output. Other authors construct variants in which higher inflation rates affect steady-state capital output ratios but not growth rates.[10]

In this section we analyze a class of endogenous growth models in an attempt to better understand the empirical results presented in the previous section. The regression results presented there implicitly ask what the growth response will be to a change in long-run monetary policy that results in a given percentage point change in the long-run rate of inflation. Thus our goal here is to describe models in which monetary policy has the potential for affecting long-run growth. Three elements are obviously necessary in a candidate model: It must generate long-run growth endogenously, it must have a well-defined role for money, and it must be explicit about the fiscal consequences of different monetary policies.

In contrast to the neoclassical family of exogenous growth models, the feature necessary for a model to generate long-run growth endogenously is that the rate of return on capital inputs does not go to zero as the level of inputs is increased, when the quantities of any factors that are necessarily bounded are held fixed. Stated another way, the marginal product of the reproducible factors in the model must be bounded away from zero.[11]

We report results for four types of endogenous growth models[12]:

- A simple, one-sector model with a linear production function (Ak)
- A generalization of the linear model that endogenizes the relative price of capital (two-sector)
- A model that emphasizes human capital accumulation (Lucas)
- A model with spillover effects in the accumulation of physical capital (Romer)

To generate a role for money in these models, a variety of alternatives is available. We report results for three models of money demand:

- A cash/credit goods model in which a subset of goods must be purchased with currency [cash in advance (CIA) in consumption]
- A shopping time model in which time and cash are substitute inputs for generating transactions (shopping time)
- A CIA model in which all purchases must be made with currency, but in which cash has a differential productivity between consumption and investment purchases (CIA in everything)

Although these models are only a subset of the available models, we think that the combinations of the various growth and money demand models represent a reasonable cross-section.

Finally, we must specify how the government expands the money supply. We restrict attention to policy regimes in which households are given lump-sum transfers of money. In all the models we examine, the growth effects of inflation that occur when money is distributed lump sum are identical to those which occur when the growth of the money supply is used to finance government consumption, as long as the increased money supply is not used to fund directly growth-enhancing policies. Alternative assumptions about the uses of growth of the money supply may lead to different conclusions about the relationship between inflation and growth. For example, using the growth of the money supply to subsidize the rate of capital formation or to reduce other taxes may stimulate growth. Since the evidence suggests that inflation reduces growth, we restrict attention to lump-sum transfers.

The growth and money demand models just listed give us 12 possible models. Rather than give detailed expositions of each of the 12 models, we will discuss the Lucas model with CIA in consumption. Full details of the balanced growth equations for each of the 12 models are presented in Chari, Jones, and Manuelli (1995).

A REPRESENTATIVE MODEL OF GROWTH AND MONEY DEMAND

We consider a representative agent model with no uncertainty and complete markets. In this model, there are two types of consumption goods, called cash goods and credit goods, in each period. Cash goods must be paid for with currency. Both of these consumption goods, as well as the investment good, are produced using the same technology. The resource constraint in this economy is given by

$$c_{1t} + c_{2t} + x_{kt} + x_{ht} + g_t \leq F(k_t, n_t h_t), \tag{1}$$

where c_{1t} is the consumption of cash goods; c_{2t} is the consumption of credit goods; x_{kt} and x_{ht} are investment purchases in physical capital and human capital, respectively; k_t is the stock of physical capital; n_t is the number of hours worked; h_t is the stock of human capital; g_t is government consumption; and F is the production function. Physical capital follows $k_{t+1} \leq (1 - \delta_k)k_t + x_{kt}$, where δ_k is the depreciation rate, while human capital follows $h_{t+1} \leq (1 - \delta_h)h_t + x_{ht}$, where δ_h is the depreciation rate on human capital.

Trading in this economy occurs as follows: At the beginning of each period, a securities market opens. In this market, households receive capital and labor income from the previous period, the proceeds from government bonds, and any lump-sum transfers from the government. At this time, households pay for credit goods purchased in the previous period. Finally, households must choose how much cash they will hold for the purchase of cash goods in the next period.

The consumer's problem is to

$$\max \sum_{t=0}^{\infty} \beta^t u(c_{1t}, c_{2t}, 1 - n_t), \tag{2}$$

subject to

$$m_{t-1} + b_{t-1} \leq v_t \tag{3}$$

$$p_t c_{1t} \leq m_{t-1} \tag{4}$$

$$\begin{aligned} v_{t+1} \leq {} & (v_t - b_{t-1} - m_{t-1}) + (m_{t-1} - p_t c_{1t}) - p_t c_{2t} - p_t x_{kt} - p_t x_{ht} \\ & + p_t r_t k_t (1 - \tau) + p_t w_t n_t h_t (1 - \tau) + [1 + (1 - \tau) R_t] b_{t-1} + T_t \end{aligned} \tag{5}$$

$$k_{t+1} \leq (1 - \delta_k)k_t + x_{kt} \tag{6}$$

$$h_{t+1} \leq (1 - \delta_h)h_t + x_{ht}, \tag{7}$$

where β is the discount factor; u is the consumer's utility; v_t is wealth at the beginning of period t; m_{t-1} is money holdings at the beginning of period t; b_{t-1} is bond holdings at the beginning of period t; R_t is the nominal interest rate paid on bonds during period t; r_t is the rental price of capital during the period; τ is the tax rate on income (assumed constant); T_t

is the size of the transfer to the household delivered at the end of period t; and w_t is the real wage rate. Note that we have adopted the standard assumption from the human capital literature that firms hire effective labor $n_t h_t$ from workers and pay a wage of w_t per unit of time.[13] Since all four goods available in a period (c_1, c_2, x_k, and x_h) are perfect substitutes on the production side, they all sell for the same nominal price p_t.

On the production side, we assume that there is a representative firm solving the static maximization problem

$$\max p_t[F(k_t,n_t h_t) - r_t k_t - w_t n_t h_t]. \tag{8}$$

Let M_t be the aggregate stock of money and μ be the (assumed constant) rate of growth of the money supply.

Equilibrium for the model requires maximization by both the household and the firms, along with the following conditions:

$$c_{1t} + c_{2t} + x_{kt} + x_{ht} + g_t \leq F(k_t,n_t h_t) \tag{9}$$

$$m_t = M_t \tag{10}$$

$$T_{t+1} = M_{t+1} - M_t = (\mu - 1)M_t \tag{11}$$

$$g_t = \tau F(k_t,n_t h_t). \tag{12}$$

The first two of these conditions are market-clearing in the goods market and the money market, respectively. Conditions 11 and 12 describe the characteristics of policy in the model. Condition 11 says that the increase in the money supply enters the system through a direct lump-sum transfer to the household. Finally, condition 12 says that government purchases are financed by a flat-rate tax on income. An implication of conditions 11 and 12 is that the government's budget is balanced on a period-by-period basis.

To study the long-run behavior of the model, we use the solutions to the maximization problems of the household and the firm together with equilibrium conditions 9 through 12 to calculate what are known as the balanced growth equations. Along a balanced growth path, output grows at a constant rate. In general, for the economy to follow such a path, both the production function and the preferences must take on special forms. On the production side, a sufficient condition is that F(k,nh) is a Cobb-Douglas production function of the form $Ak^\alpha(nh)^{1-\alpha}$, where A and α are parameters. On the preference side, the consumer, when faced with a stationary path of interest rates, must generate a demand for constant growth in consumption. This requirement is

$$U(c_{1t},c_{2t}) = [(c_{1t}^{-\lambda} + \eta c_{at}^{-\lambda})^{-1/\lambda}]^{(1-\sigma)} (1 - n_t)^{\psi(1-\sigma)}/(1 - \sigma), \tag{18}$$

where η,λ,σ, and Ψ are preference parameters. With these assumptions, we can show that the dynamics of the system converge to a balanced growth path.[14]

For this model, the balanced growth equations of the system are

$$c_2/c_1 = \{\eta[1 + (1 - \tau)R]\}^{1/(1+\lambda)} \tag{14}$$

$$\gamma^\sigma = \beta[1 - \delta_k + \alpha A n^{1-\alpha}(h/k)^{1-\alpha}(1 - \tau)] \tag{15}$$

$$\gamma^\sigma = \beta[1 - \delta_h + (1 - \alpha)An^{1-\alpha}(h/k)^{-\alpha}(1 - \tau)] \tag{16}$$

$$\gamma^\sigma \pi = \beta[1 + (1 - \tau)R] \tag{17}$$

$$[(1 - n)/n^\alpha]\,(h/k)^{1-\alpha}(1 - \alpha)A = (c_1/k)\Psi[1 + \eta(c_2/c_1)^{-\lambda}][1 + (1 - t)R] \tag{18}$$

$$\pi\gamma = \mu \tag{19}$$

$$\lambda = 1 - \delta_k + (x_k/k) \tag{20}$$

$$\lambda = 1 - \delta_h + (x_h/k)(k/h) \tag{21}$$

$$(c_1/k) + (c_2/k) + (x_h/k) + (x_k/k) + (g/k) = An^{1-\alpha}(h/k)^{1-\alpha}, \tag{22}$$

where $\pi = p_{t+1}/p_t$ is the steady-state level of inflation; $\gamma = c_{1t+1}/c_{1t} = c_{2t+1}/c_{2t} = x_{kt+1}/x_{kt} = x_{ht+1}/x_{ht} = k_{t+1}/k_t$ is the growth rate of output; $c_2/c_1 = c_{2t}/c_{1t}$ is the steady-state ratio of credit consumption to cash consumption; c_1/k, c_2/k, x_k/k, x_h/k, and h/k are the long-run ratios of the respective parts of output relative to the size of the capital stock; and n is the balanced growth level of the labor supply. This system of nine equations in nine variables π, γ, R, c_1/k, c_2/k, x_k/k, x_h/k, h/k and n can be solved given values of the parameters and the policy variables μ and τ to trace the long-run reaction of the system to a change in policy.

Consider the effect of an increase in the growth rate of money μ. Note that the right side of equation 15 (or equation 16) can be interpreted as the after-tax rate of return on savings. Thus equation 15 relates the long-run rate of growth to the equilibrium after-tax rate of return r on capital. If either time spent working n or the human capital-to-physical capital ratio h/k is affected by changes in μ, then the growth rate of the economy depends on μ. As a special case, consider what happens when $\delta_k = \delta_h$. Here, equations 15 and 16 can be used to solve for h/k and to show that it is given by $(1 - \alpha)/\alpha$, which is independent of the rate of inflation. In this case, it follows that the growth rate γ is affected by changes in μ only if n is affected. In this model, inflation acts as a tax that distorts the consumption of cash goods relative to credit goods. This distortion can in turn distort the labor/leisure choice and thus affect time allocated to work n. (See equation 18.)

Given that h/k is constant (since we have assumed that $\delta_k = \delta_h$), the steady-state, after-tax real rate of return on capital is affected by changes in the steady-state value of n. This is true here because n represents the rate of usage of the productive capital good h. A higher n corresponds to a more intensive use of the stock and hence a higher marginal product of capital (when h/k is held fixed). In this case, if n decreases in response to an increase in μ, then the equilibrium long-run rate of growth in the economy will decrease as μ is increased.

Although one would expect an increase in μ to decrease n and hence decrease γ, this is not always true. In fact, the exact behavior of this system of equations depends critically on the substitutability between cash goods and credit goods. For example, still assuming $\delta_k = \delta_h$, we can show that if the two types of consumption goods are complements (that is, $\lambda > 0$), then the growth rate falls monotonically in μ and approaches the lowest feasible rate in this economy: $1 - \delta$. However, if the two goods are substitutes (that is, $\lambda < 0$), then we can show that the relationship between the steady-state values of γ and μ is not monotone. At low levels of μ, γ is a decreasing function of μ, but eventually γ becomes an increasing function of μ as the system is demonetized. That is, if μ is high enough, c_1/c_2 goes to zero, and the growth rate converges to that of the system when monetary expansion is at its optimal rate.[15]

Computations

Next, we provide estimates of the quantitative magnitudes of the growth effects of inflation for our 12 models.

To provide these estimates, we must have parameter values for each of these 12 models. We select parameter values for each of the models using a combination of figures from previous studies and facts about the growth experience of the U.S. economy between 1960 and 1987. Throughout the calibrations, we assume that a period is 1.5 months, that is, the length of time it takes one dollar to produce one transaction for the cash good.[16] We assume that the discount factor $\beta = 0.98$ at an annual rate.[17] We also assume that the intertemporal elasticity of substitution $\sigma = 2.0$, that the preference parameter $\lambda = -0.83$,[18] that the fraction of time spent working $n = 0.17$,[19] that the capital share parameter $\alpha = 0.36$,[20] that the depreciation rate on human capital $\delta_h = 0.008$ at an annual rate,[21] and that the tax rate on income $\tau = 0.22$.[22] The rest of the parameters are estimated using the steady-state equations of the models so as to make them hold exactly. We use the following auxiliary relationships based on the U.S. economy's experience during 1960–87:

The average annual growth rate in per capita gross national product (GNP) is 2.06 percent.[23]
The average annual rate of inflation is 5.08 percent.[24]
If we ignore the fraction of cash held in banks and outside the country, cash in the hands of the public averages 2.04 percent of annual GNP.[25]
Investment in physical capital as a fraction of GNP averages 16.69 percent.[26]

These facts, along with the parameter values given, are used in conjunction with the balanced growth equations to obtain values for the other (nonspecified) parameters of the models and for the balanced growth endogenous variables of the system.

For example, in the Lucas model with CIA in consumption, the parameter values obtained are $A = 0.08$, $\delta_k = 0.04$, $\eta = 1.03$, and $\Psi = 8.22$. The values for the endogenous variables are $\mu = 1.07$, $R = 15$ percent, $c_1/k = 0.007$, $c_2/k = 0.01$, $x_k/k = 0.007$, $x_h/k = 0.01$, and $h/k = 2.31$. All variables are in annualized terms. To get some feel for these numbers, note that the fitted growth rate of money $\mu(1.07)$ is higher than the observed value of the growth rate of the monetary base in the period (1.0684), but only slightly. (That is, equation 19 does not hold exactly at the true μ, π, and γ combination because velocity is not constant in the data.) These numbers also imply a capital/output ratio in this model of 2.8, which is close to that used by Chari, Christiano, and Kehoe (1994). The implied value of 0.43 for $c_1/(c_1 + c_2)$ is roughly the same as the Nilson Report's (1992) estimate of 0.41 for the ratio of cash purchases to other purchases in the U.S. economy. Finally, the value of 23.54 percent for x_h as a fraction of GNP is close to the sum of the values of health care expenditures and education expenditures in the United States.[27]

Thus the model does well mimicking the U.S. economy along some dimensions. Note that the implied pretax nominal rate of return is 15 percent, probably high by most standards. This is a common feature of the endogenous growth models without uncertainty (given our assumptions that $\sigma = 2.0$ and $\beta = 0.98$). A detailed description of the calibration method for each model is contained in Chari, Jones, and Manuelli (1995).

We compute solutions to the balanced growth equations assuming that $\pi = 1.1$ and $\pi = 1.2$. This increase of 10 percentage points in the inflation rate allows us to easily compare the changes in the growth rates predicted by the models with those found in the data, as discussed. We choose a baseline of $\pi = 1.1$ because this is close to the average rate of inflation in the samples from across countries analyzed by empirical researchers. Note that from a purely formal point of view, the balanced growth equations describe the relationship between the growth rate and the rate of monetary expansion, μ. However, since this is not the regression that empirical researchers have run, we did the experiment by changing μ by however much is necessary to guarantee that the inflation rate is increased by 10 percentage points per year. The findings of this experiment are displayed in Table 1.[28]

Table 1 gives the percentage change in the growth rates when the inflation rate is increased 10 percentage points.[29] The results of this experiment produce several notable

Table 1 A Small Inflation Effect on Growth*

	Money demand models		
Growth models	CIA in consumption	Shopping time	CIA in everything
Ak	0	0	–.011
Two-sector	0	0	–.009
Lucas	–.009	–.005	–.027
Romer	–.007	.128	–.024

* Percentage point change in growth rate when inflation increases 10 percentage points.

features. The most important is that the predicted change in the growth rate across all of the models is an order of magnitude smaller than that of around 0.5 found in the empirical literature. Another notable feature is that there is no guarantee, in general, that an increase in the inflation rate will necessarily decrease the growth rate, although this is generally true. [Jones and Manuelli (1990) show that in the Lucas model with CIA in consumption, the relationship between inflation and growth is not monotone.] Note, however, that just because the growth rate increases as μ increases (in some regions of the parameter space), this increase does not mean that welfare increases. On the contrary, this is not true in general: Increasing levels of inflation induce welfare-decreasing substitutions from c_1 to c_2. A third notable feature is that in the Ak and two-sector models of growth in combination with the CIA in consumption and shopping time models of money demand, one can show theoretically that the growth effect of inflation is exactly zero. In these models, inflation has no effect on the after-tax real return to savings. (In this sense, these models are Fisherian.) It follows, therefore, from the analogue of equations 15 and 16, that γ is unaffected by μ.

In summary, the results of this section show that constructing models in which inflation affects growth is fairly straightforward. However, in general, these models predict a very small effect of inflation on growth.

MODELS WITH BANKS, GROWTH, AND INFLATION

In this section we study an alternative way of introducing money into the model. The 12 models already analyzed have the feature that all money is held in the hands of the public for carrying out transactions in consumption of one form or another. In fact, banks hold a significant fraction of the monetary base in the United States and other countries. Here we construct a simple model of financial intermediation in which banks are subject to reserve requirements. The equilibrium portfolio of a typical depositor is thus necessarily part capital and part money. Therefore, changes in the real rate of return on money (through inflation) reduce the real after-tax return on savings and thus affect growth. In this model, we repeat the previous computations and again find that the quantitative effect of changes in μ is much smaller than that seen in the data.

Given these conclusions, we turn to the possibility that our notion of monetary policy is too narrow. A broader and more realistic description of monetary policy allows for changes both in the growth rate of the money supply and in banking regulations. To the extent that increases in inflation rates are driven by needs for seigniorage, one would expect these increases to be accompanied by measures designed to increase the demand for the monetary base. In our model of financial intermediation, these measures are increases in reserve requirements.

We find in the data that inflation and the fraction of the monetary base held by banks are positively correlated. This correlation opens the possibility that a measure of monetary policy such as reserve requirements could be an important variable missing in the existing empirical work. To explore this possibility, we consider monetary policy experiments that consist of simultaneously changing the reserve requirements and the growth rate of the money supply in a way consistent with the empirical evidence. We find that when this change is made, existing models of growth and money demand can approximately reproduce the quantitative effects of inflation on growth found by empirical researchers.

A Simple Model with Banks

We study a model in which the banking system plays an essential role in facilitating production and capital accumulation.[30] In our model, two types of capital are used in the production of final output, both of which are essential. One of these two types of capital must be intermediated as loans through the banking system, while the other is financed through conventional equity and debt markets. Finally, we assume that smooth substitution takes place between the two so that the amount of this banking type of capital can be altered across different policy regimes. To make loans, banks are required to hold reserves.[31]

We denote the two types of physical capital by k_1 and k_2. The first type, k_1, is intermediated through capital markets. The second type, k_2, must be intermediated through banks. That is, for k_2 to be used in production, consumers must place deposits in the banking system and firms must borrow these deposits in the form of bank loans to finance purchases of k_2. Banks are required to hold reserves against their deposits. We assume that no resources are used to operate the banking system. Here then an intermediary is simply a constraint (the reserve requirement relating the amount of base money that must be held in the banking system to the amount of capital of type 2 that is to be financed). We consider only two kinds of growth models here, the Ak and the Lucas versions. For the Lucas model, the production function is

$$Y_t = K_{1_t}^{\alpha_1} k_{2_t}^{\alpha_2} (n_t h_t)^{1-\alpha_1-\alpha_2}. \tag{23}$$

Reserve Requirements

For this version of the model, the consumer's problem is to

$$\max \sum_{t=0}^{\infty} \beta^t u(c_{1t}, c_{2t}, 1 - n_t), \tag{24}$$

subject to

$$p_t c_{1t} \leq m_{1t-1} \tag{25}$$

$$\begin{aligned} d_t + m_{1t} + b_t \leq{}& (m_{1t-1} - p_t c_{1t}) - p_t c_{2t} - p_t x_{k1t} - p_t x_{ht} + p_t r_t k_t (1 - \tau) \\ &+ p_t w_t n_t h_t \times (1 - \tau) + [1 + (1 - \tau) R_{dt}] d_{t-1} + [1 + (1 - \tau) R_t] b_{t-1} + T_{t+1} \end{aligned} \tag{26}$$

$$k_{1t+1} \leq (1 - \delta_1) k_{1t} + x_{k1t} \tag{27}$$

$$h_{t+1} \leq (1 - \delta_h)h_t + x_{ht}, \tag{28}$$

where m_{1t-1} reflects the consumption transactions demand for money (that is, CIA for c_1) and d_t is deposits in the banking system. Arbitrage implies that $R_{dt} = R_t$.

The financial intermediary accepts deposits and chooses its portfolio (that is, loans and cash reserves) with the goal of maximizing profits. The intermediary is constrained by legal requirements on the makeup of this portfolio (that is, the reserve requirements), as well as by feasibility. Then the intermediary solves the problem

$$\max_{Lt,dt,m2t}(1 + R_{Lt})\, L_t + m_{2t} - (1 + R_{dt})d_t, \tag{29}$$

subject to

$$m_{2t} + L_t \leq d_t \tag{30}$$

$$m_{2t} \geq \varepsilon d_t. \tag{31}$$

where m_{2t} is cash reserves held by the bank, d_t is deposits at the bank, L_t is loans, and ε is the reserve requirement ratio. The reserve requirement ratio is the ratio of required reserves (which must be held in the form of currency) to deposits.

The firm rents capital of type 1 directly from the stock market (that is, the consumer) and purchases capital of type 2 using financing from the bank. Thus the firm faces a dynamic problem:

$$\begin{aligned}\max \sum_{t=0}^{\infty} \rho_t \{&(1 - \tau)[p_tF(k_{1t},k_{2t}n_th_t)\, p_tw_tn_th_t - p_tr_tk_{1t} \\ &- R_{Lt-1}L_{t-1}] + L_t - p_tx_{k2t} - (1 + R_{Lt-1})L_{t-1}\},\end{aligned} \tag{32}$$

subject to

$$p_{t-1}k_{2t} \leq L_{t-1} \tag{33}$$

$$k_{2t+1} \leq (1 - \delta_2)k_{2t} + x_{k2t}, \tag{34}$$

where ρ_t is the subjective discount factor used by firms. Note that constraint 33 implies that from the firm's point of view, it may as well be renting k_2 from the bank itself. Because of this situation, the firm can be seen as facing a static problem; hence, one of the implications of the equilibrium conditions for this version of the model is that the choice of ρ_t is irrelevant.

To gain some intuition for the role of reserve requirements in this model, consider the intermediary's problem. The solution to its problem is given by

$$(1 + R_{Lt})(1 - \varepsilon)d_t + \varepsilon d_t - (1 + R_{dt})d_t = 0. \tag{35}$$

Simplifying this, we obtain that in equilibrium

$$R_{Lt} = R_{dt}/(1 - \varepsilon). \tag{36}$$

Reserve requirements thus induce a wedge between borrowing rates and lending rates for the intermediary.

Next, from consumer optimization, we have that the consumer must be indifferent between holding a unit of deposits and holding a unit of capital. This indifference implies that the after-tax real returns on the two ways of saving must be equal. That is,

$$1 + (1 - \tau)R_{dt-1} = (p_t/p_{t-1})[1 - \delta_1 + (1 - \tau)r_t]. \tag{37}$$

Production firms set their after-tax marginal products of the two types of capital equal to their after-tax real rental rates. Therefore,

$$F_1(t) = r_t, \tag{38}$$

and

$$(p_t/p_{t-1})[(1 - \tau)F_2(t) + (1 - \delta_2)] = 1 + (1 - \tau)R_{Lt-1}, \tag{39}$$

where $F_1(t)$ and $F_2(t)$ denote the marginal products of the two types of capital. Substituting, we obtain

$$1 + (\{(p_t/p_{t-1})[(1 - \tau)F_1(t) + 1 - \delta_1] - 1\}/(1 - \varepsilon)) \\ = (p_t/p_{t-1})[1 - \delta_2 + (1 - \tau)F_2(t)]. \tag{40}$$

Inspection of this equation reveals that increases in the reserve requirements (higher ε) or increases in the inflation rate have the effect of raising F_2 relative to F_1. That is, higher reserve requirements or higher inflation rates distort the mix of the two types of capital. The reason for this distortion is that financial intermediaries are required to hold non–interest-bearing assets in their portfolios. This requirement introduces a wedge between the rental rates on the two types of assets, and this wedge distorts the capital mix. It can also be seen that the increased distortion in the capital mix induced by a change in the inflation rate is greater with higher reserve requirements. Thus in this model, inflation acts as a tax on capital, the effect of which is magnified by higher reserve requirements.

Distortions and Financial Intermediation

Many countries impose a variety of impediments to the smooth functioning of the financial intermediation system. Examples of these impediments include portfolio restrictions, taxes, and requirements that loans to favored industries and individuals be made at below-market interest rates. To some extent these impediments can be thought of as introducing a wedge between the interest rate goods-producing firms pay banks and the rate banks receive on their loans. We can incorporate this wedge into our model as follows. Let Θ denote the wedge. Let R_{1t} denote the interest rate paid by goods-producing firms so that $R_{1t}(1 - \Theta)$ is the interest rate received by banks. Note that the wedge Θ acts as a tax on the interest receipts of banks. The financial intermediary's problem is now

$$\max_{Lt,dt,m2t}(1 + R_{1t}(1 - \Theta))L_t + m_{2t} - (1 + R_{dt})d_t \tag{41}$$

subject to constraints 30 and 31. The solution to this problem implies that in equilibrium we have

$$R_{Lt} = R_{dt}/((1 - \varepsilon)(1 - \Theta)). \tag{42}$$

Thus, not surprisingly, a tax on the receipts of financial intermediaries introduces the same kind of wedge between lending and borrowing rates as does the imposition of reserve requirements. In this sense a wide variety of government interventions reduce growth rates in exactly the same way as do reserve requirements. In particular, these interventions reduce both growth rates, as well as the size of the financial intermediation sector. We can use this observation as a test of the plausibility of our model. Suppose the only difference between countries is in these policy wedges and suppose, as seems reasonable, that direct measures of the policies inducing distortions are not available. Our models imply a positive association between growth rates and the size of the financial intermediary sector. The quantitative magnitude of this association can be compared with the relevant association in the data as a test of our model. We perform such an exercise below.

Computations

We begin by computing the effect of changing the growth rate of the money supply so that the annual inflation rate increases 10 percentage points. This computation is done for two calibrated models: the Lucas model and an Ak version of the model.

To do the calibration, we use data on the actual holdings of money in both the banking and non-banking sectors, along with measures of assets intermediated by banks. After taking account of money held outside the United States,[32] we find that the fraction of money held as reserves by banks (denoted by m_h) is 0.46. We use assets of commercial banks minus their holdings of U.S. government securities, consumer credit, vault cash, reserves at Federal Reserve Banks, and deposits of nonfinancial businesses to obtain a measure of the capital stock intermediated through banks. We obtain these data from the flow of funds accounts published by the Board of Governors of the Federal Reserve System. The average of the ratio of this measure to GDP from 1986 to 1991 is 0.39. We use these facts (along with the assumption that $\delta_1 = \delta_2$) to calibrate the models and obtain estimates of the parameter ε and k_2's share of output (relative to k_1).

The parameters from this calibration for the Lucas version of the model are

$$A = 0.095, \delta_1 = \delta_2 = 0.02, \delta_h = 0.016, \alpha_1 = 0.306, \alpha_2 = 0.054, \beta = 0.98,$$
$$\eta = 1.03, \lambda = -0.83, \sigma = 2.0, \Psi = 6.412, \text{ and } \varepsilon = 0.042.$$

Again, all parameters are expressed in annualized terms.

Of course, alternative measures of ε could be taken directly from banking regulations. The difficulty with that approach is that reserve requirements differ greatly among the different types of accounts held in banks. Depending on which types of accounts, average reserve requirements on banks could be anywhere from 2.5 percent to 12 percent.

Findings

Given this calibration, we find that increasing μ in order to increase π from 1.1 to 1.2 on an annual basis decreases the annual growth rate of output by 0.009 percentage points for the Ak model and by 0.021 percentage points for the Lucas model. Thus, although these effects are quantitatively larger (for the Lucas model) than those we have seen in the models with transactions demand for money, they are still too small by a factor of roughly 20 than the regression results reported in the literature. [Haslag (1994) finds growth effects of up to 0.4 percentage points.]

Given that the effects on the growth rate of changing μ are still small, we now explore the effects on the growth rate of changing ε—the other aspect of monetary policy in the model. For this exploration, we use the Lucas model. We run two experiments. In the first, we hold constant the rate of inflation at $\pi = 1.1$ and increase ε. The rate of growth of money is determined by the balanced growth equation. In the second, we hold the growth of money fixed and increase ε. The inflation rate is determined by the balanced growth equation. First, consider the effect on the growth rate of holding π constant at 1.1 and adjusting the reserve requirement parameter ε. The results of these experiments are shown in Figure 1.

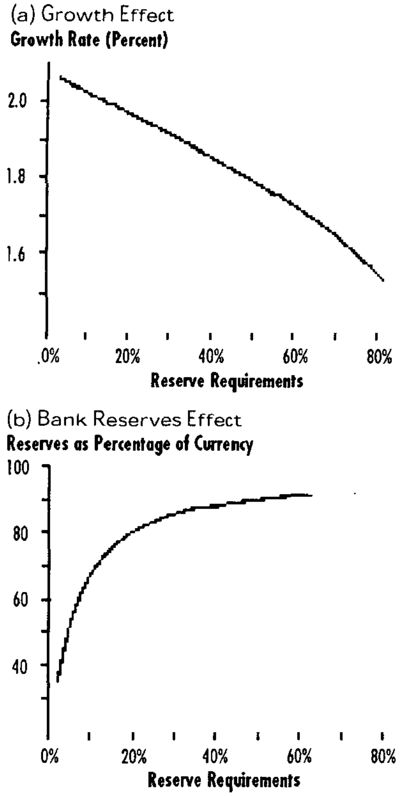


Figure 1 Effects of increasing reserve requirements in the Lucas model (inflation fixed at 10 percent and money growth adjusted).

As the charts in Figure 1 show, even moderate increases in the reserve requirements can produce the observed changes in the growth rate. For example, an increase from the calibrated level of $\varepsilon = 0.04$ to $\varepsilon = 0.35$ will give the desired effect. We show the implied money holdings (in reserves) by banks for this experiment in the right chart in Figure 1. Note that the result is highly nonlinear and, even at very low levels of ε, the resulting equilibrium changes in m_b are quite severe.

Next, consider the effect on the growth rate of increasing ε and letting π adjust, while holding μ constant. The impact on γ and m_b, respectively, is shown in Figure 2. The results of this experiment are qualitatively similar to those when π is held fixed. The growth effects of changing ε are quite large even for quantitatively reasonable changes. Note that it follows from this discussion that we cannot generate the observed correlation between growth and inflation without simultaneously adjusting ε and μ. That is, from the results of holding μ fixed and adjusting ε, it follows that the correlation between π and γ is positive: As ε is increased, both π and γ decrease.

Does this class of models show quantitative potential? That is, can we explain, through simultaneous adjustments in μ and ε, the observed relationship between growth and inflation? If we do not restrict that question further, the answer is yes. This answer is misleading, however, since the implied relationship between μ and ε may be quite different from that in the data. To subject the model to a more rigorous test, therefore, we must use data on actual countries' performances to get some feel for the magnitude of the relationship between actual changes in μ and in ε.

To do this, we collected data from 88 countries from the International Monetary Fund's International Financial Statistics (IFS).[33] Since measures of ε are not readily available, we instead gather data on m_b that in turn—conditional on the model—allow us to estimate ε. To estimate the size of the combined money growth and reserve requirement effects, we estimate the relationship between π and m_b from the data and use this estimated effect in comparing computed balanced growth path results. That is, we compute the implied change in the growth rate when the inflation rate is increased 10 percentage points and, at the same time, the reserve requirement is increased so as to change the observed m_b as is seen in the data. To do this computation, we first give the regression result concerning the relationship between π and m_b:

$$m_b = -0.220 + 0.460\pi, \tag{43}$$

where m_b is the time-series average, by country, of the fraction of the monetary base held in banks, while π is the time series average, by country, of the inflation rate. (The t-ratio for the coefficient on π is 5.98.) For this sample, the mean value of π is 1.16 (that corresponds to an inflation rate of 16 percent), and its standard deviation is 0.18. The mean value of m_b is 0.32 with a standard deviation of 0.16. Thus an increase of 0.1 in π produces an

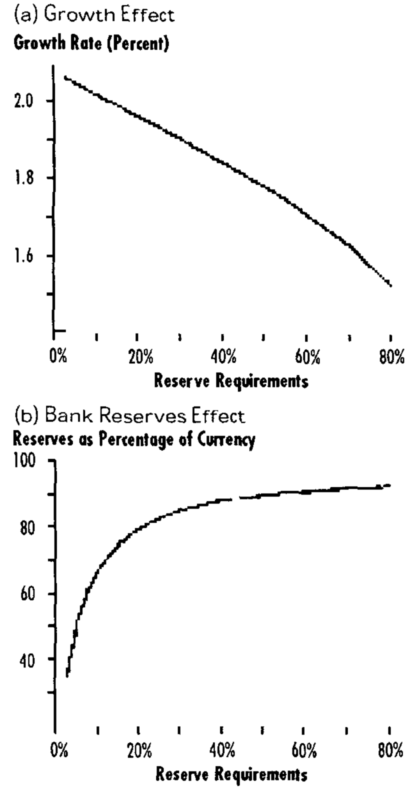


Figure 2 Effects of increasing reserve requirements in the Lucas model (money growth fixed at 12.2 percent and inflation adjusted.)

increase of approximately 0.046 in m_b. These results are similar to those found in Brock (1989). They are consistent with the view that in high inflation countries, governments choose high reserve requirements to enhance the base of the inflation tax.

The experiment we perform is to increase π from 1.1 to 1.2 and simultaneously to increase m_b by about 0.046. (We will actually change m_b by 0.05.) The size of the equilibrium growth response depends critically on the initial value of m_b because the relationship between ε and m_b is very nonlinear, as documented in the charts on the right in Figures 1 and 2. Therefore, in Table 2 we report the results for several initial values of m_b.

Experiment 1 uses the regression results from the IFS data to estimate the level of m_b at $\pi = 1.1$. Here, the increase of 0.05 in m_b is associated with only a small change in ε (less than 0.005) and hence a small change in the growth rate results. In this experiment, the predicted change in the growth rate is smaller by a factor of 10 than the regression results in the empirical studies. At higher initial levels of m_b, however, the predicted growth effects of the same experiment are substantially higher. At $m_b = 0.7$, even a relatively small increase in ε (from 0.121 to 0.176) gives a growth effect that is one-fifth as large as that found in the empirical studies. Finally, for substantial initial levels of the reserve requirements ($m_b = 0.8$), a 10 percentage point increase in inflation decreases the annual growth rate approximately 0.2 percentage points. This estimate—although lower than the average value of 0.5 found in different studies—is similar to the lower bound of 0.20 reported in Barro in this issue and elsewhere.

These results suggest that, although higher than those in the United States, reserve requirement values are within a plausible range. The model that allows for simultaneous changes in both money supply and reserve requirements therefore comes close to matching the estimated impact of inflation on growth.

Next, we used the actual values of the relative amount of currency held in the banking system, m_b, and the inflation rate, π, in the data for each country to calculate the implied value of the reserve requirement ratio, ε, as well as all the other model variables. The implied values of the growth rate and the inflation rate for all countries in our sample with the exception of Israel are reported in Figure 3. We exclude Israel here because in our model there is no combination of inflation and reserve requirements that can rationalize the relative amount of cash held by the banking system in Israel in the data. A regression of growth on inflation in Figure 3 yields the result that a 10 percentage point increase in the inflation rate results in a reduction in the growth rate of 0.02 percentage points.

Table 2 How Growth Changes in a Model with Banks When Inflation Increases 10 Percentage Points*

Value of bank base money (m_b)			Growth rate (γ)			Reserve requirements (ε)		
Experiment	Initial	New	Initial	New	Change percentage points	Initial	New	Change
1	.286	.332	1.0206	1.0204	–.02	.020	.024	.004
2	.600	.650	1.0203	1.0198	–.05	.076	.010	.066
3	.700	.750	1.0200	1.0192	–.08	.121	.176	.055
4	.800	.850	1.0195	1.0175	–.20	.217	.426	.209

* In each experiment, the inflation rate is increased from 10 percent to 20 percent.

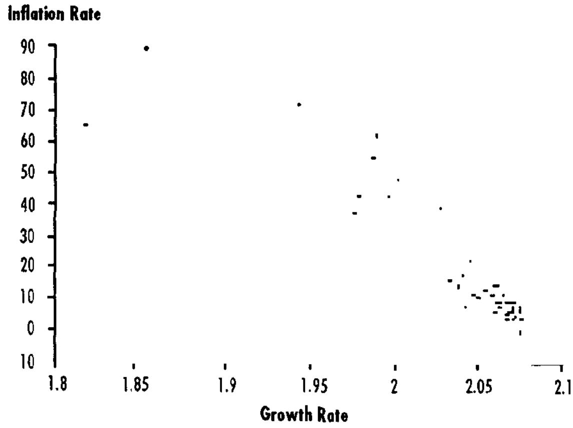


Figure 3 Baseline parameterization (excluding Israel).

We re-analyzed by including a proxy for the Israel observation. We computed solutions for the model at several different high values of m_b and π. We ran a regression at these values of the reserve requirements on m_b and extrapolated the implied value of reserve requirements for our model. The plot of inflation and growth rates constructed in this manner is shown in Figure 4. A regression of growth on inflation in Figure 4 shows that a 10 percentage point rise in inflation is associated with a 0.05 percent reduction in the growth rate. Alternative parameterizations of our model yield substantially greater evidence of the effects of inflation on growth.

We experimented with other parameter choices. In Figures 5 and 6 we report on the analogues of Figures 3 and 4 for values of $\lambda = -0.5$ and $\sigma = 1.5$. A regression of growth on inflation in Figure 5 shows that an increase in the inflation rate of 10 percentage points results in a fall in growth of 0.04 percent. In Figure 6 the fall in growth rates is 0.08 percent. In this sense our theoretical models account for anywhere between 10 percent and 40

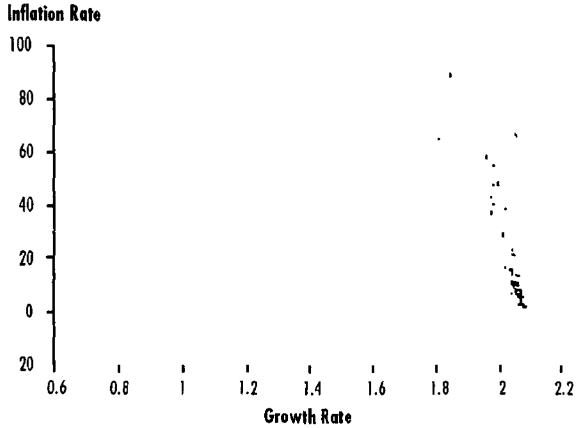


Figure 4 Baseline parameterization (including Israel).

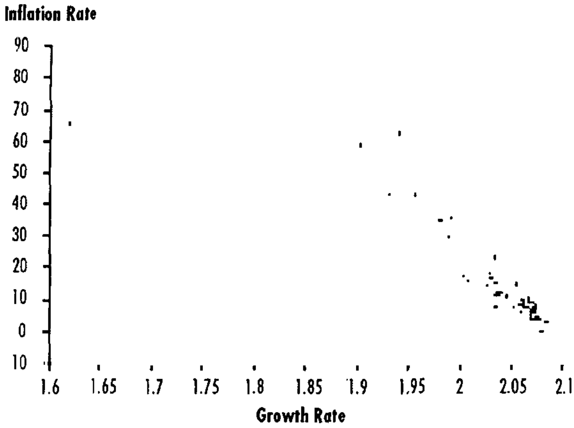


Figure 5 Alternative parameterization (excluding Israel).

percent of the association in the data. The remainder, it is plausible to suppose, arises from the fact that countries which adopt one kind of growth-reducing policy typically adopt other kinds of growth-reducing policies.

We also examined the relationship between the size of the financial system and growth rates implied by our model. King and Levine (1993) regress growth rates from 1960 to 1989 on the ratio of claims on the nonfinancial sector to GDP, a measure of the size of the financial intermediary sector, and obtain a coefficient of 0.032. The analogous measure of the size of this sector in our model is k_2/y. We varied our measure of distortions and calculated the relationship between the size of this sector and growth rates implied by our model. We obtained a coefficient of 0.01. The fact that our model did relatively well in the sense that the order of magnitude is correct at mimicking this relationship in the data increases our confidence in it.

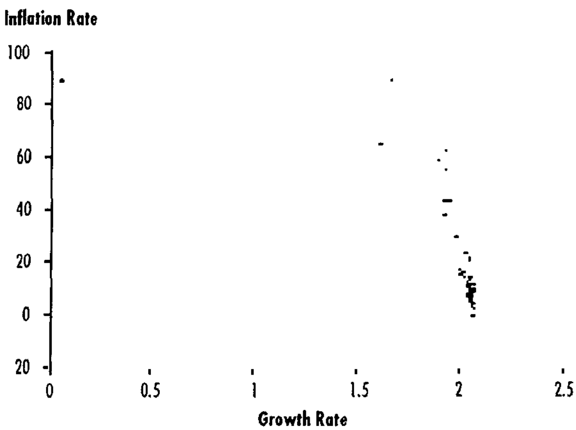


Figure 6 Alternative parameterization (including Israel).

CONCLUSIONS

Empirical researchers have found that the average long-run rate of inflation in a country is negatively associated with the country's long-run rate of growth. Moreover, the statistical relationship uncovered by these researchers is large. Roughly, increasing the inflation rate by 10 percentage points in a country otherwise similar to the United States decreases the growth rate of per capita output by 0.5 percentage points. We have examined a variety of models with transactions demand for money and have seen that none produce results anywhere near this large.

This finding leads us to reconsider our view of monetary policy to include changes in financial regulations, as well as changes in the money supply. In the data, we document a high correlation between the rate of inflation in a country and the fraction of the currency in the economy that is held in the commercial banking system. We interpret this to mean that monetary authorities who raise inflation rapidly also require banks to hold more currency. (That is, in those countries, reserve requirements are also higher.) After taking account of this extra dimension of monetary policy, we find that existing models of growth and money demand can come much closer to reproducing the results found by empirical researchers. In addition, we find that the relationship between changes in reserve requirements and growth rates is highly nonlinear. Thus the estimated effects depend sensitively on the level of the reserve requirements.

Our analysis suggests that inflation rates per se have negligible effects on growth rates, but financial regulations and the interaction of inflation with such regulations have substantial effects on growth. This analysis suggests that researchers interested in studying the effects of monetary policy should shift their focus away from printing money and toward the study of banking and financial regulation.

APPENDIX TECHNOLOGY AND PREFERENCES IN THE MODELS

Here we describe the production functions and the preferences used in the growth and money demand models discussed in the chapter.

Models of Growth

Ak Model

The resource constraint is

$$c_{1t} + c_{2t} + g_t + x_{kt} = Ak_t. \tag{A1}$$

Two-sector Model

The production function in the investment sector is

$$x_{ht} = A(k_t - k_{1t}), \tag{A2}$$

and in the consumption sector it is

$$c_{1t} + c_{2t} + g_t = Bk_{1t}^{\alpha} n_t^{1-\alpha}, \tag{A3}$$

where k_{1t} is the amount of capital used in the production of consumption goods.

Lucas Model

The production function is

$$c_{1t} + c_{2t} + g_t + x_{kt} = Ak_t^{\alpha}(n_t h_t)^{1-\alpha}. \tag{A4}$$

Romer Model

The production function is

$$c_{1t} + c_{2t} + g_t + x_{kt} = Ak^{\alpha}n^{1-\alpha}\bar{k}^{1-\alpha}, \tag{A5}$$

where $\bar{k}$ is the aggregate capital stock. Preferences are given by

$$[c_{1t}^{-\lambda} + \eta c_{2t}^{-\lambda}]^{(1-\alpha)/\lambda}(1 - n)^{\Psi(1-\sigma)}/(1 - \sigma). \tag{A6}$$

Models of Money Demand

CIA in Consumption Model

Cash goods purchases must satisfy the constraint

$$p_t c_{1t} \le m_t, \tag{A7}$$

where m_t denotes cash balances.

Shopping Time Model

Time allocated to nonleisure activities n_t is allocated to shopping time n_{ct} and market activity n_{ft} so that

$$n_t = n_{ct} + n_{ft}. \tag{A8}$$

The technology for purchasing cash goods for all models of growth except the Lucas model is

$$p_t c_{1t} \le Bm_t n_{ct}^{\varepsilon}. \tag{A9}$$

For the Lucas model, the shopping time technology is

$$p_t c_{1t} \le Bm_t^{\varepsilon}(p_t n_{ct} h_t)^{1-\varepsilon}. \tag{A10}$$

CIA in Everything Model

The cash-in-advance constraint is given by

$$p_t(c_{1t} + \varepsilon c_{2t} + \varepsilon x_{kt}) \le m_t. \tag{A11}$$

NOTES

1. See, for example, Jones and Manuelli (1990), Barro (1991), and Rebelo (1991).
2. Fischer (1991).
3. Jones and Manuelli (1990) and Gomme (1991) are exceptions.
4. See, for example, Fischer (1993).
5. The cross-sectional average of the time-series average rates of per capita income growth in the Summers and Heston (1991) data is around 1.92 percent per year.
6. Some studies do not arrive at this conclusion. McCandless and Weber (1995) find no correlation between inflation and the growth rate of output. See also Levine and Renelt (1992).

7. See also Grier and Tullock (1989).
8. Although we do not study the relationship between inflation volatility and growth here [as does Gomme (1991), theoretically], empirical studies have found that more volatile monetary policies also have depressing effects on growth rates. See Kormendi and Meguire (1985), Fischer (1993), and Easterly et al. (1994). One must be careful interpreting this relationship, however, since there is a high correlation between the average inflation rate experienced over the period in a country and the volatility of the inflation rate. This correlation is reported to be 0.97 in Levine and Renelt (1992).
9. Although these are important differences, one must be careful in interpreting this evidence. As discussed in Levine and Renelt (1992), a high degree of multicollinearity exists between many of the regressors that authors include in these studies. Hence, most of the empirical findings are nonrobust in the Learner sense.
10. See Stockman (1981) and Cooley and Hansen (1989).
11. For a detailed development of the key issues, see Jones and Manuelli (1990) and Rebelo (1991).
12. See the Appendix for a description of the technologies and preferences.
13. See Rosen (1976).
14. See Benhabib and Perli (1994) and Ladron-de-Guevara, Ortigueira, and Santos (1994).
15. See Jones and Manuelli (1990) for details.
16. Chari, Christiano, and Eichenbaum (1995).
17. Chari, Christiano, and Kehoe (1994).
18. Chari, Christiano, and Kehoe (1991).
19. Jones, Manuelli, and Rossi (1993).
20. Chari, Christiano, and Kehoe (1994).
21. Jones, Manuelli, and Rossi (1993).
22. We run several experiments to test the robustness of our results to our choice of parameters. For these experiments, we use the Lucas model of growth along with the CIA in everything model of money demand. First, we estimate the length of a period using the Nilson Report's (1992) numbers on the fraction of transactions that are completed using cash. The Nilson Report does not say exactly what transactions are included in its measure of all transactions. We calibrate the model two different ways: We assume that transactions on x_h are and are not included in the calculations. These calibrations produce estimates of the period length of 1.63 months and 1.02 months, respectively. In addition, we (as did Chari, Christiano, and Kehoe, 1994) try lowering our parameter that determines the elasticity of the labor supply ψ to the level 2, while allowing the potential workday to vary. Finally, we change the elasticity of substitution between cash goods and credit goods from –0.83 to –0.2. None of these experiments results in a significant change in the growth effect of inflation. Details of these experiments are available from us upon request.
23. Economic Report of the President (1994).
24. Economic Report of the President (1994).
25. Porter (1993).
26. Economic Report of the President (1994).
27. See issues of the U.S. Department of Commerce's Survey of Current Business (1992).
28. For the purposes of calibration, our Ak model is a version of the Lucas model in which the labor supply is inelastic. This model has all the important qualitative features of the Ak model, but it allows labor share and investment rates to be chosen so as to be close to those seen in the U.S. time series. Chari, Jones, and Manuelli (forthcoming) has details.
29. For the CIA in everything versions of the models, we assume that all of c_1 and a fraction ϵ of the c_2 and x_k expenditures used are subject to the CIA constraint. For the results presented in Table 1, we use $\varepsilon = 0.2$, since most investment transactions do not use cash directly. We experiment with increasing ε over an appreciable range and, although the growth effects are larger with larger ε, they still fall short of the effect seen in the data. In the next section, we discuss a model in which cash is used indirectly for these transactions through the banking system.

30. See Greenwood and Smith (forthcoming) for a survey of the theoretical work in this area. For recent empirical work, see Roubini and Sal-i-Martin (1992), King and Levine (1993), and Ireland (1994).
31. Our model is similar to the one analyzed by Haslag (1994), but ours is more realistic along two dimensions. First, Haslag assumes that all capital must be intermediated through banks, while we allow the share of bank assets to be endogenous. Second, Haslag uses money only to meet reserve requirements, while we use money to facilitate consumption transactions as well. See also Valentinyi (1994).
32. Porter (1993).
33. For details, see Chari, Jones, and Manuelli (1995).

REFERENCES

Barro, R. J. Economic Growth in a Cross Section of Countries. Quarterly Journal of Economics (May 1991), pp. 407–43.

Benhabib, J., and Perli, R. Uniqueness and Indeterminacy: On the Dynamics of Endogenous Growth. Journal of Economic Theory (June 1994), pp. 113–42.

Brock, P. L. Reserve Requirements and the Inflation Tax. Journal of Money, Credit and Banking (February 1989), pp. 106–21.

Chari, V. V., Christiano, L. J., and Eichenbaum, M. Inside Money, Outside Money, and Short Term Interest Rates. Journal of Money, Credit and Banking (November 1995), pp. 1354–86.

Chari, V. V., Christiano, L. J., and Kehoe, P. J. Optimal Fiscal and Monetary Policy: Some Recent Results. Journal of Money, Credit and Banking (August 1991), pp. 519–39.

Chari, V. V., Christiano, L. J., and Kehoe, P. J. Optimal Fiscal Policy in a Business Cycle Model, Journal of Political Economy (August 1994), pp. 617–52.

Chari, V. V., Jones, L. E., and Manuelli, R. E. The Growth Effects of Monetary Policy. Federal Reserve Bank of Minneapolis Quarterly Review vol. 19, no. 4 (Fall 1995), pp. 18–32.

Cooley, T. F., and Hansen, G. D. The Inflation Tax in a Real Business Cycle Model. The American Economic Review (September 1989), pp. 733–48.

Easterly, W., King, R., Levine, R., et al. Policy, Technology Adoption and Growth. National Bureau of Economic Research Working Paper No. 4681 (March 1994), p. 15.

Economic Report of the President. U.S. Government Printing Office, 1994.

Fischer, S. Growth, Macroeconomics and Development. In NBER Macroeconomics Annual 1991, Vol. 6, Olivier Jean Blanchard and Stanley Fischer eds. MIT Press National Bureau of Economic Research, 1991, pp. 329–64.

Fischer, S. The Role of Macroeconomic Factors in Growth. National Bureau of Economic Research Working Paper No. 4565 (December 1993), p. 21.

Gomme, P. Money and Growth Revisited: Measuring the Costs of Inflation in an Endogenous Growth Model. Journal of Monetary Economics (August 1993), pp. 51–77.

Greenwood, J., and Smith, B. Financial Markets in Development and the Development of Financial Markets. Journal of Economic Dynamics and Control vol. 21, no. 1 (January 1997), pp. 145–181.

Grier, K. B., and Tullock, G. An Empirical Analysis of Cross-National Economic Growth 1951–80. Journal of Monetary Economics (September 1989), pp. 259–76.

Haslag, J. H. The Effects of Monetary Policy in a Model with Reserve Requirements. Federal Reserve Bank of Dallas Working Paper 94–15 (October 1994).

Ireland, P. N. Economic Growth, Financial Evolution, and the Long-Run Behavior of Velocity. Journal of Economic Dynamics and Control (May–July 1994), pp. 815–48.

Jones, L. E., and Manuelli, R. E. A Convex Model of Equilibrium Growth: Theory and Policy Implications. Journal of Political Economy (October 1990), pp. 1008–38.

Jones, L. E., Manuelli, R. E., and Rossi, P. E. Optimal Taxation in Models of Endogenous Growth. Journal of Political Economy (June 1993), pp. 485–517.

King, R. G., and Levine, R. Finance, Entrepreneurship and Growth: Theory and Evidence. Journal of Monetary Economics (December 1993), pp. 513–42.

Kormendi, R. C., and Meguire, P. G. Macroeconomic Determinants of Growth: Cross-Country Evidence. Journal of Monetary Economics (September 1985), pp. 141–63.

Ladron-de-Guevara, A. Ortigueira, S., and Santos, M. Equilibrium Dynamics in Two-Sector Models of Endogenous Growth. Cuadernos Economicos de ICE (Vol. 58, pt. 3, 1994), pp. 107–35.

Levine, R., and Renelt, D. A Sensitivity Analysis of Cross-Country Growth Regressions. The American Economic Review (September 1992), pp. 942–63.

McCandless, G. T. Jr., and Weber, W. E. Some Monetary Facts. Federal Reserve Bank of Minneapolis Quarterly Review (Summer 1995), pp. 2–11.

Nilson Report. Oxnard, California, 1992.

Porter, R. D. Foreign Holdings of U.S. Currency. International Economic Insights (November–December 1993), p. 5.

Rebelo, S. Long-Run Policy Analysis and Long-Run Growth. Journal of Political Economy (June 1991), pp. 500–21.

Rosen, S. A Theory of Life Earnings. Journal of Political Economy (August 1976), pp. S45–67.

Roubini, N., and Sala-i-Martin, X. A Growth Model of Inflation, Tax Evasion, and Financial Repression. National Bureau of Economic Research Working Paper No. 4062 (May 1992), p. 30.

Sidrauski, M. Rational Choice and Patterns of Growth in a Monetary Economy. The American Economic Review: Papers and Proceedings (May 1967), pp. 534–44.

Stockman, A. C. Anticipated Inflation and the Capital Stock in a Cash-in-Advance Economy. Journal of Monetary Economics (November 1981), pp. 387–93.

Summers, R., and Heston, A. The Penn World Table (Mark 5): An Expanded Set of International Comparisons 1950–88. Quarterly Journal of Economics (May 1991), pp. 327–68.

Survey of Current Business, various issues (1992).

Valentinyi, A. How Financial Development May Affect Growth. European University Institute Working Paper, Florence, Italy (1994).

This Page Intentionally Left Blank

46
Controlling Inflation

Fred Furlong
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

Bharat Trehan
Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

People generally agree that uncertainty about the future level of prices is undesirable since it leads to a less efficient allocation of resources. For instance, the more uncertain people are about future inflation, the less willing they will be to write long-term contracts or to make long-term investments. In principle, the monetary authority can minimize such uncertainty by controlling inflation. However, what this means for the actual conduct of policy is subject to considerable debate. Some argue that for inflation control to be effective, monetary policy should be aimed at bringing the inflation rate down to zero. Others argue that certainty about future prices can be achieved merely by keeping the inflation rate steady at its prevailing level. With steady inflation, prices would rise, but the overall level of prices still would be easy to predict.

This chapter discusses these two alternative approaches to controlling inflation in the context of the monetary authority's need to convince the public that it is committed to that goal. This need arises because policymakers generally have multiple, and sometimes conflicting, objectives for policy. Thus, the public cannot always be sure which goal has top priority, since economic circumstances, and therefore the priorities attached to different goals, can change over time. This is a critical issue since individuals' perceptions of policy bear heavily on their economic decisions. Therefore, the monetary authority must not only strive for price stability (predictability), but also find effective ways to let the public know that it places a high weight on achieving that goal.

MULTIPLE POLICY GOALS AND THE INFLATION RATE

In most countries, controlling the inflation rate is not the only goal of the monetary authority. In the U.S., for instance, the Congress has mandated that the Federal Reserve conduct policy to achieve such goals as high levels of employment and economic growth, along with stable prices.

Reprinted from: *FRBSF Weekly Letter* (Feb. 8, 1991).

Having multiple goals necessarily limits how much policymakers can do to reduce price uncertainty, because achieving one objective often requires compromise on another. For example, suppose the monetary authority reacts to a supply shock, like a rise in the price of oil, by following a more stimulative monetary policy, in an attempt to offset any slowdown in the economy. If the stimulative policy continues long enough, it will raise the rate of inflation. However, since monetary policy cannot influence the level of output in the long run, output will eventually return to the level that would have prevailed without such a policy. Thus, in a world where the monetary authority is known to attach a high weight to stabilizing output, random shocks can move the inflation rate around; this makes predicting inflation, and, thus, the future price level, more difficult.

With multiple policy goals, then, the predictability of future prices depends upon the priority attached to inflation *versus* other objectives. In addition, if a high priority on price predictability is to be reflected fully in private economic decisions, individuals must be sure about how much weight is being given to controlling inflation relative to other objectives. Individuals also need to determine how much the priority placed upon controlling inflation might change as, say, the economy moves into a recession.

HISTORICAL PATTERNS

Data for the post World War II period provide ample evidence of how policy priorities can change. Figure 1 shows that while inflation was relatively low and stable from the mid-1950s to the mid-1960s, it has been much higher and more variable since then. Indeed, statistical tests indicate that from the mid-1960s through the early 1980s there was no tendency for inflation to return to some level or even to a trend. Such behavior makes it extremely difficult to predict what the price level will be in the future.

In part, the apparent difference in price uncertainty in the two periods may be due to the nature of the nonmonetary shocks to the economy, such as the oil price shocks in the 1970s. Nevertheless, we do have some evidence suggesting that progressively higher levels of inflation were tolerated after the mid-1960s in order to keep unemployment rates low. For instance, the behavior of the federal funds rate (adjusted for inflation) over this period has generally been seen as providing information about the stance of monetary policy, with a low funds rate indicating easier policy. It turns out that the inflation-adjusted federal funds rate was noticeably low during much of the 1970s, and was actually negative for several quarters both in the middle and towards the end of the decade.

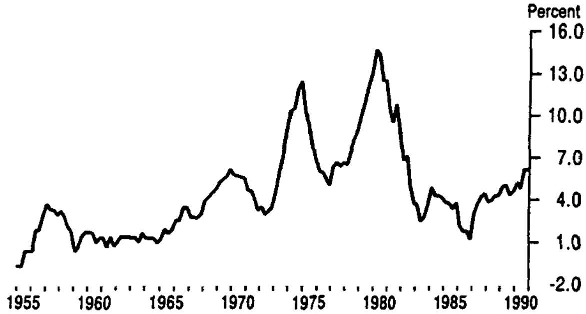


Figure 1 U.S. consumer price inflation (growth over the previous 4 quarters).

This pattern in priorities apparently was reversed in the beginning of the 1980s. At that time, the unemployment rate rose sharply and output contracted, which led to a drop in the inflation rate noticeably below the levels that prevailed in the 1970s. While the rate of inflation has varied since then, so far it has not displayed a sustained upward trend.

RETAINING CREDIBILITY

The lowering of inflation rates in recent years probably has convinced the market that policymakers place more weight on controlling inflation now than in the late 1970s. Consequently, we should have less price uncertainty today than, say, ten years ago.

But, as economic circumstances change, the public cannot be sure that the priority attached to controlling inflation will not change. Thus, while the market may believe that policy will be less likely to allow inflation to approach the peak levels observed in the early 1980s, there still may be considerable uncertainty about the range over which the inflation rate could vary. The fact that the rate of inflation has been gradually creeping up over the last few years also may be contributing to this uncertainty.

Under these circumstances, how does the Federal Reserve retain credibility as an inflation fighter?

ZERO INFLATION. . .

One answer is embodied in a House Resolution proposed over a year ago by Congressman Stephen L. Neal, and supported by many analysts and policymakers, including the Chairman of the Federal Reserve Board Alan Greenspan and the President of the Federal Reserve Bank of San Francisco Robert Parry (see *Weekly Letter* March 2, 1990). This resolution calls upon the Federal Reserve to conduct policy with a view to bringing the inflation rate down effectively to zero over a five-year horizon, and to maintaining price stability thereafter.

Passage of such a resolution would give Congressional endorsement to a policy of trading off some real growth in the economy in the near term for lower inflation in the long run. Effective implementation of the Neal Resolution would virtually eliminate uncertainty about future prices. Recent evidence to support this assertion is available, for example, in work by economists Lawrence Ball and Stephen Cecchetti. They point out that statistical tests on U.S. data reveal a positive correlation between the level of inflation and the degree of price uncertainty. Such a correlation suggests that commitments by a monetary authority to control inflation should be more believable the lower it keeps the rate of inflation. The Federal Reserve's support for the Neal resolution last year, then, can be seen as a way of emphasizing that it does not intend to let inflation get out of control.

. . . OR STABLE INFLATION?

What if the monetary authority aimed at an alternative policy prescription—stabilizing inflation at its current level? Part of the argument in favor of this option rests on the principle that the degree of price predictability need not be related to the level of inflation. If inflation stays at about the same rate, people can be confident about the future level of prices in making long-run contracts.

In that case, many argue, it does not make sense for the U.S. to incur the risk of the high economic costs, in terms of higher unemployment and lost output, required to bring the inflation rate down toward, say, zero. This argument has some appeal, especially because there is little evidence to show that the relatively modest levels of inflation that have prevailed in the U.S. over the past few years impose significant costs on the economy.

IS THE COMMITMENT CONVINCING?

The problem with a policy based on the premise that "it is too costly to reduce inflation today" is that it provides individuals little reason to believe that policymakers' would act on their commitment to control inflation in the future. Individuals have no way of knowing if policymakers will ever be willing to incur the costs of reducing inflation: if it is too costly to reduce inflation today, it will be too costly to reduce inflation tomorrow as well. In other words, the logic behind accepting today's rate of inflation would support accepting any future rate of inflation. Thus, such a policy cannot ordinarily be distinguished from one that would do nothing to control inflation.

Consequently, it seems unlikely that merely pledging to maintain the current rate of inflation would be very convincing to the market. This is particularly true given the likely political pressures that would tend to emphasize short-run output and employment considerations over longer-run price stability. The historical record, which we have already examined, lends some support for such skepticism. The steady increase in inflation over the late 1960s and the 1970s appears consistent with an acceptance of the prevailing rate of inflation and an aversion to imposing short-term costs on the economy in order to obtain the long-term benefits of a lower, more predictable inflation rate.

THE OIL SHOCK AND THE WAR

Under normal circumstances, it is difficult to counter skepticism that a policy to hold inflation at its current level is really a do-nothing policy. If policymakers are not seen as making hard choices, it is not easy to convince the public that policymakers remain committed to controlling inflation. However, it is possible to distinguish the two policies under some conditions; for instance, inflationary shocks will call forth a vigorous response from a monetary authority committed to stabilizing inflation.

Circumstances are obviously far from normal right now. The price of oil has risen substantially since Iraq's invasion of Kuwait (even though it has fluctuated wildly of late). The recession-induced increase in the unemployment rate brings with it the possibility that (as in the 1970s) policymakers will switch objectives and temporarily de-emphasize inflation control. Finally, and perhaps most important, there is the uncertainty created by the "of the war with Iraq. These events have created the potential for an increase in inflation. Containing inflation in the near future may turn out to be a difficult task, depending upon factors such as the price of oil and the duration of the war in the Middle East. Keeping inflation in check in the face of such pressures will have the effect of making the Fed's commitment to inflation control more convincing.

CONCLUSIONS

We have argued that uncertainty about future inflation will remain in check if the monetary authority is seen as placing a relatively high weight on controlling inflation. Under normal circumstances, a policy pledged to reducing the rate of inflation should be more persuasive than a policy of maintaining the current inflation rate, since the latter cannot be distinguished from a policy that would do nothing to control inflation.

At this time, however, there is some risk that inflation will accelerate. Thus, stabilizing inflation in the current economic environment could enhance the Federal Reserve's credibility. Even so, any gain in credibility is unlikely to be permanent; sooner or later, when economic circumstances are more normal, the issue of retaining credibility will have to be faced anew.

This Page Intentionally Left Blank

47
Ending Inflation

John P. Judd
Vice President and Associate Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

Brian Motley
Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

It now is widely recognized in the United States that price stability is the most appropriate long-term goal of monetary policy. Last year, Representative Stephen Neal proposed that this goal be recognized formally in legislation that would instruct the Federal Reserve to lower the inflation rate to zero within five years and maintain constant prices thereafter. This proposal was endorsed by Federal Reserve Chairman Alan Greenspan and a number of Reserve Bank Presidents (Parry 1990).

But in addition to the long-term goal of price stability, the Fed also has the goal of averting or minimizing the effects of downturns in economic activity. These objectives call for differing, and sometimes conflicting, policy actions, which raises the issue of which should take precedence at any particular time. Currently the Fed decides this issue on a discretionary basis, responding to a wide range of indicators of economic and financial activity and of inflation.

Many economists argue that a discretionary approach produces an inflationary bias in policy, as the goal of price stability tends to be sacrificed to short-term stabilization goals. These economists recommend abandoning discretion in favor of a monetary policy rule that explicitly binds policymakers to actions that stabilize prices. Proponents of discretion respond that, although rules may enhance price stability, they would be likely to involve unacceptably large losses of economic output when inflation was being reduced or when inflation shocks were encountered.

In this chapter, we assess the merits of a type of policy rule that was designed to address the concerns of both camps in this debate—ensuring price stability while attempting to limit disruptions of economic activity.

TARGETS

Choosing a policy target is the first step in moving away from discretion and toward a rules-based approach to monetary policy. Economists disagree over which variable would be the

Reprinted from: *FRBSF Weekly Letter* (May 31, 1991).

most effective target mainly because of differences in their views of the structure of the economy.

Traditionally, the money supply has been considered an appealing target because inflation is caused by excessive monetary growth in the long run. Moreover, the money stock is closely related to the actions the central bank takes to implement policy. These considerations were instrumental in the Federal Reserve's decision to establish target ranges for several monetary aggregates beginning in the mid-1970s.

In the last decade, however, the relation between the money supply, on the one hand, and economic activity and prices, on the other, has become unstable, and thus difficult to predict in the short to intermediate run. This instability, which apparently is related to the deregulation of the financial system, has greatly limited the usefulness of money as a target. As a result, although the Fed still establishes monetary ranges, they no longer play the key role in policy.

Problems with the stability of the monetary aggregates have led to increased interest among economists in nominal GNP as a policy target. This variable measures the current dollar value of economic output (in other words, real GNP times the aggregate price level). Given this definition, it would be easy to calculate the path of nominal GNP that would be consistent with long-run price stability, so long as real GNP had a predictable long-run trend.

However, there is some debate over whether real GNP follows such a trend. Until the 1980s, it was widely believed that real GNP did have a stable trend that was determined by gradually evolving movements in the labor force and productivity. Recently, however, a number of researchers have argued that real GNP does not have a stationary trend, but rather is buffeted by technological and labor supply shocks that permanently alter its level. If this "real business cycle" view is correct, the price level could drift over time under nominal GNP targeting. Statistical tests cannot distinguish very well between these two hypotheses, casting doubt on the effectiveness of nominal GNP as a target in stabilizing prices.

In part because of this concern, some economists have argued that the Fed should target prices directly. Regardless of how real GNP behaves, direct price-level targeting could avoid long-term price-level drift. But the effectiveness of this target also is subject to debate, because of differing views of the flexibility of prices. Proponents of price-level targeting believe that prices are relatively flexible—a critical feature of the real business cycle theory.

Keynesian economists, on the other hand, believe that prices are "sticky," so that policy has its principal immediate effects on output, and only later affects prices. Hence, attempts by the central bank to achieve a predetermined path for prices might involve high short-run volatility in output. Moreover, many of these economists downplay the role of disruptions to an otherwise stable trend in real GNP, and thus believe that nominal GNP makes more sense as a target than prices.

OPERATIONAL RULES

Even if the appropriate policy target were readily apparent, such knowledge would not suffice to establish a rules-based monetary policy procedure. Most target variables cannot be directly controlled by the central bank. Thus, a policy rule also must be chosen that defines the actions the central bank will take in response to movements in the variable it targets.

Those actions are defined in terms of movements in a variable called the policy instrument, which must be under the direct control of the central bank.

McCallum (1988) has proposed specific monetary policy rules that merit consideration. These rules have nominal GNP or the price level as the policy target, and the monetary base as the instrument. They are operational, in the sense that they specify precisely how the central bank should change the monetary base in response to nominal GNP or the price level.

Another appealing feature of these rules is that their implementation does not depend upon any particular economic model. Thus they sidestep the theoretical debates about the "right" structure of the economy. The rules involve adjusting monetary policy to observed misses of a variable from its target value, and do not depend upon any forecast that would require a particular macroeconomic model to be used.

In particular, these rules have two components. First, they specify a long-run equilibrium growth rate for the monetary base that is estimated to be consistent with price stability. This component allows the base to expand enough to accommodate the (noninflationary) trend rate of expansion of economic output (that is, in this component, growth in the base is equal to growth of potential GNP plus trend growth of the velocity of the base).

Second, the rules specify responses to deviations between the actual and desired values of the target variable. For example, if nominal GNP is above its target in any quarter, the nominal GNP rule calls for the Fed to reduce the growth rate of the monetary base by some fraction of the target miss. The strength of this response could be raised or lowered depending on how aggressively the central bank wished to pursue its target. Presumably, a more aggressive response would tend to involve more precise control of the target variable in the short run, but possibly with more volatility in economic activity, while a more gradual response would produce the opposite result. Thus the McCallum rules can be tailored to the objectives of the central bank. However, even with a gradual response, the rules are designed to ensure price stability in the long run.

The choice of the monetary base as the policy instrument is controversial. The advantage of this variable is that it can be accurately controlled by the Fed. Moreover, data on it are widely available, so that the public can easily observe the central bank's adherence to its rule, and thus hold the central bank accountable for its actions. The disadvantage of the of the base is that its relationship to nominal GNP has been less stable since the deregulation of the financial system.

Two features of the McCallum rules tend to mitigate this problem. First, they include a term that gradually adjusts the equilibrium base growth rate for changes in the relationship between the base and nominal GNP. Second, the structure of the rules ensures that changes in that relationship automatically are offset by policy. For example, if nominal GNP rises above its target because of an unexpected change in that relationship, the rule will require the Fed to reduce growth in the base until nominal GNP is brought back on target. Instability in the base could therefore cause temporary problems for monetary policy, but eventually they would be corrected under the rule.

WOULD THE MCCALLUM RULES WORK?

Given the uncertainties about the structure of the economy, it would be risky to adopt a rule unless it were likely to work well in a variety of different economic environments. Do Mc-

Callum's rules satisfy this criterion? In a study for the Bank's *Economic Review* (Judd and Motley 1991), we analyzed this issue by conducting counterfactual simulations of how the economy would have evolved if the Fed had used the rules in the past.

We conducted these tests in the context of several models that incorporated the various features of the economy that are crucial for the performance of alternative rules. Thus we tested models with fixed prices, flexible prices, a stationary, and a nonstationary trend in real GNP. To measure the uncertainty associated with using the rules, we conducted repeated (stochastic) simulations for each rule under alternative sets of shocks, and constructed ranges of outcomes (confidence intervals) for the economy.

We found that the nominal GNP rule was successful at achieving price level targets in all of the models tested. Moreover, that rule appears capable of holding prices close to the target level even in the face of the kinds of unexpected developments typically experienced in the U.S. economy. The price-level rule, by contrast, produced wide swings in real GNP and prices in the model that assumed sticky prices. Given the uncertainty about which model is most appropriate, this result argues for nominal GNP, and against prices, as the target.

Using the nominal GNP rule, we simulated the possible effects of moving from 4 1/2 percent inflation rate in 1990 to zero inflation in 1995 as specified under the Neal Amendment. The results suggest that this rule had a high chance of success under all the models. Moreover, none of the models suggested that the disinflationary process would increase the probability of a recession significantly compared with that experienced over the past 35 years under actual policy. Under the most likely outcome, annual real GNP growth would not fall below 1 1/2 percent.

We did find one problem with the nominal GNP rule. It raised the short-run volatility of real GNP relative to that observed under current discretionary policies. However, it is important to bear in mind that our estimates may represent upper bounds on the detrimental effects on income volatility of following the rule. Adopting a rule should enhance Fed credibility, which would reduce uncertainty in the economy and could have beneficial effects on volatility.

Overall, the McCallum-type nominal GNP rule shows promise as a way to eliminate inflation without greatly increasing the chances of a recession. The rule also could ensure price stability thereafter, possibly at the expense of more volatility in real GNP. Whether this rule seems worth trying depends on the importance placed on achieving and maintaining zero inflation compared with the risk that the economy would be more volatile.

REFERENCES

Judd, J. P., and Motley, B. 1991. Nominal Feedback Rules for Monetary Policy. Federal Reserve Bank of San Francisco *Economic Review* (Summer).

McCallum, B. T. 1988. Robustness Properties of a Rule for Monetary Policy. *Carnegie-Rochester Conference Series on Public Policy* 29.

Parry, R. T. 1990. Price Level Stability. *FRBSF Weekly Letter* (March 2).

48

Has the Fed Gotten Tougher on Inflation?

John P. Judd
Vice President and Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

Bharat Trehan
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

The federal funds rate has risen from about 3 percent in February 1994 to 6 percent today. Statements of Federal Reserve officials suggest that these actions can be interpreted as a "preemptive strike" against inflation—that is, policy was tightened in response to *current* indications that suggest higher inflation in the *future.* Many observers seem to have concluded that the Fed reacted more strongly to a threat of inflation than in the past. In this chapter, we attempt to shed light on this issue by comparing the Fed's behavior since early 1994 to its behavior at different points in the past. We find that policy in 1994 was similar to policy since the early 1980s, and maintained a focus on keeping inflation low while remaining responsive to the business cycle.

BENCHMARKS

Developments since the mid-1960s provide convenient benchmarks to gauge the recent tightening of monetary policy. From 1965 to 1979, consumer price inflation rose from less than 2 percent to above 12 percent. While developments such as the oil price shocks of the 1970s were responsible for some of this increase, analysts generally agree that a quicker and more forceful response by the Fed would have kept inflation from picking up as much as it did over this period.

The years from 1979 to 1993 provide a sharp contrast. The Fed raised rates sharply from late 1979 to mid-1981 in an effort to bring inflation under control. Tighter policy was accompanied by two back-to-back recessions in the early 1980s; however, they were followed by the longest expansion in the post-war period. The Fed's concern about inflation was evident once again in the late 1980s and early 1990s, when it raised the funds rate as inflation began to rise above 5 percent.

Reprinted from: *FRBSF Weekly Letter,* No. 95-13 (Mar. 31, 1995).

Below, we will use two different measures to summarize Fed policy in each of these periods and then use these measures to gauge policy in 1994. Our use of alternative measures reflects the inherent difficulty in finding ways to characterize policy. Any method will be subject to ambiguity and controversy.

A REACTION FUNCTION

Our first measure involves the use of a *reaction function.* Reaction functions attempt to *describe* the behavior of the Fed in conducting monetary policy. They measure the relationship between the Fed's policy instrument (such as bank reserves or the federal funds rate) and key variables that the Fed typically responds to in adjusting the stance of policy (such as inflation, real GDP and the unemployment rate). Thus, for example, a reaction function can suggest how the Fed would respond if the inflation rate were to rise by 1 percent. Reaction functions can change over time, depending in part on the weight the Fed places on various objectives of policy. For example, if the Fed becomes more determined to control inflation, then it will react more strongly to an increase in inflation.

A large number of reaction functions have been estimated by economists. We use one by Mehra (1994). In this function, the *real* (or inflation-adjusted) funds rate tends to move toward its observed long-run average level of just under 2 percent. However, the Fed moves the *nominal* rate in response to a variety of factors. Consistent with our discussion above, Mehra finds that these factors differ in the periods before and after 1979. From 1979.Q4 to 1992.Q4, the Fed reacted to business cycle developments by "leaning against the wind." In other words, the funds rate tended to go up when real GDP was above the level of activity that could be sustained in the long run (the estimated level of trend GDP), and went down when real GDP was below this level. In addition, the funds rate also rose in response to an acceleration in inflation, and to a rise in the long-term bond rate, which Mehra interprets as an increase in expected inflation.

The Fed's response over the 1965.Q4–1979.Q3 period was markedly different. Over this period the funds rate did not appear to respond significantly either to changes in the inflation rate or the long-term bond rate. And while the Fed also had a policy of leaning against the wind over this period, its response to cyclical developments was smaller than its response after 1979.

To determine whether the Fed changed its reactions in setting policy in 1994, we use Mehra's reaction function to see what the funds rate would have been if the Fed had continued to act as it typically did during 1979–1992. These estimates are compared to the actual funds rate in Figure 1. Assuming behavior that was typical in 1979–1992, the funds rate would have reached about 4 percent in the fourth quarter of 1994, roughly 1 percent below the rate that actually prevailed over that period. Thus, the Fed seems to have reacted more strongly than its earlier behavior would suggest. However, the difference is *not* statistically significant, which is to say that errors of this magnitude are not uncommon when trying to explain interest rates during the period over which the equation is estimated. At best, these results provide weak evidence that the Fed was more aggressive in heading off a surge in inflation than it has been in the past. A more tenable conclusion is that the Fed's reactions in 1994 were within a normal range of its typical behavior in the post-1979 period.

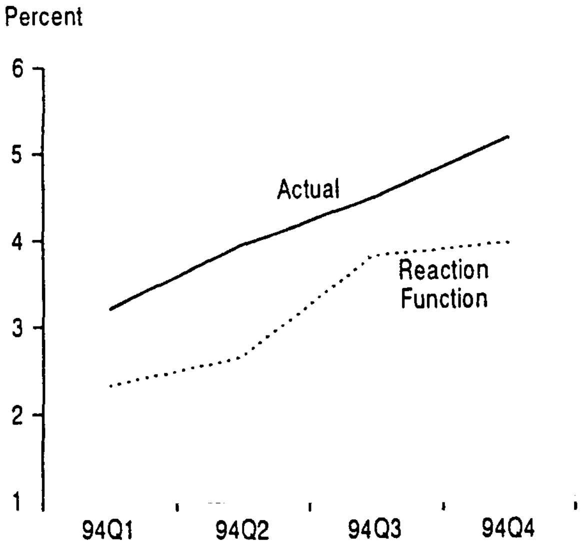


Figure 1 Federal funds rate: actual vs. "typical" post-1979 Fed behavior.

A POLICY RULE

For our second measure we use a policy rule proposed by Taylor (1993). While Mehra's reaction function describes Fed behavior, Taylor's rule *prescribes* it. The rule is designed to ensure that monetary policy would achieve an inflation target of 2 percent in the long run (measured in terms of the GDP deflator), while still being responsive to the business cycle. Under the rule, the real funds rate would rise relative to its long-run average of 2 percent whenever inflation rises relative to its 2 percent target and real GDP rises relative to an estimate of its long-run trend level. (The similarity with Mehra's reaction function is not coincidental, since his work draws upon Taylor's rule.) Like any simple rule, there is room for disagreement about the underlying assumptions, such as the appropriate long-run (or equilibrium) real funds rate, or the trend level of real GDP. (For Figure 2, we estimated the trend using the method of Braun (1990), in which trend growth declines from about $3\frac{1}{2}$ percent to about $2\frac{1}{2}$ percent in several steps over 1965–1994.)

What makes this rule interesting from our perspective is that it provides a remarkably accurate description of the funds rate since the mid-1980s. Taylor had already shown that his rule closely fit the actual path of the funds rate from 1987 (when Alan Greenspan became Fed Chairman) to 1992 (when Taylor did his study). Figure 2 shows that the same close relationship continued to hold over 1993 and 1994 as well. Given that the rule incorporates an inflation rate of 2 percent, this result suggests that policy in 1994 was consistent with maintaining low inflation.

The figure also extends the rate suggested by the rule back to 1965 and plots it together with the actual funds rate. (Note that the line in the figure labeled "rule" uses *actual* values for inflation and real GDP, and does not incorporate the effects of following the rule on those variables.) A comparison of the two lines in the figure shows how the policy regime has changed over time. In making this comparison, it is useful to keep in mind that our use of the rule is somewhat different from what Taylor had intended. We are not tak-

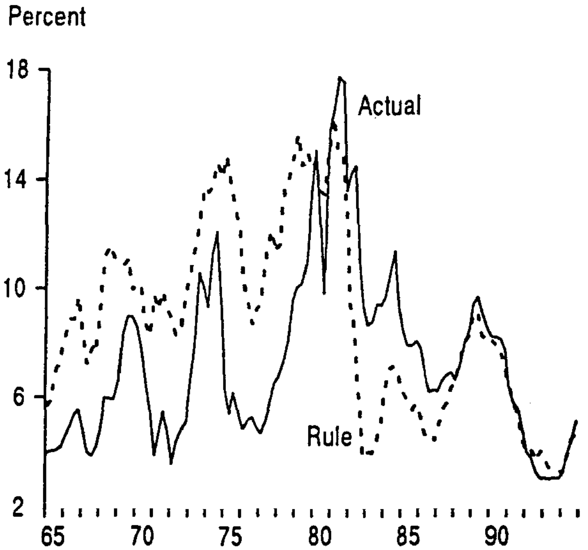


Figure 2 Federal funds rate: actual vs. rule's prescription for Fed behavior.

ing a stand on whether the funds rate should have followed the rule during this period. Instead, we use the rule as a tool for comparing actual policy across different regimes.

Figure 2 shows that the actual funds rate was considerably lower than what the rule would have prescribed from the mid-1960s to the late 1970s. Since this was a period of rising inflation, the results simply confirm that the rule would have done what it was designed to do—produce lower inflation than resulted from actual policy during this high inflation period. The figure also shows what the rule would have suggested during the Volcker era (1979–1987), when the Fed significantly reduced inflation from double digit rates to around 4 percent. In contrast to the pre-1980 period, the funds rate over this period was consistently higher than what the rule recommends, suggesting that the Fed was more aggressive in reducing inflation than the rule would have been.

Taylor's rule, then, also confirms the shift in policy that occurred during the late 1970s, and demonstrates that since the end of 1979 the funds rate has been more responsive to changes in output and inflation than it was earlier. It also suggests that the behavior of the funds rate in 1994 was little different from its behavior over the preceding six or seven years. To the extent that the behavior of the funds rate has changed, the break seems to lie around 1987.

Before concluding, it is worth discussing why the Taylor rule matches Fed policy so well in 1994. At first glance this match appears surprising: The rule does not contain direct measures of expected future inflation, while—according to statements by Fed officials—the funds rate increases since 1994 have been designed to head off future increases in inflation. The rule does, however, include an estimate of the GDP gap—that is, real GDP relative to its long-run trend. Measures of the productive capacity of the U.S. economy, whether represented by trend GDP, industrial capacity utilization or the "natural" rate of unemployment, appear to figure prominently in Fed forecasts of future inflation (Greenspan, 1995, pp. 2–3). Thus, in a sense the rule does include a forecast of future inflation.

CONCLUSION

In this chapter we have shown that Fed policy during 1994 was more concerned about inflation than it was prior to 1980. However, it is harder to make the case that the Fed's behavior over 1994 was different from the disinflationary regime established in the 1980s. And even though the two measures we have employed in this *Letter* suggest some differences in the way policy was conducted from 1979 to 1994, the most reasonable conclusion appears to be that the focus of policy has not changed very much over this period. The Fed's policy reactions in 1994 appear to be generally consistent with the desire for low inflation that it has exhibited since late 1979.

REFERENCES

Braun, S. N. 1990. Estimation of Current-Quarter Gross National Product by Pooling Preliminary Labor-Market Data. *Journal of Business and Economic Statistics,* (July) pp. 293–304.

Greenspan, A. 1995. Testimony before the Committee on Banking, Housing, and Urban Affairs, United States Senate (February 22).

Mehra, Y. P. 1994. A Federal Funds Rate Equation. Mimeo. Federal Reserve Bank of Richmond (March).

Taylor, J. B. 1993. Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, pp. 195–214.

This Page Intentionally Left Blank

49

Is Inflation Dead?

Roger E. Brinner

Managing Director and Chief Economist, The Parthenon Group, Boston, Massachusetts

In the past few years the United States has enjoyed the unique economic duet of very low unemployment and declining price inflation. For decades, we have come to associate tight labor markets with accelerating wages and prices. But in 1997, the unemployment rate sank below 5 percent and neither wage nor price inflation became a problem. Have our inflation processes fundamentally changed for the better? Are we in a new era of permanently better economic performance due to new behavior by our citizens? Or are we simply enjoying good luck in the form of positive supply shocks?

The overwhelming majority of research economists had previously estimated that an unemployment rate of about 6 percent was the lowest level that could be sustained. Below the vicinity of 6 percent, employers would compete aggressively for scarce labor by offering wage increases exceeding prevailing norms. The lower the unemployment rate, the greater the percentage wage gain would likely be. As wage increases pushed above existing price inflation and productivity gains, producers would raise prices to cover higher costs. The ensuing wage-price inflation spiral would not be broken until unemployment once more moved above 6 percent, a development typically caused by rising interest rates, as financial markets and the Federal Reserve reacted to the growing inflation problem.

The experience of the late 1980s seemed to confirm the modern validity of this decades-old perspective. The national unemployment rate fell from 6.2 percent in 1987 to 5.5 percent in 1988 and 5.3 percent in 1989; in response, wage increases jumped from 3.1 percent to 3.7 percent and then 4.2 percent. Wholesale price inflation rose from 1.9 percent in 1987 to 2.4 percent in 1988 and then to 3.9 percent in 1989; consumer price inflation jumped from 3.7 percent to 4.1 percent and on to 4.8 percent. Rising prices for imported goods, including oil, contributed significantly to the nation's inflation problems, but most of the blame was placed on excessive tension in domestic labor markets. In other words, lessons learned in the high-inflation 1970s had apparently been reinforced.

At first glance, the experience of recent years seems different. Price inflation remained in check from 1996 through 1998 even though unemployment rates fell further and further below 6 percent. It appeared natural to suspect that the balance point for U.S. labor markets should be recalibrated. Perhaps workers' insecurity was so intensified by the heav-

Reprinted from: *New England Economic Review, Federal Reserve Bank of Boston* (Jan./Feb. 1999) 37–49.

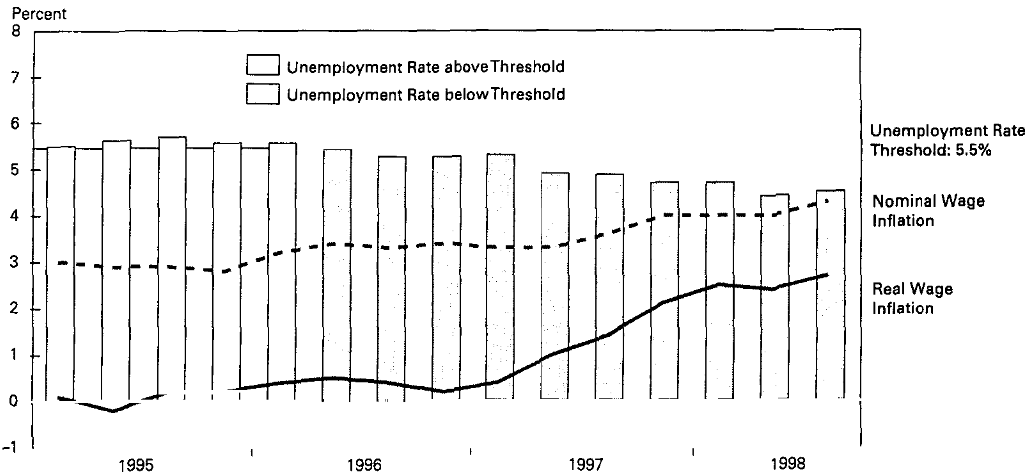


Figure 1 Real wages accelerated as usual in 1997, as unemployment fell below 5.5 percent.

ily publicized layoffs of the early 1990s that a lower unemployment rate was required before existing employees would demand norm-breaking pay increases.

Unfortunately, a more careful reading of the full inflation story reveals a different conclusion. *Nominal* wage inflation has been subdued by exceptionally modest price inflation. As can be seen in Figure 1, *real,* or price-adjusted, wage inflation has been increasing in response to low unemployment, just as in past decades. Price inflation has been held down by a set of "supply shocks," including a strong dollar, falling energy prices, and a cost-reducing regime shift in the health care industry. Moreover, most of these supply shocks are not novel ingredients of the inflation process in the United States. They have been studied and successfully used to understand price behavior for decades.

Inflation is not dead. However, recent supply shocks have shifted wage and price inflation to a lower zone. According to a revalidated, standard model of U.S. inflation, inflation can stay this low only if the unemployment rate rises to between 5.5 and 6 percent over the next year. Otherwise, as supply shocks shift to neutral or worse, tight labor markets will create a traditional inflation problem.

THE EVOLUTION OF THE STANDARD MODEL OF INFLATION

The original "Phillips Curve" (of the 1960s) presented a simple, inverse association between wage inflation (w) and the unemployment rate (u).

$$w = a0 - a1 * U \qquad \text{(where } a1 > 0\text{).} \tag{1}$$

In equation (1) and subsequent notation, a lower-case letter refers to the percentage change in the variable and an upper-case letter to the level. Thus, "w" is the percentage change in the hourly wage and "p" is the percentage increase in the aggregate price level. The constant term "$a0$" includes the average effect of all other influences on wages.

This model of wage inflation was informally paired with a simple view of price inflation. If prices are a markup on wages adjusted for trend productivity growth (q), then price inflation (p) equals wage inflation minus productivity growth. If, in addition, the markup is sensitive to the business cycle and, thus, to the unemployment rate, the price inflation model takes the form

$$p = w - q - b1 * U \qquad (\text{where } b1 > 0). \tag{2}$$

Equations (1) and (2) can be combined to create a reduced form model of price inflation driven by only the unemployment rate.

$$\begin{aligned} p &= a0 - a1 * U - q - b1 * U \\ &= (a0 - q) - (a1 + b1) * U \\ &= c0 - c1 * U. \end{aligned} \tag{3}$$

The rule-of-thumb derived from estimates of Equation (3) told policymakers in the 1960s that they had to choose between low inflation and low unemployment. The actual economic performance, displayed in Figure 2, was offered as good evidence of this trade-off. It seemed that those responsible for making the big economic decisions for the country—the President, the Congress, the Federal Reserve System—had to make a choice. If they wanted to hold inflation down to 2 or 3 percent, they had to accept an unemployment rate of 6 percent. If they wanted the unemployment rate near 4 percent, they had to accept inflation of 6 percent.

In the early 1970s, cheaper computing power and more extensive time series data permitted more complete multiple regression analysis, and applied economists were able to improve these simple models. Economic theory argues that real wages, not nominal wages, should adjust to the excess demand for labor. Indeed, Phillips and Lipsey, in their original works, had agreed that prices mattered, although they did not include prices in their final quantitative models. However, in the early 1970s, researchers added approximations of expectations of future price inflation to Equation (1), transforming it to (1a).

$$w - p(exp) = q + a0 - a1 * U. \tag{1a}$$

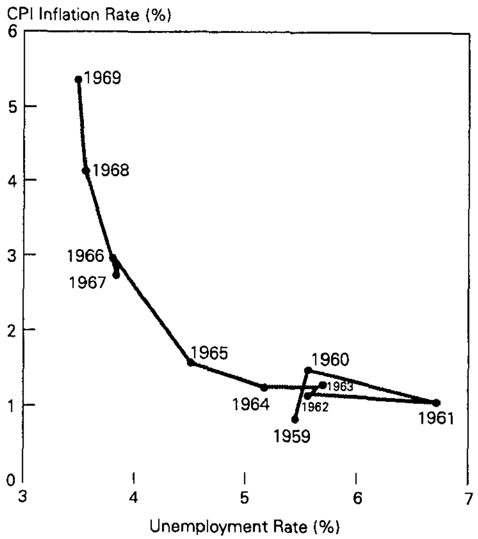


Figure 2 Apparent policy options in the 1960s: the simple Phillips Curve.

Although considerable creative debate took place over how to represent inflation expectations, many researchers adopted the simple approach of making expected inflation equal to inflation in the recent past. This approach was found to fit the data well, while being consistent with the view that workers are not subject to "money illusion" and that excess demand affects real rather than nominal wages.

The simplest form of this relationship is shown in Equation (1b).

$$p(exp) = p(t-1), \tag{1b}$$

in which $(t-1)$ refers to the prior year.

In practice, expected inflation is generally represented as a weighted sum of inflation rates over the past several years, with the estimated weights constrained to equal one.

Shifts in trend productivity growth (q) also can reasonably be expected to affect real wage growth. The new econometricians took demographic shifts into account as well: The entry into the labor force of the baby boom generation and rising female involvement in paid work focused attention on the potential importance of these shifts. As described in the accompanying Box, the typical approach to addressing changing demographics involved adjusting the reported unemployment rate to remove the influence of changes in the composition of the labor force. For the purpose of this discussion, consider (U) to include the effect of such adjustments.

Demographic Influences on the "Natural Rate of Unemployment"

In the regressions in Table 1 (below), the degree of excess demand pressure in the labor market is measured by the official civilian unemployment rate minus an adjustment for the changing demographic composition of the labor force. Even when the economy is at full employment, not all age–sex groups have the same unemployment rates. Therefore, when defining an indicator of excess labor demand for wage inflation equations, it is desirable to remove shifts in the unemployment rate caused solely by the changing composition of the labor force.

Unemployment exists when a new person enters the labor force before finding a job, when an existing employee quits a job to search for a better position, or when an individual is fired or laid off. Young people therefore have the highest characteristic unemployment rates and experienced, middle-aged workers the lowest. Women historically have had higher unemployment rates than men of the same age because of discrimination and more frequent movements into and out of the official work force. As their societal roles and labor force participation rates have changed, women's unemployment rates have dropped to near-equality with those of similarly aged men.

Given these differences in normal unemployment rates, two time periods sharing a 6 percent unemployment rate could have very different pressures on wages. If one period had a higher proportion of younger or female workers, the 6 percent unemployment would actually imply a tighter labor market. George Perry and others proposed simple adjustments to reflect this, which the author and colleagues at then DRI/McGraw-Hill updated using the following methodology.

The years 1962 and 1963 can reasonably be described as fitting the definition of full employment with stable inflation. The unemployment rate was 5.6 percent in both years, and CPI inflation was 1.2 percent followed by 1.3 percent. Moreover, the bordering years reinforce this image of stability. A time series is constructed by multiplying the unemployment rates of specific age–sex groups in 1962 and 1963 by the labor force shares of these groups in prior and subsequent quarters. Changes in the characteristic unemployment rates of each group are also varied through time based on regressions that estimate the difference between group rates and the rate for men aged 35 to 54 as a function of the group's population share and participation rate. This new series portrays how the unemployment rate varies through time relative to the early 1960s, based on changes in basic demographics.

These demographic shifts do not have a huge impact. The highest "full-employment" unemployment rates (equivalent to the 5.6 percent unemployment rate in 1962 and 1963) are 6.2 percent in 1980 and 6.1 percent in the years from 1974 to 1979. The lowest equivalent rates are 5.3 to 5.4 percent in the early 1950s and the years 1990 to 1993. The equivalent for 1997 is 5.5 percent. Most of the variation can be traced to fluctuations in the share of young adults in the labor force (Figure 3). The fraction of workers aged 16 to 24 rose from 16 percent of the labor force in the mid 1950s to 24 percent in the early 1970s, and then fell back toward 16 percent by the mid 1990s. These demographic changes allow the economy to operate at a lower unemployment rate, 5.5 percent today, compared to 6.1 to 6.2 percent 15 years ago, with the same inflation pressure arising in the labor market.

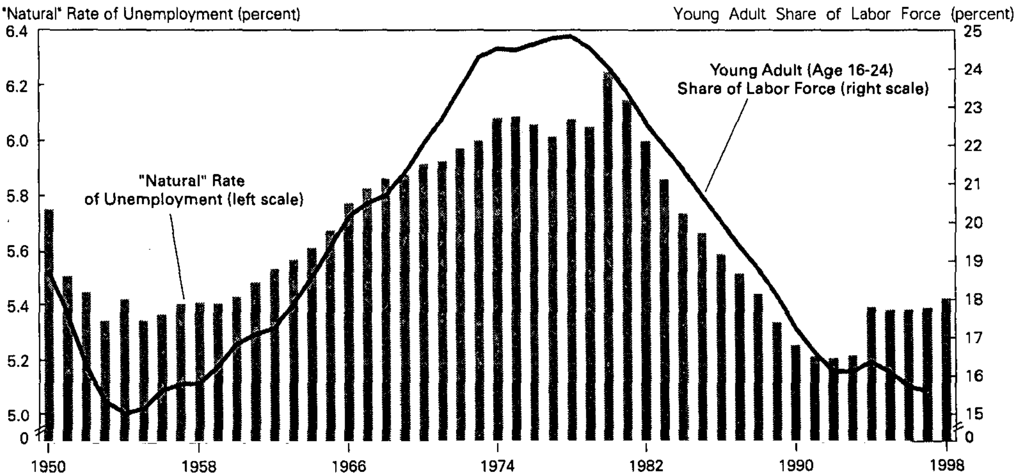


Figure 3 Changes in the unemployment rate due to demographic shifts.

The price inflation model was extended to recognize the influences of certain non-labor costs such as energy and imported goods, producing Equation (2a). The inclusion of changes in energy prices relative to changes in other domestic costs (*pe*) was natural following the radical shift of this cost element after the OPEC oil embargo sent energy prices soaring in the mid 1970s. Similarly, the end of the fixed exchange rate regime in the early 1970s made it appropriate to separately include shifts in the cost of imports relative to domestic goods (*pm*), as an indicator of an important component of domestic production costs and the price pressure presented by global competitors.

$$p = w - q - b1 * U + b2 * pe + b3 * pm. \qquad \text{(2a)}$$

This richer model produced a strikingly different view of long-run policy options. Specifically, only one unemployment rate is consistent with stable wage and price inflation, holding other factors constant.

Substituting (1b) into (1a) produces (1c):

$$w - p(t - 1) = q + a0 - a1 * U. \qquad \text{(1c)}$$

And substituting (1c) into (2a) produces (2b):

$$p - p(t - 1) = a0 - (a1 + b1) * U + b2 * pe + b3 * pm. \qquad \text{(2b)}$$

Equation (2b) describes what is referred to as an "accelerationist" Phillips' Curve, because the acceleration in prices ($p - p(t - 1)$)

has replaced price inflation (p as the phenomenon that responds to variations in the unemployment rate.

The balance point for the economy, defined as the rate of unemployment that preserves existing inflation, is found by setting (2b) equal to zero and solving for the unemployment rate. This unemployment rate has been called the "NAIRU" (*n*on-*a*ccelerating-*i*nflation *r*ate of *u*nemployment).

If inflation is unchanging, then $p - p(t - 1) = 0$, and

$$NAIRU = (a0 + b2 * pe + b3 * pm)/(a1 + b1). \qquad \text{(4)}$$

If the relative prices of energy and imports are not changing, then NAIRU $= a0/(a1 + b1)$. Most researchers who have estimated these relationships conclude that the NAIRU is close to 5.5 percent today, given the current demographic composition of the labor force.

Equation (4) also indicates that the economy can achieve decelerating price inflation, that is, disinflation, even with an unemployment rate below 5.5 percent—if the relative prices of energy or imported goods are falling sufficiently. In these circumstances, which we have enjoyed in 1997 and 1998, the NAIRU is *temporarily* below 5.5 percent. However, when these favorable supply shocks end, the unemployment rate must increase to 5.5 percent or else inflation will rise.

Similarly, (2a) can be substituted into (1b) and this new expression can then be substituted into (1a) to obtain an accelerationist model of *wage* inflation, (1d):

$$w - w(t - 1) = a0 - (a1 + b1) * U + b2 * pe + b3 * pm. \qquad \text{(1d)}$$

As will be shown below, this model is not at all surprised by the current phenomenon of relatively moderate nominal wage increases despite low unemployment. Small increases in *nominal* wage inflation are fully explained by the supply shocks that have damped down price inflation increases; larger *real* wage increases have occurred as anticipated by this standard model—in use for almost three decades.

STATISTICAL VALIDATION OF THE WAGE INFLATION MODEL

Most of the recent controversy about a new economy without the risk of inflation has focused on an allegedly lower NAIRU. However, careful testing demonstrates that the magnitudes of the critical coefficients of the standard model are quite stable over time, suggesting no improvement. Table 1 presents regressions that test the basic model and that demonstrate a high degree of explanatory power and coefficient stability over recent decades.

Table 1 Regressions Results for Wage Inflation

Dependent variable: Employment cost index inflation[a]

Sample period	(1) Full sample 1976:Q2–1998:Q1		(2) Full sample with estimated coefficient shifts 1976:Q2–1998:Q1		(3) First half of full sample 1976:Q2–1986:Q4		(4) Second half of full sample 1987:Q1–1998:Q1	
Explanatory variables	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic
Constant	.65	4.41	.69	2.59	.61	1.79	.74	4.78
Adjusted unemployment rate	–.58	–5.86	–.60	–4.01	–.55	–2.85	–.67	–5.38
Change in adjusted unemployed rate	–.67	–5.45	–.54	–3.79	–.56	–.322	–.98	–5.45
Output price inflation	.76	12.07	.78	10.74	.75	7.69	.91	7.08
Consumer price inflation (price inflation sum must = 1)	.24		.22		.25			
			Tests for Shifts:					
'87–'98 Filter (= 0 '76–'86; = 1 '87–'98)			–.06	–.22				
'87–'98 Filter × adjusted unemployment rate			–.01	–.03				
'87–'98 Filter × change in adjusted unemployment rate			–.47	–1.79				
R-bar squared	.765		.765		.786		.690	
Durbin-Watson statistic	2.160		2.260		2.260		2.320	
Standard error of the regression	.700		.700		.840		.531	
Sum of squared residuals	41.00		39.44		27.55		11.56	

[a] Notes on variable definitions: All inflation rates are calculated as annual rate of change equivalents, using changes in logarithms of the underlying variables. "Employment Cost Index" is calculated as 400 * log (eci/eci(t – 1)), where "eci" is the employment cost index for wages and salaries. "Output price inflation" is calculated as 400/14 * log (price(t – 1)/price (t – 15)), where "price" is the nonfarm private output deflator excluding housing, government, and excise tax components. "Consumer price inflation" is calculated as 100 * log (CPI(t – 1)/CPI (t – 5)). The dependent variable equals "wage inflation" minus two constrained effects: 1) "consumer price inflation" and 2) a special adjustment for changes in the minimum wage (estimated as 3 percent times the quarterly increase in the real minimum wage). The "adjusted unemployment rate" is a moving average over the current and four prior quarters of the official civilian unemployment rate minus the demographic adjustment described in the text. The "change in adjusted unemployment rate" is the current level minus the average of this, lagged two and three quarters. The time structures, such as the 14-quarter average of output price inflation, were chosen from earlier full-sample regression using polynomial distributed lags. To simplify presentation and verification of coefficient stability, the results reported here use the results of such regressions to select the length of the appropriate simple oving averages.

The sensitivity of wage inflation to the unemployment rate ($b1$) has been found to be roughly 50 percent by many researchers over many years. In other words, for each 1 percentage point that the unemployment rate falls below the NAIRU for one year, wage inflation rises by roughly 0.5 percentage point. Many researchers also include the change in the unemployment rate in their equations, but this inclusion changes only the short-run dynamics, not the long-run relationship produced by the core model.

The regressions presented in Table 1 follow a form that the author has used for forecasting purposes over many years with considerable success. Wage inflation, represented by inflation in the wage and salary component of the employment cost index, is estimated as a function of the unemployment rate adjusted for demographic shifts, the change in the adjusted unemployment rate, and two measures of price inflation. The regressions show that wage inflation increases by somewhat more than 0.5 percent for each percentage point decline in the level of the unemployment rate, consistent with most previous estimates.

Two tests of the stability of the coefficients appear in Table 1. The first test introduces a filter variable equal to 0 in the first half of the sample (1976 to 1986) and 1 in the second half. This filter is introduced in three dimensions: as a constant term and as cross-products with both the level and the change in the unemployment rate. The results appear in the second pair of columns of Table 1. As can be seen, the t-statistics of all three of these coefficients are too low to indicate that a statistically significant change has occurred since the mid 1980s. Moreover, the absolute magnitudes of the coefficients are generally small, although the effect of the change in the unemployment rate is larger in the later period.[1]

An alternative test examines the stability of the coefficients for all terms by running separate regressions for each half of the time period. These regression results are reported in column pairs (3) and (4). Although the coefficients may look somewhat different, using the F-statistic to compare the sums of squared residuals in the two separate equations with the sum of squared residuals in the original equation (where the coefficients are restricted to be the same throughout the period) yields a value of 0.39, well below the critical threshold that would indicate that a change had taken place.[2]

Finally, Figure 4 compares the actual path of wage inflation with that estimated by the base equation. The comparison gives no indication that historic relationships have bro-

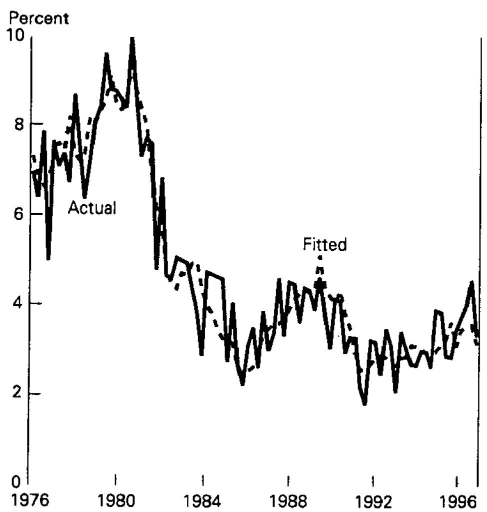


Figure 4 Actual and fitted values of wage inflation.

ken down in recent years. Thus, any novelty in the U.S. inflation process is not centered in the labor markets. Worker demands and employer responses appear to be unchanged.

Expectations of Price Inflation

According to the regressions in Table 1, price inflation enters the wage adjustment process as a blend of past output prices received by employers and consumer prices paid by employees. The former represent the value of work to the demander of labor; the latter represent the cost of living to the supplier of labor. Wage inflation appears more sensitive to changes in output prices than to changes in consumer prices. This is consistent with a fairly inelastic supply of labor compared with a relatively elastic demand for labor.

The sum of the coefficients for these two inflation rates is forced to equal one by subtracting the CPI inflation rate from both the wage inflation dependent variable and the output price inflation explanatory variable. This constraint reflects the assumption that real wages, rather than nominal wages, respond to prevailing unemployment conditions.[3]

The high explanatory power of this simple model of wage inflation, plus its consistency through time, gives credibility to the underlying assumption that expectations of price inflation are based on the momentum of inflation in the previous three to four years. During the 1970s and 1980s, considerable theoretical discussion focused on the importance of "rational expectations," the view that economic actors look forward in time and will adjust their expectations to incorporate likely policy actions. Some of the arguments for rational expectations challenged the validity of traditional inflation models, but little empirical testing was reported. Fortunately, a good source of directly measured inflation expectations can shed light on this issue: the University of Michigan Survey Research Center's time series of median expected inflation rates, derived from a sample of 500 households.

Table 2 presents a regression of inflation expectations from the University of Michigan's survey on past rates of inflation. Expected inflation, as measured by the survey, closely matches the momentum of inflation (Figure 5). The sum of the lagged inflation coefficients is 0.93. Expectations are also affected by the unemployment rate and by interest rates, measured here by the Treasury bill rate. Apparently, the public implicitly believes in the Phillips Curve. The estimated impact of the unemployment rate on expectations of price inflation is 0.44, very similar to the estimate of the response of price inflation to unemployment that we develop below.

The negative effect of interest rates can be interpreted as evidence that the public views monetary policy as effective in reducing inflation. According to the regression, each percentage point boost in interest rates is thought to cut inflation by 0.2 percentage point relative to its momentum. (This expected inflation response is a little higher than most econometric models would estimate for the first year after monetary tightening.)

Thus, the process whereby the American public forms expectations of inflation closely matches the approach assumed in conventional econometric models. These tests of a simple model of the public's inflation expectations also tend to validate the core wage inflation model just presented.

COMPLEMENTARY MODEL OF PRICE INFLATION

The absence of a significant change in the relationship between the unemployment rate and wage inflation suggests that the explanation for our recent favorable inflation experience

Table 2 Inflation Expectations of American Households[a]

Dependent variables: Median expected inflation over the next 12 months

Explanatory variables	Coefficient	Standard error	T-Statistic
Constant	1.25	.14	8.78
Lagged CPI inflation (polynomial distributed lag, 4th order, 30 potential lags, no other constraints)			
Sum of impacts	.93	.05	15.93
Current quarter	.18	.02	
First year sum	.52		
Second year sum	.19		
Change in inflation (current quarter at annual rate less prior year CPI inflation)	.06	.03	1.96
Current treasury bill rate	–.19	.04	–4.89
Current unemployment rate (adjusted)	–.44	.10	–4.58
R-bar squared	.964		
Durbin-Watson statistic	1.433		
Standard error of the regression	.408		

[a] University of Michigan Survey Research Center Quarterly Data, 1978:I to 1998:1.

lies with supply shocks directly affecting prices. The supply shocks that have been cited most frequently in this regard are as follows:

1. Fluctuations in oil prices. After rising sharply in 1990, oil prices declined in the early 1990s. They jumped up in 1996 but retreated in 1997 and plummeted in 1998.

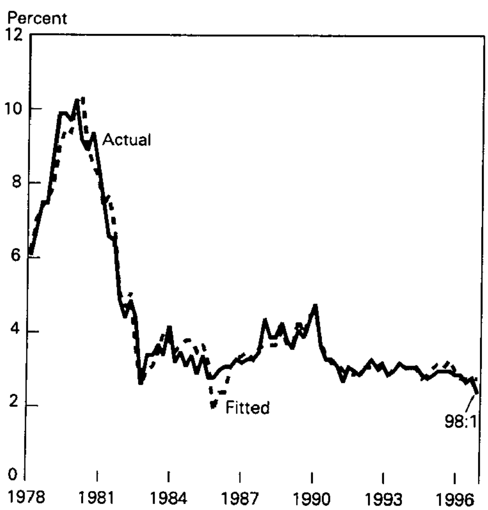


Figure 5 Inflation expectations: actual survey versus regression-fitted.

2. Lower costs for imported goods because of a strong U.S. dollar. Besides their direct effect, lower import prices also cut component costs and increase competitive pressure on domestic producers.
3. The favorable effect of a rising stock market on pension costs for employers providing defined benefit pension plans.[4]
4. Reduced inflation in health care costs because of structural changes in the industry arising from increased competition and pressures from employers and government.

Equation (2a), developed earlier, has been estimated to quantify the size of these shocks and to test the stability of the price inflation process (Table 3). The measure of price inflation used in the regressions in Table 3 is "Finished Goods Wholesale Price Inflation." This measure is closely correlated with both the consumer price index (CPI) and the GDP deflator-based measure of output prices in Table 1. In the author's judgment, however, it provides a cleaner measure of underlying inflationary pressures than the other measures. It has the virtue, relative to the CPI, that its methodology has not been subject to periodic improvements and, thus, it provides a more consistent measure of inflation over time.[5] It also does not include services, the prices of which are generally acknowledged to be measured imperfectly. The same conceptual model has been estimated for CPI and GDP deflator inflation rates, yielding wholly consistent conclusions. The estimated impacts of energy and imported goods prices are smaller, logically so, given the composition of the goods and services covered by these alternative inflation indexes.

"Finished Goods Wholesale Price Inflation" has been even lower in recent years than CPI inflation. Nevertheless, the results in Table 3 indicate that this moderation is readily explicable. As was true for wages, there is no evidence of a novel inflation process at work. Figure 6 compares actual inflation with the fitted values from the full sample regression.

For reader reference, the equation took the form:

$$p = w - q - b1 * U + b2 * pe + b3 * pm. \tag{2a}$$

The coefficient $b2$ measures the sensitivity of wholesale price inflation to relative energy prices. Its value of 0.11 means that a 10 percent rise in relative energy prices adds 1 percent to prevailing inflation. No significant lag was found. The coefficient $b3$ measures the sensitivity of inflation to relative import prices. A 10 percent rise in the relative cost of imported goods adds 2 percent to average wholesale prices; the impact is spread over a year.

Since the sensitivity of inflation to shocks from energy and import prices should relate to the relative size of these inputs in domestic costs and to competitive pricing calculations, the coefficients could vary through time. There is no indication of a shift for energy: The coefficient is 0.11 with a small standard error in all time periods tested. Consistent with the U.S. economy's rising openness to trade, however, the impact of changes in imported goods prices increases from 0.18 in the first half of the sample period to 0.25 in the second half. The speed of response is also faster, with more impact felt in the current quarter.

Prices reflect total labor costs, not just wages. Therefore, any surprise reduction in the cost of fringe benefits relative to base wages would also trim price inflation. To take account of supply shocks from health care, pensions, and other fringe benefits, the regression includes the inflation rate for total compensation relative to that for wages only. Any gap between inflation in total compensation and base wage inflation should translate into an equal-sized percentage point change in the rate of wholesale price inflation. The freely es-

Table 3 Regressions Results for Price Inflation

Dependent variable: Finished goods wholesale price inflation[a]

Sample period	(1) Full sample 1979:Q1–1998:Q1		(2) Full sample with estimated coefficient shifts 1979:Q2–1998:Q1		(3) First half of full sample 1979:Q2–1987:Q4		(4) Second half of full sample 1988:Q1–1998:Q1	
Explanatory variables	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic
Constant	−.66	−2.58	−.81	−1.87	−.76	−1.53	−.45	−1.13
Inflation rate of employment cost index	1.00		1.00		1.00		1.00	
Inflation rate of fuel and Power WPI	.11	13.09	.11	12.85	.11	7.33	.11	10.18
Inflation rate of imported goods prices (chain-weight deflator, goods except oil and computers)	.20		.22		.18		.25	
Current quarter	.11	3.65	.12	2.17	.05	.63	.20	2.62
Prior quarter	.09	2.66	.12	2.72	.13	2.68	.05	.80
Adjusted unemployment rate	−.36	−2.00	−.23	−.90	−.39	−1.23	−.50	−1.59
Inflation rate of total compensation relative to wages	1.06		1.38		1.32		.99	
Current quarter	.56	2.07	.63	2.21	.47	1.14	.77	2.01
Prior quarter	.50	1.83	.75	4.21	.85	2.00	.22	.53
			Test for Shifts					
'88–'98 Filter (= 0 '79–'87; = 1 '88–'98)			.30	.53				
'88–'.98 Filter × adjusted unemployment rate			−.33	−.82				
R-bar squared	.800		.795		.817		.756	
Durbin-Watson statistic	2.031		2.045		2.018		2.045	
Standard error of the regression	1.502		1.516		1.628		1.451	
Sum of squared residuals	157.9		156.3		76.9		71.6	

[a] Notes on variable definitions: All inflation rates are calculated as annual rate of change equivalents, using changes in logarithms of the underlying variables. "Finished Goods Wholesale Price Inflation" is calculated as 400 * log (wpi/wpi(t − 1)), where "wpi" is the quarterly average of monthly prices. The inflation rates for fuel and power, for imported goods, and for total compensation including fringes are the difference between the gross inflation rates for these categories and the inflation rate of the employment cost index. In other words, these terms reflect the shocks from these sources beyond prevailing inflation. The "adjusted unemployment rate" is a moving average over the four prior quarters of the official civilian unemployment rate minus the demographic adjustment described in the text.

timated coefficient of the gap between total compensation inflation and base wage inflation is 1.06, right on the mark.

Cyclical variations in the markup of prices over costs are captured in Table 3 by the same demographically adjusted unemployment rate used in the wage model. A 1 percentage point shift in the unemployment rate is estimated to produce a 0.36 percentage point re-

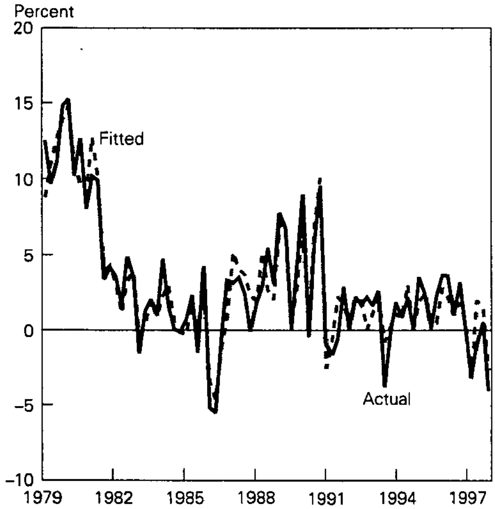


Figure 6 Actual and fitted values of inflation (wholesale price index, finished goods).

verse shift in prices relative to wages. An insignificant increase in this sensitivity (from 0.4 percent to 0.5 percent) is detected in the second half of the sample period.

The unemployment rate is used here to conform to the compact model of wage-price inflation developed earlier. However, another common indicator of excess demand in the goods market, used in many price inflation regressions in the literature, is the "vendor performance index." This is the percentage of purchasing managers surveyed who report slower deliveries. Substitution of a distributed lag of this indicator for the adjusted unemployment rate produced no significant change in any of the other coefficients.

SUMMARY AND CONCLUSION

Figure 7 summarizes the estimated roles of the key factors shifting inflation during the past two decades, while Figure 8 provides a close-up view of recent years. The bold line shows the total impact of cycles in excess demand, as indicated by variations in the unemployment rate multiplied by the coefficients in the wage and price regressions. The dotted, dashed, and ordinary lines are the similarly estimated effects of shocks: The dotted line shows the impact of imported goods prices, the dashed line indicates the effect of energy costs, and the thin solid line shows the influence of surges in fringe benefits relative to wages.

The bar is the sum of all effects. As can be seen, the bar is often close to the bold line. In other words, for most periods, unemployment is the dominant influence on inflation. This close relationship explains why public and professional discussion often focuses on this indicator to the exclusion of others.

However, the inflation surge in the late 1980s, described earlier and generally seen as validating the concept of the NAIRU, was actually due to a confluence of adverse inflation shocks from all other identified sources. The drop of unemployment to 5.3 percent had only a small impact.

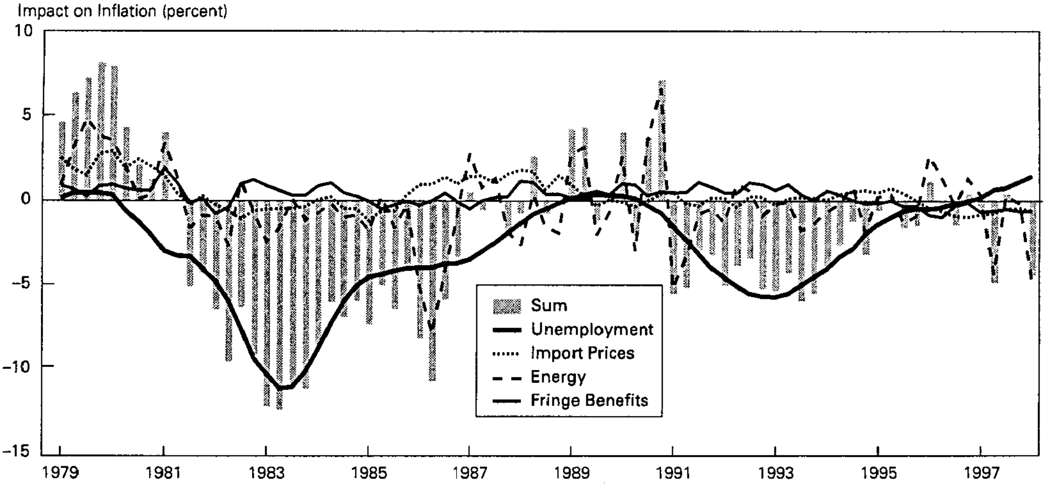


Figure 7 Key factors affecting inflation, first quarter 1979 to first quarter 1998 (wholesale price index, finished goods).

Conversely, the moderate inflation of recent years is due to a confluence of beneficial shocks from all factors other than unemployment. This can be seen most clearly in Figure 8. Tight labor markets in 1997 and early 1998 were tending to add a full percentage point to prevailing inflation. But declining prices for imported goods and energy were each pushing inflation sharply down. Slower growth in the cost of fringe benefits was also tending to relieve inflationary pressures. In other words, the increasingly tight labor market of 1997 and 1998 would have produced accelerating inflation, were it not for declining energy and import prices and the slow growth in fringe benefit costs.

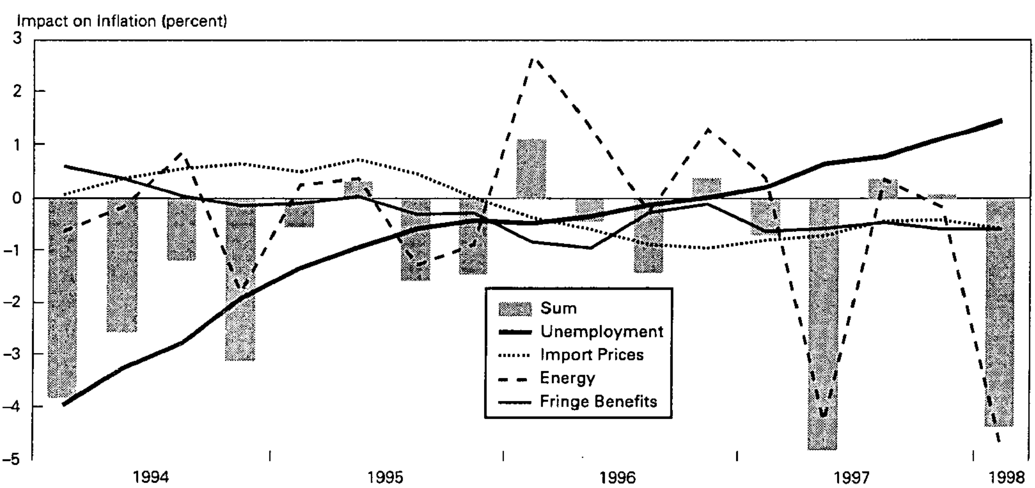


Figure 8 Key factors affecting inflation: close-up on first quarter 1994 to first quarter 1998 (wholesale price index, finished goods).

Although the combination of low unemployment and low inflation in the mid and late 1990s was remarkable, it did not herald a new economy, forever destined to enjoy high growth and low inflation. Rather, it reflected an unusual confluence of favorable supply shocks operating through traditional channels.

NOTES

1. The constant term for the second half is estimated to be just 0.06 lower than the first half, with a standard error of 0.26. The coefficient for the level of the unemployment rate is estimated to be very slightly smaller (0.005 higher, with a standard error of .0.18). The coefficient for the change is estimated to be stronger by 0.48 in the second half, with a standard error of 0.27.
2. The critical value of the F(4.80)-statistic at the 95 percent confidence level is 5.7.
3. A test of this constraint, not reported in Table 1, added CPI inflation to the regression. The estimated coefficient was–0.05 with a standard error of 0.04. In other words, the freely estimated sum of the inflation coefficients is 0.95 and is insignificantly different from 1.0.
4. Rising stock prices mean that employers do not have to contribute as much to fund their plans as they would otherwise.
5. At times changes have been made to the methodology used to measure the Consumer Price Index. In the early 1980s, for example, a switch was made to the calculation of the cost of home ownership, changing to a rental equivalence cost from an estimate of the cost of home purchase. Such changes are thought to have made the CPI a more accurate measure of "true" inflation. However, because the CPI is not revised historically to reflect these methodological enhancements, it does not provide a measure of inflation that is consistent over time.

REFERENCES

Brinner, R. E. 1977. The Death of the Phillips Curve Reconsidered. *The Quarterly Journal of Economics,* vol. 91, no. 3 (August), pp. 389–418.

Eckstein, O. and Brinner, R. 1972. The Inflation Process in the United States. *A Study Prepared for the Use of the Joint Economic Committee,* February.

Eckstein, O. and Wilson, T. 1962. The Determination of Money Wages in American Industry. *Quarterly Journal of Economics,* vol. 76 (August), pp. 379–414.

Friedman, M. 1968. The Role of Monetary Policy. *The American Economic Review,* vol. 58 (March), pp. 1–17.

Gordon, J. 1972. Wage-Price Controls and the Shifting Phillips Curve. *Brookings Papers on Economic Activity,* 2, pp. 385–421.

Keynes, J. M. 1939. Relative Movements of Real Wages and Output. *Economic Journal,* vol. 49 (March), pp. 34–51.

Lipsey, R. G. 1960. The Relation between Unemployment and the Rate of Change of Money Wages Rates in the United Kingdom, 1862–1957: A Further Analysis. *Economica,* vol. 27 (February), pp. 1–31.

Lucas, R. E., Jr. 1972. Econometric Testing of the Natural Rate Hypothesis. *The Econometrics of Price Determination Conference,* O. Eckstein, ed. Board of Governors of the Federal Reserve System and Social Sciences Research Council.

Muth, J. F. 1961. Rational Expectations and the Theory of Price Movements. *Econometrica,* vol. 29 (July), pp. 315–35.

Perry, G. 1970. Changing Labor Markets and Inflation. *Brookings Papers on Economic Activity,* 3, pp. 411–48.

Perry, G. 1970. Inflation and Unemployment. The Brookings Institution, Reprint no. 188, December.

Phelps, E. S. 1967. Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time. *Economica,* vol. 34 (August), pp. 254–81.

Phelps, E. G. 1968. Money-Wage Dynamics and Labor Market Equilibrium. *Journal of Political Economy,* vol. 76 (July/August), pp. 678–711.

Phillips, A. W. 1958. The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1857. *Economica,* vol. 25 (November), pp. 283–99.

Theil, H. 1971. *Principles of Econometrics.* New York: John Wiley & Sons, Inc.

50
Central Bank Inflation Targeting

Glenn D. Rudebusch
Research Officer, Federal Reserve Bank of San Francisco, San Francisco California

Carl E. Walsh
Professor of Economics, University of California, Santa Cruz, California

This chapter summarizes the papers presented at a conference on Central Bank Inflation Targeting held March 6–7, 1998, under the joint sponsorship of the Federal Reserve Bank of San Francisco and Stanford University's Center for Economic Policy Research.

The five conference papers (listed at the end) were centered around measuring or evaluating the degree to which inflation should be the focus of the operating framework used to implement monetary policy. Explicit inflation targeting has been adopted by a number of central banks around the world. In practice, inflation targeting is best described as an operational framework for policy decisions in which the central bank makes an explicit commitment to conduct policy to meet a publicly announced numerical inflation target within a particular time frame. All the conference papers—like much of the research literature—model this framework by focusing on "policy rules," that is, specific formulas for adjusting the policy instrument in response to inflation (or forecasts of inflation) and, in some cases, to the state of the economy as measured by the gap between GDP and potential. The research reported in these papers suggested that these simple inflation targeting rules do a good job of capturing the way many central banks behave and that these rules perform well in achieving a balance between output variability and inflation variability.

Four of the conference papers compare the performance of various policy rules in model simulations: The better a policy rule's performance, the more stable output and inflation. This focus reflects a common emphasis in recent policy discussions on the tradeoff between variability in output and variability in inflation (Walsh 1998).

The fifth paper takes a different tack. Rather than simulating data from a model, this paper uses historical data to estimate the policy rules that central banks have used. This exercise, by providing evidence on how central banks have actually implemented policy, served to complement the results from the first four papers on which rules perform best.

Reprinted from: *FRBSF Economic Letter,* No. 98-17 (May 22, 1998).

POLICY RULES FOR INFLATION TARGETING

The Rudebusch-Svensson paper uses a small empirical model of the U.S. economy to examine the performance of policy rules that are consistent with a policy regime of inflation targeting. Two special features of this research are the choice of an empirical model and the large set of policy rules explored.

The authors choose their model based on three considerations: it is a small linear model with fairly clear and tractable results, it exhibits properties that accord with the spirit of many policy discussions (e.g., a plausible interest rate elasticity), and it is able to reproduce the salient dynamic features of the data relative to an unrestricted model (such as a Vector AutoRegression). However, unlike the other three papers evaluating policy rules, the Rudebusch-Svensson model has no forward-looking expectational terms in the equations for price and spending determination. While the Rudebusch-Svensson model appears to fit the data quite well without these forward-looking terms, their absence raises some questions about the robustness of the policy evaluation results.

The authors find that some simple policy rules that use inflation forecasts do remarkably well in minimizing output and inflation variability. In particular, an implicit rule that adjusts the policy instrument so that the forecasted inflation rate matches the target rate at a horizon of about three years does surprisingly well. Such a rule also appears to be close to the actual decision framework of many central banks under inflation targeting.

OPERATIONAL POLICY RULES

The McCallum-Nelson paper provides a different example of the type of small, econometrically estimated model that can be used to investigate monetary policy questions. Their model had two distinguishing features. First, savings and portfolio decisions are consistent with optimizing behavior by households. This is important, since it serves to incorporate forward-looking expectations into the model. The expected effects of monetary policy in the future can then have real effects in the present as decisions about consumption, savings, and money holdings adjust in response to these expectations. Second, McCallum and Nelson employ two alternative specifications designed to capture the sluggish adjustment of prices to changes in macroeconomic conditions.

In addition to differing from the Rudebusch-Svensson paper in the choice of model, the McCallum-Nelson paper differs in the approach taken to evaluating policy rules. Rudebusch-Svensson generally start by specifying an objective function for the policymaker and then deriving an optimal policy rule. In contrast, McCallum and Nelson start with policy rules that seem to capture actual central bank behavior and then report the value of output and price level volatility for different values of the response coefficients in the policy rules. The general type of policy rule they evaluate is of the form suggested by Taylor in which the policy instrument (a nominal rate of interest) is adjusted in response to inflation and the output gap. In addition to Taylor type rules, the McCallum-Nelson paper includes an analysis of interest rate and base money growth rate rules designed to target nominal income.

One advantage of the McCallum and Nelson approach is that it avoids the need to assume an objective function. A disadvantage is that there is no natural way to rank the resulting outcomes with different policy rules. One interesting finding, though, was that policy rules that incorporated some degree of interest rate smoothing seemed to lower both price and output variability. The nature of the inflation-output variability trade-off also appears to be sensitive to the specification of the price adjustment equation.

IMPLEMENTING PRICE STABILITY

Like the McCallum-Nelson paper, the Tetlow-Williams paper also examines simulations of a forward-looking model. However, the model Tetlow and Williams used—which is the Federal Reserve Board staff's main model of the U.S. economy—is large (containing about 30 behavioral equations and several hundred identities). Consistent with the other conference papers, monetary policy is represented by a rule for setting a nominal short-term interest rate, in this case, the federal funds rate.

The Tetlow-Williams paper analyzes two issues related to inflation targeting. The first is whether target bands for inflation can improve macroeconomic performance. Under a target band system, the near-term inflation target is allowed to fluctuate but must remain between upper and lower limits. The role of such target bands is of interest since some inflation targeting countries have used bands. Tetlow and Williams conclude that with forward-looking expectations, bands can be useful in concentrating the public's expectations of future inflation. To serve this purpose, however, the bands need to be reasonably narrow.

The second issue Tetlow and Williams consider involves the constraint that nominal interest rates cannot be less than zero. This lower bound reflects the fact that no one would lend money with the sure prospect of getting less of it back in nominal terms (a negative nominal interest rate) because just holding on to the cash (say, under a mattress) will ensure a zero nominal interest rate. With central banks typically using the level of short-term interest rates as the instrument of monetary policy, the zero bound on nominal interest rates might affect the ability of central banks to implement monetary policy, and, particularly, to lower interest rates sufficiently to stimulate the economy from recession. This constraint is more likely to be a factor for very low inflation targets (recall that the nominal interest rate is approximately equal to the sum of the inflation rate and the real interest rate). Tetlow and Williams suggest that, except with an inflation target of zero, the lower bound on interest rates is unlikely to be a serious problem.

WHEN ECONOMIC BEHAVIOR CHANGES

The Amano-Coletti-Macklem paper also examines monetary policy rules with a large forward-looking econometric model—in this case, a model of the Canadian economy developed at the Bank of Canada. The authors focus on how changes in economic behavior, and hence the equations of the model, change the nature of the optimal monetary policy rule. The authors consider three changes in economic behavior that are motivated by developments in the 1990s: an increase in monetary policy credibility, a flattening of the Phillips curve linking inflation and unemployment, and a greater degree of counter-cyclical activism of fiscal policy.

The policy rules analyzed are "inflation-forecast-based" (IFB) rules. According to this class of rules, the central bank raises (lowers) short-term interest rates whenever the rule-consistent inflation forecast is above (below) the target for inflation. As Amano, Coletti, and Macklem note, this type of rule plays an important role in policy analysis at two leading inflation targeting central banks, the Bank of Canada and the Reserve Bank of New

Zealand. (IFB rules are similar in spirit to the implicit rule in the Rudebusch-Svensson paper, but they are too restrictive to perform well in that analysis.)

Perhaps the most interesting results in this analysis concern changes in central bank credibility, that is, changes in the degree to which the public believes the central bank will meet its inflation target. The central bank is assumed to follow a rule that alters short-term interest rates by some proportion of the difference between the two-year-ahead inflation forecast and the inflation target. As credibility increases, the central bank can attain more stable inflation and output, but this is typically true only if the central bank adjusts its rule and changes the amount by which it reacts to inflation forecasts. Thus, the best rule for policy may have to adjust to changes in the macroeconomy, even within a framework of inflation targeting.

SOME INTERNATIONAL EVIDENCE

In contrast to the other four papers, the conference paper by Clarida, Galí, and Gertler provides empirical evidence on the way policy actually has been implemented since 1979 by the central banks of Germany, France, Italy, Japan, the United Kingdom, and the United States. Clarida, Galí, and Gertler begin by assuming policy responds to expected future inflation and to the current expected output gap. For the Bundesbank, the Bank of Japan, and the Federal Reserve, the empirical evidence supports the view that each responds to movements in expected inflation, with a rise in expected future inflation causing a contractionary shift in policy. In each case, however, there is evidence that policy responds to output conditions as well. They also find that the Bundesbank and the Bank of Japan seem to respond to exchange rate movements and to the U.S. federal funds rate, although these effects are small.

Estimating the policy rule followed by the Banks of England, France, and Italy is complicated by their participation in the European Exchange Rate Mechanism (ERM). From 1990 until the European currency crisis in September 1992, these countries abandoned independent monetary policy in order to fix their exchange rates with the Deutschmark. In an interesting approach to understanding monetary policy during this period, the authors conduct a counterfactual experiment in which they estimate the domestic interest rate that would have occurred in each country if policy had been conducted using the same rule as employed by the Bundesbank. Thus, they estimate the interest rate the Bank of England, for example, would have set if it had responded to domestic inflation and output as the Bundesbank did without being constrained to maintain a fixed exchange rate. The difference between this estimated policy setting and the actual interest rate needed to maintain the fixed exchange rate provides a measure of the economic stress in each country as a result of ERM membership. For England, France, and Italy, Clarida, Galí, and Gertler find that their stress measure provided evidence of growing stress in the period immediately preceding the exchange rate crisis in September 1992.

CONFERENCE PAPERS

Amano, R., Coletti, D. and Macklem, T. Monetary Rules when Economic Behavior Changes. Bank of Canada.

Clarida, R., Galí J. and Gertler, M. Monetary Policy Rules in Practice: Some International Evidence. Columbia University.
Rudebusch, G., and Svensson, L. E. O. Policy Rules for Inflation Targeting. Federal Reserve Bank of San Francisco.
McCallum, B. T., and Nelson, E. Performance of Operational Policy Rules in an Estimated Semi-Classical Structural Model. Carnegie-Mellon University.
Tetlow, R., and Williams, J. Implementing Price Stability: Bands, Boundaries and Inflation Targeting. Federal Reserve Board.

REFERENCE

Walsh, C. E. 1998. The New Output-Inflation Trade-off. *FRBSF Economic Letter* 98-04 (Feb. 6).

This Page Intentionally Left Blank

51

A Better CPI

Allison Wallace

Research Associate, Federal Reserve Bank of San Francisco, San Francisco, California

Brian Motley

Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

The monthly consumer price index (CPI) is the most oft-cited measure of inflation and one of the most important and closely watched statistics in the U.S. economy. It is an indicator of how well the Federal Reserve is doing in achieving and maintaining low inflation, and it also is used to determine cost-of-living adjustments for many government programs, collective bargaining contracts, and individual income tax brackets.

Since 1995, the Bureau of Labor Statistics (BLS) has been eliminating biases that cause the index to overstate inflation, and further changes will come in January 1999. These changes are expected to create a more reliable index and by 1999 will have lowered measured CPI inflation by more than half a percentage point. Although this may seem like a small change, the effect of these changes is permanent so that measured inflation will be lower by this amount in all future years.

It is important that the CPI should measure inflation accurately or that the degree of bias be known. Macroeconomic policymakers such as the Fed then can take approriate steps to keep inflation low, and the public can be informed about their successes and failures in achieving their goal. Also, if the CPI does not measure inflation correctly, cost-of-living adjustments based on it will have different effects from those desired when the commitments to make these adjustments were made. For example, adjusting Social Security benefits based on an upwardly biased CPI may shift spending power from the young toward the old.

This chapter will explain the types of biases that cause the CPI to overstate inflation, BLS actions to remove these biases, and the possible implications for monetary policy.

SOURCES OF BIAS IN THE CPI

The BLS has been studying possible biases in the CPI for a long time. The issue gained national prominence in 1996 when the Congress commissioned a panel of experts on price measurement issues, chaired by Michael Boskin of Stanford University, to examine biases

---

Reprinted from: *FRBSF Economic Letter,* No. 99-05 (Feb. 5, 1999).

in the CPI. Their report, "Toward a More Accurate Measure of the Cost of Living," identified four major sources of bias and estimated that they caused the CPI to overstate inflation by 1.1 percentage points per year at that time.

Substitution Bias

Substitution bias occurs because the CPI measures the price changes of a *fixed basket* of goods and services and thus does not capture the savings that households enjoy when they change their spending in response to relative price changes of goods and services. For example, a rise in the price of beef leads people to buy more chicken in order to keep their food costs down. The Boskin report identifies two types of substitution bias. The first, estimated to raise measured inflation by 0.25 percentage point annually, is *lower-level* substitution bias and occurs when consumers substitute between similar items within a category (e.g., substituting between pippin and gala apples). The second type, estimated to boost inflation by 0.15 percentage point annually, is called *upper-level* substitution bias and occurs when consumers substitute between items from different categories (computers for television sets, for example) in response to price changes.

Currently, the market basket that is priced is updated approximately once a decade. The new basket, based on consumers' purchases in 1993–1995, was introduced into the index earlier last year. As we move further away from this date, upper-level substitution bias may increase as spending patterns move away from the basket on which the present CPI is based.

Outlet Bias

This type of bias is similar to substitution bias, but refers to *where* households shop rather than to *what* they purchase. Over the past 15 years, for example, the growth of discount stores has helped consumers lower their expenditures by offering high-volume purchases at reduced prices. The expansion of these establishments has not been adequately represented in the CPI, thus creating an upward bias of prices estimated at 0.1 percentage point per year. A similar problem may arise in the future as shopping online becomes more widespread.

New Product Bias

This bias occurs because new products, such as VCRs and cellular phones, are not introduced into the index until they are commonplace items. This means that the substantial price decreases and quality increases that occur during the early years following introduction are not captured by the index. A problem of dealing with this bias is that the BLS can never know in advance which of the many new products introduced each year will be successful and hence worthy of inclusion in the CPI.

Quality Bias

This bias arises because some of any increase in the price of an item may be due to an improvement in quality, rather than being a pure price increase. For example, when car prices rise, this may be due to the addition of seat belts, air bags, or anti-smog devices, or to pure price inflation. In the case of cars, the BLS often uses the price of the new item as an optional feature before it becomes standard equipment as an indicator of what the improvement is worth to consumers. Quality improvements in other areas—such as medical care—are more difficult to measure so that bias is more likely to occur. And features of a product that become mandatory—such as seat belts, which buyers are forced to purchase even if they would prefer not to—are particularly difficult to handle.

The combination of quality bias and new product bias was estimated by the Boskin Commission to boost measured inflation by 0.6 percentage points annually. Any estimate of this magnitude, though, is inherently subjective and subject to debate.

CHANGES IN THE CPI SINCE 1995

The BLS began to address the bias in the CPI even before the Boskin Commission was convened. For example, in 1995 the BLS introduced a new sampling procedure to determine which outlets to visit to obtain price data for specific items and what weights to apply to those item prices. The old procedure put too much weight on items that were temporarily cheap at that outlet, so when their prices rose back to their normal level, this registered as an increase in inflation. That same year, the BLS also revised sampling methods to remove the effects of substituting between brand drugs and generic drugs. In 1997, the BLS adopted some of the procedures used to measure hospital prices that are used in the producer price index.

Spurred by the work of the Boskin Commission, the BLS introduced further changes to confront substitution and outlet bias. The BLS has sought added funds to update the commodity and outlet samples more frequently and to do so at lower cost. Updating the commodities and the outlets more often should reduce substitution bias by allowing the published index to include more of household's responses to observed price changes.

There also have been attempts to reduce quality bias. For example, the BLS is expanding the use of hedonic regressions to compare quality differences. Hedonic regressions attempt to estimate econometricaly the value that households put on quality differences. These methods are currently used for measuring quality distinctions in the categories of apparel, rent, and computers and peripheral equipment, and as of January 1999, they will be used for television prices. Research is underway to extend this technique to other categories.

PLANNED CHANGES

Future changes will address substitution and new product bias further. Beginning next January, the BLS will attempt to reduce lower-level substitution bias by using a geometric mean formula to calculate price changes for many of the basic categories of the CPI. The geometric mean formula assumes the household spends the *same proportion of its outlays* on each category; the arithmetic mean, which is now in use, assumes the household always buys the *same quantity* of each item. Using the geometric mean implies that if the price of pippin apples rises 10%, the quantity of pippins bought decreases 10%, so that the average household spends the same amount on pippins. This assumption is, of course, arbitrary, but it may give a better overall result than the assumption of no substitution. The BLS has estimated the monthly CPI using both methods over the past few years and concluded that adopting the geometric mean formula will reduce measured inflation by about 0.2 percentage point annually (the Boskin Commission estimated that lower-level substitution bias raised measured inflation by about 0.25 percentage point). The new formula will be used to calculate inflation in most categories in the CPI except those in which consumers cannot easily substitute between alternatives when there is a relative price change: for example, no change in relative prices would cause a person with a heart problem to consider buying a hearing aid. The new approach will be used only to aggregate price changes of individual commodities into broad commodity groups; these groups will continue to be aggregated into the overall index using fixed-quantity weights. After this change has been completed, the reduction in measured inflation due to methodological changes since 1995 is estimated to be 0.6 percentage point.

The BLS will update more frequently the expenditure weights obtained from the Consumer Expenditure Survey. The BLS is in the process of deciding how often the weights will be updated and wishes to increase the sample size of the survey to allow the use of only two years of expenditure data to construct the weights. These changes will make the market basket more representative of what consumers are actually purchasing and will introduce new products in a more timely manner.

To tackle upper-level bias more extensively, the BLS will produce an official "superlative index" starting in 2002 in addition to the CPI. This term—originally coined by Canadian economist Erwin Diewert—refers to an index that approximately removes all substitution bias for most assumptions about household preferences. One superlative index is the Fisher ideal index, which uses a combination of weights from both the original market basket and the current market basket to take into account changes in consumer spending patterns. A version of this method is currently used in constructing the national income and product accounts. A superlative index would be subject to revision unless it were published with a time lag, because it takes time to gather data on current expenditures; thus, the CPI cannot be converted to a superlative index. However, recent research suggests that if the current market basket were initially represented by the basket for the previous year, the degree of subsequent revision required would be small.

IMPLICATIONS FOR PUBLIC POLICY

Since the Fed uses the CPI as an indicator of price inflation, a more accurate index should make anti-inflationary monetary policy more effective. The public will have a better indicator to check how well the Fed is doing its job. The effect on policy is not likely to be large, however, because the Fed already takes account of the best available estimate of the remaining biases in the published data. On the other hand, when inflation was higher, modest errors were less important since it was *always* appropriate to make policy with a view to reducing inflation. But now that we are closer to zero inflation, an accurate measure is more important especially if policymakers wish to avoid a situation where actual inflation is negative. Under the Boskin Commission estimate of a total upward bias of 1.1 percentage points, a goal of zero inflation would be equivalent to an actual goal of –1.1%.

Some economists argue that negative inflation is undesirable in the long run. If it is difficult to reduce nominal wages, then it may not be possible to lower individual workers' real wages if prices are not rising. This may mean that it is difficult to provide appropriate incentives to move unneeded workers into other lines of work where they are more useful. Similarly, declining prices may cause the real value of debts to rise, which could cause some otherwise sound businesses to fail. Finally, if we want our tax and transfer system to be invariant to inflation, an accurate CPI is essential, so that the task of adjusting tax and transfer payments to price changes can be done quickly, easily, and without undue dispute.

CONCLUSION

Ongoing research is necessary to identify biases in the CPI. Changes to this index are inevitable as the BLS strives to maintain an accurate measure of inflation in our dynamic economy. By 1999, about half of the bias identified by the Boskin Commission will have been removed. Although the changes may be inconvenient—because they make the current index less comparable with the past index—they will lead to an improved measure of actual inflation and, thus, a better CPI.

This Page Intentionally Left Blank

52

Is There an Inflation Puzzle?

Cara S. Lown
Research Officer, Federal Reserve Bank of New York, New York

Robert W. Rich
Economist, Federal Reserve Bank of New York, New York

Historically, inflation has followed a fairly predictable course in relation to the business cycle. Inflation typically rises during an economic expansion, peaks slightly after the onset of recession, and then continues to decline through the first year or two of recovery. During the present U.S. expansion, however, inflation has taken a markedly different path. Although more than six years have passed since the 1990–91 recession, inflation in the core CPI (the consumer price index excluding its volatile food and energy components) has yet to accelerate (Figure 1). Moreover, during the last three years, inflation has remained stable despite projections of higher expected inflation from the Blue Chip Consensus forecast and contrary to traditional signals such as the run-up in commodity prices experienced from late 1993 to early 1995.

Economists and policymakers have referred to the restrained behavior of prices during this long expansion as an "inflation puzzle." In a recent interview, Robert T. Parry, president of the Federal Reserve Bank of San Francisco, commented, "I have a question mark, and it leads me to recommend vigilance with regard to inflation, but I do have to note that things have turned out well. . . . [We've] either been lucky, in which case the old relationships will reassert themselves, or [we've] got a new regime under way. And I don't think we know enough at this point to know which of those two things is operative."[1] As Parry suggests, two different types of explanations could account for the recent behavior of inflation. The failure of inflation to accelerate may reflect the effects of temporary factors unique to this expansion. Alternatively, the unexpectedly low level of inflation may indicate a permanent change in the way inflation reacts to economic growth and other related variables.

Each of these explanations holds important implications for the conduct of monetary policy. The Phillips curve, the principal tool used by economists to explain inflation, has been subject to systematic overprediction errors during the past few years. If these errors reflect the influence of temporary factors, then the Phillips curve relationship should ultimately regain its stability. However, if these errors reflect a permanent change in the dy-

Reprinted from: *FRBNY Economic Policy Review*, v. 3, n. 4 (Dec. 1997) 51–69.

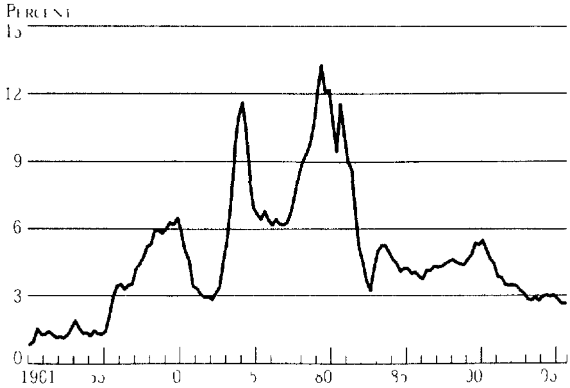


Figure 1 Core CPI (percentage change from a year ago)

namics of the inflation process, then economists could no longer view the Phillips curve as a reliable guide in forecasting inflation.

Because labor costs are an important factor in determining prices, the recent slowdown in compensation growth has been cited in both types of explanations for the inflation puzzle. Some commentators argue that this slow-down in compensation growth, attributable largely to declining benefit costs, has acted as a supply shock and has temporarily lowered inflation relative to its historical proximate determinants. Others contend that a permanent change in compensation growth, resulting from heightened job insecurity and its constrictive effect on wage growth, has led to a fundamental shift in the inflation process.

This article explores the inflation puzzle and investigates whether compensation has acted as either a temporary restraint on inflation or as the underlying source of a new inflation regime.[2] After reviewing the recent behavior of inflation, we specify and estimate a traditional price inflation Phillips curve model over the 1965–96 period. Our results show that in late 1993 the model begins to systematically overpredict inflation and appears to break down.

We then modify our traditional Phillips curve specification by incorporating compensation growth as an additional determinant of inflation. With this variable, the model's explanatory power improves significantly, and it tracks inflation much more accurately over the current expansion. The restored stability of the model appears to rule out the view that inflation's recent behavior reflects a fundamental shift in the inflation process.

Finally, we specify and estimate a wage-inflation Phillips curve model quantifying the restraint in compensation growth over the post 1991 period. Our findings indicate that compensation growth has been weak during this expansion, especially from late 1992 through early 1995, a period that corresponds to the observed breakdown in our traditional Phillips curve specification. This coincidence further supports our conclusion that compensation's slow growth has temporarily restrained inflation during this expansion.

THE EMERGENCE OF THE INFLATION PUZZLE

Contrary to expectations, inflation has not accelerated since the end of the 1990–91 recession. Yet variables commonly regarded as inflation indicators have remained at levels that usually coincide with an inflation pickup. The level of the actual unemployment rate relative to the nonaccelerating inflation rate of unemployment (NAIRU) is one such variable. The NAIRU represents the rate of unemployment that is consistent with stable inflation. Unemployment rates below (above) the NAIRU are thought to signal higher (lower) inflation in wages and prices. As the upper panel of Figure 2 shows, the unemployment rate has been below 6 percent—the consensus estimate of the NAIRU at the beginning of this expansion—since late 1994. Even if the NAIRU has declined below 6 percent during the 1990s, as some analysts argue, there is little direct evidence suggesting that it has tracked the unemployment rate or fallen low enough to be consistent with the level of inflation observed since 1995.[3]

Like the NAIRU, the capacity utilization rate has stayed at levels that typically signal higher future inflation (bottom panel of Figure 2). In the past, capacity utilization rates in excess of 82 to 84 percent were associated with rising inflation because of the onset of supply shortages and bottlenecks in production (Boldin 1996). Capacity utilization has moved down from its peak of almost 85 percent; still, it has stayed above or close to 83 percent since 1994.

Consistent with these two indicators, the Blue Chip Consensus forecast overpredicted inflation from 1992 to 1995 by progressively larger margins of error each year (Figure 3). Estimated price-inflation Phillips curves have also systematically overpredicted inflation in the past couple of years. The Phillips curve's recent failure in forecasting price changes contrasts sharply with its long-standing reliability in predicting short-run movements in inflation. We now turn to a discussion of the Phillips curve and its recent record in forecasting inflation.

A TRADITIONAL PRICE-INFLATION PHILLIPS CURVE

The origin of the Phillips curve can be traced back to the 1950s, when A. W. Phillips documented an inverse relationship between the rate of change of nominal wages and the level

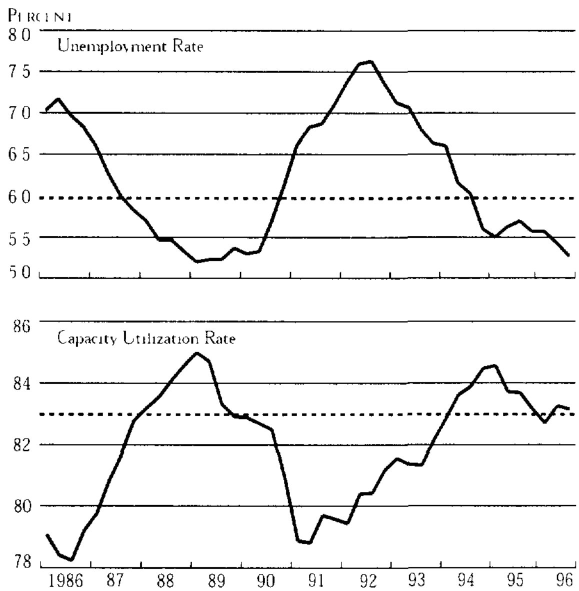


Figure 2 Unemployment and capacity utilization rates.

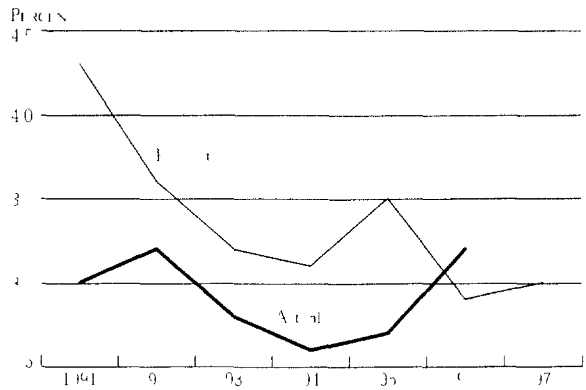


Figure 3 CPI inflation actual and forecast

of unemployment in the United Kingdom His findings were interpreted as establishing a wage adjustment process in which low levels of unemployment represent tight labor markets that signal or coincide with, accelerating wage growth Although the term "Phillips curve" still refers to the posited relationship between nominal wage or price changes and various indicators of real economic activity the econometric modeling of this relationship has changed considerably over the years [4]

Modern versions of the Phillips curve incorporate several features that differentiate them from earlier descriptions of the behavior of nominal wages and prices [5] For example, in current models the output gap (the log ratio of actual to potential real GDP) and the unemployment gap (the difference between the actual rate of unemployment and the NAIRU) figure importantly as measures of excess aggregate demand pressure in the economy In addition, current models recognize the role that expected inflation plays in wage bargaining and price setting and typically include past rates of inflation as a proxy for this expectation [6] Finally, modern Phillips curve models include variables to control for supply shocks such as the oil price increases of the 1970s As Fuhrer (1995) notes, many of these developments were anticipated by Phillips in his original discussion

We begin our empirical analysis by specifying a traditional price-inflation Phillips curve model The model allows for a more formal investigation of the stability of the Phillips curve relationship during the current expansion In addition, the model will serve as a benchmark to evaluate compensation growth s role in explaining recent movements in inflation

Our traditional Phillips curve model is given by

$$INF_t = \alpha_0 + \alpha_1 GDPGAP_{t-1} + \alpha_2(\Delta GDPGAP_{t-1}) + \sum_{i=1}^{3} \alpha_{2+i} INF_{t-i} + \sum_{i=1}^{2} \alpha_{5+i} OILGAP_{t-i} + \varepsilon_t, \qquad (1)$$

where

INF = inflation measured by the growth rate of the core CPI,

$GDPGAP$ = the output gap measured by the log ratio of actual to potential real GDP,

$\Delta GDPGAP$ = the first difference or change in the output gap,

$OILG^{+}$ = the net positive change in the real price of oil, and
ε = a mean zero, serially uncorrelated random disturbance term.

Equation 1 provides a general specification for the rate of change in prices and is similar to other models currently used in the Phillips curve literature.[7] In the terminology of Gordon (1996), the specification embodies the "triangle" model of inflation: the set of explanatory variables is meant to capture the effects of demand, inertia, and supply considerations on inflation.

The model uses the output gap (the percentage deviation of real GDP from potential GDP), shown in Figure 4, as a measure of excess aggregate demand pressure.[8] A positive (negative) output gap indicates that the economy is operating above (below) potential GDP and would thus generate upward (downward) inflationary pressure on prices. Following the methodology in Gordon (1977, 1996) and Fuhrer (1995), we also include the quarterly change in the output gap variable to allow for a rate-of-change effect so that the pressure on prices depends on how quickly the output gap narrows or widens.

The remaining basic determinants of inflation include its own lagged values and oil prices. To incorporate price inertia effects, we include lagged inflation terms in the model. In the past, researchers used lagged inflation rates as a proxy for expected inflation. In modern versions of the Phillips curve, however, this interpretation has been deemed overly restrictive (Gordon 1996). Instead, past inflation rates are viewed as capturing the dynamics of price adjustment related to expectations formation as well as the importance of institutional factors such as wage and price contracts and delivery lags in the economy.

Our benchmark model also includes a measure of the net positive change in real oil prices to account for the influence of supply shocks.[9] This oil price variable is the only notable departure from other conventional Phillips curve specifications and allows for an asymmetric effect of oil price changes on inflation (Figure 5). In other words, while oil price *increases* appear to affect inflation, oil price *decreases* do not seem to be important.[10] The construction of the supply shock variable follows the approach in Hamilton (1996) and is designed not only to model the asymmetric effects of oil price changes, but also to account for the observed increase in the volatility of oil prices over the post-1986 period. Because the core CPI has no energy price component, our supply shock variable attempts to capture any indirect effect of oil price increases on inflation.

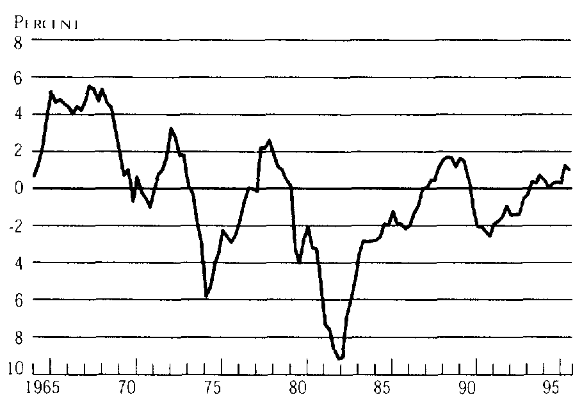


Figure 4 The output gap (percentage difference between actual and potential GDP).

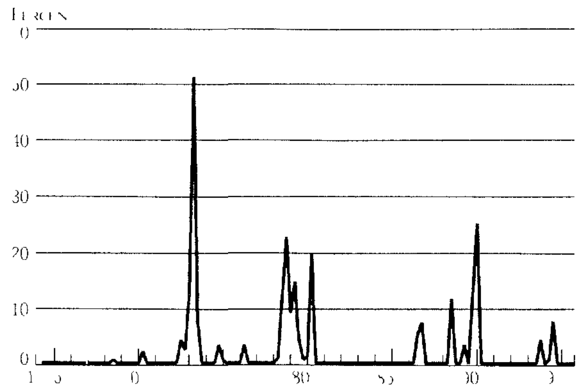


Figure 5 Net positive change in real oil prices

Although our traditional price inflation Phillips curve takes real oil prices as exogenous, we include only lagged values of the output gap as regressors in order to avoid simultaneity bias arising from the endogeneity of this variable The lag lengths in equation 1 are selected by maximizing adjusted R^2 (a measure of the model's ability to explain inflation), by searching over one to four lags of inflation and the output gap, and by searching over zero to four lags for the net positive change in the real price of oil [11]

Model Estimation

We estimate equation 1 using the method of ordinary least squares (OLS) for quarterly data from the first quarter of 1965 to the third quarter of 1996 Parameter estimates are presented in Table 1 For the full sample period, the value of the adjusted R^2 indicates that the model can explain a high proportion of the variation in inflation In addition, the Ljung-Box (1978) Q-test statistic—a general test for serial correlation in the regression residuals—does not reveal any evidence of model misspecification

The estimation results also indicate that both the level of the output gap variable and the rate-of-change effect are highly significant and have the expected positive signs The two lagged values of the net positive change in the real price of oil are also highly significant with the anticipated positive signs The three lags of the inflation rate are generally significant and we are unable to reject the hypothesis that the sum of the coefficients equals unity ($\alpha_3 + \alpha_4 + \alpha_5 = 1$) at conventional significance levels The latter restriction follows from the natural rate hypothesis and has been previously imposed in the estimation of Phillips curves to make the level of potential output (or the unemployment rate) independent of inflation in the long run

Model Stability over the 1992–96 Period

We conduct two exercises to examine the stability of the model from 1992 to 1996 First, we apply Chow (1960) split-sample tests to test the null hypothesis of constant parameters against the alternative hypothesis of a onetime shift in the parameters at some specified date One test compares the estimates obtained using the data from one subperiod (1965–91) with the estimates using the full sample [12] Another test employs dummy variables for the entire parameter vector for one subperiod (1992–96) and then tests the joint significance of the dummy variables.[13] As shown by the reported value of the two test statistics in Table 2, we fail to reject the null hypothesis of parameter stability for the post-1991 period at conventional significance levels.[14]

As a second exercise, we construct dynamic out-of-sample forecasts from the traditional price-inflation Phillips curve. This simulation provides a more stringent test of model stability by relying on lagged predicted values of inflation rather than the lagged actual values of inflation to construct the subsequent one-quarter-ahead forecasts of inflation. In addition, the Chow tests may suffer from low power because they are conducted over a relatively small part of the sample period (1992–96). For this part of the analysis, we estimate equation 1 using data from the first quarter of 1965 through the fourth quarter of 1991. We then use the estimated equation to forecast inflation over the 1992–96 period.

Table 1 Traditional and Modified Price-Inflation Philips Curve Models

	Traditional model		Modified model	
Variable	Estimate	p-Value	Estimate	p-Value
Constant	0.0786	0.3146	0.0532	0.4601
	(0.0782)		(0.0720)	
$GDPGAP_{t-1}$	0.0339**	0.0016	0.0190	0.0783
	(0.0107)		(0.0108)	
$\Delta\ GDPGAP_{t-1}$	0.1452**	0.0045	0.2620**	0.0000
	(0.0511)		(0.0537)	
INF_{t-1}	0.4080**	0.0007	0.2610*	0.0142
	(0.1209)		(0.1064)	
INF_{t-2}	0.1296	0.2672	0.1252	0.2312
	(0.1168)		(0.1046)	
INF_{t-3}	0.3487**	0.0045	0.2913**	0.0040
	(0.1227)		(0.1011)	
$OILG_{t-1}$	0.0186**	0.0009	0.0167**	0.0003
	(0.0056)		(0.0046)	
$OILG^{+}_{t-2}$	0.0242**	0.0007	0.0228**	0.0001
	(0.0071)		(0.0058)	
$UNITG^{+}_{t-1}$	—	—	0.1901**	0.0000
			(0.0380)	
$UNITG_{t-2}$	—	—	0.0732	0.0609
			(0.0390)	
Memo:				
Adjusted R^2	0.776	0.815		
Q-test statistic	22.731	27.572		
	(0.859)	(0.643)		

Notes: Asymptotic standard errors for the parameter estimates are reported in parentheses and are computed using the procedure of White (1980). The Ljung-Box Q-test statistic for serial correlation of the regression residuals is distributed asymptotically as χ^2 with thirty-one degrees of freedom.

Probability values for the test statistics are reported in parentheses.

* Significant at the 5 percent level.

** Significant at the 1 percent level.

Table 2 Traditional and Modified Phillips Curve Models (Chow Test Results for 1992–96)

Model	F Statistic	Likelihood ratio statistic
Taditional Phillips curve	0.192	4.539
	(0.999)	(0.999)
Modified Phillips cuve	0.244	5.860
	(0.999)	(0.998)

Note Probability values for the test statistics are reported in parentheses

The dynamic simulation provides strong evidence of instability in the traditional price inflation Phillips curve during the current expansion (Figure 6). Specifically, the out-of-sample forecasts systematically overpredict inflation beginning in the third quarter of 1993. In addition, the forecasted inflation series is characterized by a rising trend and generates prediction errors that increase over time. This exercise is robust to the choice of starting dates.[15]

The results of our dynamic simulation appear to show a shift in the Phillips curve relationship and are consistent with commentators' claims that inflation has remained unexpectedly low during this expansion. We now examine the role of compensation growth in the recent behavior of inflation.

EXAMINING THE ROLE OF COMPENSATION GROWTH

Because labor costs represent about two-thirds of the total cost of production, some economists have suggested that inflation's recent behavior may be linked to movements in compensation growth and its two components, benefits and wages (Figure 7). Since the end of the 1990–91 recession, the growth rates for total compensation, benefits, and wages have not only failed to display any significant acceleration, but have generally displayed a down-

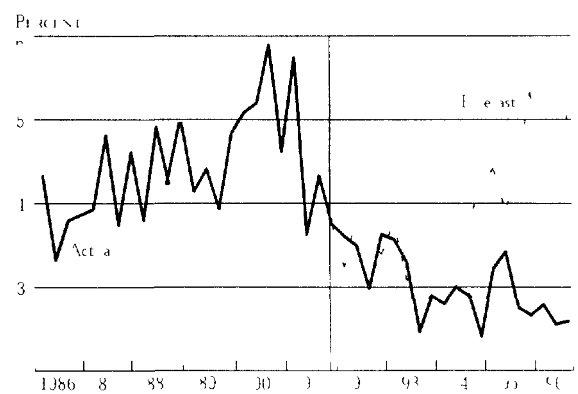


Figure 6 Out of sample forecast of core CPI inflation (traditional Phillips curve model)

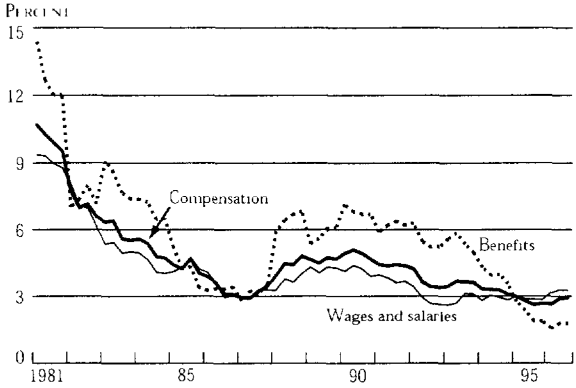


Figure 7 Employment cost index for private industry (percentage change from a year ago).

ward trend. This downward trend is particularly apparent for benefit costs, where the four-quarter change has fallen from 6 percent to about 2 percent during the 1990s. These observed patterns support the view that labor costs may be a key factor in understanding recent movements in inflation.

Meyer (1997), for example, poses two explanations relating compensation growth to inflation's puzzling behavior. First, he suggests that declining benefit costs have caused a temporary slowdown in compensation growth, which has acted as a supply shock. By lowering the increase in overall labor costs, this shock has reduced the pressure on firms to raise prices. Because most price-inflation Phillips curves exclude the effects of compensation growth altogether, their forecasting ability appears to break down and the models overpredict inflation.

Alternatively, Meyer suggests, the slowdown in compensation growth may reflect a long-term change in the behavior of the labor market. In particular, Meyer questions whether heightened job insecurity has permanently diminished workers' ability to obtain wage increases and has consequently altered the link between changes in compensation (and other macroeconomic variables) and price changes. According to this view, the recent breakdown in price-inflation Phillips curves reflects a fundamental shift in the inflation process emanating from the labor market.[16]

Although we do not look at the decline in benefit costs or the behavior of wages individually, we investigate the role of total compensation growth in restraining inflation.[17] Our methodology is designed to evaluate whether this role has been temporary or permanent in nature.

If compensation growth has acted as a temporary supply shock, we would expect the forecasting performance and the stability of the Phillips curve over the current expansion to be restored by incorporating the effects of compensation growth. Moreover, because a "shock" implies an unexpected event, we would also likely observe some evidence of unusual restraint in the recent behavior of compensation growth. However, if a change in the behavior of compensation growth has permanently altered the Phillips curve relationship, we should find evidence of a breakdown, rather than stability, in the relationship between the inflation process and compensation growth during the current expansion. We now turn to our modified Phillips curve equation.

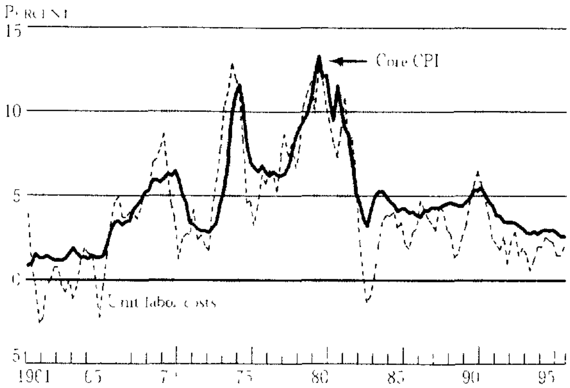


Figure 8 Core CPI and unit labor costs (percentage change from a year ago).

MODIFYING THE TRADITIONAL MODEL

Within our Phillips curve framework, we include the growth rate of unit labor costs—compensation (benefits and wages) divided by productivity—as an additional determinant of inflation. Unit labor costs provide a measure of compensation that controls for the effects of productivity.[18]

During this expansion, growth in unit labor costs has been weak and a persistent gap has been evident between unit labor cost growth and core CPI inflation (Figure 8). The decline in unit labor cost growth could suggest either falling compensation growth or rising productivity growth. As Figure 9 shows, however, productivity growth has not been unusually strong in the current expansion. Although from late 1991 to early 1992 the series rose at roughly a 3 percent rate, contributing to weaker growth in unit labor costs, since then productivity has typically grown at rates below 1 percent.

By contrast, compensation growth fell to around 2 percent fairly early in the expansion and hovered around that rate for more than two years before showing signs of a modest pickup. This 2 percent growth rate is below any rate recorded in the past thirty-five

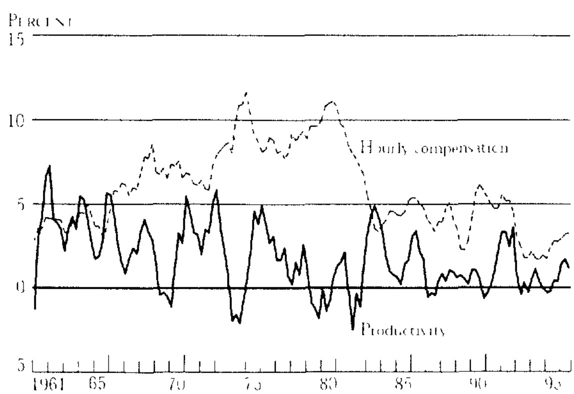


Figure 9 Productivity and hourly compensation (percentage change from a year ago).

years. Thus, we can conclude that the growth rate of unit labor costs over the post-1991 period has been primarily driven by slow compensation growth rather than high productivity growth. This finding ensures that our approach will pick up the effect of slow compensation growth, not the effect of high productivity growth, on inflation during this expansion.

Our modified price-inflation Phillips curve model is given by:

$$INF_t = \alpha_0 + \alpha_1\, GDPGAP_{t-1} + \alpha_2(\Delta GDPGAP_{t-1}) + \sum_{i=1}^{3} \alpha_{2+i} INF_{t-i} + \sum_{i=1}^{2} \alpha_{5+i} OILG^{+}_{t-1} + \sum_{i=1}^{2} \alpha_{7+i} UNITG_{t-i} + \varepsilon_t, \tag{2}$$

where *UNITG* is the growth rate of unit labor costs in the nonfarm business sector. In our modified model, unit labor costs provide an explicit channel by which slow compensation growth may have acted to offset other sources of inflationary pressures over the current expansion, resulting in lower inflation rates than those predicted using the traditional model.[19]

Model Estimation

We estimate equation 2 by the method of OLS using quarterly data from the first quarter of 1965 to the third quarter of 1996. Parameter estimates are presented in Table 1. The two lagged values of unit labor cost growth enter with the anticipated positive sign. The inclusion of the unit labor cost terms improves the fit of the model over the full sample period by almost 5 percent relative to the traditional model, and the Q-test statistic does not suggest evidence of model misspecification.

The results for all other explanatory variables are broadly similar across the traditional and modified models, although the modified Phillips curve suggests that the output gap has a smaller level effect and a larger rate-of-change effect on core CPI inflation. Like the traditional model, the estimated version of the modified model does not constrain the sum of the coefficients on lagged inflation to equal unity ($\alpha_3 + \alpha_4 + \alpha_5 = 1$). As shown in the Equation Appendix, however, we can eliminate compensation growth from the system consisting of equation 2 and our estimated wage-inflation Phillips curve to yield a reduced form of a price-inflation Phillips curve. The resulting model is characterized by coefficients on lagged inflation whose sum is not statistically different from unity, and it associates an acceleration in inflation with a positive output gap and a negative unemployment gap.

Model Stability over the 1992–96 Period

Does the inclusion of unit labor costs and the effects of compensation growth correct the instability of our bench-mark model over the post-1991 period? An examination of the dynamic simulation for the modified price-inflation Phillips curve suggests that it does (Figure 10).[20] Once we incorporate the effects of unit labor costs in the model, the simulated values track inflation closely over the post-1991 period and display no significant sign of model instability. Despite a notable error in the fourth quarter of 1995, the equation regains its predictive accuracy over the next two quarters.[21] Because the dynamic simulation uses forecasted values of inflation, however, the error in the fourth quarter of 1995 continues to affect the subsequent quarters' forecasts and contributes to the error in the third quarter of 1996.

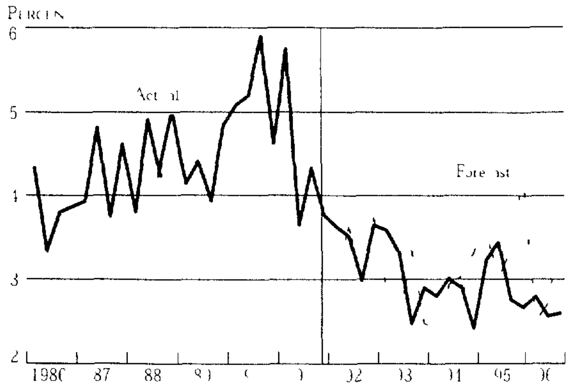


Figure 10 Out of sample forecast of core CPI inflation (modified Phillips curve model)

Overall, the evidence from the modified price-inflation Phillips curve is compelling Indeed slow compensation growth appears to be a key force in restraining inflation over the current expansion By including unit labor costs as an additional explanatory variable, the multiperiod forecast performance of the model improves dramatically, and we seem to eliminate the sharp divergence between actual and predicted inflation Thus the restored stability of the model resulting from the inclusion of unit labor costs appears to rule out the view that inflation's recent behavior reflects a fundamental shift in the Phillips curve relationship The analysis however, has yet to provide any specific insights into compensation growth and its recent behavior We explore these issues in the next section

THE BEHAVIOR OF COMPENSATION GROWTH

The results from our modified price-inflation Phillips curve reveal compensation growth's role in lowering inflation since 1991 In this section, we analyze compensation's level of restraint compared with expected levels during the present expansion The comparison allows us to determine if the recent slowdown in compensation growth has been particularly severe We show that while restraint in compensation growth appears to be easing, compensation growth was unexpectedly low from late 1992 to early 1995

To analyze the behavior of compensation growth, we specify a model that represents a modified version of the wage inflation Phillips curve proposed by Englander and Los (1983)

$$LXNG_t = \beta_0 + \sum_{i=1}^{2} \beta_i LXNG_{t-i} + \beta_3 U_{t-1} + \sum_{i=1}^{3} \beta_{3+i} INF_{t-i}$$
$$+ \beta_7 SOC_t + \beta_8 UIR_{t-1} + \beta_9 DUM_t + \eta_t, \qquad (3)$$

where

$LXNG$ = the growth rate of compensation per hour in the nonfarm business sector,
U = the unemployment rate for males aged twenty-five to fifty-four,
INF = inflation measured by the growth rate of the CPI (all items, urban consumers)
SOC = the change in employer Social Security contributions

UIR = the income replacement ratio from unemployment insurance benefits,
DUM = dummy variable for the wage and price controls of the 1970s, and
η = a mean zero, serially uncorrelated random disturbance term.

Equation 3 principally links the movements in compensation growth to the unemployment rate and other labor market variables.[22] The unemployment rate of prime-age males is used as a measure of labor market tightness. We enter the variable in its level form and thereby abstract from any explicit discussion of the NAIRU, except to note that the specification can be viewed as implicitly assuming a constant value for the NAIRU over the sample period.[23] Equation 3 does not include a rate-of-change effect for the unemployment rate; the estimated coefficient on a second lag of the unemployment rate was found to be quantitatively and statistically insignificant and therefore was omitted from the specification.[24]

The remaining determinants of compensation growth include the change in employer Social Security tax contributions, a component of hourly compensation. The income replacement ratio from unemployment insurance benefits attempts to capture changes in compensation growth related to job search. A dummy variable accounts for the restraining effect of wage and price controls in the fourth quarter of 1971 and for the rebound effect after the relaxation of the controls in the first quarter of 1972.[25] We include lagged values of compensation growth and price inflation to incorporate wage and price inertia effects. Finally, we include only lagged values of the unemployment rate and inflation rate as regressors because of endogeneity considerations.

Model Estimation and Model Stability over the 1992–96 Period

We estimate equation 3 using the method of OLS for quarterly data from the second quarter of 1967 to the third quarter of 1996. The parameter estimates are presented in Table 3. As the table indicates, the lagged values of both compensation growth and price inflation are generally significant. The unemployment rate is highly significant and has the expected negative sign. Further, the variables reflecting other labor market conditions are all significant with the expected signs. The adjusted R^2, although not quite as high as the values reported in Table 1, also indicates that the estimated equation fits the data quite well over the full sample period. In addition, the regression residuals display little evidence of serial correlation over the full sample period.

We also conduct Chow tests and a dynamic simulation. The Chow tests do not reject the null hypothesis of parameter stability at conventional significance levels (Table 4). For the dynamic simulation, we estimate equation 3 from the second quarter of 1967 to the fourth quarter of 1991; we then use the estimated equation to generate predicted values for compensation growth over the 1992–96 period.

The evidence from the dynamic simulation indicates that compensation growth has displayed unexpected restraint during this expansion. The out-of-sample forecasts consistently overpredict compensation growth beginning in the fourth quarter of 1992 (Figure 11). In addition, the size of the errors at times is quite large. For example, our dynamic simulation predicts that compensation growth should have been about 2 percent higher from the end of 1992 through the end of 1994. After 1994, however, the size of the forecast errors begins to diminish, a pattern that supports the temporary supply shock hypothesis. If a permanent change in compensation growth had occurred, we would expect the large disparity between the model's simulated values and actual growth to continue, as it did in the traditional price-inflation Phillips curve model.

Table 3 Wage-Inflation Phillips Curve Model for Compensation Growth

Variable	Estimate	p-Value
Constant	0.3884	0.0715
	(0.2155)	
$LXNG_{t-1}$	0.1359	0.1144
	(0.0861)	
$LXNG_{t-2}$	0.2621**	0.0001
	(0.0689)	
U_{t-1}	−0.0672**	0.0021
	(0.0218)	
INF_{t-1}	0.2018**	0.0036
	(0.0692)	
INF_{t-2}	0.0175	0.8332
	(0.0832)	
INF_{t-3}	0.1257	0.0720
	(0.0698)	
SOC_t	0.0849**	0.0000
	(0.0186)	
UIR_{t-1}	1.4288*	0.0321
	(0.6666)	
DUM	−0.7442**	0.0000
	(0.0790)	
Memo		
Adjusted R^2	0.709	
Q-test statistic	28.109	
	(0.838)	

Notes: Asymptotic standard errors for the parameter estimates are computed using the procedure of White (1980) and are reported in parentheses. The Ljung-Box Q test statistic for serial correlation of the regression residuals is distributed asymptotically as χ^2 with twenty-nine degrees of freedom. Probability values for the test statistics are reported in parentheses.
* Significant at the 5 percent level.
** Significant at the 1 percent level.

Table 4 Compensation Growth Model (Chow Test Results for 1992–96)

Model	F Statistic	Likelihood ratio statistic
Compensation growth Phillips curve	0.879	20.287
	(0.609)	(0.377)

Note: Probability values for the test statistics are reported in parentheses.

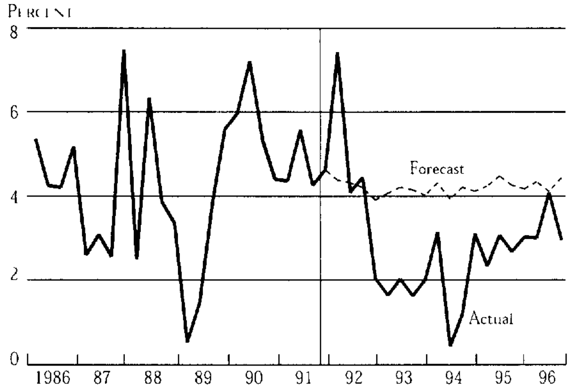


Figure 11 Out-of-sample forecast of compensation growth.

Evidence from the dynamic simulation corroborates our earlier finding that the modified price-inflation Phillips curve model, which incorporates the effects of compensation growth, appears to resolve the inflation puzzle. The slowdown in compensation growth is most pronounced from the end of 1992 to early 1995, the same period during which the traditional Phillips curve starts to display evidence of model instability. Thus, not surprisingly, variables and relationships that ignore compensation growth's influence (such as the inflation indicators in Charts 2 and 3 and the traditional Phillips curve) begin to break down in late 1993 and 1994.

CONCLUSION

Contrary to its behavior in previous expansions, price inflation has not accelerated in the six years since the 1990–91 recession. This article focuses on compensation's role in the inflation puzzle, investigating whether a temporary slowdown in compensation growth has lowered the level of inflation or if a more permanent change in compensation growth has fundamentally altered the inflation process. We present two pieces of evidence suggesting that slow compensation growth has acted as a temporary restraining force on inflation.

We begin our investigation by estimating a traditional price-inflation Phillips curve model over the 1965–96 period. Although the model tracks inflation quite well over most of the period, it begins to break down in late 1993. We then modify the traditional Phillips curve model to include the effects of compensation growth. With this addition, the model tracks inflation much more accurately over the current expansion and displays no significant evidence of instability. This finding provides the first piece of evidence suggesting that no fundamental change in the inflation process has occurred.

To arrive at the second piece of evidence supporting the notion that the low level of inflation has resulted from a temporary slowdown in compensation growth, we look at compensation growth itself. By estimating a wage-inflation Phillips curve model, we find that compensation growth showed unusual restraint from late 1992 to early 1995. This period of restraint appears to be temporary and coincides with the observed breakdown in the traditional Phillips curve model and in other inflation indicators. Thus, taking compensation growth into account appears to explain inflation's behavior during the current expansion Still uncertain, however, is the reason for the dramatic slowdown in compensation growth during the early 1990s The solution to this puzzle must await further investigation

EQUATION APPENDIX: DERIVATION OF THE ACCELERATIONIST PHILLIPS CURVE MODEL

This appendix briefly examines the derivation of the accelerationist model of the Phillips curve from equations 2 and 3 The key features of this model can be illustrated by examining the relationship between the output gap (and the unemployment gap with a constant NAIRU) and the inflation rate Abstracting from the influence of other terms, we note that the system of equations 2 and 3 can be rewritten as

$$INF_t = \alpha_1 GDPGAP_{t-1} + \sum_{i=1}^{3} \alpha_{2+i} INF_{t-i} + \sum_{i=1}^{2} \alpha_{5+i}(LXNG_{t-i}), \tag{4}$$

and

$$LXNG_t = \frac{\beta_3 U_{t-1} + \sum_{i=1}^{3} \beta_{3+i} INF_{t-i}}{(1 - \beta_1 L - \beta_2 L^2)}, \tag{5}$$

where we substitute for the definition of the growth rate of unit labor costs (compensation growth less productivity growth) in equation 4, and L denotes the lag operator in equation 5 such that $L^k X = X_{t-k}$

We can substitute equation 5 into equation 4 to obtain an expression relating current inflation to the output gap the unemployment gap, and past rates of inflation If the sum of the coefficients on lagged inflation equals unity, then there is a natural rate value of the output gap (and unemployment gap) of zero that is consistent with a constant rate of inflation Alternatively the model would associate a permanent positive value for the output gap with an ever-accelerating inflation rate Within our system of equations the condition that the sum of the coefficients on lagged inflation equals unity is given by

$$\alpha_3 + \alpha_4 + \alpha_5 + \left[\frac{(\alpha_6 + \alpha_7)(\beta_4 + \beta_5 + \beta_6)}{(1 - \beta_1 - \beta_2)}\right] = 1 \tag{6}$$

The hypothesis that the coefficients on lagged inflation sum to unity can be tested using the OLS estimates of equations 2 and 3 to construct estimates for the expression on the left-hand side of equation 6 and its standard error The standard error is the standard error of a function of several estimated parameters and can be computed using the delta method approximation (Greene 1993, p 297)

$$SE[g(\theta)] = \sqrt{\frac{\partial g}{\partial \theta'} \; VAR(\theta) \; \frac{\partial g}{\partial \theta}},$$

where θ denotes the parameters in equation 6, $g(\theta)$ is the function of the parameters in 6, and $VAR(\theta)$ is the variance-covariance matrix of those parameters

Because of the slight disparity in the sample periods for Tables 1 and 2 we estimate equation 2 and equation 3 from the second quarter of 1967 to the third quarter of 1996 The estimate for the expression on the left-hand side of equation 6 is 0 87, with an estimated standard error of 0 08 Thus we are unable to reject the null hypothesis that the sum of the coefficients in equation 6 is equal to unity at the 5 percent significance level

DATA APPENDIX

This appendix defines the variables and the data sources used to estimate our traditional Phillips curve model, modified Phillips curve model, and compensation growth model. All data in our analysis include revisions through August 12, 1997.

Inflation Equation Variables

INF = the growth rate of the core CPI for all urban consumers as reported by the Department of Labor, Bureau of Labor Statistics. Data are released monthly and are seasonally adjusted.

UNITG = the growth in unit labor costs for the nonfarm business sector as reported by the Department of Labor, Bureau of Labor Statistics. Data are released quarterly and are seasonally adjusted.

GDPGAP = the logarithmic ratio of *GDP* to *POTGDP*, where *GDP* equals quarterly real gross domestic product and *POTGDP*, quarterly potential GDP. Both variables are in 1987 dollars until the third quarter of 1987. They are in chain-weighted 1992 dollars from the fourth quarter of 1987 to the present. The GDP data are from the National Income and Product Accounts. Potential GDP is a Federal Reserve Bank of New York staff estimate.

$OILG^{+}$ = the net positive change in the real price of oil, calculated as the percentage change in the current real price of oil from the previous year's maximum (if that change is positive, zero otherwise). Data for the price of oil are an extension of Mork's (1989) series, which reflects corrections for the effects of price controls during the 1970s. The real price of oil is defined as the nominal oil price index deflated by the GDP deflator.

Compensation Equation Variables

LXNG = the growth rate of compensation per hour for the nonfarm business sector as reported by the Department of Labor, Bureau of Labor Statistics. Compensation comprises wages and salaries for workers plus employers' contributions for Social Security insurance and private benefit plans. The series also includes an estimate of wages, salaries, and supplemental payments for self-employed workers. Data are released quarterly and are seasonally adjusted.

INF = the growth rate of the CPI for all urban consumers as reported by the Department of Labor, Bureau of Labor Statistics. Data are released monthly and are seasonally adjusted.

U = the unemployment rate for males aged twenty-five to fifty-four as reported by the Department of Labor, Bureau of Labor Statistics. Data are released monthly and are seasonally adjusted.

UIR = unemployment insurance per job loser, normalized by the average annual earnings of a manufacturing worker. This variable can be thought of as a replacement ratio, that is, the fraction of earnings of manufacturing workers replaced by unemployment insurance payments. Manufacturing workers are the most likely workers to collect unemployment insurance. *UIR* is constructed as (*YPTU*/*LUJL*)/(*YPWF*/*LAMANU*), where

YPTU = government unemployment insurance benefits according to the National Income and Product Accounts. Data are reported quarterly and are seasonally adjusted.

LUJL = job losers and persons who have completed temporary jobs as reported by the Department of Labor. Bureau of Labor Statistics. Data are released monthly and are seasonally adjusted.

YPWF = wage and salary disbursements in manufacturing according to the National Income and Product Accounts. Data are reported quarterly and are seasonally adjusted.

LAMANU = nonfarm payroll employees in manufacturing as reported by the Department of Labor, Bureau of Labor Statistics. Data are reported monthly.

SOC = a measure of the direct effect of changes in payroll tax rates for Social Security and Medicare. The quarterly data are Federal Reserve Bank of New York staff estimates.

DUM = 1 in the fourth quarter of 1971, −0.6 in the first quarter of 1972, and 0 elsewhere. This variable accounts for the restraining effect of the wage and price freeze in the fourth quarter of 1971 and the rebound effect after the wage and price controls were relaxed in the first quarter of 1972.

ACKNOWLEDGMENTS

The authors are grateful to J.S. Butler, Gabriele Galati, Steve Kamin, Jonathan McCarthy, Richard Peach, Charles Steindel, and two anonymous referees for helpful comments. We also benefited from the suggestions of conference participants at the Bank for International Settlements. Beethika Khan provided excellent research assistance.

NOTES

1. Dow Jones News Service, January 7, 1997.
2. Our analysis expands on results that we presented in two earlier papers. See Lown and Rich (1997a, 1997b).
3. Gordon (1996), however, obtains an estimate of 5.3 percent for the NAIRU starting in 1996.
4. Gordon's work (1970, 1975, 1977, 1982, 1990) is prominent in the literature on the estimation of the Phillips curve.
5. See King and Watson (1994), Tootell (1994), Fuhrer (1995), King, Stock, and Watson (1995), and Gordon (1996).
6. The estimation of "expectations-augmented" Phillips curves is the result of work by Phelps (1967) and Friedman (1968), who developed the natural rate hypothesis and drew the distinction between the short-run and long-run Phillips curve trade-off.
7. For detailed definitions and sources of data, see the Data Appendix.
8. The results are little affected when the unemployment rate instead of the output gap is used to measure aggregate demand pressure. Potential GDP measures the full-employment level of output or the output level at which there is no tendency for inflation to accelerate or decelerate. The level of potential GDP grows over time because of the increased availability of resources (land, labor force, capital stock, and the level of technology). Because potential GDP is not directly observable, several techniques have been developed to calculate estimates of the series. A complete review of these techniques and an evaluation of the alternative potential GDP series are

beyond the scope of this paper As noted in the Data Appendix, we employ a staff estimate of potential GDP to construct the output gap variable

9 Commodity prices and/or an exchange rate term have been used as supply shock variables in some price-inflation Phillips curve models We do not include these terms in our specification, however, because we found their effects to be small and statistically insignificant The absence of a strong link between commodity prices and inflation is consistent with evidence presented by Blomberg and Harris (1995), who document a recent decline in the predictive power of commodity prices for inflation

10 We exclude the net negative real oil price change variable from equation 1 because the variable displays quantitatively and statistically insignificant effects

11 The compensation growth Phillips curve described later in the text includes dummy variables to capture the effects from the imposition and relaxation of wage and price controls during the 1970s We exclude these dummy variables from the traditional price-inflation Phillips curve because they were found to be statistically insignificant Alternative dating schemes for the dummy variables (Gordon 1982) also proved to be unimportant in explaining the dynamics of inflation during the 1971–75 period

12 This test yields an F-statistic, which is distributed asymptotically as F with $(m, n-k)$ degrees of freedom under the null hypothesis The values of n and $n+m$ refer to the number of observations in the first subperiod and the total sample, respectively The value of k refers to the number of parameters in the model

13 This test yields a likelihood ratio statistic, which is distributed asymptotically as chi-square with k degrees of freedom under the null hypothesis

14 We also looked for evidence of parameter instability using the CUSUM and CUSUMSQ tests proposed by Brown, Durbin, and Evans (1975) The tests are based on recursive residuals, with the CUSUM test primarily used to detect gradual structural change and the CUSUMSQ test used to detect sudden structural change The tests provided no evidence of parameter instability

15 The dynamic simulation yielded similar results for the 1994–96 period

16 Meyer (1997) notes that the declines in computer prices and import prices over the current expansion may also be acting as temporary supply shocks helping to restrain inflationary pressures in the economy Moreover, as an additional explanation for the inflation puzzle, he cites firms' inability to raise prices because of increased international competitive pressures We do not address these factors in this paper and instead restrict our attention to the two explanations that concern labor market phenomena Further, while our analysis is not exhaustive, we nevertheless believe that it is instructive to evaluate these explanations before considering alternative hypotheses

17 Our focus on compensation growth is also motivated by the idea that the pricing decision of a firm should be based on a consideration of its total labor costs rather than the behavior of the wage and benefit components of these costs In addition, the data preclude us from obtaining observations on wages and benefits separately over the full sample period The employment cost index, which provides measures of wages and benefits, is only available beginning in 1980 for the nonfarm sector

18 We modify the traditional price-inflation Phillips curve to include unit labor costs rather than compensation per hour because it is the behavior of compensation growth *relative* to productivity growth that is relevant for describing the dynamics of the inflation process That is, greater productivity growth will act to offset the inflationary pressure on prices arising from an increase in compensation growth

19 Note that our model does not allow us to examine whether a shift in the Federal Reserve's inflation fighting credibility has changed the inflation process by directly altering inflation expectations Such an examination is beyond the scope of this paper and would involve estimating a separate equation for inflation expectations and including some measure of Federal Reserve credibility as an explanatory variable Previous evidence, however, suggests that such

a shift has not taken place Blanchard (1984) notes that similar types of Phillips curves remained stable even after the 1979 change in Federal Reserve operating procedures

20 As the value of the test statistics in Table 2 indicates the Chow tests fail to reject the null hypothesis of parameter stability at conventional significance levels However this result is not particularly informative because the Chow tests also failed to reject the null hypothesis of model stability for the traditional Phillips curve

21 The increase in the forecasted value for inflation primarily reflects the influence of a change in the output gap and the oil price variable

22 For definitions of the data and their sources see the Data Appendix

23 For example we could follow the approach of Fuhrer (1995) who assumes a value of 6 percent for the NAIRU and use the unemployment gap (the difference between the actual level of un employment and the NAIRU) instead of the unemployment rate as an explanatory variable in equation 3 This approach however would not affect the regression results other than to change the estimated value of the constant term

24 Fuhrer (1995) also finds an absence of significant rate of change effects for the unemployment rate in wage inflation Phillips curve models

25 The definition of the dummy variable is from Englander and Los (1983)

REFERENCES

Blanchard O J 1984 The Lucas Critique and the Volcker Deflation American Economic Review 74 no 2 211–5 Papers and Proceedings of the 96th Annual Meeting of the American Economic Association December 1983

Blomberg S B and Harris E S 1995 The Commodity Consumer Price Connection Fact or Fable? Federal Reserve Bank of New York Economic Policy Review 2 no 2 21–38

Boldin M D 1996 Problems in Determining the Capacity Inflation Link Unpublished manuscript the Conference Board

Brown R L Durbin J and Evans J M 1975 Techniques for Testing the Constancy of Regression Relationships over Time Journal of the Royal Statistical Society 37 149–92

Chow G C 1960 Tests of Equality between Sets of Coefficients in Two Linear Regressions Econometrica 28 591–605

Eggert R ed 1991–95 Blue Chip Economic Indicators Capitol Publications Inc

Englander S A and Los C A 1983 The Stability of the Phillips Curve and Its Implications for the 1980s Federal Reserve Bank of New York Research Paper no 8303 February

Friedman M 1968 The Role of Monetary Policy American Economic Review 58 1–17

Fuhrer J 1995 The Phillips Curve Is Alive and Well Federal Reserve Bank of Boston New England Economic Review March April 41–56

Goodfriend M 1993 Interest Rate Policy and the Inflation Scare Problem 1979–1992 Federal Reserve Bank of Richmond Quarterly Review 79 no 1 1–24

Gordon R J 1970 The Recent Acceleration of Inflation and Its Lessons for the Future Brookings Papers on Economic Activity no 1 8–41

Gordon R J 1975 The Impact of Aggregate Demand on Prices Brookings Papers on Economic Activity no 3 613–62

Gordon R J 1977 Can the Inflation of the 1970s Be Explained? Brookings Papers on Economic Activity no 1 253–77

Gordon R J 1982 Inflation Flexible Exchange Rates and the Natural Rate of Unemployment In Baily M N ed Workers Jobs and Inflation pp 88–152 Washington D C Brookings Institution

Gordon R J 1990 U S Inflation Labor s Share and the Natural Rate of Unemployment In Konig Heinz ed Economics of Wage Determination pp 1–34 Berlin and New York Springer Verlag

Gordon, R. J. 1996. The Time-Varying NAIRU and Its Implications for Economic Policy. National Bureau of Economic Research Working Paper no. 5735, August.

Greene, W. H. 1993. Econometric Analysis. New York: Macmillan Publishing Company.

Hamilton, J. D. 1996. This Is What Happened to the Oil Price–Macroeconomy Relationship. Journal of Monetary Economics 38, no. 2: 215–20.

King, R. G., and Watson, M. W. 1994. The Post-War U.S. Phillips Curve: A Revisionist Econometric History. Carnegie-Rochester Conference Series on Public Policy 41 (December): 157–219.

King, R. G., Stock, J. H. and Watson, M. W. 1995. Temporal Instability of the Unemployment-Inflation Relationship. Federal Reserve Bank of Chicago Economic Perspectives 19, no. 3 (May-June): 2–12.

Ljung, G. M., and Box, G. E. P. 1978. On a Measure of Lack of Fit in Time Series Models. Biometrika 67: 297–303.

Lown, C. S., and Rich, R. W. 1997a. A Look at the U.S. Inflation Puzzle. In Monetary Policy and the Inflation Process. Bank for International Settlements Conference Papers 4 (July).

Lown, C. S. and Rich, R. W. 1997b. Is There an Inflation Puzzle? Federal Reserve Bank of New York Research Paper no. 9723.

Meyer, L. H. 1997. The Economic Outlook and Challenges for Monetary Policy. Remarks delivered to the Charlotte Economics Club on January 16, Charlotte, N.C.

Mork, K. A. 1989. Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results. Journal of Political Economy 97, no. 3: 740–4.

Phelps, E. S. 1967. Phillips Curves, Expectations of Inflation, and Optimal Unemployment over Time. Economica 34: 254–81.

Phillips, A. W. 1958. The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. Economica 25: 283–300.

Tootell, G. M. B. 1994. Restructuring, the NAIRU, and the Phillips Curve. Federal Reserve Bank of Boston New England Economic Review, September–October: 31–44.

White, H. 1980. A Heteroskedasticity-consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity. Econometrica 48, no. 4: 817–38.

This Page Intentionally Left Blank

53

Inflation and Growth

Brian Motley

Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

In recent years, the Federal Reserve has stressed its long-run commitment to maintaining low inflation because it believes that persistent inflation imposes burdens that reduce economic welfare. A variety of such burdens have been identified in the economics literature. This chapter examines the argument that a lower rate of inflation increases the potential growth rate of the economy, and discusses the possible size of this effect. Since a policy to reduce inflation generally has some short-run costs because it requires a temporary slowing of economic activity, it is useful to estimate its benefits in the form of higher long-run growth.

HOW INFLATION MIGHT LOWER GROWTH

Inflation might affect potential output in a number of ways. First, inflation may interfere with the efficiency of the price system and make it more difficult for households and firms to make correct decisions in response to market signals. It is often argued that when most prices are rising, economic agents find it harder to distinguish between changes in relative prices that require them to reallocate resources and changes in the overall price level that require no such microeconomic response. A widget-producing firm that observes that its customers are bidding higher prices for its widgets may interpret this as indicating a rise in the demand for its products, when it actually represents the effects of generalized inflation in which the prices of competing products (wadgets, wodgets, etc.) also are rising at the same pace.

Second, inflation imposes various costs on the economy that would disappear if prices were stable. The costs of changing prices and wage rates frequently, the search costs imposed on buyers and sellers when prices change often, and the costs of economizing on holdings of non-interest-bearing money ("shoe-leather" costs) are familiar examples.

Inflation also has differing effects on individuals. For example, the incomes of wage and salaried workers generally are adjusted for inflation only annually, whereas self-employed workers can alter the prices of their services more frequently. Similarly, inflation, especially when it is unexpected, tends to benefit borrowers at the expense of lenders. Finally, because some parts of the tax code are indexed for inflation whereas others are not,

Reprinted from: *FRBSF Weekly Letter,* No. 93-44 (Dec. 31, 1993).

generalized increases in prices have differing effects on individual tax-payers As a result of these considerations inflation often is perceived as causing unfairness, since some households and firms benefit and others are harmed

However whether or not these differential effects of inflation are unfair they do impose real costs on society at large They frequently add to the uncertainties that households and firms face, which may be undesirable even for those that turn out to benefit And many activities that seek to reduce the impact of inflation on individuals may hurt the overall economy but yield no corresponding overall benefits In an inflationary economy, for example talented persons may devote their energies to developing strategies to avoid the deleterious consequences of inflation for themselves rather than to inventing new products and processes that would raise overall living standards Unfortunately many of these activities that aim to mitigate the effects of inflation are counted as additions to measured GDP, even though they may not add to welfare in any meaningful sense

Finally, inflation may affect investors' saving and investment decisions, reducing the proportion of GDP devoted to investment and causing the economy to accumulate less productive capital For example when inflation is high, it usually tends to be more variable and so harder to forecast Uncertain inflation makes it more difficult to deduce the real returns on investments from available market information As a result, savers and investors often are less willing to enter into long-term nominal contracts or to invest in long-term projects The reduced stock of productive capital that results from decreased investment will, in turn, imply lower levels of future GDP

These considerations suggest that there are sound *a priori* reasons why persistent inflation might reduce the growth rate of GDP in the long run A number of studies have investigated the sources of long-term growth using data from a cross-section of countries, and several have examined whether differences in growth among countries are related to differences in average rates of inflation The results have been mixed, perhaps because it is difficult to isolate the impact of inflation on long-run growth

WHY THE EFFECTS OF INFLATION ARE HARD TO DISCERN

The most obvious difficulty is that inflation is only one of many factors that may affect a country's long-run growth For example, a country that saves a large share of its output and devotes it to investment in productive capital or to educating its workers is likely to enjoy a higher and more rapidly growing GDP than one that devotes most of its output to current consumption Conversely a country with a rapidly increasing population is likely to have lower GDP per worker because more of its saving is needed just to provide the *existing* levels of education and capital to the new entrants to its work-force and so less is available to increase the stocks of human and tangible capital *per worker* Finally many empirical studies find that countries with initially low levels of output tend to grow more rapidly than advanced countries, perhaps because they find it easier to adopt technologies that already are in use in more advanced economies or because their output levels are below the long-run equilibrium associated with their rates of saving and population growth

To isolate the impact of inflation on growth, we also must estimate the effects of these (and other) systematic influences, in order to judge what if any additional explanatory power inflation provides Unfortunately, there is no fully accepted model of economic growth that can serve as a baseline for estimating the additional effects of inflation Although most economists agree that factors such as the saving rate and the population growth rate are important determinants of overall growth, they are far from unanimous on the precise mechanisms through which their effects are felt.

In addition to variables that received economic theory suggests should be related to long-term growth, there is a vast array of factors that might plausibly have an influence. Political stability, the size of the financial sector, openness to world trade, and the size of the government sector all have been suggested as potential influences on growth. Each of these factors may be measured in a variety of ways. Econometricians may be tempted to search the data to find the set of variables that were most closely related to growth in the past. The problem is that statistical relationships that are uncovered by such "data-mining" but that have no genuine causal basis are unlikely to hold up in the future. Moreover, some of these "plausible" variables may be statistically related to inflation so that their inclusion may obscure the link between inflation and growth.

Another serious problem is that even if inflation does reduce long-run economic growth, this effect may be difficult to detect in the data because there also may be short-run links that obscure the long-run relation. In the short run, real growth may be *positively* related to inflation because fluctuations in GDP reflect variations in aggregate demand. In a cyclical upswing, increasing demand may raise real output while it also bids up prices by putting greater pressure on resources. As a result, we may observe a positive relation between real GDP growth and inflation in the short run, even though there is a negative relation over the long haul.

Any long-run relation between inflation and growth also may be obscured if the world is affected by supply shocks that influence prices and/or real output by differing amounts in different countries. Indeed, the observed negative across-country correlation between long-run inflation and growth might be due to a small number of major supply shocks that affected the levels of prices and of output in a significant group of countries. For example, if a worldwide increase in the price of oil were to permanently raise the *level* of prices and lower the *level* of output by more in some countries than in others, a cross-country comparison over a longer period that included the oil shock years might detect a negative correlation between the average *growth rates* of prices and output, even if there were no causal relation between these two growth rates.

DEALING WITH THE STATISTICAL PROBLEMS

Recent research at this Bank (Motley 1993) has attempted to deal with these problems in a variety of ways. This research used a model of economic growth that economists previously have found helpful in understanding differences in experience among countries (Solow 1956, Mankiw, Romer, and Weil 1992). This model emphasizes the key role of the rates of saving and population growth as the fundamental determinants of long-run growth, since these factors ultimately determine the supplies of capital and labor.

This model was expanded to include the effects of differences in rates of inflation. Inflation was assumed to influence long-run growth by affecting the pace of "technological change," a port-manteau term that includes the effects of all variables apart from the supplies of the factors of production. This choice of theoretical framework implies that although technological change is affected by inflation, it is otherwise exogenous and so is independent of the saving and population growth rates.

No other variables apart from those suggested by this theoretical model were included in the empirical analysis. This limitation avoids the temptations of data mining, but it implies the risky assumption that economic growth is driven by the same fundamental forces in countries with widely diverse environments.

The statistical problem of separating the longer-run effects of inflation on potential GDP from any short-run business cycle relation between prices and output was dealt with by examining growth and inflation over a long period that spanned several business cycles. This makes it more likely that the results represent a long-run relationship. At the same time, the possibility that any long-run correlation between inflation and growth actually represents the influence of a few supply shocks was examined by estimating cross-section equations over a series of shorter time periods in addition to the cross section covering a longer span. Finding a negative cross section relation between inflation and real growth in several time periods of varying length might constitute evidence that it represents a true causal relation rather than the effect of supply shocks that affected countries differently. However, this procedure raises the possibility that the equations estimated over the shorter time periods may be contaminated by the business cycle problem mentioned earlier.

FINDINGS

The results of this research suggest that countries with lower inflation rates do tend to exhibit higher rates of long-run growth. The cross country comparisons indicate that a 5 percentage point decrease in the average rate of inflation is associated with an increase in annual growth per capita of about 0.2 percentage point. Since inflation in the U.S. between 1982 and 1992 was about 5 percentage points lower than during the previous ten years, this result suggests that the lowering of inflation will add about 0.2 percentage point to long-run growth. This is sufficient to increase the discounted lifetime income of a typical worker by an amount equal to about one year's income, assuming a 40 year working life and a 3 percent real discount rate. This estimated benefit exceeds the costs of lowering inflation which typically are estimated as amounting to at most one fourth of one year's GDP for a 5 percentage point reduction in inflation (Ball 1993, Mankiw 1992, p. 309). It must be recognized, however, that although the benefits exceed the costs in present value terms, much of the benefit will accrue in the relatively distant future, whereas the costs must be borne today.

As I have emphasized, there are a number of statistical problems that always will make it difficult to demonstrate conclusively that lower inflation will lead to faster growth. However, the evidence in this study, while not conclusive, not only is consistent with such a relationship but also suggests that the quantitative magnitude of this effect is sufficiently large that policymakers should lean toward low inflation.

REFERENCES

Ball, L. 1993. How Costly Is Disinflation? The Historical Evidence. Federal Reserve Bank of Philadelphia *Business Review* (November/December).

Mankiw, N. G. 1992. *Macroeconomics*. New York: Worth Publishers.

Mankiw, N. G., Romer, D., and Weil, D. N. 1992. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* (May).

Motley, B. 1993. Growth and Inflation: A Cross-Country Study. Working Paper 93-11. Federal Reserve Bank of San Francisco.

Solow, R. M. 1956. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* (February).

54

The New Output-Inflation Trade-Off

Carl E. Walsh

Professor of Economics, University of California, Santa Cruz, California

One of the hallmarks of economic analysis is the recognition that choice involves trade-offs. Whether it's a consumer deciding if the roominess of a sports utility vehicle is worth the lower gas mileage, or a firm deciding whether lower wages of an overseas production facility compensate for the lower worker productivity, or Congress deciding whether a new expenditure program justifies the higher taxes needed to finance it, trade-offs must be faced. The same is true in the conduct of monetary policy. Acting too slowly to head off inflation may risk an increase in expected inflation that will make subsequent moves to reduce inflation more costly, while acting too quickly may run the risk of slowing economic growth prematurely.

During the 1960s and early 1970s, many economists and policymakers believed a central bank could achieve permanently lower unemployment by accepting permanently higher inflation. Attempts to exploit such a trade-off to gain the benefits of lower unemployment were, unfortunately, self-defeating. As unemployment fell and inflation rose, individuals began to expect that inflation would be higher. Workers demanded more rapidly rising money wages to compensate for expected price increases, and firms were willing to agree to these wage demands as they expected to be able to pass through their increased costs by raising prices. Rather than remaining stable at a new higher level, the inflation rate continued to increase as long as unemployment remained below the economy's natural rate. That experience has convinced most policymakers that no such trade-off exists. Instead, most agree that the average level of the unemployment rate and the long-run rate of real economic growth are determined by such fundamentals as technological change, population growth, labor market institutions, and the skills of the work force. These factors are unrelated to the economy's average rate of inflation, so allowing average inflation to rise brings no long-run benefit in the form of faster growth or lower average unemployment.

This does not mean, however, that central banks do not face unemployment-inflation trade-offs as they implement monetary policy. In fact, recent research in macroeconomics has focused increasingly on an important trade-off involving output and inflation. Unlike the short-run trade-off between the level of output or unemployment and the level of the inflation rate that was a focus of earlier policy debates, the new emphasis is on the choice between the variability of output and the variability of inflation. The research on this vari-

Reprinted from: *FRBSF Economic Letter*, No. 99-04 (Feb. 6, 1998).

ability trade-off suggests that attempting to keep inflation within a very narrow band may increase fluctuations in real output and employment. Conversely, attempts to smooth business cycle fluctuations more actively will lead to wider fluctuations in inflation. The nature of this trade-off, and even whether it really exists, is a subject of debate among economists. This chapter discusses the new output-inflation trade-off and its implications for the design of monetary policy.

TWO PROPOSITIONS

Today, most economists and central bankers accept the proposition that there is no long-run trade-off between the rate of inflation and the level of unemployment. At the same time, many believe that policies designed to help stabilize inflation do have real consequences. John Taylor (1996) has summarized current thinking about these issues in the form of two propositions:

> The *first* proposition, about which there is now little disagreement, is that there is no *long-run* trade-off between the rate of inflation and the rate of unemployment (p. 186).
>
> The *second* proposition, and there is more disagreement here, is that there is a *short-run* trade-off between inflation and unemployment. I think that the short-run trade-off is best described in terms of a trade-off between the *variability* of inflation and the *variability* of unemployment; that is, in terms of the short-run fluctuations in the variables rather than their levels over time (p. 186).

It is this trade-off between the variability of output and inflation that represents the new policy trade-off. It is easiest to understand why such a trade-off might arise by considering the economic impact of an adverse aggregate shock such as a rise in the price of oil or private sector expectations about future inflation. The direct result of either would be an increase in inflation. If policy acts to bring inflation back on target quickly, inflation will be less variable, but output will fluctuate more around trend. If policy acts more slowly to bring inflation back on target, then output will fluctuate less while inflation becomes more variable. Acting to offset the inflationary impact of supply shocks leads output and unemployment to fluctuate more in the short run, while stabilizing output leads actual inflation to fluctuate more.

SIMULATION EVIDENCE

Much of our knowledge of variability trade-offs comes from simulations of models designed to mimic the behavior of the major industrialized economies. These models incorporate realistic inflation and output adjustment so that they can be used to study the variability trade-off implied by different rules for conducting monetary policy. Fuhrer (1997) provides an example of this type of research, employing a model of the U.S. economy. The evidence from simulations can be used to determine the nature of the volatility trade-off that arises under a particular policy rule and to evaluate alternative policy rules. For example, Taylor (1993) has suggested that recent Fed behavior is characterized by a rule that describes how the federal funds rate is adjusted in response to movements in inflation and the output gap. Using such a rule for determining the funds rate, together with particular values for how much the funds rate is adjusted as inflation and the output gap change, the implied variability of inflation and output can be determined. By then changing how much the funds rate is adjusted in response to inflation and the output gap, a different combination of inflation variability and output variability will be implied. Linking together the different combinations of inflation and output variability, a trade-off emerges. In a similar manner, the frontier associated with a different rule for adjusting the funds rate, such as one that responds to nominal income movements, can be derived. In addition, a change in the nature of the underlying economic disturbances would shift the trade-off frontier; an increase in the volatility of energy prices, for example, would lead to more inflation *and* output variability.

Figure 1 illustrates the output-inflation variability trade-off for two hypothetical policy rules. The rule that produces the dashed trade-off frontier can be described as inefficient; for any given output volatility, the policy rule that produces the solid line results in lower inflation volatility. Once the efficient trade-off frontier has been found, policymakers then must weigh the relative costs of output variability versus inflation variability in choosing a point on the frontier. If inflation variability is viewed as more costly than output variability, a point such as A might be optimal, while point B would be optimal if the costs of output variability are assessed more highly. This two-step approach, finding the efficient frontier and then deciding which point to pick, is useful in separating two distinct aspects of policy choice. On the one hand, the structure of the economy and the nature of economic disturbances that affect it will define the efficient frontier. On the other hand, the factors that determine which point on the frontier to choose depend on an assessment of the relative costs of different forms of economic variability.

OTHER EVIDENCE?

The notion that focusing more on limiting fluctuations in real output will lead to more inflation variability is fairly intuitive. But does such a variability trade-off actually exist?

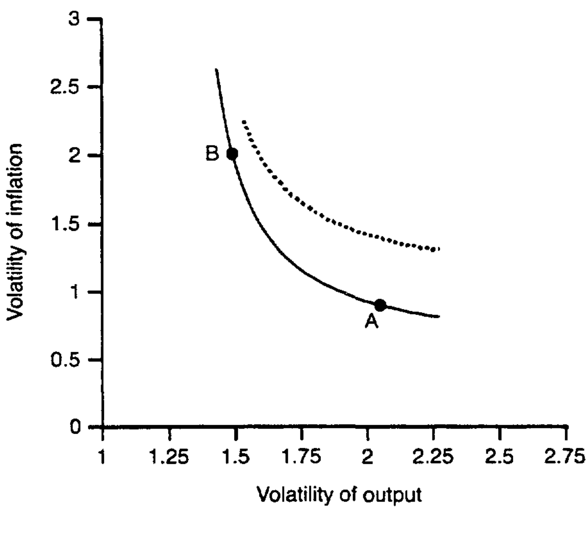


Figure 1 Illustrative volatility trade-off.

Simulations of economic models reveal such a trade-off, but economists disagree about which model best captures the true behavior of the economy, and these disagreements mean that there is no consensus about the true trade-off faced by policy makers. It is also difficult to find evidence of the trade-off in the data from actual economies. There are several reasons why the empirical evidence is inconclusive. The chief problem is that each point on the trade-off frontier is associated with a specific way of conducting monetary policy. If policy has been conducted in a stable and efficient fashion over several years, then the observed volatility of output and inflation would provide an observation on a single point on the trade-off frontier. Evidence on just a single point does not provide information on the entire trade-off frontier.

One way around this problem is to look at the experiences of many different countries. If countries have similar economic structures, have faced similar disturbances, and have operated on the efficient frontier but have differed in the choices policymakers have made between output and inflation stability, then historical patterns of different countries would provide evidence on the output and inflation variability trade-off. Unfortunately, actual economies have different economic structures, have experienced different disturbances, and have conducted policy in different ways. Thus, it is difficult to identify a variability trade-off using the historical experiences of a cross-section of countries.

RECENT DEBATES

Evaluating alternative policies in terms of their implications for the trade-off between output volatility and inflation volatility offers useful insights into some recent monetary policy debates. For example, the widely held consensus that monetary policy cannot have permanent effects on the level of the unemployment rate or the rate of real economic growth has led some to advocate that central banks focus only on maintaining low inflation. As a position about the long run, few economists would disagree. But, as Taylor's second proposition suggests, many economists would argue that a single-minded focus on maintaining inflation within a very narrow band may lead to undesired real economic fluctuations. And conversely, attempts to smooth real fluctuations too actively will lead to excessively volatile inflation.

The variability trade-off is also important for those countries that have moved to an inflation targeting policy regime since it is critical for determining the appropriate width of the inflation target. New Zealand, for example, initially defined its inflation target as 0–2% inflation. In 1997, however, this was widened to 0–3%. The Bank of England has a target inflation band of plus or minus 1% around its target of 2.5% inflation. The output-inflation variability trade-off is one of the key factors in determining the effects of changing the width of the inflation band. If the trade-off frontier is steep, for example, then reducing the variability of inflation causes little increase in output variability. In this case, a narrow target inflation band would be appropriate. A recognition of the variability trade-off shifts the focus from the level of inflation (which should be low) to questions of how wide the target band should be.

Realizing that the long-run effects of monetary policy determine average inflation, not average unemployment or the economy's real rate of growth, is critical to maintaining a successful policy. However, central banks still face trade-offs as they balance short-run inflation variability against short-run output variability, and this also needs to be recognized when evaluating the contribution of monetary policy to achieving macroeconomic goals.

REFERENCES

Fuhrer, J. C. 1997. Inflation/Output Variance Trade-offs and Optimal Monetary Policy. *Journal of Money, Credit and Banking* 29 (May) pp. 214–234.

Taylor, J. B. 1996. How Should Monetary Policy Respond to Shocks while Maintaining Long-Run Price Stability?—Conceptual Issues. In *Achieving Price Stability,* Federal Reserve Bank of Kansas City, pp. 181–195.

Taylor, J. B. 1993. Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, pp. 195–214.

This Page Intentionally Left Blank

55

Do Rising Labor Costs Trigger Higher Inflation?

David A. Brauer
Economist, Domestic Research Function of Research and Market Analysis Group, Federal Reserve Bank of New York, New York

Policymakers have followed with great interest the search for a reliable early indicator of inflation. One measure that has received considerable attention is the changes in hourly compensation—wages and benefits—of workers. On average, compensation represents about two-thirds of the total cost of production, and economic theory suggests that an increase in the rate of compensation growth will lead to accelerating price inflation unless the increase is offset by greater productivity growth or a squeeze on profits.

Throughout the 1990s, both the employment cost index (ECI)—which measures the rate of change in employers' hourly costs of providing wages, salaries, and benefits—and the consumer price index (CPI)—the most widely watched inflation measure—have risen at moderate rates relative to the previous decade. More broadly, movements in the two measures, both of which are produced by the U.S. Department of Labor's Bureau of Labor Statistics (BLS), have often mirrored one another. However, the evidence that compensation growth developments *lead* overall CPI inflation has thus far been inconclusive.

This chapter attempts to clarify some of the ambiguity surrounding the links between labor costs and price inflation over the post-1982 period. By breaking down compensation and prices into their various components, this study demonstrates that compensation growth in the service-producing segment of the private sector can help predict prices for a specific group of services and that these prices can in turn help predict movements in goods prices. However, compensation growth among goods producers in the private sector is shown to have little predictive power for goods prices. Overall, these results suggest that by concentrating on compensation developments in private sector services, we can obtain important information about the future path of inflation.

PREVIOUS STUDIES OF LABOR COSTS AND INFLATION

Several economists have attempted to sort out statistically the links between compensation and prices, with mixed results. For example, Gordon (1988) finds that for the 1954–87 period, wages and prices are determined by separate processes. More recently, Emery and

Reprinted from: *Federal Reserve Bank of New York Current Issues in Economics and Finance,* v. 3, n. 11 (Sept. 1997) 1–5.

Chang (1996) show that unit labor costs (compensation growth minus productivity growth) have no forecasting power for CPI inflation during the 1990s. On the one hand, Mehra (1991) and Huh and Trehan (1995) find that although unit labor costs and output prices move in tandem over the long run, prices lead labor costs but not vice versa. On the other hand, Mehra (1993) shows that changes in consumer prices and labor costs help to predict one another. Finally, Lown and Rich (1997) conclude that unit labor costs play a role in predicting inflation over the 1965–96 period.

In general, all of these results tend to be sensitive to the period of analysis, the exact question asked, and the specific labor cost and price measures used. Significantly, none of the studies uses the ECI—the best measure of labor cost inflation currently available—because the relevant series of the ECI have only been issued since the late 1970s or early 1980s.[1] Moreover, none of these studies uses a disaggregated approach to examine labor costs or price inflation.

LABOR COSTS AND PRICE DETERMINATION IN THE GOODS AND SERVICES SECTORS

Theoretically, a link between compensation and prices could run in either direction. For instance, firms whose compensation costs are rising more rapidly than their productivity growth could at some point be expected to attempt to raise product prices. However, higher price inflation could itself trigger more rapid compensation growth through explicit or implicit contractual arrangements (such as cost-of-living allowances) or through the influence of inflation expectations on the wage-setting process. Consistent with theory, when we first examine the relationship between private sector compensation and prices on an *aggregate* level, we, like the other researchers, arrive at mixed results. We find that the ECI and the core CPI (which excludes the often volatile food and energy components) frequently move in tandem, with only a slight tendency for labor cost developments to precede price movements (Figure 1).

These mixed findings suggest that a disaggregated approach to explaining CPI inflation would be fruitful. We break down the CPI into various components in order to isolate those goods or services whose prices are most likely to respond to changes in labor costs. In theory, one would expect labor costs to have the strongest direct impact on prices when these costs represent a substantial portion of the cost of production and when producers

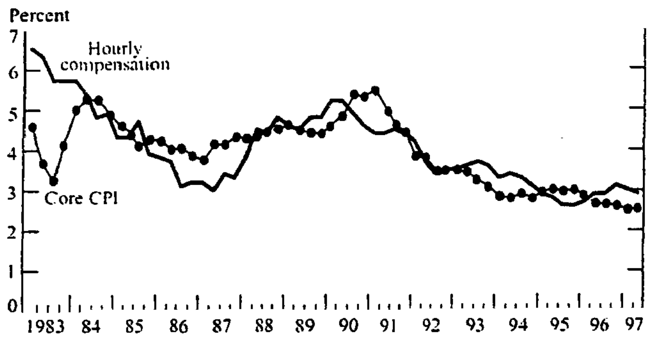


Figure 1 Labor costs and core inflation (four-quarter changes).

have some control over the setting of prices. Thus, we look for categories of goods or services that meet both of these criteria.

We find that factors other than labor costs dominate price determination for many of the CPI's components. In fact, these components (identified as "other expenditure categories" in Table 1) represent a significant share of consumer spending. For example,

Table 1 CPI Expenditure Categories

Labor-cost-sensitive services		Labor-cost-sensitive goods		Other expenditure categories	
Category	Relative importance[a]	Category	Relative importance[a]	Category	Relative importance[a]
Food away from home	5.9	New vehicles	5.0	Homeowners' costs	20.1
Personal and educational services	4.1	Apparel commodities	5.0	Food at home	9.9
Other private transportation services[b]	4.0	House furnishings	3.4	Medical care	7.4
Entertainment services	2.4	Entertainment commodities	2.0	Rent, residential	5.8
Other renters' costs[c]	2.2	Used cars	1.3	Fuels	3.8
Private transportation: maintenance and repairs	1.5	Housekeeping supplies	1.1	Other utilities and public services[c]	3.2
Housekeeping services	1.5	Toilet goods and personal care appliances	0.6	Motor fuel	2.9
Personal care services	0.6	Oher private transportation commodities[d]	0.6	Tobacco and smoking products	1.6
Apparel services	0.5	School books and supplies	0.3	Alcoholic beverages	1.6
Sheter: maintenance and repair sercices	0.1	Shelter: maintenance and repair commodities	0.1	Public transportation	1.5
Total	22.8	Total	19.4	Total	57.7

Sources: Names of categories and estimates of relative importance are drawn from the U.S. Department of Labor, Bureau of Statistics. The division of the BLS categories into "labor-cost-sensitive services." "labor-cost-sensitive goods." and "other expenditure categories" is based on the author's evaluation of the responsiveness of various types of expenditures to wage changes.

[a] The BLS defines relative importance as the share of total expenditures for which a category would account given current prices and a level of consumption unchanged from the base period 1982–84. The figures reported in the table are based on 1995 prices.

[b] Includes auto insurance, finance charges, and fees.

[c] Includes out-of-town lodging, lodging at school, and tenants insurance.

[d] Includes auto parts, products, and equipment.

[e] Includes telephone, cable television, water, and trash collection services.

prices of energy and food eaten at home, which together make up about 17 percent of the CPI, often fluctuate because of weather conditions, international political developments, or other temporary factors unrelated to producers' labor costs. Housing costs, which account for about 26 percent of the CPI, are also little affected by current labor costs, because the short-run supply of housing is essentially fixed, with rents determined much more by land values and the cost of materials than by the labor input into current housing services.[2] Moreover, in the case of several components of the CPI, the government plays a major role in setting prices, so even if labor costs were an important part of production we would not expect them to affect prices directly. Among these components are utilities, public transportation, and medical care—services in which the government frequently either regulates prices or provides the services itself—and alcoholic beverages and tobacco, whose price movements often reflect shifts in federal or state taxation.

For the purpose of our analysis, the rest of consumer spending can be divided into two classes: one in which product prices are determined through global markets and one in which price determination is much more local. In the goods sector, the presence of global competition arguably limits firms' ability to raise product prices in response to higher labor costs. This effect is strongest in industries that have high and rising import-penetration ratios (apparel is a prominent example), because imports lead directly to lower prices. Even in other goods-producing industries, however, the threat of competition from low-cost producers—foreign or domestic—can act to restrain price increases. Some firms respond to higher labor costs by taking steps aimed at boosting productivity growth, thus avoiding the need to raise prices in order to maintain profitability. Firms that cannot either raise prices or increase productivity will suffer reduced profits. Therefore, for most goods we would expect labor costs to have only a limited direct impact on prices.

In the services sector, however, providers of consumer services typically are much less subject to global competition than goods producers are, such that any cost increase can be more easily passed on to consumers. For example, convenience based on location is an essential feature of the dry-cleaning industry in New York City. The existence of lower cost producers of the same service in other cities will therefore have no bearing on the price-setting mechanism for this service in New York City. Barring the entry of lower cost providers into the same neighborhood, we would thus expect any local increase in wages to lead quickly to higher prices. Moreover, the ability of service providers to improve productivity in response to higher labor costs (at least in ways that can be measured) appears to be considerably more limited than it is for goods producers: Because many services are based on direct interaction with customers, any increase in the number of customers served per hour may diminish service quality.

Taking these descriptions of the goods and services sectors into account, we have decomposed the CPI into three broad groups. The table below lists the components of these groups, together with their December 1995 "relative importances"—a BLS term that is roughly equivalent to expenditure shares.[3] The first group, "labor-cost-sensitive services," is one in which labor costs are important and prices are set primarily in local markets. It consists of food away from home and a variety of other services. The second group, "labor-cost-sensitive goods," is one in which labor costs may be important but price determination is global, such as motor vehicles, clothing, and a number of other goods. The third group, described earlier, consists of all other expenditure categories, including food at home, energy, and housing. The rest of this article will examine the first two groups, which together account for about 40 percent of consumer expenditures.

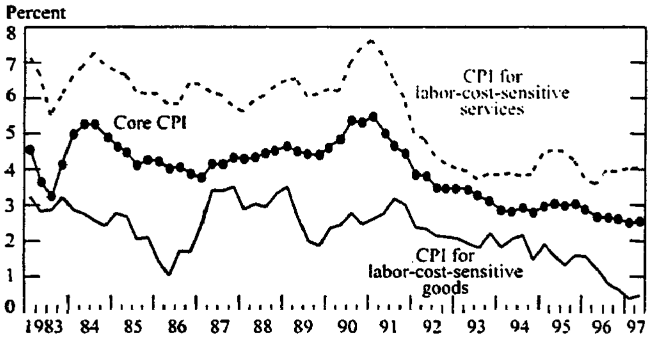


Figure 2 Price inflation in major expenditure categories (four-quarter changes).

TESTING THE COMPENSATION-PRICE LINKS

To test our hypothesis that the effect of any acceleration in labor costs on inflation is concentrated in the services sector, we calculate quarterly estimates of the CPI for the two groups beginning in 1983 (Figure 2). The estimates are constructed by taking averages of the indexes for the individual components, with the averages weighted by the relative importance of the components. Overall, we find that service prices throughout this period have risen relative to goods prices. Using our decomposed CPI and data from the sector-specific ECI, we now examine the link from compensation to prices in the goods-producing and service-producing sectors separately.

In the goods sector, increases in the rate of compensation growth do not appear to lead increases in price inflation (Figure 3). Most notably, the uptick in the CPI in 1986–87 preceded the acceleration in compensation growth, and in both 1989 and 1996 the CPI and the ECI moved in opposite directions. This result is confirmed by regression tests estimated over the fourth quarter of 1983 through the fourth quarter of 1996 using seasonally adjusted data. These tests indicate that past changes in the ECI for goods-producing industries do not improve on forecasts of changes in the CPI for labor-cost-sensitive goods—an unsurprising result given the limited ability of many goods producers in highly competitive markets to raise prices in response to higher costs.[4]

A rather different picture emerges in the services sector (Figure 4). We see that both the acceleration in labor costs during the late 1980s and the deceleration beginning in 1990

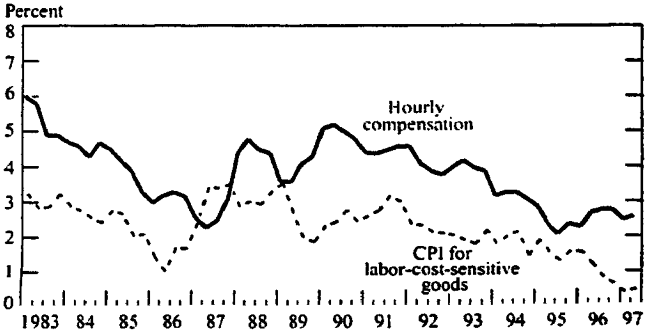


Figure 3 Compensation and price inflation: goods (four-quarter changes).

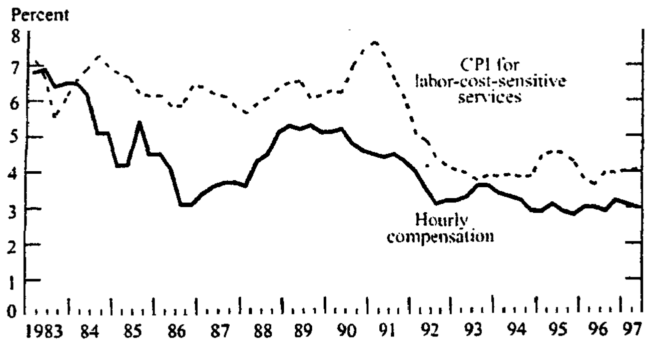


Figure 4 Compensation and price inflation: services (four-quarter changes).

preceded similar movements in the CPI for labor-cost-sensitive services and that both the CPI and the ECI have been quite stable since 1993. Running the same regression tests conducted for the goods sector, we are unable to reject the hypothesis that changes in the ECI for service-producing industries help predict changes in the CPI for labor-cost-sensitive services.[5] Over a three-quarter period, the estimated cumulative impact of hourly labor costs on the inflation rate is about 0.4 percentage point for each percentage point change in the labor cost inflation rate.[6] Because labor-cost-sensitive services represent nearly 23 percent of the CPI (see Table 1), this result implies that such an increase in labor costs directly adds about 0.1 point to overall inflation (0.4*0.23). This finding is consistent with the view that providers of services have a greater ability than goods producers to pass on higher costs to their customers.

We conduct additional regression tests to determine whether accelerating labor costs in the services sector can increase inflation indirectly, through their effects on the price of consumer goods. Current services, after all, play a role in the production and distribution of consumer goods. For example, if manufacturers are forced to pay more for financial and legal services, or for the shipping of goods from factories to stores, then these increases are likely to be reflected in the price of the goods. Although a full simulation of the impact of any increase in labor costs is beyond the scope of this article, our tests of this indirect influence indicate that each percentage point increase in the inflation rate for labor-cost-sensitive services adds roughly 0.4 percentage point to inflation in the other components of the CPI.[7] This finding implies that in addition to labor costs' direct effect on inflation, the original percentage point increase in labor costs has an indirect impact on overall inflation of about 0.1 point (0.4*0.4*0.77). When we test the effect of a price change in labor-cost-sensitive goods on the price of labor-cost-sensitive services, however, we find no evidence of a link, most likely because newly purchased goods play a fairly minor role in the provision of services.

Although our results suggest that rising labor costs can increase inflation both directly and indirectly, we are careful to note that other kinds of supply shocks can trigger an inflationary spiral. For instance, a large and sustained increase in food or energy price inflation can cause workers to increase their wage demands, a development that could in turn force up prices of other items through the mechanisms discussed above. However, even in these cases, services sector labor costs are likely to be an important part of the inflationary transmission mechanism.

CONCLUSION

The results presented here confirm a link from services sector wages and prices to overall inflation. We find that if compensation growth accelerates in the service-producing sector, that growth is likely to show up directly as more rapid inflation in service prices. Moreover, higher hourly labor costs in services can, through their contribution to the production and distribution of goods, indirectly affect goods prices. Given earlier researchers' findings showing a link from prices to wages, even these modest initial effects may therefore be enough to set off an inflationary spiral. Since no such effects are found to arise from an acceleration of labor cost increases in the goods-producing sector, policymakers seeking to prevent a resurgence of inflation may wish to pay particular attention to hourly labor costs in the service-producing private sector.

NOTES

1. The main advantage of the ECI over two other common compensation measures—the monthly average hourly earnings series (which excludes benefits) and the nonfarm business sector compensation measure—is that it controls for the impact of shifts in the mix of jobs by measuring wages and benefits for a fixed set of industries and occupations. It is also designed to remove the influence of changes in the volume of overtime worked on reported hourly compensation. One disadvantage of the ECI is that it does not adjust for changes in productivity growth; however, a related measure—unit labor costs—which does control for productivity changes, cannot be reliably decomposed into goods and services.
2. Over time, however, rents are influenced by *general* wage increases in the sense that the resulting higher consumer incomes boost the demand for housing.

   In the CPI, the cost of owner-occupied housing is expressed in terms of "owners' equivalent rent," meaning the imputed value of the services provided by the housing unit. In practice, the estimation of owners' equivalent rent is largely based on observed rents paid for comparable housing units in similar locations.
3. To be precise, the BLS (U.S. Department of Labor 1997, p. 170) defines a category's relative importance as the "share of total expenditures" for which the category would account "if quantities consumed were unaffected by changes in relative prices and actually remained constant [at the fixed 1982–84 weight]." Thus, over time, relative importance increases for categories in which prices are rising faster than the overall CPI and decreases for categories in which relative prices are falling. In January 1998, the BLS plans to introduce updated weights January 1998, the BLS plans to introduce updated weights reflecting 1993–95 spending patterns.
4. Specifically, we performed Granger-causality tests in which the seasonally adjusted change in the CPI for each quarter was regressed on lagged values for the quarterly change in the ECI, as well as its own lagged values, and then conducted an F-test for joint significance of the lagged ECI variables. In the case of goods using three lags, the F-statistic on the lagged ECI variables was .69, with a significance level of 56.4 percent. Tests using different numbers of lags yielded similar results.
5. For three lags, the F-statistic was 6.31, with a significance level of 0.1 percent.
6. The standard error of this estimate is 0.17 percentage point.
7. For two lags, the F-statistic for a link from labor-cost-sensitive services to labor-cost-sensitive goods was 4.98 (significance level 1.09 percent), with a .34 point cumulative impact (standard error .14). In the case of a link from services to all other items, the F-statistic was 5.42 (significance level .75 percent), with a cumulative impact of .40 point (standard error .22).

REFERENCES

Emery, K. M., and Chang, C-P. 1996. Do Wages Help Predict Inflation? Federal Reserve Bank of Dallas *Economic Review,* first quarter: 2–9.

Gordon, R. J. 1988. The Role of Wages in the Inflation Process. *American Economic Review* 78, no. 2 (May): 276–83.

Huh, C. G., and Trehan, B. 1995. Modeling the Time-Series Behavior of the Aggregate Wage Rate. Federal Reserve Bank of San Francisco *Economic Review,* no. 1: 3–13.

Lown, C. S., and Rich, R. W. 1997. Is There an Inflation Puzzle? Federal Reserve Bank of New York Research Paper no. 9723.

Mehra, Y. 1991. Wage Growth and the Inflation Process: An Empirical Note. *American Economic Review* 81, no. 4 (September): 931–7.

Mehra, Y. 1993. Unit Labor Costs and the Price Level. Federal Reserve Bank of Richmond *Economic Review* 79, no. 4: 35–52.

U.S. Department of Labor. Bureau of Labor Statistics. 1997. *BLS Handbook of Methods.*

56

Wage Inflation and Worker Uncertainty

Mark E. Schweitzer
Economist, Federal Reserve Bank of Cleveland, Cleveland, Ohio

According to a recent article in the *New York Times,* the leading explanation of why inflation has been so limited these last three years—despite low unemployment rates—is that wage demands have been held down by an unusually high degree of "worker uncertainty."[1] Substantial effort has gone into identifying (and disputing) the sources of this presumed insecurity in the face of a rather buoyant labor market. The most commonly mentioned reasons are the threat of middle-management layoffs, competition from foreign workers, and less unionization, all of which are believed to reduce wage inflation by making workers think twice before requesting higher pay—even if their firms' balance sheets have improved.

As an alternative to that approach, I begin by reviewing the salary adjustment policies favored by human resource managers, the people who propose and justify pay increases in most large U.S. firms. Typically, these managers report that they use local cost-of-living increases and the wages paid by other employers to guide their wage-scale adjustments. Though potentially compatible with many economic theories (including the idea that firms pay workers the prevailing market price), these policies suggest that wage changes are driven by inflation instead of driving it. At a macroeconomic (economywide) level, this means that wage-setting policies should tend to tie pay increases to inflation and productivity growth on a lagging or contemporaneous basis, reversing the order of events described by the worker uncertainty hypothesis.

To examine the relationship between nominal wage growth and inflation, I use an unusually long, detailed time series (including more than 40 years of data) which shows that wages have generally moved with the sum of prices and productivity. Furthermore, this relationship is contemporaneous, at least for annual data. These results are largely confirmed by data from the more commonly used, but far shorter, Employment Cost Index (ECI) time series.

The subtle distinctions between these explanations for the recent restraint in wage and price growth are critical, because evidence for the worker uncertainty hypothesis has diminished as the expansion has continued. The "human resource policy" view implies that

Reprinted from: *Economic Commentary Federal Reserve Bank of Cleveland* (Aug. 15, 1997).

future wage growth will remain modest because recent inflation has been quite subdued. In contrast, the worker uncertainty story suggests that there will be a run-up, first of wages and then of prices, when the cloud of insecurity finally lifts.

FIRMS' WAGE-SETTING DECISIONS

Human resource directors (or compensation managers) typically use wage-setting procedures that seem alien to economists' models. In adjusting their firms' nominal wage structure, they look foremost at the wages paid by other firms, then at cost-of-living indexes, and last at their own firms' financial results.[2]

Many managers survey local companies on their salary rates for comparable occupations. Unemployment rates, an obvious indicator of labor market tightness, rarely enter the calculation directly. In a recent survey, only 13 percent of the compensation managers questioned said they would alter their companies' wages even if the jobless rate in their own industry rose five percentage points.[3] The combined impact of diverse firms setting wages along these lines has not been fully explored, but the procedures they use suggest that wages are likely to adjust either contemporaneously or following inflation rate changes.

One interpretation of this behavior is that firms adjust their real wages to keep them even with their workers' alternatives at comparable firms, even though few employees may in fact contemplate switching. This is surprising on several counts. When managers look outward at the wages currently being *paid* by other employers and at cost-of-living indexes, they ignore other wage rates that might prove acceptable to unemployed workers with appropriate skills. Furthermore, the low priority given to firms' inside financial information suggests that the link between firms' own prices and their wages is weaker than simple microeconomic explanations indicate. Overall, compensation managers' procedures show that firms are operating with limited direct information about their workers' marginal products; otherwise, they would simply pay those rates period by period.

We know that individual firms primarily use outside information to determine wage adjustments, but what guides aggregate changes? While any number of factors might be relevant, I will focus on the simplest determinants. In a competitive market, economic theory says that wages will equal workers' marginal revenue product, which is calculated as their marginal physical product times the market price of the goods or services they produce.[4] Expressing this relationship in aggregate rates of change produces an equilibrium condition in which wage growth should approximate the inflation rate plus productivity growth.[5] My earlier description of how human resource managers adjust pay rates questioned firms' ability to pinpoint marginal physical product at the individual worker level. We know that aggregate movements in wages relative to prices and productivity would alter labor's share of production. However, the stability of labor's share of output confirms the relevance of the microeconomic relationship, even at the aggregate level. Therefore, including productivity growth is simply a statement that *real* wage growth can also change nominal wages.

THE LONG-RUN EMPIRICAL RELATIONSHIP

This description of processes does not answer two important empirical questions: Are wage changes closely associated with inflation and productivity growth? And, if they are, what is there relative timing of these comovements? To examine the relationship between inflation and wage changes, the data set must either specify occupations or adjust for the changing structure of employment. This is because even in a five- or 10-year period, the economy's occupational mix will shift substantially, creating an illusion of wage inflation or deflation. For example, an increase in the number of highly paid professionals, even if their wages are held constant, will cause an apparent increase in wage rates. The effect can be large in periods like 1983–94, when the share of the workforce employed as managers or professionals rose more than four percentage points (from 23 to 27 percent of total employment).

The source of the long-term data used here is unique in that it includes wage rates within defined occupations over the last 40 years. The data set is constructed from the Community Salary Survey (CSS), conducted by the Federal Reserve Bank of Cleveland (FRBC) since 1965 for use in its own salary administration in Cleveland, Pittsburgh, and Cincinnati.[6] The survey asks employers their wage and salary levels (including bonuses but not fringe benefits). The measured changes are averaged, accounting for both the firm and occupation, to provide the mean wage change line plotted in Figure 1.

This figure confirms that CSS wage changes are generally synchronized with CPI+ (an abbreviation for the sum of productivity gains and inflation). The CSS mean wage adjustments' correlations with inflation and CPI+ are fairly high (0.815 and 0.642, respectively). Throughout the period shown, productivity growth varies substantially more than either the inflation rate or average wage growth, which obscures the relationship. However, productivity growth must be included, because over the whole period, wages grew 0.35 percentage point faster than prices on average. Over the last three years, average wage growth has been much closer to the inflation rate (wage growth led by only 0.05 percentage point), yet wage gains have been closer to CPI+ (0.3 percentage point lower) than in the full sample (1.5 percentage points lower). This suggests that wage growth in the last three years was fairly strong, considering the weakness of measured productivity gains.

As for the timing question, CSS data are gathered annually, so wages and prices are best described as changing contemporaneously (their correlation is 0.81). The correlations between inflation (or CPI+) and wage growth are considerably lower for wage growth, leading inflation by one year (0.54) or two years (0.35). The alternative—that wage growth

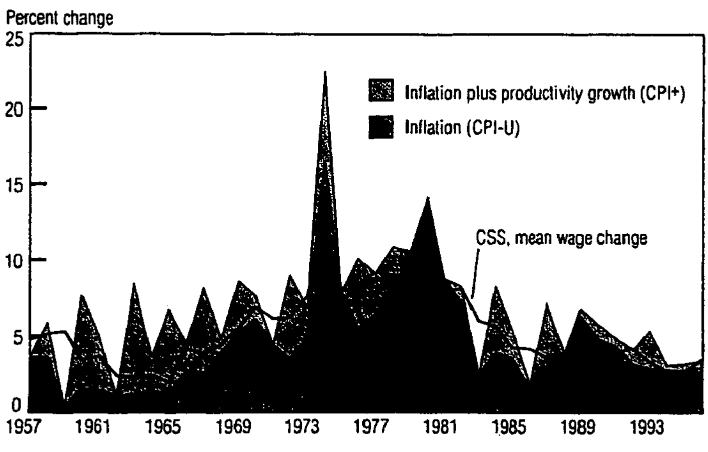


Figure 1 Community salary survey data.

follows inflation or CPI+ by one or two years—is similarly not supported. It is also clear that wage growth exceeded inflation or CPI+ growth during particular periods, with or without subsequent increases in the inflation rate. Overall, this source of detailed wage data is not consistent with the implication that current wage growth is unusually weak (at least relative to inflation and productivity growth) or that wage growth leads inflation.

THE EMPLOYMENT COST INDEX

The ECI addresses some of the potential problems of CSS data (it covers all regions and most occupations and includes information on benefits), but it has only been compiled since 1982. Total compensation, which includes both wages and benefit costs, is the best measure of the firm's labor-cost structure. Benefits, which now amount to 28 percent of total compensation, have generally risen as a fraction of compensation.[7] To keep the index reflecting actual compensation-rate increases for like work, the ECI holds the occupational composition of the work-force constant and isolates changes in hours (for example, increased use of overtime) from changes in the standard rate of pay.[8] These careful statistical controls make the ECI similar to the CPI. Repeating the comparison for the more limited sample, Figure 2 shows quarterly year-over-year growth rates in the CPI+ and the ECI, both of which are more variable; however, like CSS mean wage changes, ECI total compensation has typically grown faster than inflation but more slowly than the sum of inflation and productivity growth. Over the whole period, total compensation growth lagged growth in CPI+ by 0.44 percentage point on average. More recently, this difference has been smaller—only 0.26 percentage point lower since 1993. Again, this is primarily due to weak productivity growth, since the inflation rate has remained remarkably stable during this period.

The timing of inflation and wage growth, as measured by the ECI, is less clear. The peak correlation (0.72) between CPI+ and wage growth occurs when wage growth leads by three quarters. Interestingly, inflation, on its own, is most correlated (0.66) when wage growth *follows* inflation by three quarters. When wage growth leads inflation, the correlations are far lower (for example, 0.34 for a three-quarter lead). The reason that increases in

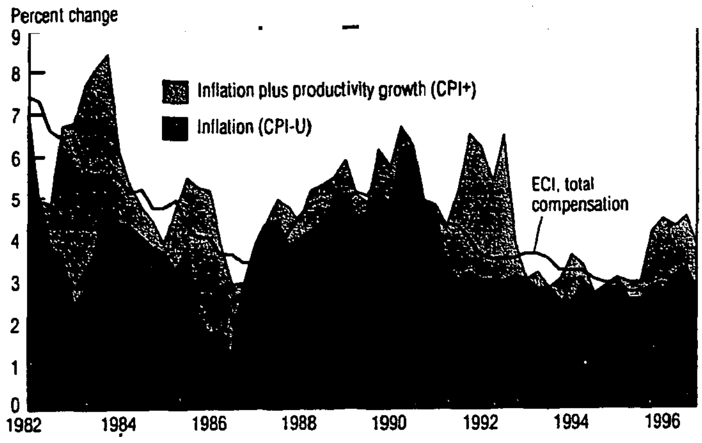


Figure 2 Employment cost index data.

CPI+ follow wage growth is the relatively strong correlation between wage growth and productivity increases three or four quarters out (0.540). Ultimately, the timing of these correlations lends limited support to either story. One of the problems with this analysis may be the shorter period for which ECI data are available. While CSS mean wage changes are not shown (to avoid overcomplicating Figure 2), they are very comparable to yearly changes in the ECI, suggesting that the CSS data do match national wage trends.

CONCLUSION

This chapter is not intended to refute the idea that worker uncertainty accounts for today's low inflation rates. Instead, it seeks to provide an alternative explanation for the "low" wage increases of the last three years that is at least as consistent with long-run *and* recent wage, price, and productivity data. Indeed, I do not attempt to analyze the unemployment rate's empirical importance in determining wage growth and inflation. Nonetheless, a preliminary look at two data sources yields evidence on the relationship between inflation and nominal wage growth that is consistent with the "human resource policy" explanation for our current low rates of wage growth. Further research will be required to determine how wage offerings at the firm level relate to firms' own price decisions and aggregate data on inflation and productivity.

Why worry about which story provides the more accurate explanation of recent wage gains? Because different answers suggest different paths for future inflation. It has become increasingly difficult to fathom why workers should feel insecure about their jobs, when surveys show they believe employment opportunities are abundant. If this is the case, inflation in wage demands and prices must be just around the corner. On the other hand, if human resource managers expect today's low inflation, low productivity gains, and low wage growth in other firms to continue, there is no reason to anticipate pay increases that could set off an inflationary spiral. Because the policy implications of these stories are so divergent, it is important to investigate alternatives to the worker uncertainty hypothesis.

NOTES

1. See P. Passell, A Pulse that Lingers, *New York Times,* July 22, 1997, p. A1.
2. See A. Freedman, The New Look in Wage Policy and Employee Relations, Report No. 865, The Conference Board, 1985.
3. See D. I. Levine, Fairness, Markets, and Ability to Pay: Evidence from Compensation Executives, *American Economic Review,* vol. 83, no. 5 (December 1993), p. 1248.
4. The marginal physical product is the addition to output made by the last worker hired.
5. Several simplifications are implicit in this move from microeconomic wage determination to the aggregate relationship. Notably, both the price and marginal products must be correctly aggregated for the simple relationship to hold, yet the aggregate data series are not even intended for this purpose. For example, productivity is measured on an average, rather than a marginal, basis. These are reasons to examine how well the relationship performs using actual data over an extended period.
6. The FRBC chooses participants in each city as representative of the area's employers. Although the survey has been conducted annually, the month for which data are collected has changed several times since 1955. All data, including the CPI and productivity figures, refer to the period between the preceding survey and the one conducted that year. In most cases, this is a 12-month

span, but occasionally the interval is less or more than a year. How well do the CSS wages reflect national trends? The year-to-year changes usually follow the national pattern closely (when U.S. wage-change data are available), but characteristics specific to this region have also caused its wage levels to change relative to the nation's. In general, Cleveland, Cincinnati, and Pittsburgh are more urban, have more cyclically sensitive employment, and have undergone more industrial restructuring than the United States as a whole. Before the 1980s, wages in these three cities were higher than the national average, but now they are on par with the rest of the country.

7. This has resulted from increases in both mandated benefits (such as the employer's contribution to the Social Security fund and unemployment insurance premiums) and voluntary benefits (such as health care insurance and paid vacation days).
8. Occupations in which overtime is a standard, continuing feature of the compensation structure are assigned a rate that exceeds the base wage, on the basis of typical overtime levels within the occupation.

57

U.S. Inflation Targeting
Pro and Con

Glenn D. Rudebusch
Research Officer, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Carl E. Walsh
Professor of Economics, University of California, Santa Cruz, California

In recent years, monetary economists and central bankers have expressed growing interest in inflation targeting as a framework for implementing monetary policy. Explicit inflation targeting has been adopted by a number of central banks around the world, including those in Australia, Canada, Finland, Israel, New Zealand, Spain, Sweden, and the United Kingdom. In the United States, there has been little public debate over inflation targeting, although some bills have been introduced in Congress to mandate the use of explicit targets for inflation.

Issues and questions surrounding inflation targeting formed a major focus of a recent conference on Central Bank Inflation Targeting jointly sponsored by the Federal Reserve Bank of San Francisco and the Center for Economic Policy Research at Stanford University (Rudebusch and Walsh 1998). In this chapter, we set out some of the arguments for and against adopting inflation targeting in the United States discussed at the conference. (For further discussion, see Bernanke and Mishkin 1997, Bernanke, et al., forthcoming, and Keleher 1997.)

WHAT WOULD INFLATION TARGETING MEAN IN THE U.S.?

There has been some ambiguity about the precise definition of an inflation targeting policy regime, in part because certain institutional arrangements have differed from one inflation targeting country to another—most notably with regard to how the inflation target is set and how deviations from the target are tolerated. For our discussion, we define inflation targeting to be a framework for policy decisions in which the central bank makes an explicit commitment to conduct policy to meet a publicly announced numerical inflation target within a particular time frame. For example, at the start of 1993, Sweden's central bank announced an inflation target for the consumer price index of between 1 and 3 percent by 1995. Simi-

Reprinted from: *FRBSF Economic Letter,* No. 98-18 (May 29, 1998).

larly, if the Federal Reserve wanted to adopt inflation targeting, it would publicly commit to achieving a particular numerical goal for inflation (a target point or range) within a set time span of, say, a couple of years. Also, as part of an ongoing policy framework for targeting inflation, the Fed's semiannual Humphrey-Hawkins report, which currently provides a near-term (one-year) outlook for inflation, could be augmented to include a discussion of whether the medium-term (two or three-year) inflation forecast is consistent with the announced medium-term inflation target. Inflation targeting would not impose a rigid simple rule for the Fed; instead, policy could employ some discretion to take into account special shocks and situations. However, the organizing principle and operational indicator for monetary policy would be focused on inflation and (in light of lags in the effects of policy) inflation forecasts.

Given some earlier ambiguity, it is important to be clear about how we interpret the ultimate goals of inflation targeting monetary policy. In particular, an inflation targeting central bank need not care only about inflation. Indeed, most inflation targeting central banks continue to recognize multiple goals for monetary policy with no single primary one. (In New Zealand—an exception—the start of inflation targeting coincided with a legislative mandate that the central bank's "primary function" was price stability, while in Canada—as is more typical—such legislation was never passed.) Accordingly, an operational policy framework of inflation targeting would still be consistent with the Fed's current legislated objectives of low inflation and full employment. An inflation targeting regime can accommodate a goal of output stabilization by having wide inflation target bands, long inflation target horizons, and explicit exemptions for supply shocks. Thus, the adoption of an inflation targeting regime does not necessarily require that price stability or low inflation be the preeminent goal of monetary policy. As with monetary targeting, inflation targeting is an operational framework for monetary policy, not a statement of ultimate policy goals. (For further discussion, see Rudebusch and Svensson 1998.)

ARGUMENTS PRO

By focusing attention on a goal the Fed can achieve, by making monetary policy more transparent and increasing public understanding of the Fed's strategy and tactics, by creating institutions that foster good policy, and by improving accountability, the adoption of inflation targeting would represent a desirable change in the U.S. monetary policy.

1. The announcement of explicit inflation targets for the Fed would provide a clear monetary policy framework that would focus attention on what the Fed actually can achieve. Bad monetary policy often has resulted from demands that central banks attempt to achieve the unachievable. Most notably, few macroeconomists believe that monetary policy can be used to lower the average rate of unemployment permanently, but central banks often are pressured to achieve just that through expansionary policy; such policy instead only results in higher average inflation without leading to a systematically lower average rate of unemployment. In contrast, implementing explicit inflation targets would help to insulate the Fed from such political pressure.

2. Transparent inflation targets in the U.S. would help anchor inflation expectations in the economy. When making real and financial investment decisions and planning for the future, businesses and individuals must form expectations about future inflation. Inflation targets would provide a clear path for the medium-term inflation outlook, reducing the size of inflation "surprises" and their associated costs. Inflation targets also likely would boost

the Fed's credibility about maintaining low inflation in the long run, in part, because they mitigate the political pressure for expansionary policy. Since long-term interest rates fluctuate with movements in inflation expectations, targeting a low rate of inflation would lead to more stable and lower long-term rates of interest. Together, the reduced uncertainty about future medium-term and long-term inflation would have beneficial effects for financial markets, for price and wage setting, and for real investment.

3. The establishment of inflation targets in the United States would help institutionalize good monetary policy. Recent U.S. monetary policy has been generally considered excellent, but earlier in the postwar period, monetary policy clearly failed by allowing inflation to ratchet up significantly several times. To some extent, the quality of policy over time has reflected the skills and attitudes of the people involved in the policy process. Monetary policy is an area in which it is especially important to implement institutional structures that will help to avoid bad policies. Inflation targets can provide this institutional structure and help ensure that monetary policy is not dependent on always having the good luck to appoint the best people.

4. In the current system, there is some ambiguity about how and why the Fed operates. For example, although monetary aggregates play a very modest role in the policy process, they are the only variables that the Fed is required to set target ranges for and report about to Congress. As noted above, inflation targets would focus discussion on what the Fed actually could achieve. Furthermore, an inflation target provides a clear yardstick by which to measure monetary policy. Given forecasts of future inflation, it is easy to compare them to the announced inflation target and hence judge the appropriate tightness or looseness of current monetary policy. Also, on a retrospective basis, an explicit target allows Fed performance to be easily monitored. Thus, Congress and the public will be better able to assess the Fed's performance and hold it accountable for maintaining low inflation.

ARGUMENTS CON

Inflation targeting, even without imposing a rigid rule, would unduly reduce the flexibility of the Fed to respond to new economic developments in an uncertain world. Furthermore, publicly committing solely to an inflation target would not enhance overall accountability or transparency given the multiple objectives of monetary policy.

1. The purpose of inflation targeting is to focus the attention of monetary policy on inflation. However, concentrating on numerical inflation objectives (even with caveats or escape clauses) also reduces the flexibility of monetary policy, especially with respect to other policy goals. That is, inflation targets place some constraints on the discretionary actions of central banks. Such constraints can be quite appropriate in countries where monetary policy has performed poorly, exhibiting sustained unproductive inflationary tendencies; however, this is not the case in the United States. U.S. monetary policy has operated quite well for almost two decades, so limiting the flexibility and discretion of the Fed to respond to new economic developments would be ill-advised. Why change a system that is working? Certainly, adept policymakers are one reason for the good performance of recent monetary policy, but there is also a strong institutional structure—stronger than existed at the start of the 1970s—that is already in place at the Fed that fosters good monetary policy.

2. Monetary policy requires the careful balancing of competing goals—financial stability, low inflation, and full employment—in an uncertain world. There is uncertainty about the contemporaneous state of the economy, the impact policy actions will have on fu-

ture economic activity and inflation, and the evolving priority to be given to different policy objectives. However, because monetary policy actions affect inflation with a lag, inflation targeting means, in practice, that the Fed would need to rely heavily on forecasts of future inflation. Given the uncertainties the Fed faces, an inflexible and undue reliance on inflation forecasts can create policy problems. For example, most forecasts in the mid-1990s of inflation in the late 1990s overestimated the inflation we are currently experiencing. If the Fed had been inflation targeting in the mid-1990s, it might well have raised the funds rate based on its inflation forecasts. Yet with today's low inflation and robust economy, it is difficult to argue that the Fed was too expansionary and that the more contractionary policy implied by inflation targeting would have produced a better outcome. As in this instance, it seems unlikely that a mechanical dependence on inflation forecasts to achieve inflation targets will improve policy.

3. Proponents of inflation targeting argue that it promotes accountability. However, as is generally agreed, low inflation is only one of the objectives of monetary policy. While monetary policy may not affect average real growth or unemployment over time, it does have an important role to play in helping to stabilize the economy. Even if average inflation is the one thing the Fed can control in the long run, it does not follow that the Fed should be held accountable only for its inflation record. Inflation targeting actually could reduce the Fed's overall accountability by allowing it to avoid responsibility for damping short-run fluctuations in real economic activity and unemployment. Making the Fed publicly accountable for only one policy goal may make it harder for Congress and others to monitor the Fed's contribution to good overall macroeconomic policy.

4. Similarly, with regard to the transparency and public understanding of policy, inflation targeting highlights the inflation objective of central banks but tends to obscure the other goals of policy. Just as uncertainty about future inflation impedes good economic decision making, so does uncertainty about the future level of output and employment. Inflation targeting sweeps the latter concerns under the rug (often by adjusting the amount of time that deviations are allowed from the inflation target). Given the multiple legitimate goals of policy, the single public focus of inflation targeting does not enhance overall transparency.

SUMMARY

The debate about the appropriateness of inflation targets in the U.S. continues, but it is likely that the actual experiences of inflation targeting countries will provide the most convincing evidence. The recent record of inflation targeting countries has been good, but many other countries also have reduced inflation and maintained low rates of inflation even without employing a formal targeting framework. The generally benign macroeconomic environment of the past few years still leaves much unknown about how best to reconcile sufficient policy flexibility with the maintenance of low inflation. The oldest inflation targeting regime (New Zealand) is barely eight years old, and there is still much to learn.

REFERENCES

Bernanke, B. S., and Mishkin, F. S. 1997. Inflation Targeting: A New Framework for Monetary Policy? *Journal of Economic Perspectives* 11 (Spring) pp. 97–116.

Bernanke, B. S., Laubach, T., Mishkin, F. S., and Posen, A. 1999. *Inflation Targeting: Lessons from the International Experience.* Princeton, NJ: Princeton University Press.

Keleher, R.E. 1997. Establishing Federal Reserve Inflation Goals. Joint Economic Committee Study (April). Available at http://www.house.gov/jec.

Rudebusch, G.D., and Svensson, L.E.O. 1998. Policy Rules for Inflation Targeting. NBER Working Paper 6512.

Rudebusch, G. D. and Walsh, C. E. 1998. Central Bank Inflation Targeting. *FRBSF Economic Letter* 98-17 (May 22).

This Page Intentionally Left Blank

58

Historical U.S. Money Growth, Inflation, and Inflation Credibility

William G. Dewald
Research Director, Federal Reserve Bank of St. Louis, St. Louis, Missouri

INTRODUCTION

Although many forces affect individual prices in the short run, the historical record shows that in the long run changes in the general level of prices, i.e., inflation, have been linked systematically to changes in the quantity of money. The Federal Reserve uses as its principal monetary policy target an overnight inter-bank interest rate, the federal funds rate, which it manipulates by open market operations that change its portfolio of government securities, which in turn influences monetary growth. Economists both inside and outside the Federal Reserve monitor a wide range of indicators so as to judge the appropriateness of a monetary policy target relative to the goals of achieving a stable price level and sustained real growth. For many years presidents of the Federal Reserve Bank of St. Louis and many of its economists have called attention to research showing that long-term growth of monetary aggregates is among the more important of these indicators. They also have championed the preeminence of price level stability as a monetary policy goal to provide the best environment for sustained economic growth in a market economy.

The historical data reveal a consistent correlation between long-term growth rates in broad monetary aggregates, spending, and inflation in the United States, but not between such nominal variables and real output. Data from the bond market show that, despite inflation being at its lowest level in decades, the Fed has not regained fully the inflation credibility that it lost in the 1960s and 1970s.

NOMINAL AND REAL GROWTH, AND INFLATION

The financial press and the public often seem to believe that the way to contain inflation is to pursue policies that reduce real economic growth. The view that it necessarily takes lower real growth, or even a recession, to slow inflation is an improper reading of historical data. It fails to differentiate between the short run and the long run. Sustained real out-

Reprinted from: *Federal Reserve Bank of St. Louis Review* (November/December 1998) 13–23.

put growth depends on increases in the supply of labor and capital, and increases in the productivity of such inputs. Growth in demand for output certainly influences what is produced, but fundamental scarcities limit the aggregate amount of how much can be produced on a sustained basis. Furthermore, as the level and variability of inflation increase, price signals become fuzzier and decisions are distorted, which would tend to decrease real gross domestic product (GDP). Thus, one should not expect an increase in demand growth to increase real growth on a sustained basis, but if at all, only in the short run. An examination of the historical record supports this proposition.

Figure 1 is a plot of percentage changes over 1959–1997 in nominal GDP, real GDP, and the GDP price index. The chart includes year-over-year and 10-year moving averages. The four-quarter changes—the fine lines—remove the high frequency noise from the data. The 10-year changes—the heavy lines—remove the business cycle fluctuations as well. The gap between nominal and real GDP growth rates approximates inflation, a s measured by percentage changes in the GDP price index, which are shown in the bottom panel of Figure 1.[1]

Figure 1 reveals several regular patterns:

Inflation trended up through about 1980, and down since then.

Annual growth rates in nominal GDP and real GDP went up and down together.

Annual growth rates in nominal and real GDP were also more volatile than annual inflation rates.

Recessions—marked by the shaded bars—occurred more frequently from the late 1960s through the early 1980s when inflation was high and rising than since the early eighties when it trended down.

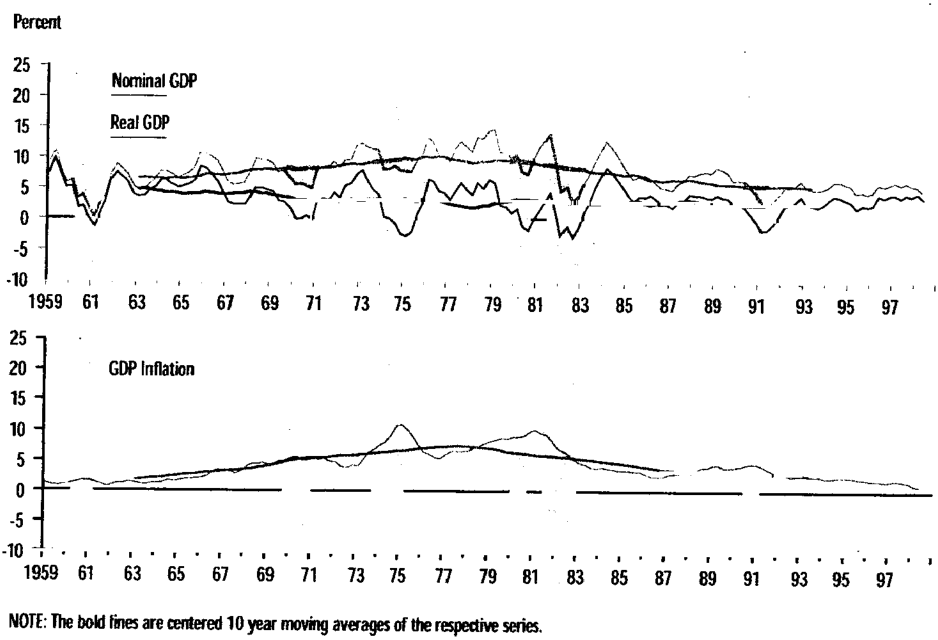


Figure 1 Inflation and real growth (year/year and 10-year averages).

Inflation typically accelerated before cyclical peaks, but then decelerated beginning in recessions and extending into the early phase of recoveries.

In recent years inflation has been its lowest and most stable since the late 1950s and early 1960s, and, atypically, it has continued to decelerate in the seventh year of the expansion.

The year-over-year movements in nominal and real GDP are matched closely in the short run, an observation seemingly suggesting that policies to increase nominal GDP growth would increase real GDP growth, too. That short-run relationship, however, does not hold up in the long run. From the 1960s through the early 1980s the increase in 10-year average nominal GDP growth was associated with a matching increase in 10-year average inflation, but, if anything, a decrease in 10-year average real GDP growth. Thus, increased average nominal GDP growth in the long run was not associated with increased real GDP growth, but only with inflation.

International evidence supports this finding that inflation harms long-run growth. Studies of other countries have identified a small negative effect of even moderate inflation on real growth.[2] Small differences amount to a lot over long periods because of compounding. Thus, it may not be an accident of history that the most highly industrialized economies with the highest per capita income today have had comparatively low inflation over extended periods. With respect to countries that have experienced inflation of 40 percent a year or more, the evidence is unambiguous: High inflation reduces real growth.[3]

High inflation also has been linked to cyclical instability. There was a deep recession in 1981 and 1982. This recession was associated with a genuinely restrictive monetary policy and interest rates at unprecedented levels. The rate of unemployment built up to more than 10 percent and inflation fell far more sharply than most forecasters had expected. Despite some relapse in the late 1980s, inflation has trended down since the early 1980s, and real GDP growth has averaged somewhat less than it did in the 1960s, but this is because of lower productivity growth and not because of recessions and unemployment. In fact, the U.S. economy has performed very well relative to its potential and better than ever in terms of cyclical stability. The 29-quarter expansion from 1991:Q1 through 1998:Q2 had not yet lasted as long as the record 34-quarter expansion of the 1960s. However, as inflation decreased from the end of the recession in 1982:Q4 through 1998:Q2, there were 63 expansion quarters and only three contraction quarters, an unprecedented era of cyclical stability in U.S. history. It surpassed the record of 1961:Q1 through 1973:Q4, which included 47 positive growth quarters and four contraction quarters. The record was not too shabby in either case, but there was a difference. The 1960s and early 1970s were a period of accelerating inflation, which laid a foundation for the instabilities that followed. The 1980s and so far the 1990s have been a period of decelerating inflation, which has lain a foundation for stable price level credibility and efficient resource utilization. The next figures bring monetary growth into the picture.

MONETARY GROWTH AND INFLATION

M2 is a measure of money that Milton Friedman and Anna Schwartz trace in their *Monetary History of the United States.*[4] It is a broad measure made up of assets having a common characteristic: Each is either issued by the monetary authorities, for example, currency and coin, or is an obligation of a depository institution legally convertible into such stan-

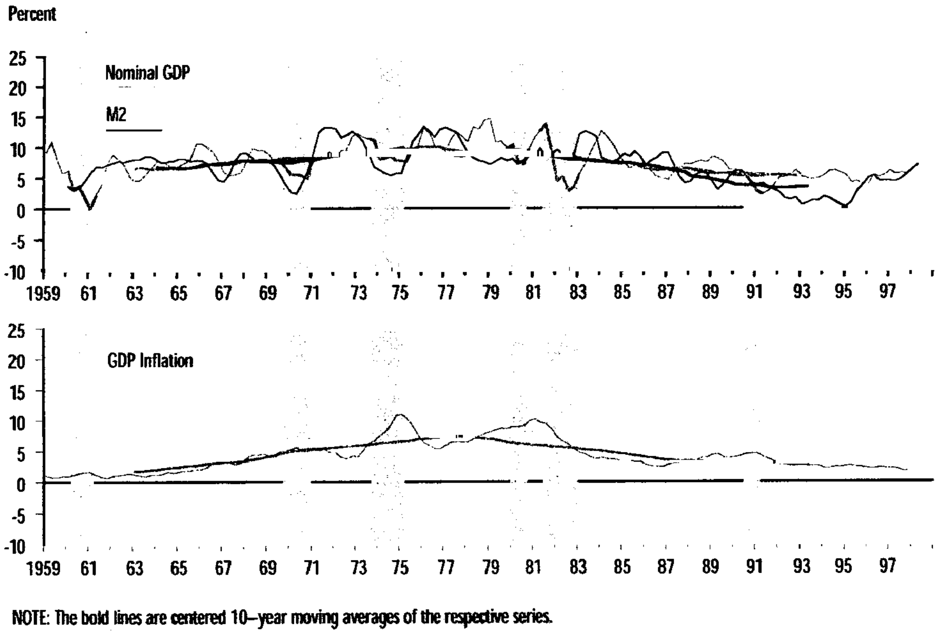


Figure 2 M2 and nominal GDP growth, and inflation (year/year and 10-year averages).

dard monetary units. M2 assets can be divided into M1 and non-M1 categories. M1 components can be used to make payments directly (currency, travelers checks, and checking accounts). Non-M1 components, which can be readily turned into M1 assets, include savings deposits, money-market mutual fund balances, and short-term time deposits. Such non-M1 components of M2 have become increasingly accessible to depositors for payments in recent years. M2, as a broad monetary aggregate, represents the essence of "liquidity," i.e., a way station between income receipts and expenditures for both households and non-financial businesses, and, as such, a variable that would be expected to be related to total national spending in current dollar terms, i.e., nominal GDP.

Figure 2 plots growth rates in M2, nominal GDP, and the GDP price index. It reveals some regular short-term patterns in the year-over-year data:[5]

M2 and nominal GDP growth rates slow before and during the initial stages of a recession.

M2 growth turned down many more times than the number of cyclical peaks.

M2 growth turned up during each recession and early recovery except during the most recent instance when it continued to slow.

M2 and nominal GDP have been growing at similar rates between 1995 and 1998.

This figure reveals the major reason why M2 has been discredited as an indicator of the stance of monetary policy in recent years—in the short run, movements in the monetary aggregates, nominal GDP, and inflation sometimes appear to be unrelated. For example, whereas M2 growth slowed dramatically between 1992 and 94, nominal GDP growth accelerated. That discrepancy produced the largest and most persistent deviation between the growth rates in M2 and nominal GDP in many years. This deviation has led the Federal

Open Market Committee (FOMC) and the public to place less emphasis on the money supply targets.

Nevertheless, giving up on the aggregates might be a mistake. The reason is that there has been a close *long-term* fit between M2 growth and nominal GDP growth and, in turn, inflation. Figure 2 shows the general upward trend in 10-year average M2 growth, nominal GDP growth, and inflation during the 1960s and 1970s, and the general downward trend in these 10-year averages during the 1980s and so far in the 1990s. Such a longer-term historical relationship is presumably a reason why M2 is one of the monetary variables for which the Federal Reserve continues to announce a target range in the Congressional Humphrey-Hawkins hearings twice a year. In his Humphrey-Hawkins testimony in February 1998, and again in July, Chairman Alan Greenspan noted that M2 growth might be back on track as an indicator of nominal GDP growth and inflation, after it appeared to have been off track earlier in the expansion.

Observations about M2, nominal GDP growth, and inflation over the long run support Milton Friedman's dictum that "inflation is always and everywhere a monetary phenomenon." The looseness of the short-term association supports his dictum that "lags are long and variable." Figures 3 and 4 makes these points with data going back to 1875.

Figure 3 shows that short-run, year-over-year changes in these historical series are very noisy. Yet, even on a year-over-year basis, the association of large movements in M2 with large movements in nominal gross national product (GNP) and inflation is apparent, (for example, the contraction of monetary growth and nominal GDP growth in the early 1930s and the associated deflation). In less turbulent times such as recent decades, however, there is no clearly discernable systematic short-run association between broad money growth, nominal GNP growth, and inflation. Of course, a change in the price level over a year or two is not really what is meant by inflation unless it is substantial enough to change the price level a lot.

Figure 4 shows that over the past 35 years the long upward and downward cycle in M2 and nominal GNP growth rates and inflation is only one of a series of comparable long

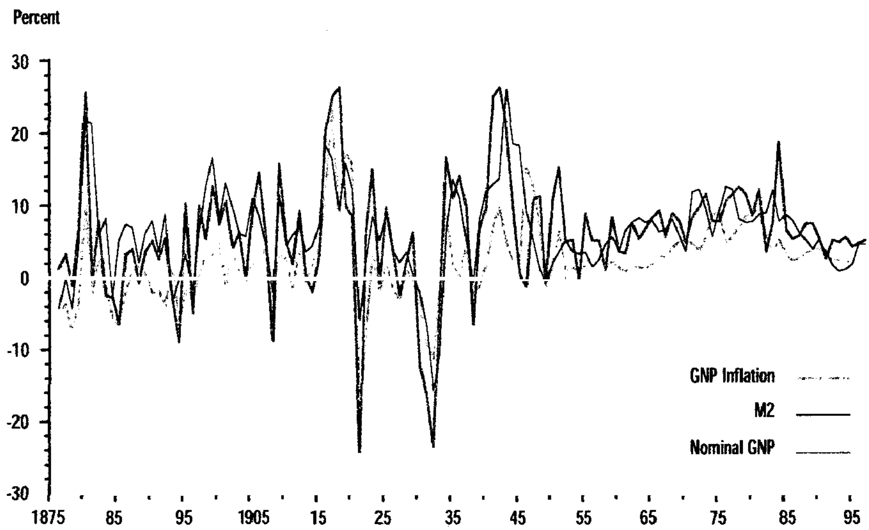


Figure 3 M2, nominal GNP growth, and inflation (annually).

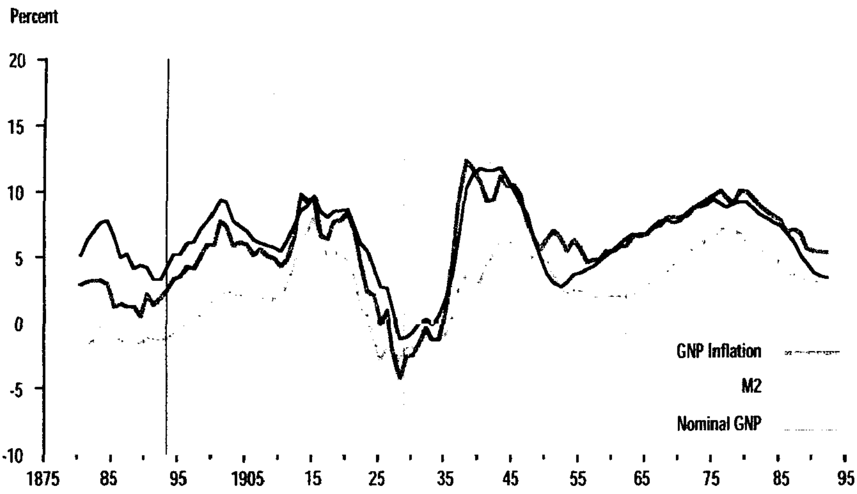


Figure 4 M2, nominal GNP growth, and inflation (10-year averages).

cycles in U.S. history. Following a period of low M2 growth and deflation in the 1870s and 1880s, there have been four long inflation-disinflation cycles. They are marked on the figure by troughs in centered 10-year average inflation in 1893, 1909, 1928, and 1962. In 1998, it is not known yet whether the last 10-year average plotted was a trough.

Because M2 growth tracks all previous inflation-disinflation cycles, it goes a long way to avert the suspicion that the relationship between monetary growth and inflation is spurious. Monetary historians such as Milton Friedman and Anna Schwartz have recognized that the mere association of monetary growth with inflation does not establish the direction of causality. To confirm that monetary growth causes inflation, they cite the evidence that the long-term relationship between monetary growth and inflation has remained much the same throughout history, including periods when we know that monetary growth resulted from supply-side factors. For example, when monetary growth accelerated in the 1890s, as engineering advances increased gold output, there was an associated inflation. Gold was then a standard into which currencies could be converted. When monetary growth collapsed in the early 1930s because of bank failures, there was an associated deflation.

The historical record also includes episodes when demand pressures led the Fed to support monetary increases. In both World Wars I and II, Fed policies to help the government finance its debt stimulated monetary growth. What followed were substantial increases in inflation. Nevertheless, even in wartime, there is reason to think that the Fed could have kept a damper on inflation. When federal deficits rose in the 1980s, but the rate of monetary growth fell, inflation did not rise. It fell. The historical evidence is that when the Fed has held interest rates down in the face of demand pressures by stimulating monetary growth, inflation has accelerated. However, in periods such as the 1980s, when monetary growth has not accelerated, inflation has not accelerated.

Every major acceleration in M2 growth has been associated with a major acceleration in inflation. Likewise, every major deceleration in M2 growth has been associated with a major deceleration in inflation. Accordingly, policy makers might be making a serious mistake if the noisy short-term movements in M2 and inflation persuaded them that money does not matter anymore. At a minimum, policy makers and the public might be wise to monitor monetary growth, mindful that inflationary demand pressures do not cause money growth unless the monetary authorities passively allow that to happen. Since the long run consists of an accumulation of short runs, it follows that sustained shifts in M2 growth are worth noting when formulating monetary policy. Keeping longer-term average M2 growth and nominal GDP growth in the neighborhood of longer-term real growth remains a practical guide for achieving a stable price level environment.

INFLATION AND INTEREST RATES

Readers might be surprised that monetary policy has been discussed to this point without much reference to interest rates. This approach was not an oversight.

Interest rates compensate lenders for giving up current purchasing power and taking some risk. One risk is that borrowers might default. Another is that what they pay back might have less purchasing power than what was lent.

Despite the conventional wisdom to the contrary, interest rates often have not been a good measure of the thrust of monetary policy on demand growth and inflation.[6] Because increases in expected inflation would tend to raise rates, rising nominal interest rates do not necessarily signal an anti-inflationary (tighter) monetary policy. Correspondingly, falling nominal interest rates are not necessarily a measure of a more inflationary (easier) monetary policy.

Nominal interest rates are highly sensitive to inflation and inflationary expectations. High inflation expectations lead lenders to demand compensation for the expected depreciation in the purchasing power of the money they lend, and borrowers are forced to add an inflation premium to the interest rates they pay.

Apart from default and inflation risk premiums, real (inflation adjusted) interest rates depend largely on underlying real factors such as domestic saving and investment and international capital flows, not on monetary growth and inflation. Thus, regardless of monetary growth and inflation, higher real interest rates generally reflect increased investment opportunities or decreased saving. That real interest rates reflect underlying real factors is another reason why interest rates are not a reliable measure of the stance of monetary policy.

In this regard, technological change in the 1990s, coupled with the long expansion, may have increased the return to capital investment in the U.S. economy, and hence the demand for capital relative to historical experience, which would tend to increase real interest rates. In such circumstances, there is a monetary policy risk in under-estimating the upward pressures on real interest rates that result from an increase in real investment demand. Any attempt to attenuate such pressures by stimulating monetary growth would risk a build up of inflationary pressures.

Fundamentally, monetary policy is tighter or easier not in terms of whether nominal *or* real interest rates are rising or falling, but in terms of whether inflationary pressures are falling or rising. As the historical figures have demonstrated, inflation in a longer-term sense is associated with high monetary growth. Figure 5 shows that increases and decreases in inflation trends are reflected in major increases and decreases in nominal interest rate levels.

Figure 5 plots the federal funds rate, the 10-year Treasury bond rate, and annual changes in the Consumer Price Index. When inflation held in the range of 1 1\2 to 2 per-

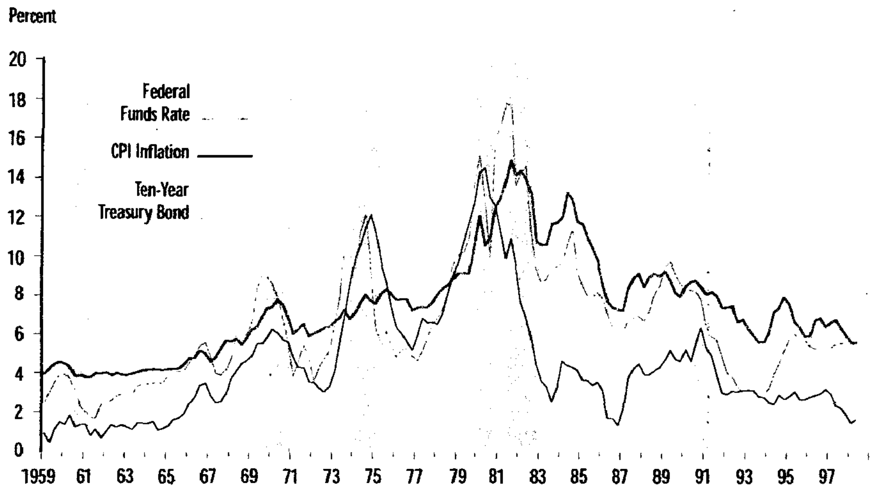


Figure 5 Inflation and selected interest rates.

cent during the late 1950s and early 1960s, 10-year Treasury bonds yielded about 4 percent. From the mid-1960s through the early 1980s, inflation trended up and so did both short- and long-term nominal interest rates. Since then, inflation has trended down and so have both short and long term nominal interest rates. Thus, the events of recent decades tend to confirm that high inflation is associated with high nominal interest rates and low inflation with low nominal interest rates. As a corollary, Federal Reserve policies that increase the growth of the monetary aggregates, and thereby inflation, would in due course also increase nominal interest rates, despite the myopic view that expansionary monetary policies lower nominal interest rates. Federal Reserve policies cannot lower nominal interest rates permanently except by actions that lower inflation.

INFLATION CREDIBILITY AND INTEREST RATES

Given the propensity to save, average real (inflation-adjusted) interest rates would tend to rise with an increase in trend real GDP growth. The reason is that measured real GDP growth is associated with increased real rates of return on investment. Average nominal interest rates tend to deviate from the real rate of return on investment by an amount that reflects expectations of inflation and inflation risks. The greater the gap between nominal interest rates and real rates of return, the lower the Fed's credibility is for keeping inflation low. Thus, the difference between nominal interest rates and trend real growth provides a crude measure of inflation expectations in the bond market, i.e., inflation credibility.[7]

Real GDP growth (averaged over 10 years to remove business cycle movements) drifted down from about 4 percent during the 1950s and early 1960s to about 2 percent during the late 1970s and early 1980s. It then rose back up to about 2.5 percent so far during the 1990s. Five-year average inflation drifted up from about 1 to 2 percent during the 1950s and early 1960s to nearly 10 percent in 1980, then back down to about 3 percent during the 1990s. The five-year average of the five-year Treasury security yield rose from 2 percent during the 1950s to 12 percent during the early 1980s; but it then fell back to about 6 percent in the 1990s. Since bond yields rose when inflation accelerated, but real GDP growth slowed, the influence of inflation outweighed the influence of real GDP growth on bond yields. Correspondingly, when inflation decelerated, bond yields fell even though real GDP remained stable.

The difference between the five-year average of the five-year Treasury security yield and the 10-year average of real GDP growth is an estimate of the bond market's five-year inflation forecast, adjusted for inflation risk. It is the height of the shaded area in the lower panel of Figure 6. This measure of inflation credibility roughly lagged inflation, indicating that bond yields have not been very forward looking in forecasting inflation. The measure of inflation credibility hovered close to zero during the 1950s and early 1960s, which was credibly a zero-inflation-expectations period. It under forecast inflation from the late 1960s until the early 1980s when inflation was rising. It over forecast inflation in the 1980s and so far in the 1990s as inflation has fallen. It peaked at about 10 percent in the early 1980s, but fell to about 4 percent in recent years.

The bond market inflation forecast (or inflation premium) over the past five years represents a substantial gain in credibility compared with the early 1980s, but a substantial loss compared with the 1950s and early 1960s. In that earlier period, actual inflation was about 2 percent, but the bond market forecast a rate close to zero. In recent years, inflation has averaged about 3 percent, but the bond market has forecast about 4 percent inflation inclusive of an inflation risk premium. Thus, despite recent inflation being the lowest and most stable in decades, bond markets have seemingly not yet been convinced that inflation is down to stay. If the inflation premium were eliminated, bond yields could fall to match trend real growth, as was the pattern in the 1950s and early 1960s. That is about 3 percent, which is considerably lower than the approximately 5 to 5 1/2 percent bond yields observed in mid 1998.

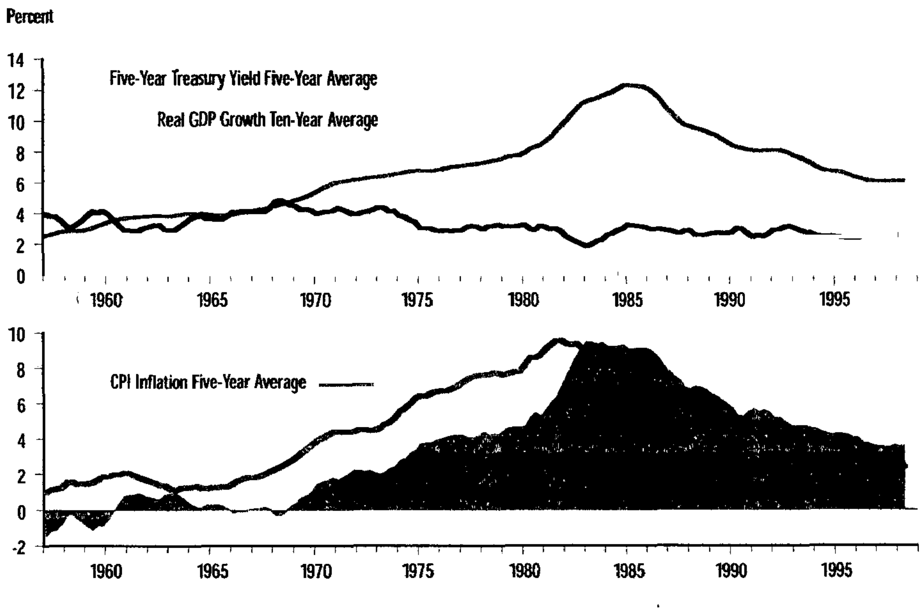


Figure 6 Long-term yields, real growth, and bond market inflation credibility.

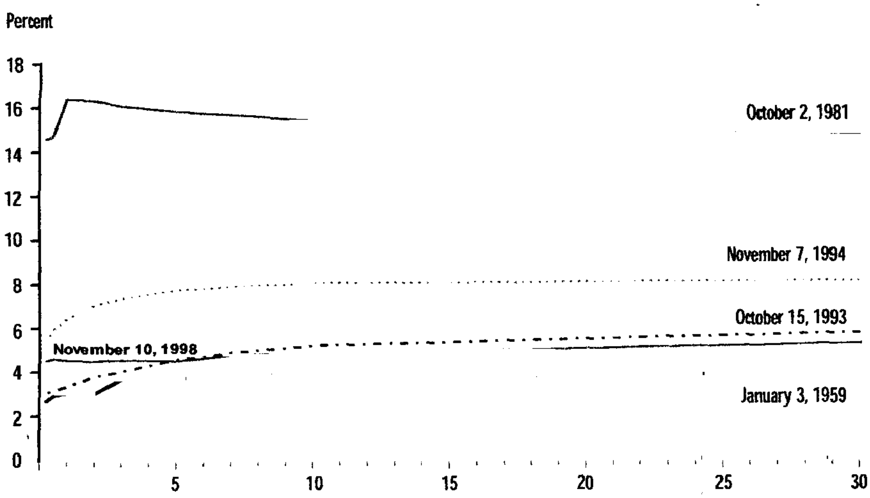


Figure 7 Government securities yield curve.

Double-digit inflation and inflationary expectations are what explain the all time peak in security yields in October 1981 as plotted in Figure 7. Since then, the entire yield curve has shifted down by roughly 10 percentage points, undoubtedly a reflection of the decline in inflation and inflationary expectations. Although markets do not expect double-digit inflation today, they do not expect price stability either. During the 1950s and early 1960s inflation was low and generally expected to stay low, a condition that was reflected in long-term rates hovering in the 3 to 4 percent range as represented by the January 3, 1959, yield curve on the figure. Despite the historical record of an unstable price level in the short run, there really was widespread expectation of longer-term price stability until inflation took off in the mid-1960s. In fact, never before the 1960s had the U.S. federal government borrowed long term at more than a $4\frac{1}{4}$ percent rate.

During the expansion that began in 1991, the yield curve touched a cyclical low on October 15, 1993. It then shifted up to a cyclical peak on November 7, 1994. Three-month bill rates had increased from 3 percent to 5.4 percent and 30-year bond rates, from 5.8 to 8.2 percent. The latter was presumably an illustration of increases in long-term interest rates indicative of rising inflationary expectations in the bond market. Although inflation, in fact, did not increase much during the 1990s expansion, bond markets may well have been anticipating a repeat of the experience of inflation accelerating as had typically occurred in the past. Historically, monetary policy often has lagged behind market interest rates in expansions and thereby added to, rather than damped, inflationary pressures. By comparison, the record during the 1990s expansion has been very good: An extended period of positive real growth with inflation held in check. Yet, with bond rates still above the real growth trend, the bond markets seemingly continue to reflect the fear that inflation will rise again.

HOW TO GET AND KEEP INFLATION CREDIBILITY

What could the Federal Reserve do to enhance its inflation credibility; and thereby allow long-term interest rates to stay low and prospectively fall further? Most important, the

Fed should continue to keep inflation low by limiting the rate of monetary growth. A practical goal would be to get back to the low inflation and low interest rates of the late 1950s and early 1960s. One way to persuade markets that low inflation is here to stay is for the FOMC to focus more sharply on the desired outcome for inflation by following several other countries that have legislated specific low inflation targets for their central banks. This list includes Australia, Canada, New Zealand, and the United Kingdom, as well as Portugal, Spain, and Sweden. Whether or not such efforts are directly responsible, the fact is that these countries have had considerable success in bringing inflation down and keeping it down.

A second proposal comes from economists who have argued that credibility would be enhanced if there were an announced policy rule (with respect to the federal funds rate or monetary growth) and the Fed acted on the basis of that rule. The advantage of a rule is that markets would know in advance how the Fed would react to deviations of nominal spending, inflation, or other variables from specified targets.

A third proposal made by Dewald (1988) is that federal budget offices base their budget projections over a 5- to 10-year horizon not on their own inflation assumptions, but on longer-term inflation forecasts from the Federal Reserve. Since the Fed has the power to influence inflation over the long term, why not relieve the budget offices of the responsibility for making an independent assessment of future inflation as they make their budget projections? Not only could the budget offices benefit, but also every business, state and local government, and household could benefit from having confidence that the Fed would act to keep inflation as low as it had forecast. Lars Svensson (1996) has proposed that the Fed make its own announced inflation forecasts an explicit policy target. By using a forecast as a guide to policy, the Fed would be focused on this objective, but not blind to other things going on in the economy that influence inflation.

An environment of credible price stability has a high payoff in a market economy. The historical evidence examined in this article supports the conclusion that risks of starting another costly inflation-disinflation cycle could be avoided by monitoring M2 monetary growth and maintaining a sufficiently tight monetary policy to keep a damper on inflation. Having achieved the lowest and most stable inflation environment in many decades, the Federal Reserve has an unusual opportunity to persuade markets that it will continue to keep inflation low and, in principle, eliminate it. An environment of credible price stability would allow the economy to function unfettered by inflationary distortions—which is all that can be reasonably expected of monetary policy, but precisely what should be expected of it.

NOTES

1. With chain-weights, price indexes and quantity indexes are calculated separately for components of GDP and therefore the difference between nominal GDP and real GDP growth is only approximately equal to the change in the GDP price index. With fixed-weights, the GDP deflator is defined as the ratio of nominal to real GDP and hence the gap between nominal and real GDP growth rates is precisely equal to the growth rate in the GDP deflator.
2. See Barro (1996) and Eijffinger, Schaling, and Hoeberichts (1998).
3. See Bruno and Easterly (1996).
4. See Friedman and Schwartz (1963).
5. Despite M2's imperfections as a cyclical indicator, the Conference Board's monthly Leading Indicators Index includes M2 relative to the price level as one of its 10 components.

6. The faultiness of interest rates as measures of monetary policy in a non-inflationary environment was evaluated in Dewald (1963).
7. The analysis rests on the assumption that the real rate of interest equals the rate of growth of real GDP, when both series are averaged over a moderately long period of time. This condition arises in theoretical models in which consumers are Ricardian and the rate of time preference is zero. The condition that the real interest rate is equal to the real output growth rate arises in theory since real GDP growth is acting as a proxy for an equilibrium rate of return on investment. It would be appropriately expressed in per capita terms. Per capita GDP growth has slowed more than overall GDP growth over the period plotted on Figure 6. Therefore, current inflation credibility would have fallen even more relative to its level in the late 1950s and early 1960s than indicated on Figure 6, if per capita real GDP growth had been used to proxy the equilibrium rate of return on investment.

REFERENCES

Barro, R. Inflation and Growth. Federal Reserve Bank of St. Louis *Review,* (May/June 1996), pp. 153–69.

Bruno, M., and Easterly, W. Inflation and Growth: In Search of a Stable Relationship. Federal Reserve Bank of St. Louis *Review* (May/June 1996), pp. 139–46.

Dewald, W. G. Monetarism is Dead: Long Live the Quantity Theory. Federal Reserve Bank of St. Louis *Review,* (July/August 1988), pp. 3–18.

Dewald, W. G. Free Reserves, Total Reserves, and Monetary Control. *The Journal of Political Economy,* LXXI, 1963, pp. 141–53.

Eijffinger, S., Schaling, E. and Hoeberichts, M. Central Bank Independence: A Sensitivity Analysis. *European Journal of Political Economy,* Vol 14 (1998), pp. 73–88.

Friedman, M., and Schwartz, A. J. *A Monetary History of the United States,* 1867–1960, Princeton University Press, 1963.

Svensson, L. Commentary: How Should Monetary Policy Respond to Shocks While Maintaining Long-Run Price Stability?—Conceptual Issues. *Achieving Price Stability, A Symposium Sponsored By the Federal Reserve Bank of Kansas City,* August 29–31, 1996, pp. 209–27.

59

The Century of Markets

Jerry L. Jordan
President and Chief Executive Officer, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Governments have long pursued policies that determined the degree to which markets have been permitted to operate. But with the rise of global capital markets, we have learned that the opposite is also true—markets can affect national economic policies.

Business people know very well that market forces do not treat kindly companies that fail to satisfy their customers. Politicians also are now learning that global capital markets treat harshly governments whose policies fail to enhance the living standards of their people. Good business practices and good government policies both are essential to sustained prosperity. But there is an important division of labor. Private firms best enhance public welfare by producing goods and services at the lowest possible prices; governments contribute to the common good by establishing well-functioning institutions within which the society operates. Good business practices cannot effectively take root without good government policies.

During the twentieth century there was a massive increase in the intrusion of governments into economic affairs, but it is becoming increasingly clear that this wave has crested; the role of the state in economic affairs has begun to diminish. In the new century and millennium, a growing share of the world will enjoy the prosperity that comes from the "century of markets."

GOVERNMENT PRESENCE IN THE ECONOMY

Just over 70 years ago, in the autumn of 1929, equity markets around the world entered a period of steep decline—so much so that the label "crash" is often used to describe the events of 1929–30. Those developments and the ensuing policies brought about worldwide economic depression. Indeed, it is now well accepted that the 1930s were a "watershed decade" in which economic depression gave rise to public support for the nationalization

Reprinted from: *Federal Reserve Bank of Cleveland Economic Commentary,* Feb. 15, 2000. This *Economic Commentary* is adapted from the John Bonython Lecture given by Jerry L. Jordan, president and CEO of the Federal Reserve Bank of Cleveland, at the Centre for Independent Studies in Sydney, Australia, in November 1999.

of entire industries, and what remained privately owned was subject to pervasive governmental regulation. For several subsequent decades, decisions about what to produce, who could produce it, where to produce it, what prices to charge, what wages to pay, and many other economic decisions about interest rates, exchange rates, and even profitability were either made by government agencies or were subject to their approval. Remnants of many of those policies haunt us still.

I suggest the 1980s were another watershed decade, marking the beginning of the state's withdrawal from economic affairs, and argue that recent trends to strengthen property rights and enhance the economic infrastructure of market economies on a global basis will endure for several decades into the future. The financial "crises" of our time largely reflect the breaking up of the old order. Moreover, the vestiges of ill-conceived government involvement in economic affairs will be under continuous attack. Social and political disturbances can be expected—the more highly industrialized countries are not immune—as the relentless pressures of global capital markets confront legacy government programs and agencies. The drive toward greater economic efficiency is an irresistible force, and governmental policies are not, in the end, immovable objects.

MARKET FORCES AT WORK

From a historical perspective, the age of capitalism is now at most a teenager, and it is already evident that the power of unfettered markets to generate wealth is building momentum. Capitalism requires mobility of resources—goods, labor, and capital—so they may find their highest valued use. But resource mobility is an idea that is more often than not resisted by governments, whether democratic or authoritarian. Governments around the globe have used a variety of methods—with varying degrees of success—to restrict either the entry or the exit of people, goods, and capital. The collapse of the Berlin Wall just 10 years ago serves as a very visible symbol of the ultimate futility of erecting artificial barriers to at least one type of mobility.

Less visible, yet more pervasive, are the countless barriers to the mobility of financial capital. These, too, have been tumbling down in recent years. The process is still in the early stages, and we have no blueprints for constructing market mechanisms to replace ossified government mechanisms. Nevertheless, just as the global political environment has changed dramatically in the decade since the Wall crumbled, the global economic environment has started to move rapidly away from government-controlled markets.

THE SEARCH FOR BEST PRACTICES

Interestingly, the idea of irresistible market forces meeting seemingly immovable objects is commonplace in the world of business. Innovations continuously bombard the economy, forcing changes in how and with whom we interact. Business leaders are used to the idea that there is a continuous, never-ending search for best practices that can better accommodate new production processes or even produce different goods as consumers' tastes change in unpredictable ways. This is unavoidable because failure to recognize and incorporate superior management processes proves fatal in the marketplace. People in business know that it is not simply the quality and price of the product that must compete at a point in time, but entire business systems. These systems must compete in getting new products to the market and then getting them to the customer—when the customer wants them, how the customer wants them, and where the customer wants them.

Workers are subjected to the same forces, as the demands for what they can do and how they do it change as business changes its way of doing things. In response to the innovations bombarding businesses, the labor market undergoes substantial churning, leading to simultaneous job creation and job destruction. Workers must learn new skills and methods to deliver their services to employers, just as business must learn new processes to deliver its product to consumers. Uncertain and unforeseeable events affect both workers and businesses. There is no escape. Economic prosperity depends on the ability to recognize and react to those forces, whether for an individual in the labor market, a firm in the business sector, or a government in today's global economy.

Current management literature asserts the existence of business maxims or "first principles" essential to success. In economics there are also maxims or first principles. One is universally used by economists to argue for the elimination of barriers to the mobility of goods. That principle—comparative advantage—holds that welfare is maximized when unfettered market forces determine where the opportunity cost of producing a good is lowest.

As trade barriers continue to erode and the principle of comparative advantage becomes universally operative, people are becoming accustomed to the idea of consuming goods produced elsewhere in the world. More recently, they have become used to the idea that various services, such as transportation, communications, and banking, may also be best provided by firms headquartered elsewhere on the globe. These trends, of course, reflect the dramatic changes in information and communications technologies that have brought ever-lower costs of comparing products and services over larger regions.

BEST PRACTICES AND THE INFORMATION REVOLUTION

We all marvel at the new products and services that come from technological innovations. But it certainly is also true that the information technology revolution has accelerated the rate of obsolescence of old ideas and old ways of doing things. The well-known phrase of the Austrian economist Joseph Schumpeter about creative destruction is something that people in business live with every day; new products and new services render obsolete—or at least reduce the economic value of—old ideas, products, services, and ways of doing things.

The half-life of knowledge is getting shorter all the time. What one knows today becomes out of date faster than ever before. The inverse is that new knowledge must be acquired and incorporated much more quickly than before in order to stay in the same relative position. Political organizations and institutions must also change at an ever-faster pace.

There was a time in the not-too-distant past when people in commerce needed to look only at competitors within their national borders—especially in very large countries like the United States. In smaller, more open economies, business people learned early on that best practices were often found in other countries and that failure to respond quickly to them produced a possibly fatal competetive threat.

For a while, the expression "multinational company" described one that operated internationally. It was a holding company, located in one place, that owned and operated businesses located in various other places around the world. However, in the early versions, there was not much more to it than ownership, since management techniques, labor market practices, input-factor sourcing, product distribution systems, and so on all remained local and distinct from place to place. Over time, though, the spread of best practices resulted in global companies succeeding over multinational companies. In other words, businesses found that what works best in one place works best in every place. The idea of local content or place of national origin became a political obstacle or burden to be overcome but not a desirable management practice.

Ultimately, it seemed to be simply untrue that there were best ways of doing things in Asia and quite different, but still best, ways of doing things in Europe, Latin America, or North America, all of which were different from each other. Instead, best practices meant simply that—it was best, with little regard for local social, cultural, or political settings.

GOVERNMENTS AND BEST PRACTICES

This trend toward borderless commerce means that local political institutions are coming under increased scrutiny as well, and the reforms we are witnessing can be thought of as the sometimes grudging adoption of best practices. For most of history, the evolution of institutional arrangements in the political sphere progressed very slowly. Certain democratic institutions have migrated around the world for hundreds of years since the signing of the Magna Carta, but even in the twentieth century most of the world was not living under what would be considered the best practices of political and economic infrastructure.

There are, of course, many local, institutional, and political reasons for the slow adoption of superior political institutions, but the persistent forces arising from capital markets have meant that reform processes accelerate, forcing many of the old structures to crumble in their path. As informational barriers fall—and indeed we have witnessed substantial declines in the cost of acquiring information—it becomes easier to identify and compare different institutional arrangements, including tax policies, regulations, guarantees, subsidies, and so on. This more intense international comparison is the additional force giving rise to institutional reforms. As the costs of acquiring information decline, it becomes more difficult to sustain bad practices. This includes more than just monetary and fiscal policies. The costs of engaging in corrupt behavior—as well as pursuing ineffective economic policies—have risen dramatically. In small villages, it has long been the case that "outlier behavior" was subject to discipline. Instant global communications extend the "village effect" into previously isolated places. Inappropriate behavior of both government ministers and business executives now results in "early retirement," and maybe disgrace, more swiftly than ever before.

Even local judicial systems are not immune. If a country does not have a well-functioning legal system in place that protects property rights, businesses must offer a higher rate of return in order to attract or hold capital in the country. This increases the cost of capital, resulting in lower rates of investment, which will affect profits and the pace of real growth. That means fewer consumption goods and lower income per capita.

As it becomes easier for people to recognize where and how resources will earn their highest return, the half-life of bad government policies will become ever shorter. That is to say, global capital markets can have a major say in determining how long before a poorly performing government is forced to reform or is turned out of office.

Countries whose futures look bleak due to bad policies, such as massive unfunded pension liabilities, double- or even triple-digit inflation, lack of well-defined property rights, and so on, will not attract or keep the resources necessary to foster significant increases in their standard of living. They are destined to fall farther and farther behind in terms of per capita wealth, until the pressures for reform become overwhelming.

CRISES AND THE NEW ORDER

In news reports, it is common to see people lament the apparent increased frequency of crises, especially in financial markets over the last decade. To repeat a point I touched upon earlier, a different interpretation of the phenomenon we are witnessing is that a crisis is a breaking down of an old order and the creation of a new one. The evolving order is conducive to the rapid adoption of new processes and institutional arrangements that are superior to those they are replacing.

In a world with highly mobile resources, the lessons learned in a crisis invariably lead to changes in behavior that prevent a repeat of the conditions that led to the crisis. Once a crisis atmosphere has subsided, we rarely see reinstitution of the practices and arrangements that created the crisis situation.

This interpretation of what we are observing would suggest that the frequency of financial crises is evidence that the pace of adoption of new and better ways of doing things has accelerated. Borrowing from Schumpeter, just as there is a creative destruction in goods and services as new and better products come onto the market, so too in political and economic matters, the replacement of obsolete arrangements with more effective practices is a wrenching process.

It is essential to understand that, in a partial sense, wealth creation simultaneously involves wealth destruction. The true meaning of the expression creative destruction is that when something new and better comes along, the old—whether goods, services, or distribution methods—loses value. That means its economic or market value declines. When an upstart firm—for example, retail- distributor.com—comes along and gets the product to the consumer in a less costly, more timely way, then old methods of distribution are less valuable, and firms engaged in the old methods lose market capitalization.

The same is true of ideas and political and economic institutional arrangements. When new and better methods compete head-on with less effective existing methods, the old institutions must evolve, or they will perish. Foreign trade will be severely hampered in countries whose courts will not enforce the contracts and protect the property of their own citizens. Banks that engage in unsound local lending practices cannot sustain the risk-adjusted rate of return sought by foreign investors—unless government guarantees are involved. Governments with unsustainable fiscal policies, such as promising overly generous pensions to citizens, will find it increasingly difficult—or impossible—to raise taxes sufficiently or issue enough new debt to meet their commitments. The discipline exerted by global financial markets is beneficial in that it erodes local resistance to more efficient domestic markets.

BRAND-NAME CAPITAL

The erosion of barriers to trade in goods and services offers clues to what we can expect in monetary affairs. Today, brand-name recognition and identification are more important than ever. When a company like Sony produces a new product—a CD player—that is better and less costly than other brands, consumers will want to buy it. Consumers everywhere are the same—they want the best product for the lowest price! Only barriers to trade will prevent a superior product from gaining global market share.

Such "brand-name" identification of goods—which has made a product's national origin irrelevant to consumers—is becoming evident in financial and monetary affairs. Lack of global specialization in the production of goods was due to governmental and technological constraints. International brand identification evolved as these constraints diminished. As we are now seeing in the monetary arena, brand identification of standards of value—money—also becomes more pervasive as falling costs of information and communications technologies make it increasingly easy to compare the quality dimension of standards of value.

While there are vested interests in maintaining local governmental monopolies over issung of the national media of exchange, history demonstrates that national currencies inevitably must compete in the international financial arena. Countries whose monetary policies have resulted in large fluctuations in the value of the currency have come under pressure to adopt a system that prevents such behavior. This is just the "brand-name" argument—people want the best product or service. Currency boards and "dollarization" are two outcomes forced on many governments by their inability to assure stable purchasing power for the domestic currency. The "brand-name" of currency used to denominate contracts and trade assets is more important than the "local content" or "national origin" of the standard of value.

It seems natural to extend such arguments to forms of government. There are a number of different models of government, just as there are many models of successful business operation. And, as best practices in governing evolve, countries that do not adopt such practices will lose "capitalization"; that is, they will fail to attract and hold a share of the world's investment capital, and the process will culminate in much lower standards of living.

The expression, "vote with their feet" is still relevant for many less developed places on earth. Oppressed and impoverished people still flee bad governments in search of an opportunity for prosperity. That long-time tradition is now supplemented by the powerful forces of capital markets.

The crumbling of the barriers that have corralled the movement of goods, labor, and capital tells us that the role of government in economic affairs continues to ebb. An economic infrastructure that best encourages entrepreneurship and wealth creation is becoming more commonplace. Integral to these changes is that fiscal and monetary policies are also becoming less activist and more predictable.

In the final analysis, sustainable long-term prosperity, whether at the global or the local level, occurs when human action is focused on converting productive resources into marketable goods. It is no longer useful to think of the government's relationship to its citizens as that of an architect, engineer, carpenter, or to use any other metaphor implying activism. Instead, the role of the state is to nurture an economic garden—cultivating the soil to allow growth to take root, warding off pests that seek to feed off the budding crop, and keeping weeds from suffocating the plant before it achieves its potential.

I predict that in the twenty-first century—the century of markets—globalization and technology will force governments to establish the infrastructure that their economies need to thrive.

60
European System of Central Banks

Mark A. Wynne
Senior Economist and Policy Advisor, Federal Reserve Bank of Dallas, Dallas, Texas

In the early hours of May 3, 1998, the leaders of the European Union (EU) took the most significant step toward European integration since the signing of the Treaty of Rome in 1957. In giving the green light to the economic and monetary union (EMU) of eleven of the fifteen EU members—the EU11—they took another major step toward a more unified Europe.[1] This latest development culminates a process of European integration that began shortly after World War II and that may one day lead to a Europe as politically and economically integrated as the United States is today. The substance of monetary union is that the countries in EMU no longer have distinct national currencies. A new currency—the euro—has replaced them, and monetary policy for the EU11 is no longer determined by their national central banks but by the European System of Central Banks (ESCB). (See Glossary.)

The unprecedented monetary union of such a large and disparate group of sovereign nations will pose enormous challenges to the ESCB, which consists of the recently established European Central Bank (ECB) and the national central banks (NCBs) of EU members. The ECB commenced operations on June 1, 1998, and assumed responsibility for the conduct of monetary policy for the euro area on January 1, 1999. The euro has replaced the national currencies of the EU11, and in 2002 the notes and coins that currently circulate in these countries will cease to be legal tender.[2]

The ESCB is conducting monetary policy on a continental scale. Table 1 presents comparative statistics for the United States, the EU, the EU11, and Japan. In terms of population and aggregate output, the EU11 is comparable to the United States. Should EMU eventually incorporate all fifteen members of the EU, its economic weight would significantly exceed that of the United States. Table 1 also compares the recent economic performance of the four groups. The most significant difference is the much poorer employment performance of the EU, whose unemployment rate is more than twice that of the United States. The consensus among economists is that the bulk of this unemployment is structural rather than cyclical and reflects the greater rigidity of Europe's labor market institutions.[3]

The extent to which the euro can credibly challenge the U.S. dollar's primacy in global finance will depend largely on the ECB's success in maintaining the euro's purchasing power and making it attractive to international investors. The structure of the

Reprinted from: *Federal Reserve Bank of Dallas Economic Review* (First Quarter 1999) 2–14.

Table 1 Comparative Statistics on the U.S., EU, EU11, and Japan

	U.S.	EU	EU11	Japan
Population	263 million	373 million	290 million	126 million
GDP (current dollars)	$6.955 trillion	$8.497 trillion	$6.809 trillion	$5.217 trillion
GDP (constant dollars)	$6.149 trillion	$7.203 trillion	$5.721 trillion	$3.168 trillion
GDP growth	3.8 percent	2.6 percent	2.4 percent	.9 percent
Inflation rate	2.3 percent	1.9 percent	1.7 percent	1.7 percent
Unemployment rate	4.9 percent	11.1 percent	11.7 percent	3.4 percent

Sources: Population and GDP for 1995: United Nations (1997). GDP growth, inflation, and unemployment for 1997: *OECD Economic Outlook* and *OECD Main Economic Indicators.*

ESCB is similar in many ways to that of another central bank charged with conducting monetary policy on a continental scale—the Federal Reserve System. In this article, I review the structure of the new central bank, sketching out key similarities to and differences from the Federal Reserve.

THE ROAD TO EMU

EMU is the latest step in the move toward greater economic (and political) integration in Europe that began with the establishment of the European Payments Union in 1950. That entity was little more than a technical device to facilitate the reconstruction of Europe following the devastation of World War II. But it can also be seen as the first manifestation of the political will to forge closer bonds between the wartime belligerents so as to preclude future conflict. A more substantive step was taken in 1951 with the formation of the European Coal and Steel Community, which created a common market for these commodities involving Germany, France, Italy, and the Benelux countries (Belgium, Netherlands, and Luxembourg). This entity was supposed to be accompanied by stronger political and military ties (including the creation of a European army), but concerns about loss of national sovereignty led to abandonment of these plans. Instead, the emphasis shifted to greater integration on the economic front, and in 1957 the Treaty of Rome created the European Economic Community (EEC), or Common Market. Coordination of economic policies was always seen as integral to the success of the Common Market, and in 1964 the Committee of Governors of the central banks of the European Community (EC) was formed to coordinate monetary policies. The central banks of Europe have had varying degrees of success coordinating their monetary policies over the past three decades.[4]

Monetary union of EC members was proposed in 1970 in the Werner Report. While this report envisioned a monetary union by 1980, two key international developments derailed the plan. The first was the breakdown of the Bretton Woods system of fixed exchange rates in August 1971; the second was the 1973 oil crisis. The EC responded to the exchange rate turbulence that followed both events with a system of quasi-fixed exchange rates, the so-called snake, but this rapidly collapsed to an arrangement involving only a few members. The second attempt to fix exchange rates, the European Monetary System, was established in 1979. It proved more durable, although it, too, experienced a number of major and minor crises. By the mid-1980s the EC had expanded to twelve members, and in 1989, renewed interest in a formal monetary union resulted in the Delors Report.

The Delors Report laid out the basic plan and timetable for monetary union that has been followed since the early 1990s. The proposals in the report were incorporated into the Treaty on European Union, which was agreed upon at a meeting of the European Council in Maastricht, Netherlands, in December 1991 and signed in February 1992. This agreement, commonly known as the Maastricht Treaty, was the most comprehensive change in the basic law of the European Community since the Treaty of Rome. It established the institutional framework for monetary policy under EMU, a timetable for the creation of a monetary union, and the criteria for countries' participation. Many academic economists and others have questioned the wisdom of a monetary union between such disparate countries, but the debate became moot with the decision to proceed.[5] However, the points made by the critics of monetary union indicate where stresses may arise in the future and the kind of challenges the ECB may face.

INSTITUTIONAL STRUCTURES

The Maastricht Treaty established the institutional arrangements for the conduct of monetary policy under EMU. The treaty provides for the formation of the ESCB, which in many ways is indirectly modeled on the Federal Reserve System.[6] At the top of the ESCB is the Frankfurt-based ECB, which has a role similar to that of the Federal Reserve's Board of Governors. The various national central banks play a role similar to that of the regional Federal Reserve Banks. Monetary policy decisions are made by the Governing Council of the ECB, which consists of the Executive Board of the ECB and the governors of the participating countries' central banks. The Executive Board consists of the president and vice president of the ECB and four other members. The president of the ECB chairs the Governing Council, in essence occupying a position similar to that of the chairman of the Fed's Board of Governors. Under the Maastricht Treaty, the Governing Council is responsible for formulating monetary policy for the single-currency area, while the Executive Board is responsible for implementing monetary policy.[7] Executive Board members are appointed for nonrenewable eight-year terms, shorter than the fourteen-year terms of Federal Reserve Governors but the same as the terms of members of the Directorate of Deutsche Bundesbank.

RELATIONSHIP OF THE ECB AND NCBS

While there are many similarities in the structures of the ESCB and the Federal Reserve System, there are also important differences. The Executive Board of the ECB will be in a permanent minority on the Governing Council, whereas the Board of Governors has a permanent majority on the Federal Open Market Committee (FOMC). All NCB governors have a vote in all policy decisions of the Governing Council, whereas with a single exception, Reserve Bank presidents participate in FOMC votes every two or three years, depending on which Bank they represent. Indeed, the relationship of the ECB and the NCBs probably bears a closer resemblance to the relationship of the Board of Governors and the Reserve Banks in the early years of the Federal Reserve System than to the situation today. For its first twenty years, power was more diffuse in the Federal Reserve System than it is now. Some critics have argued that this diffuse distribution of power and the struggle for hegemony contributed to the Fed's inability to deal with the Great Depression (see in particular Friedman and Schwartz 1963).[8]

Glossary

European Economic Community (EEC): Also known as the Common Market, established by the Treaty of Rome in 1957.

European Monetary Institute (EMI): Created by the Maastricht Treaty to carry out preparatory work for EMU; dissolved with the establishment of the European Central Bank.

European Parliament: Advises the European Commission and reviews all legislative proposals; members are elected by popular vote.

European System of Central Banks (ESCB): Responsible for conducting monetary policy for the economic and monetary union (EMU). The ESCB consists of the European Central Bank and the national central banks of all fifteen EU members. The ESCB is governed by the decision-making bodies of the ECB.

- *European Central Bank (ECB):* The central bank for the economic and monetary union. The decision-making bodies of the ECB are the Governing Council and the Executive Board.
    - *Executive Board:* Responsible for the day-to-day functioning of the ECB and the implementation of the single monetary policy.
    - *Governing Council:* Consists of the Executive Board of the ECB and the governors of the national central banks of the states participating in EMU. Responsible for the formulation of a single monetary policy.
- *National central bank (NCB):* The individual central banks of countries in the European Union.

European Community (EC): Consists of the European Coal and Steel Community, the European Atomic Energy Community, and the European Economic Community. The EC became the European Union when the Maastricht Treaty took effect on November 1, 1993.

Friedman and Schwartz (1963) contend that the distribution of power in the Federal Reserve System was a key contributor to the "ineptness" of monetary policy during the Depression. In the 1920s, the institutional structure did not present a problem as long as all regional Reserve Banks and the Board were willing to accept the leadership of the governor of the New York Bank, Benjamin Strong. But with Strong's departure in 1928, the structure became unworkable. The other Reserve Banks were no longer willing to accept the domination of the New York Bank, and the Board was not in a position to impose its will on the System. Friedman and Schwartz argue that the Board's weak position was due to the fact that it had not played a leadership role in the System in the 1920s but had instead functioned primarily as a supervisory and review body.

The distribution of power in the ESCB differs from that in the Federal Reserve System in other important respects as well. For example, the Board of Governors exercises a

European Union (EU): Established by the Maastricht Treaty to deepen economic and political links between the countries of Europe.

Council of Ministers: The primary decision-making institution of the European Union; consists of ministerial-level representatives of all EU states.

European Council: The name given to the Council of Ministers when it meets in the form of EU heads of state or government.

ECOFIN: The name given to the Council of Ministers when it meets in the form of EU economics and finance ministers.

European Commission: The executive branch of the European Union; responsible for implementing the decisions of the Council of Ministers and proposing new measures and directions for the EU.

Maastricht Treaty: More formally, the Treaty on European Union; signed in 1992 by the EU heads of state, it established the framework for economic and monetary union in Europe.

Statute of the European System of Central Banks and of the European Central Bank: The statute, appended to the Maastricht Treaty, detailing the structures and mandates of the ESCB and the ECB.

lot more power in the Federal Reserve System than the ECB exercises within the ESCB. One source of the Board's power is its authority to supervise the Reserve Banks' activities and approve their budgets and the appointment of their presidents. Furthermore, the Board of Governors appoints three of the nine directors of the regional Reserve Banks, one of whom is designated chairman of the board of directors and Federal Reserve agent.[9] The Board also appoints the deputy chairman of the board of each regional Bank.

By contrast, the Maastricht Treaty gives the Governing Council control over the Executive Board:

> The terms and conditions of employment of the members of the Executive Board, in particular their salaries, pensions and other social security benefits shall be the subject of contracts with the ECB and shall be fixed by the Governing Council on a proposal from a Committee comprising three members appointed by the Governing Council and three members appointed by the Council. The members of the Executive Board shall not have the right to vote on matters referred to in this paragraph. (Statute of the European System of Central Banks and of the European Central Bank, Article 11)

Also, the principle of subsidiary (whereby EU decisions are supposed to be made at the lowest possible level of political authority) may pose a fundamental obstacle to centralization of power with the ECB. The importance of this principle in EU decision making should not be underestimated; it is even articulated in the preamble to the Maastricht Treaty. The national central banks could use subsidiarity as a weapon to prevent the ECB from developing expertise in areas the NCBs feel are properly their province.

The distribution of power within the ECB differs slightly from that in the Board of Governors. It is generally accepted that the chairman of the Board of Governors is more powerful than any of the other Board members.[10] Maisel (1973) attributes the power of the chairman to a number of factors. The first is his role as titular head of the Federal Reserve System and his role as its spokesman; only the chairman speaks for the System as a whole. The second is the role of the chairman as the representative of the System in other forums. The third is the inherent power of the chairman to set the agenda for FOMC meetings. Fourth, the Board of Governors' delegation of much of its supervisory power over the staff and the System to the chairman enhances the position's authority within the System. And finally, the chairman has the ability to attract votes simply by virtue of office.

The president of the ECB also possesses powers beyond those of other Executive Board or Governing Council members. The president chairs meetings of both bodies and casts the deciding vote in the event of a tie.[11] He also represents the ECB externally.[12] Under Article 109b of the treaty, the president of the ECB may be invited to participate in Council meetings and presents the ECB's annual report to the Council and the European Parliament. Perhaps the only additional source of power potentially but not currently available to the president is full control over the ECB's staff. Each of the six members of the Executive Board oversees some areas of the ECB's operations.[13] The Economics and Research Directorates, which employ the bulk of the ECB's professional economists, do not report to the president but to another board member.

APPOINTMENT PROCESS

All seven members of the Fed's Board of Governors are appointed by the president of the United States and are subject to Senate confirmation. The Federal Reserve Act requires that "in selecting the members of the Board, not more than one of whom shall be selected from any one Federal Reserve district, the president shall have due regard to a fair representation of the financial, agricultural, industrial, and commercial interests, and geographical divisions of the country."[14] Regional Reserve Bank presidents are nominated by the boards of directors of those banks, but their final appointment is subject to approval by the Board of Governors.

The ECB Executive Board is appointed by the European Council, with nominees subject to confirmation by the European Parliament. The Maastricht Treaty does not contain any provisions about the national composition of the Executive Board analogous to those in the Federal Reserve Act. However, the reality of EU politics is such that it would be unthinkable for there to be more than one national of any EU country on the board. While the European Parliament confirms nominees to the Executive Board, in reality Parliament has little real power to reject a nominee. The governors of the NCBs are appointed by their national governments and are not subject to approval by the ECB's Executive Board.

MONETARY POLICY OBJECTIVES

The Maastricht Treaty is unambiguous about the objective of monetary policy:

> The primary objective of the ESCB shall be to maintain price stability. Without prejudice to the objective of price stability, the ESCB shall support the general economic policies in the Community with a view to the achievement of the objectives of the Com-

> munity as laid down in Article 2. The ESCB shall act in accordance with the principle of an open market economy with free competition, favoring an efficient allocation of resources, and in compliance with the principles set out in Article 3a.[15] (Maastricht Treaty, Article 105)

This mandate is qualified by an obligation to "support the general economic policies in the Community," but this support should be "without prejudice to the objective of price stability."[16] Treaty provisions dealing with the objectives of the ESCB are modeled on those in the Bundesbank Act but, interestingly, are a lot more specific about the ultimate objectives of monetary policy than is the German legislation. That act requires that "the Deutsche Bundesbank shall regulate the amount of money in circulation and of credit supplied to the economy—with the aim of safeguarding the currency" (Deutsche Bundesbank 1995, 23). Arguably, the act could be seen as giving the Bundesbank the freedom to choose between stabilizing the internal value of the currency—that is, the price level—or the external value of the currency, as reflected by the exchange rate.

The ESCB's mandate to pursue price stability contrasts with the Federal Reserve's more ambiguous mandate:

> The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long run growth of the monetary and credit aggregates commensurate with the country's long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices and moderate long-term interest rates. (Federal Reserve Act, Section 2A.1)

Maisel (1973, 66) contends that the Fed has traditionally placed more emphasis on achieving price stability than on its other mandated objectives. Many economists believe that price stability is a precondition for the objectives of sustainable growth and high employment.

In recent years there have been calls for a clearer price stability mandate for the Federal Reserve System. For example, Hetzel (1990) has supported a mandate that stipulates price stability as the Fed's primary goal. Hetzel favored the Neal Resolution (House Joint Resolution 409), introduced in September 1989, which would have required that "the Federal Open Market Committee of the Federal Reserve System . . . adopt and pursue monetary policies to reduce inflation gradually in order to eliminate inflation by not later than 5 years from the date of the enactment of this legislation and shall then adopt and pursue monetary policies to maintain price stability." More recently, the Mack-Saxton bill, introduced in 1995 and reintroduced in 1997, would have made long-term price stability the primary goal of the Federal Reserve System. While several Reserve Bank presidents and the chairman of the Board of Governors have testified before Congress in support of legislation to mandate price stability as the Fed's sole objective, this legislation has not gotten very far. The reasons for this are unclear: it may simply be that inflation is not currently perceived to be a major problem in the United States and that in the current low-inflation environment it would be undesirable for the Fed to de-emphasize output stabilization in its policy decisions.

INDEPENDENCE

The ESCB is probably the most independent central bank in the world.[17] The source of this independence is manifold. At the most basic level, the fact that the charter of the ESCB is an international treaty that can only be changed with the unanimous consent of its signatories makes it very difficult to exert political pressure on the ESCB. Furthermore, the Maastricht Treaty explicitly addresses the relationship between the ESCB and the political authorities in the EU:

> When exercising the powers and carrying out the tasks and duties conferred upon them by this Treaty and the Statute of the ESCB, neither the ECB, nor a national central bank, nor any member of their decision-making bodies shall seek or take instructions from Community institutions or bodies, from any government of a Member State or from any other body. The Community institutions and bodies and the governments of the Member States undertake to respect this principle and not to seek to influence the members of the decision-making bodies of the ECB or of the national central banks in the performance of their tasks. (Maastricht Treaty, Article 107)

The reason for granting such strong independence to the ESCB is the overwhelming evidence that independent central banks tend to deliver relatively better inflation performance (that is, lower rates of inflation) at no cost in terms of slower real output growth or higher unemployment. Banaian, Laney, and Willett (1983) were among the first to examine the relationship between central bank independence and inflation outcomes. Subsequent work by Alesina and Summers (1993) shows that the better inflation performance delivered by independent central banks comes at no cost in terms of real economic performance. Numerous other studies (see, for example, Cukierman, Webb, and Neyapti 1992 and Eijffinger and De Haan 1996) confirm these findings.

Other provisions of the Maastricht Treaty further reinforce the independence of the ESCB. First, Executive Board members and governors of the NCBs are appointed for relatively long terms. Executive Board members have nonrenewable eight-year terms, while NCB governors are appointed for a minimum of five years.[18] The terms of the first appointees to the Executive Board were staggered from four to eight years so that subsequent terms will also be staggered. Second, the treaty states:

> Overdraft facilities or any other type of credit facility with the ECB or with the central banks of the Member States (hereinafter referred to as "national central banks") in favor of Community institutions or bodies, central governments, regional, local or other public authorities, other bodies governed by public law, or public undertakings of Member States shall be prohibited, as shall the purchase directly from them by the ECB or national central banks of debt instruments. (Maastricht Treaty, Article 104)

This prohibition is restated in Article 21 of the statute of the ESCB and ECB.

It is worth noting that some authors have recently challenged the causal interpretation of the relationship between central bank independence and inflation outcomes. Specifically, Posen (1993) has argued it is popular opposition to inflation that leads to independent central banks and low inflation outcomes. The corollary is that simply granting independence to a central bank is insufficient to generate good inflation performance unless the central bank has significant political support. Attaining such support is one of the greatest challenges facing the ESCB. There is little doubt the Bundesbank's success in pursuing price stability has been helped considerably by strong public support for the central bank's policies. The ESCB—at least initially—will not enjoy anything like the same degree of support, and this may complicate the political environment in which it has to operate.[19]

Central bank independence takes many forms. Fischer (1994) distinguishes between goal independence and instrument independence.[20] In his taxonomy, "a central bank whose goals are imprecisely defined has goal independence" (Fischer 1994, 292). Since the Maastricht Treaty makes price stability the primary goal of the ESCB without defining what is meant by price stability, the ESCB enjoys considerable goal independence. Thus, the ESCB could define price stability to mean a stable price level, or a specific (low) inflation rate, or as prevailing when "inflation ceases to be a factor in the day to day decisions of households and businesses." The European Monetary Institute (EMI) argued that a public announcement of a quantified definition of price stability should be an integral component of whatever monetary strategy the ECB pursues. In October 1998 the ECB announced that "price stability shall be defined as a year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2%."[21] The ECB also announced that "price stability is to be maintained over the medium term," without defining medium term. Fischer considers a central bank to have instrument independence "when it has full discretion and power to deploy monetary policy to attain its goals." By that definition, the ESCB enjoys full instrument independence.

Under Article 109b of the Maastricht Treaty, the president of the Council of Ministers and a member of the European Commission (the executive branch of the EU) have the right to participate as nonvoting members in meetings of the ECB's Governing Council. Furthermore, the president of the Council of Ministers has the right to submit motions for deliberation by the Governing Council. Article 109b also stipulates that "the President of the ECB shall be invited to participate in Council meetings when the Council is discussing matters relating to the objectives and tasks of the ESCB."

This arrangement echoes somewhat the early structure of the Federal Reserve System. Until 1935, the secretary of the Treasury and the comptroller of the currency were ex officio members of the eight-member Federal Reserve Board, with the secretary of the Treasury also acting as Board chairman. The Banking Act of 1935 removed both officials from the Board. As noted above, the motivation for centralizing power at the Board of Governors was to eliminate perceived ambiguities about the distribution of power in the System that were believed to have contributed to the failure of Fed policy during the Depression.[22] While the president of the Council of Ministers does not have the same degree of formal power to influence ECB deliberations as the Treasury secretary had over the Federal Reserve System in its early years, his presence at Governing Council meetings may influence the course of debate in ways that are difficult to anticipate. The ECB arrangement more closely mirrors the provision in Article 13.2 of the Bundesbank Act, whereby representatives of the government have the right to attend (without voting) meetings of the Bundesbank Council. In practice, the only time government ministers do attend Bundesbank meetings is when the annual money supply targets are being set.

The independence of the ESCB, or at least its ability to pursue price stability, is circumscribed somewhat by the fact that exchange rate policy remains the province of the Council of Ministers. Article 109 of the Maastricht Treaty stipulates that "in the absence of an exchange-rate system in relation to one or more non-Community currencies—the Council, acting by a qualified majority either on a recommendation from the Commission and after consulting the ECB or on a recommendation from the ECB, may formulate general orientations for exchange-rate policy in relation to these currencies." While the article goes on to state that "these general orientations shall be without prejudice to the primary objective of the ESCB to maintain price stability," it remains to be seen how a conflict between the two goals—fixed exchange rates and price stability—would be resolved. A decision by EU political authorities to fix the exchange rate of the euro vis-à-vis, say, the dollar would seriously compromise the ESCB's ability to conduct a monetary policy targeted solely at price stability in the euro area.[23] A similar situation prevailed in Germany before the establishment of EMU. Apparently, an understanding between the Bundesbank and the German government (the so-called Emminger Letter) temporarily released the Bundesbank from its obligation to intervene to support fixed exchange rates in the European Monetary System if such intervention threatened price stability in Germany.[24] This understanding was invoked in September 1992 when intervention to support the Italian lira threatened the Bundesbank's ability to hit its money growth targets.

However, some argue that laws can only go so far in ensuring a central bank's independence. What politicians give, they can just as easily take away. Others contend that in relationships between central banks and political authorities, personalities matter as much if not more than laws. Friedman and Schwartz (1963, 228) suggest as much in discussing the early relationship of the Fed and the Treasury. Giovannini (1993) makes a more compelling case along these lines. He argues that the ESCB's independence, as codified in the treaty and statute, is a necessary but insufficient condition for the successful pursuit of low inflation. Substantial and consistent political support is also required.

The Federal Reserve System enjoys significant independence, but, notes Maisel (1973, 24), it is "ill-defined and circumscribed." The Constitution gives Congress the right "to coin money and regulate the value thereof." Congress has delegated this authority to the Federal Reserve but could, in principle, revoke it at any time. The Federal Reserve Act has been amended and supplemented several times since its passage in 1913, although typically the changes have given the Fed greater operational independence while simultaneously increasing its accountability to Congress.

ACCOUNTABILITY

As noted above, the ESCB is the most independent central bank in the world. The Maastricht Treaty ensures that the ESCB will not be torn between pursuing multiple objectives or subject to political pressure to take what it views as inappropriate policy actions. However, the quid pro quo of central bank independence in a democratic society is that there should be adverse consequences for the central bank if it fails to achieve its objectives. Some critics have argued that independent central banks are fundamentally inconsistent with democratic principles (see, for example, Friedman 1962).

The Maastricht Treaty imposes minimal reporting obligations on the ECB, requiring only that the ECB submit an annual report to the European Parliament (and ECOFIN—the council of economics and finance ministers, the European Commission, and the European Council). The president of the ECB has indicated his willingness to testify before the European Council). The president of the ECB has indicated his willingess to testify before the European Parliament up to four times a year. The treaty notes that the ECB may decide to publish its decisions, recommendations, and opinions but does not impose any obligation in this regard. The treaty also provides for the ECB president and other Executive Board members to be heard by the relevant committees of the European Parliament.

Some prominent members of the European Parliament have called for the ECB to exceed its treaty obligations in communicating with the public. Randzio-Plath argues that

> in addition to publishing its annual and quarterly reports the ECB should be required to make public its decisions and the reasoning behind its monetary policy actions. The decisions of the Executive Board meetings should be made public on the same day. The Bank should explain why the decision has been taken as well as how the decision is linked to, and affects other policies. Minutes should be published, as should the voting

> behavior of the members, on the day of the subsequent meeting of the Executive Board. Detailed minutes should be published at the latest five weeks after the meeting. The reasons for decisions should be clear and public. Transparency is needed in a democracy. (Randzio-Plath 1997–98, 24)

There is little doubt that transparency is crucial to the success of a central bank. However, how best to achieve this is not always obvious. The FOMC's current practice is to announce policy changes as soon as they are made. Immediately after each meeting the FOMC issues a statement that a decision was made to lower or raise rates, or merely noting that the meeting ended, if the decision is to leave rates unchanged. The FOMC publishes the minutes of each meeting shortly after the subsequent meeting.

The ECB does things differently. For a variety of reasons, there is considerable resistance to publishing minutes and the voting records of Governing Council members. Perhaps the most important reason is the need to insulate Council members from domestic political pressures. While the Executive Board and the NCBs have statutory independence from domestic and Community political institutions, publication of voting records and the minutes of Council meetings may lead to pressure to vote along national lines rather than in the interests of the euro area as a whole. Issing (1998) argues that the Maastricht Treaty requires keeping the votes of the Governing Council confidential. He cites Article 10 of the statute, which states that "the proceedings of the meetings [of the Governing Council] shall be confidential. The Governing Council may decide to make the outcome of its deliberations public." Issing contends that insofar as the votes of individual Governing Council members can be considered part of the proceedings rather than part of the outcome, the treaty prohibits their publication. Others argue that the votes could just as easily be considered part of the outcome rather than part of the proceedings and that publication of votes would enhance Council members' ability to resist domestic political pressures.[25] It remains to be seen whether the ECB's decisions on confidentiality will foster or impede the development of its credibility.

STRATEGY

A strategy for monetary policy can be defined as a rule whereby a central bank responds to developments in the economy to attain its final objective. After much preparatory work on possible strategies, the EMI concluded that the only realistic options for the ESCB were monetary targeting or inflation targeting.

According to the EMI, one of the key attractions of monetary targeting is "that it clearly indicates a responsibility of the central bank for developments that are more directly under its control." An additional attraction of monetary targeting is that this strategy was successfully pursued by the Bundesbank before EMU. Adopting monetary targeting might therefore help the ESCB inherit some of the Bundesbank's credibility.[26] The strategy's primary drawback is the high degree of uncertainty about the likely behavior of monetary aggregates in the euro area following the start of monetary union.

Inflation targeting is attractive because ultimately price stability is the responsibility of the central bank.[27] Indeed, many newly independent central banks, such as the Bank of England and the Reserve Bank of New Zealand, have opted for inflation targets as a means of rapidly acquiring credibility for their commitment to price stability. The primary drawback of an inflation-targeting strategy is the difficulty of forecasting inflation at the relevant horizons. Because monetary policy actions only affect inflation with a long and variable lag (of eighteen months to two years), accurately forecasting inflation at long horizons is crucial to the success of an inflation-targeting strategy.

However, the two strategies overlap significantly in their implications for the day-to-day conduct of monetary policy. Both strategies are forward-looking in their emphasis and aim to control inflation by acting preemptively. Where they differ most is in their implications for the ESCB's communications policy—that is, how the ESCB goes about explaining its actions to the general public. Under a monetary-targeting strategy, the ESCB would explain and justify its actions primarily by reference to the behavior of the money stock vis-à-vis some target range. Under inflation targeting, the ESCB would explain and justify its actions by reference to the forecasted behavior of inflation vis-à-vis some target level.

Because of the many uncertainties accompanying the start of EMU, it is not surprising that the ESCB opted for a mixed strategy that combines elements of inflation targeting and money targeting. This is the "stability-oriented monetary policy strategy" announced by the ESCB in October 1998, whose key elements are a quantitative definition of price stability, a prominent role for money with a reference value for the growth of a monetary aggregate, and a broadly based assessment of the outlook for future price developments.

Adoption of a mixed strategy might seem to defeat the purpose of articulating a strategy in the first place. One of the most important reasons for formulating and adhering to a strategy is that doing so makes monetary policy actions more transparent and easier to communicate to the public. The simpler the strategy, the easier that communication. Under a rigid monetary-targeting strategy, a central bank need only point to money growth in excess of its target to justify increases in interest rates. Under a mixed strategy, the situation would be more complicated because the central bank would have to spell out in detail how it would respond to different scenarios. In particular, the central bank would need to explain what it would do if growth in the money stock were signaling that a tightening of monetary policy would be appropriate, while the inflation forecast was signaling that an easing of monetary policy was appropriate. Having to detail all these contingencies makes it considerably harder to communicate with the general public, and it is only a short step from this to the look-at-everything, respond-to-everything approach to policy.

Again, the contrast with the way the Federal Reserve conducts monetary policy is instructive. At present the Fed does not employ either a monetary-targeting or inflation-targeting approach. Monetary targets have not played an important role in U.S. monetary policy since at least the early 1990s. And the Fed has never formally adopted inflation targeting as a strategy, at least not to the extent that, say, the Bank of England has. However, the Fed is a lot more forward-looking in its deliberations than it was in the 1970s.[28]

Why the Fed does not feel the need to articulate a strategy for monetary policy is an open question. One reason may be that the Fed has done reasonably well controlling inflation and building credibility without a formal strategy, and as long as that continues, it sees no need to change. This is consistent with the view that debates about strategy are most intense in central banks that need to rapidly acquire credibility for their commitment to price stability.

MONETARY POLICY TOOLS

The ESCB has three instruments available for the conduct of monetary policy. It engages in open market operations, offers standing facilities, and requires credit institutions to hold minimum reserves. Open market operations play a central role in the conduct of monetary policy. The ESCB has five types of instruments available for the conduct of open market operations, the most important of which is reverse transactions. The ESCB also has the option of using outright transactions, issuing debt certificates, making foreign exchange swaps, and collecting fixed-term deposits. Open market operations are initiated by the ECB but are conducted through the NCBs. The ECB decides on the instrument to be used in all open market operations and on the terms and conditions for their execution. This highly decentralized approach to monetary policy operations is in marked contrast to the Fed's practice of conducting all operations through the New York Reserve Bank.

The ESCB offers standing facilities to provide and absorb overnight liquidity, signal the general stance of monetary policy, and bound overnight market interest rates. These facilities—a marginal lending facility and a deposit facility—are available to eligible counterparties on their own initiative as long as they fulfill the relevant conditions. Only financial institutions subject to the reserve requirement may access the standing facilities (and participate in open market operations based on standard tenders). The Fed does not provide comparable facilities.

The ESCB transacts in a wide range of financial assets in conducting monetary policy operations. These assets are not necessarily restricted to the debt liabilities of national governments, but they are required to satisfy certain criteria so as to protect the ESCB from the risk of losses on its monetary policy operations.

Finally, the ECB has set a reserve requirement ratio at 2 percent, with the reservable components of the liability base consisting of overnight deposits, deposits with agreed maturity up to two years, deposits redeemable at notice up to two years, debt securities with agreed maturity up to two years, and money market paper. The ECB allows financial institutions to deduct a lump-sum allowance of 100,000 euros from their reserve requirement. The ECB renumerates reserve holdings at an interest rate corresponding to the rate of its main refinancing operations, with interest paid on the first business day after the end of the reserve maintenance period.

CONCLUSIONS

The launching of EMU is probably the single most important development in international monetary relations in the past fifty years. If monetary union succeeds, the euro may one day challenge the U.S. dollar's dominance in international transactions. The sheer size of the single-currency area will fundamentally alter international monetary arrangements. How the euro fares against the dollar will depend on the relative performances of the ESCB and the Federal Reserve in maintaining price stability in their respective territories. The ESCB starts with the advantage of an unambiguous mandate for price stability based on an international treaty that can only be altered with the consent of all its signatories. However, the diffuse distribution of power within the ESCB may make it difficult to resolve the conflicts of national interests that some academic critics of EMU believe doom the undertaking to failure. By contrast, the Federal Reserve System does not have as strong a mandate for price stability, but its more centralized decision-making structure arguably enhances the monetary policy process in the United States.

NOTES

1. The fifteen members of the EU are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the UK. All but four members are participating in the first round of EMU. Denmark, Sweden, and the UK are not participating for domestic political reasons; Greece failed to meet the convergence criteria laid down in the Maastricht Treaty but intends to join as soon as possible.
2. EMU countries' notes and coins will continue to circulate until 2002; however, they no longer exist as currencies in their own right but as nondecimal denominations of the euro. There are several reasons for the three-year transition before the euro acquires a physical form. First, it will take time to adapt the physical payments infrastructure in each of the participating countries to the new notes and coins. In 1995, there were some 3.15 million vending machines and 130,000 ATMs in the EU; such machines will have to be recalibrated to accept the new currency. Second is the magnitude of the task of replacing national currencies. Printing enough banknotes and minting enough coins to replace all the existing notes and coins will take time. In 1994, more than 12 billion banknotes and 70 billion coins circulated in the EU, with a combined weight of 300,000 metric tons. Minting of euro coins began in May 1998. Finally, the transition allows businesses and the general public to become familiar with the new currency before having to use it for all transactions. During the transition, the no-compulsion, no-prohibition principle governs the use of the euro.
3. For a recent analysis of the unemployment problem in Europe, see Ljunqvist and Sargent (1998).
4. For this article, the significance of the Committee of Governors is that the economic unit created to support the committee would subsequently form the cadre for the European Central Bank.
5. For a textbook review of the major issues, see De Grauwe (1997). See also Feldstein (1997) and Wyplosz (1997).
6. Actually, many features of the ESCB are modeled on Deutsche Bundesbank, which is modeled on the Federal Reserve System. See Deutsche Bundesbank (1995).
7. Maastricht Treaty Protocol (no. 3) on the Statute of the European System of Central Banks and of the European Central Bank, Article 12.1.
8. The FOMC in its current form, with the Board of Governors enjoying a permanent majority, did not come into being until 1935. When the Federal Reserve System was established in 1914, it was thought discount lending would be the primary tool of monetary policy, with individual Reserve Banks having considerable discretion to set discount rates. It was not until the 1920s that the potential of open market operations was discovered. In the spring of 1922 the Committee of Governors on Centralized Execution of Purchases and Sales by Federal Reserve Banks was established to coordinate the actions of the System. This committee was reconstituted as the Open Market Investment Committee (OMIC) in 1923, consisting of representatives of the Boston, New York, Philadelphia, Cleveland, and Chicago Reserve Banks, under the chairmanship of the New York Bank. The OMIC was disbanded in 1930 and reconstituted as the Open Market Policy Conference, composed of representatives from all twelve Reserve Banks. The Banking Act of 1933 established the FOMC, consisting of representatives of the twelve Reserve Banks and the seven Board of Governors members. The Banking Act of 1935 altered the FOMC's composition to give the seven Board members a vote in open market policy and, more importantly, reduce the representation of the Reserve Banks to five members. This gave the Board of Governors a permanent majority.
9. Federal Reserve Act, Section 4.20.
10. See, for example, the schematic diagrams of the informal power structure of the Federal Reserve System in any intermediate money and banking textbook.
11. Statute of the European System of Central Banks and of the European Central Bank, Articles 10 and 13.

12. Statute of the European System of Central Banks and of the European Central Bank, Article 13.
13. European Central Bank (1998b).
14. Federal Reserve Act, Section 10.1.
15. Article 2 of the treaty states that "the Community shall have as its task, by establishing a common market and an economic and monetary union and by implementing the common policies or activities referred to in Articles 3 and 3a, to promote throughout the Community a harmonious and balanced development of economic activities, sustainable and non-inflationary growth respecting the environment, a high degree of convergence of economic performance, a high level of employment and of social protection, the raising of the standard of living and quality of life, and economic and social cohesion and solidarity between the Member States." Article 3a of the treaty states: "A1. For the purposes set out in Article 2, the activities of the Member States and the Community shall include, as provided in this Treaty and in accordance with the timetable set out therein, the adoption of an economic policy which is based on the close coordination of the Member States' economic policies, on the internal market and on the definition of common objectives, and conducted in accordance with the principle of an open market economy with free competition. 2. Concurrently with the foregoing, and as provided in this Treaty and in accordance with the timetable and the procedures set out therein, these activities shall include the irrevocable fixing of exchange rates leading to the introduction of a single currency, the ECU [European currency unit], and the definition and conduct of a single monetary policy and exchange-rate policy the primary objective of both of which shall be to maintain price stability and, without prejudice to this objective, to support the general economic policies in the Community, in accordance with the principle of an open market economy with free competition. 3. These activities of the Member States and the Community shall entail compliance with the following guiding principles: stable prices, sound public finances and monetary conditions and a sustainable balance of payments."
16. Other provisions in the treaty further reinforce the mandate for price stability. First, Article 2 of the statute repeats Article 105 of the treaty. Article 2 of the treaty makes the promotion of non-inflationary growth one of the European Community's objectives. Article 3 of the treaty states that the primary objective of both monetary and exchange rate policy following the start of monetary union "shall be to maintain price stability." Article 3 of the treaty also states that achieving stable prices is one of the guiding principles of the Community.
17. Alesina and Grilli (1992) evaluate the political and economic independence of the ECB using the same criteria as other authors to construct quantitative indexes of central bank independence. They find that the ECB will enjoy the same degree of political and economic independence as the Bundesbank, which is somewhat more independent than the Fed.
18. Statute of the European System of Central Banks and of the European Central Bank, Article 14.2.
19. Posen (1993) is more sanguine about the ECB's prospects, arguing that it will have important political support from the European financial community.
20. Alesina and Grilli (1992) use the terms *political independence* and *economic independence* to refer to essentially the same things.
21. European Central Bank (1998a).
22. See, for example, Friedman and Schwartz (1963) and Timberlake (1993).
23. As Fischer (1994, 304) notes, "Monetary and exchange rate policies cannot be independent. Under floating rates, monetary policy affects the exchange rate. Thus the government cannot have control over exchange rate policy while the central bank has control over monetary policy. The government should have the authority to choose the exchange rate regime. If it chooses a fixed exchange rate regime, it has then essentially—though not completely—determined monetary policy. While a central bank can be more or less independent of the government in a fixed exchange rate regime, its independent ability to determine the rate of inflation and interest rates is sharply curtailed." See also Giovannini (1993).
24. See, for example, Ungerer (1997) and Gros and Thygesen (1998).

25. See, in particular, Buiter (1998a, 1998b).
26. Issing (1994) argues along these lines.
27. For analyses of inflation targeting as a strategy for monetary policy, see Haldane (1995), Leiderman and Svensson (1995), and Bernanke and Mishkin (1997).
28. See Goodfriend (1993).

REFERENCES

Alesina, A., and Grilli, V. (1992). The European Central Bank: Reshaping Monetary Politics in Europe. In *Establishing a Central Bank: Issues in Europe and Lessons from the United States,* ed. Matthew B. Canzoneri, Vittorio Grilli, and Paul R. Masson (Cambridge: Cambridge University Press), 49–85.

Alesina, A., and Summers, L. H. (1993). Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence. *Journal of Money, Credit, and Banking* 25 (May): 151–62.

Banaian, K., Laney, L. O. and Willett, T. D. (1983). Central Bank Independence: An International Comparison. Federal Reserve Bank of Dallas *Economic Review,* March, 1–13.

Bernanke, B. S., and Mishkin, F. S. (1997). Inflation Targeting: A New Framework for Monetary Policy? *Journal of Economic Perspectives* 11 (Spring): 97–116.

Buiter, W. H. (1998a). Independence of ECB Likely to Be Perverted If Council Voting Kept Secret. *Financial Times,* September 24, 14.

Buiter, W. H. (1998b). ECB Secrecy Is the Euro's Fatal Flaw. *Sunday Telegraph,* October 18, B6.

Cukierman, A., Webb, S. B. and Neyapti, B. (1992). Measuring the Independence of Central Banks and Its Effect on Policy Outcomes. *World Bank Economic Review* 6 (September): 353–98.

De Grauwe, P. (1997). *The Economics of Monetary Integration,* 3rd ed. (Oxford: Oxford University Press).

Deutsche Bundesbank (1995). *The Monetary Policy of the Bundesbank* (Frankfurt am Main: Deutsche Bundesbank).

Eijffinger, S., and De Haan, J. (1996). *The Political Economy of Central-Bank Independence.* Special Paper no. 19, International Finance Section, Princeton University.

European Central Bank (1998a). Press release, October 13. A Stability-Oriented Monetary Policy Strategy for the ESCB.

European Central Bank. (1998b). Press release, June 4. Statement by the Executive Board of the European Central Bank on the Distribution of Responsibilities Among Its Members for the Business Areas of the ECB.

Feldstein, M. (1997). The Political Economy of the European Economic and Monetary Union: Political Sources of an Economic Liability. *Journal of Economic Perspectives* 11 (Fall): 23–42.

Fischer, S. (1994). Modern Central Banking. In *The Future of Central Banking: The Tercentenary Symposium of the Bank of England,* ed. F. Capie, C. Goodhart, S. Fischer, and N. Schnadt (Cambridge: Cambridge University Press), 262–308.

Friedman, M. (1962). Should There Be an Independent Monetary Authority? In *In Search of a Monetary Constitution,* ed. L. B. Yeager (Cambridge, Mass.: Harvard University Press), 219–43.

Friedman, M., and Schwartz, A. J. (1963). *A Monetary History of the United States 1867–1960* (Princeton, N.J.: Princeton University Press).

Giovannini, A. (1993). Central Banking in a Monetary Union: Reflections on the Proposed Statute of the European Central Bank. *Carnegie-Rochester Conference Series on Public Policy* 38: 191–230.

Goodfriend, M. (1993). Interest Rate Policy and the Inflation Scare Problem: 1979–1992. Federal Reserve Bank of Richmond *Economic Quarterly,* Winter, 1–24.

Gros, D., and Thygesen, N. (1998). *European Monetary Integration: From the European Monetary System to Economic and Monetary Union,* 2nd ed. (London: Longman).

Haldane, A. (1995). *Targeting Inflation* (London: Bank of England).

Hetzel, R. L. (1990). A Mandate for Price Stability. Federal Reserve Bank of Richmond *Economic Review,* March–April, 45–53.

Issing, O. (1994). Monetary policy strategy in the EMU. In *A Framework for Monetary Stability,* ed. J. A. H. de Beaufort Wijnholds, S. C. W. Eijffinger, and L. H. Hoogduin (Dordrecht, Netherlands: Kluwer Academic Publishers), 135–48.

Issing, O. (1998). Open for Business. *Financial Times,* September 22, 16.

Leiderman, L., and Svensson, L. E. O. (1995). *Inflation Targets* (London: Center for Economic Policy Research).

Ljungqvist, L. and Sargent, T. J. (1998). The European Unemployment Dilemma. *Journal of Political Economy* 106 (June): 514–50.

Maisel, S. J. (1973). *Managing the Dollar* (New York: W. W. Norton).

Posen, A. S. (1993). Why Central Bank Independence Does Not Cause Low Inflation: There Is No Institutional Fix for Politics. In *Finance and the International Economy 7: The Amex Bank Review Prize Essays,* ed. Richard O'Brien (Oxford: Oxford University Press), 45–54.

Randzio-Plath, C. (1997–98). Democratic Accountability of the European Central Bank. *Central Banking* 8 (Winter): 22–25.

Timberlake, R. H. (1993). *Monetary Policy in the United States: An Intellectual and Institutional History* (Chicago: University of Chicago Press).

Ungerer, H. (1997). *A Concise History of European Monetary Integration: From EPU to EMU* (Westport, Conn.: Quorum Books).

United Nations (1997). *Statistical Yearbook 1995,* 42nd ed. (New York: United Nations).

Wyplosz, C. (1997). EMU: How and Why It Might Happen. *Journal of Economic Perspectives* 11 (Fall): 3–21.

This Page Intentionally Left Blank

61

Economic Factors, Monetary Policy, and Expected Returns on Stocks and Bonds

James R. Booth

Department of Finance, Arizona State University, Tempe, Arizona

Lena Chua Booth

Department of Business, Thunderbird-American Graduate School of International Management, Glendale, Arizona

A growing body of research has focused on forecasting stock and bond returns using economic and monetary factors. Fama and French (1988, 1989), Fama (1990), and Schwert (1990) focus on economic factors and find that three business conditions proxies, the dividend yield, default spread, and term spread, can explain significant variation in expected stock and/or bond returns. These studies generally find that the required returns that investors demand vary over the business cycle.

The majority of the research on monetary policy has focused on its impact in the real sector (see Romer and Romer 1989 and Bernanke and Blinder 1992). Less attention has been directed at the impact of monetary policy actions on stock and bond returns. Recently, Jensen, et al. (1996) used an index of the stance of monetary policy based on changes in the discount rate to show that expected stock returns are higher in expansive periods than in restrictive periods. Combining the previously used business cycle proxies with a measure of monetary policy, they find that the impact of the various business conditions proxies varies across monetary environments. Specifically, they find that the business conditions proxies have explanatory power only during restrictive periods.

In this study, we examine the impact of monetary policy on expected stock and bond returns and expand on previous work in several ways. First, we construct measures of the business conditions proxies in a slightly different way to test the robustness of the findings related to the predictability of stock returns. Second, we use two measures of monetary policy actions, the one developed by Jensen, et al. (1996) related to the directional change of the discount rate and one proxied by the federal funds rate, to determine whether there exists a direct monetary sector effect on stock and bond returns through these measures of monetary policy. Third, we èxamine a portfolio of small stocks and a portfolio of large

Reprinted from: *FRBSF Economic Review*, no. 2 (1997) 32–42.

stocks to determine whether the findings related to either the business conditions or monetary stringency have a differential impact given firm size. The motivation for this is based on the notion that smaller companies are more directly affected by changes in monetary policy due to their dependence on bank and private market financing.

We find, similar to earlier work on business conditions and expected returns, that the default spread, dividend yield, and the term spread are important in explaining expected returns on both large and small stock portfolios and on a portfolio of corporate bonds. We find that both measures of monetary policy actions have explanatory power for expected excess returns on the large stock portfolio and for the small stock portfolio in monthly returns. For the expected excess returns on corporate bonds, we find that the discount rate change measure of monetary policy stance has explanatory power. When we interact the discount rate change index with the business conditions proxies, we find that the monetary policy effect is direct and does not work through the business conditions proxies as suggested by Jensen, et al. (1996). We do find a larger monetary or business condition effect for smaller firms, consistent with a differential impact on these firms compared to large firms. Overall, these results suggest monetary policy actions can be used to forecast excess returns on stocks and bond portfolios.

RELATED RESEARCH

Business Conditions and Security Returns

The recent research on the relation between stock returns and business conditions have focused on three measures of the business environment: dividend yield, the default spread, and the term spread. Dividend yield, as a business conditions proxy, is perhaps the oldest of the measures believed to vary with expected stock returns (see Dow 1920). The intuition for this relation, provided by Fama (1990), is that stock prices are low relative to dividends when discount rates and expected returns are high, and vice versa, so $D(t)/V(t)$ varies with expected returns. Rozeff (1984), Shiller (1984), Campbell and Shiller (1987), Fama and French (1988, 1989), Fama (1990), and Jensen, et al. (1996) document that dividend yields forecast stock returns.

Evidence that the default spread is important in explaining stock and/or bond returns is more recent. Chen, Roll, and Ross (1986) argue that the spread of lower- to higher-grade bonds is a proxy for business conditions. They argue that when business conditions are poor, spreads are likely to be high, and when business conditions are strong, spreads are likely to be low. Studies by Fama and French (1989), Fama (1990), and to a lesser degree Jensen, et al. (1996), find that the default spread captures variations in expected returns in response to business conditions.

The third measure of business conditions that has been used in previous studies is the term spread. The motivation for this is that the term spread is shown to decrease near peaks of economic activity and increase near economic troughs. Consistent with this motivation, Campbell (1987), Fama and French (1989), Fama (1990), Schwert (1990), Shiller (1984), Campbell and Shiller (1987), and Jensen, et al. (1996) find that the term spread also explains similar variations in expected stock returns.

Monetary Policy and Security Returns

It has long been contended that monetary policy affects not only economic activity, but also security returns. An early examination of the link between stock returns and monetary policy by Rozeff (1984) finds a relation between stock returns and contemporaneous monetary policy developments. Additional studies by Shiller (1984), Campbell and Shiller (1987), Geske and Roll (1983), and Kaul (1987) present evidence linking the monetary sector to stock returns.

More recently, Jensen and Johnson (1995) find that stock returns are related to changes in the Federal Reserve discount rate. In Jensen, et al. (1996), this measure of monetary policy is used to show that business conditions proxies used in previous studies (as discussed above) vary dramatically across monetary environments. Their motivation for using the discount rate as a proxy for the stance of monetary policy follows from the view that the discount rate is routinely regarded as a signal of monetary and possibly economic developments. Their argument is based on Waud's (1970) suggestion that discount rate changes affect market participants' expectations about monetary policy because (1) rate changes are made only at substantial intervals, (2) they represent a somewhat discontinuous instrument of monetary policy, and (3) they are established by a public body perceived as being competent in judging the economy's cash and credit needs. Using discount rate change series as their measure of expansive and restrictive policies, they are able to show that the behavior of the business conditions proxies and their influence on expected returns is significantly affected by the monetary environment.

We reexamine the impact of monetary policy based on the measure developed by Jensen, et al. (1996) with slightly different proxies for business conditions. We also use the federal funds rate, based on evidence by Bernanke and Blinder (1992) and Laurent (1988) that the federal funds rate is a good indicator of monetary policy actions. To examine whether business conditions and monetary policy have a differential impact on small versus large stocks, we examine expected returns on a portfolio of the S&P 500 firms, a portfolio of small stocks (approximately the fifth quintile of firms on the New York Stock Exchange), and a portfolio of Aaa and Aa rated bonds. This allows us to test for a differential impact of both business conditions and monetary policy on large versus small firm returns and on bond returns.

DATA

Sample Period

We examine stock and bond returns over the period August 1954 through December 1992. This follows closely the sample period chosen by Jensen, et al. (1996) and the first availability of the federal funds rates. Even though February 1954 reflects the first change in stance through the discount rate since the Federal Reserve/Treasury accord of 1951, we start our sample from August 1954 to match the federal funds rate data. This permits us to compare the information contained in each measure.

Following the Jensen, et al. (1996) approach in constructing the discount rate series, we find this 39-year period includes a total of 99 discount rate changes, 49 increases and 50 decreases. They define a rate change series as a period of time over which discount rate changes are in only one direction, either increasing or decreasing. This results in 23 rate change series, 12 decreasing and 11 increasing. Using this framework, we accept their notion that a series reflects a period in which the Fed is operating under the same monetary policy; the next series occurs when a rate change in the opposite direction is announced. The months in which rates are announced are eliminated from the sample. This results in 439 monthly observations, 239 months following discount rate increases and 200 following discount rate decreases.

Table 1 Federal Reserve Discount Rate Change Series: February 1954 through December 1992

Series	Increasing (I) or decreasing (D)	First rate change	Number of rate changes	Monthly observations
1	D	02/05/54	2	13
2	I	04/14/55	5	30
3	D	11/15/57	4	8
4	I	08/15/58	5	21
5	D	06/30/60	2	36
6	I	07/17/63	3	44
7	D	04/07/67	1	6
8	I	11/20/67	3	8
9	D	08/16/68	1	3
10	I	12/19/68	2	22
11	D	11/11/70	5	7
12	I	07/16/71	1	3
13	D	11/11/71	2	13
14	I	01/15/73	8	22
15	D	12/09/74	7	31
16	I	08/30/77	14	32
17	D	05/29/80	3	3
18	I	09/26/80	4	13
19	D	11/02/81	9	28
20	I	04/09/84	1	6
21	D	11/23/84	7	33
22	I	09/04/87	3	38
23	D	12/18/90	7	24

Source: Extracted from Jensen, et al. (1996).

In the quarterly sample, we have 131 observations. This is 11 quarters fewer than that of Jensen, et al. (1996) because of the creation of quarters around rate changes. They drop months when the number of months in a rate change series is not divisible by 3. We use the traditional calendar quarters and eliminate the quarters in which a rate change occurred. This analysis places the monthly and quarterly data into one of two subsamples: observations that occur during increasing rate series and observations that occur during decreasing rate series. Table 1 provides the number of months and quarters in each rate change series.

Return and Macroeconomic Variables

The return and explanatory variables follow those used in previous studies, particularly Fama (1990) and Jensen, et al. (1996).

Return Variables

Large Stock Returns (LS) Monthly stock returns for the large stock portfolio are collected from Ibbotson and Associates for the sample period February 1954 through December 1992. The data comprise the total returns, including dividends, for the S&P 500 after March 1957 and for the S&P 90 stocks before 1957. These represent a portfolio of the largest market value companies in the U.S. The portfolio returns are value-weighted. To obtain a measure of excess returns, we subtract the contemporaneous monthly return on T-bills.

Small Stock Returns (SS) These are the monthly returns on the Ibbotson small stock portfolio for the same sample period. For the period February 1954 to December 1981, this portfolio was the Dimensional Fund Advisors (DFA) Small Company 9/10 (ninth and tenth) Fund. The fund is a market-value-weighted index of the ninth and tenth deciles of the New York Stock Exchange (NYSE), plus stocks listed on the American Stock Exchange (AMEX) and over-the-counter (OTC) with capitalization that is the same as or less than the upper bound of the NYSE ninth decile.

The weight of each stock within the fund is proportionate to its market capitalization; therefore, stocks with a higher market capitalization value will be weighted more than stocks with a lower market capitalization value. Since the lower bound of the tenth decile is near zero, stocks are not purchased if they are smaller than $10 million in market capitalization (although they are held if they fall below that level). A company's stock is not purchased if it is in bankruptcy; however, a stock already held is retained if the company becomes bankrupt. Stocks remain in the portfolio if they rise into the eighth NYSE decile, but they are sold when they rise into the seventh NYSE decile or higher. The returns for the DFA Small Company 9/10 Fund represent after-transactions-cost returns while the returns on other asset classes and for the pre-1982 small company stocks are before-transactions-cost returns.

For the period after 1982, the small stock portfolio is represented by the historical series developed by Banz (1981). This equals the fifth quintile of the NYSE, based on market value. Every five years the portfolio is rebalanced and the new portfolio includes the new fifth quintile of the NYSE. Excess returns are obtained by subtracting the return on the contemporaneous T-bill.

Corporate Bond Returns (CB) The corporate bond total returns are represented by the Salomon Brothers Long-Term High-Grade Corporate Bond Index. According to Ibbotson Associates, the index includes nearly all Aaa- and Aa-rated bonds. Capital appreciation returns were calculated from yields assuming a 20-year maturity, a bond price equal to par, and a coupon equal to the beginning-of-period yield. The monthly income return was assumed to be one-twelfth the coupon. The monthly return on the T-bill is subtracted to obtain excess returns.

Explanatory Variables

Dividend Yield (D/P) To obtain the dividend yield for the large stock portfolio, we use the income return calculated by Ibbotson Associates. Following Fama and French (1989), we use annual income returns as the independent variable.

Term Spread (TERM) To calculate the term spread, we use the long-term government bond return from Ibbotson Associates. For the 1954 to 1976 period, this involved using approximately 20 bonds with reasonably current coupons. For the 1977–1992 period, the return was calculated as the change in the price plus the coupon payments. To develop a measure of TERM, we subtract the contemporaneous T-bill return from the long-term government bond return. This measure differs from Fama (1990) and Jensen, et al. (1996) in that they measure the difference between the 10-year and 1-year T-bond returns.

Default Spread (DEF) The default spread is measured as the difference between the return on the corporate bond portfolio and the T-bond portfolio. Our measure is obtained by subtracting the 20-year T-bond portfolio return (approximately) from the return of a portfolio containing Aaa- and Aa-rated corporate bonds. This measure is closest to the Jensen, et al., measure of the Baa corporate bond minus the 10-year T-bond. Fama (1990) and Fama and French (1989) use the difference between a portfolio of all corporate bonds and the yield on the Aaa corporate portfolio. Schwert (1990) uses the difference in yield between Baa and Aa-rated corporate bonds.

Discount Rate Changes (DIR) This is a binary variable taking on the value of one if the previous discount rate change was an increase and zero if the previous change was a decrease.

Federal Funds Rate (FFRATE) This annualized rate equals the monthly and quarterly averages of daily federal funds rates collected from the Federal Reserve Bank of St. Louis (FRED) data series.

To obtain security returns for the analysis involving quarterly holding periods, we cumulate monthly observations. Following previous studies, we use excess returns of large stocks (LS), small stocks (SS), and corporate bonds (CB) as dependent variables. Consistent with earlier approaches, we focus on expected returns. In performing the statistical analysis, we lag the independent variables D/P, TERM, DEF, and FFRATE by one period relative to the excess returns variables.

EMPIRICAL RESULTS

Variable Means Across Monetary Environments

Table 2 presents the means of the variables used in the analysis across the sample period and during the expansive and restrictive monetary periods, based on the discount rate index constructed according to the Jensen, et al. (1996) approach. The excess return variables for our large stock portfolio, which is based on the S&P 500, are similar in magnitude to those reported for the value-weighted CRSP index in Jensen, et al. The excess returns for our small stock portfolio are slightly higher than those reported for the equally weighted CRSP index in Jensen, et al. The excess returns for our portfolio of high-grade corporate bonds

Table 2 Means of Observations of Business Conditions Proxies and Security Returns: August 1954 through December 1992

Variable	Full sample (n = 439)	Expansive periods (n = 200)	Restrictive periods (n = 239)	t Test
Security returns (monthly)				
Large stock excess returns (LS)	0.523	1.299	–0.125	3.49**
Small stock excess returns (SS)	0.885	1.932	0.008	3.42**
High-grade bond excess returns (CB)	0.088	0.418	–0.187	2.70**
Business Conditions Proxies (annualized):				
Term spread (TERM)	0.072	6.067	–4.884	12.84**
Dividend yield (D/P)	4.065	4.153	3.991	1.89
Default spread (DEF)	0.737	1.325	0.246	2.58**
Federal funds rate (annualized):	6.298	5.490	6.975	4.43**

** Statistically significant at the 0.01 level.

are consistent with the findings of Jensen, et al., and Rozeff (1984), who find that stock returns vary across the monetary policy environment.

The results on annual dividend yield are slightly lower than those reported for the CRSP index by Jensen, et al., and by Fama and French (1990). The difference across monetary policy environments is similar to that reported in Jensen, et al. Our measure of TERM differs substantially, both in construction and in results, from other studies. We use the difference between the long-term 20-year T-bond and the T-bill rates; Jensen, et al., uses the difference between the 10-year and 1-year Treasury yields, and Fama (1990), Fama and French (1989), and Schwert (1990) use the difference between corporate bond yields and the T-bill. Compared to the results in Jensen, et al., the mean of our variable is lower, and our measure shows much greater variation across different monetary regimes. We prefer it because it reflects the spread between two of the more liquid Treasury issues and does not contain any potential for a default spread, as do the measures using corporate series.

Our measure of the default spread (DEF) uses the difference between the return on the portfolio of Aaa- and Aa-rated corporate bonds and the return on long-term T-bonds. Earlier studies use the difference between high- and low-grade corporate bonds (Fama 1990 and Schwert 1990). Our measure is closer to that used in Jensen, et al. (1996), viz, the Baa-rated corporate bond minus the 10-year T-bond yield. Compared to the measure used by Jensen, et al., our measure of the default spread, DEF, has a smaller mean, and it exhibits greater variability over different monetary regimes. This is consistent with the interpretation of Jensen, et al., that there is an increasing concern about a firm's ability to service its debt during expansive periods. This is also consistent with higher risk premiums during economic downturns.

Our results for the second measure of monetary policy actions, the federal funds rate, indicate that the level of the federal funds rate is consistent with the direction of monetary policy indicated by the discount rate change measure. The correlation between the federal funds rate and the discount rate index is 0.22. Thus, they both contain unique information that may affect expected returns.

Business Conditions Proxies and Expected Returns

In Table 3, we provide regressions of business conditions on the expected returns on stocks and bonds. The results presented here are similar to earlier studies by Fama and French (1989) and Jensen, et al. (1996). We find that our measure of the term spread (TERM) has a positive coefficient and is significant in explaining returns of large stocks, small stocks, and corporate bonds for both monthly and quarterly horizons. This finding is consistent with Fama and French (1989), Fama (1990), and Jensen, et al. (1996). The dividend yield (D/P) has explanatory power for large and small stock returns but not for corporate bond returns in the monthly returns. For the quarterly horizon, D/P loses significance for large and small stocks and corporate bonds. These findings differ from those of Fama and French (1989) and the monthly returns of Jensen, et al. (1996), who find that D/P has explanatory power for corporate bond returns. For quarterly returns, we find that D/P does not have explanatory power for either stocks or bonds.

We find the default spread (DEF) has explanatory power for monthly returns of large and small stocks but not for corporate bonds. Over the quarterly return horizon, we find that DEF has explanatory power in forecasting quarterly corporate bond returns as well as large- and small-stock portfolios returns. Jensen, et al. (1996) find that the default spread is important only in explaining equally weighted stock portfolio returns. Overall, we find that

Table 3 Results of Regressions of Business Conditions on the Expected Returns of Stocks and Bonds: February 1954 through December 1992

Monthly returns

Dependent variable	Constant	Term	D/P	DEF	Adj. R^2	F_{test}
(1) Large stock portfolio	−0.011	0.269	0.404	0.404	0.02	4.62
	(−1.197)	(3.112)**	(1.809)	(2.256)*		[0.01]
(2) Small stock portfolio	−0.015	0.409	0.582	0.549	0.03	5.35
	(−1.172)	(3.446)**	(1.899)	(2.232)*		[0.01]
(3) Bond portfolio	0.003	0.116	−0.063	0.086	0.01	2.10
	(0.660)	(2.441)*	(−0.513)	(0.881)		[0.10]
(4) Large stock portfolio	−0.011	0.184	0.402		0.02	4.35
	(−1.155)	(2.354)*	(1.790)			[0.01]
(5) Small stock portfolio	−0.014	0.294	0.579		0.02	5.48
	(−1.131)	(2.735)**	(1.879)			[0.01]
(6) Bond portfolio	0.003	0.097	−0.063		0.01	2.75
	(0.674)	(2.286)*	(−0.518)			[0.06]
(7) Large stock portfolio	0.005	0.268		0.403	0.02	5.27
	(2.626)**	(3.092)**		(2.242)*		[0.01]
(8) Small stock portfolio	0.009	0.408		0.547	0.02	6.18
	(3.140)**	(3.423)**		(2.217)*		[0.01]
(9) Bond portfolio	0.001	0.116		0.087	0.01	3.01
	(0.732)	(2.446)*		(0.884)		[0.05]

Quarterly Returns

Dependent variable	Constant	Term	D/P	DEF	Adj. R^2	F_{test}
(1) Large stock portfolio	−0.033	0.558	1.114	1.122	0.10	5.66
	(−1.042)	(3.507)**	(1.444)	(2.589)**		[0.00]
(2) Small stock portfolio	−0.037	0.875	1.457	1.669	0.10	5.86
	(−0.767)	(3.676)**	(1.262)	(2.619)**		[0.00]
(3) Bond portfolio	0.010	0.240	−0.236	0.485	0.05	3.26
	(0.610)	(2.862)**	(−0.579)	(2.118)*		[0.02]
(4) Large stock portfolio	−0.032	0.411	1.128		0.06	4.92
	(−0.989)	(2.709)**	(1.432)			[0.01]
(5) Small stock portfolio	−0.035	0.654	1.478		0.06	5.13
	(−0.719)	(2.874)**	(1.253)			[0.01]
(6) Bond portfolio	0.011	0.177	−0.230		0.02	2.57
	(0.627)	(2.227)*	(−0.557)			[0.08]
(7) Large stock portfolio	0.012	0.570		1.126		7.38
	(1.726)	(3.575)**		(2.588)**	0.09	[0.00]
(8) Small stock portfolio	0.022	0.892		1.705		7.96
	(2.183)	(3.742)**		(2.622)**	0.10	[0.00]
(9) Bond portfolio	0.001	0.238		0.484	0.05	4.74
	(0.209)	(2.842)**		(2.119)*		[0.01]

* Statistically significant at the 0.05 level.
** Statistically significant at the 0.01 level.
Notes: *t* statistics in parentheses; *p* values in brackets.

the business conditions proxies have explanatory power for explaining stock and bond returns on both monthly and quarterly return horizons. Our results for the dividend yield (D/P) are not as strong as earlier studies but may reflect differences in the computation of this variable.

Monetary Sector and Security Returns

In Table 4, we add the proxies for monetary policy stance, the federal funds rate and the discount rate change series. The coefficients for the federal funds rate (FFRATE) in the monthly regressions are negative and statistically significant for the large and small stock regressions but not significant in the bond return regressions. The coefficient for DIR (value of one during restrictive periods) is negative and statistically significant at the 0.05 level for all the monthly regressions.

For the quarterly regressions in Table 4, the results are quite different. The federal funds rate (FFRATE) is important only in predicting large stock returns. The discount rate change (DIR) has explanatory power only for corporate bond returns. DIR has explanatory power for large stocks returns when FFRATE is not included.

The regressions indicate that both the changes in the federal funds rate (FFRATE) and the discount rate series (DIR) have explanatory power for predicting excess stock returns, but only the DIR measure has explanatory power for predicting excess bond returns. These results indicate that the returns on all portfolios are higher during expansive monetary periods than during restrictive periods.

We also find that the business conditions proxies have explanatory power for stock and bond returns. The addition of the proxies for monetary restrictiveness alters, to a slight degree, the explanatory power of the business conditions proxies for stock and bond portfolio returns. In particular, the coefficient and explanatory power of D/P, the dividend yield, is consistently smaller for large stock, small stock, and corporate bond portfolios. The coefficients on TERM remain statistically significant for most stock regressions. These results differ from those of Jensen, et al. (1996), who find that the introduction of the monetary policy variable causes their measure for the term spread to lose explanatory power for all stock regressions, although it is still significant in the monthly and quarterly bond portfolio regressions. The default spread (DEF) loses explanatory power, although it is still significant at the 0.10 level for the large and small stock portfolios in the monthly regressions. For the quarterly return horizons, DEF continues to be significant at the 0.05 level for the stock regressions. Thus, the introduction of the two proxies only slightly alters the results related to the business conditions proxies. This suggests the potential for a direct monetary policy effect on expected stock and bond returns.

In Table 5, we present evidence related to the stability of the slope parameters across monetary policy environments. To do this, we interact DIR with the business conditions proxies TERM, D/P, and DEF, and this is done with and without the federal funds rate (FFRATE) included. In the monthly regressions for the large stock portfolio, the coefficient on TERM*DIR is positive but not statistically significant at traditional levels. However, the addition causes the statistical significance of TERM to be reduced. DEF continues to be significant, but DEF*DIR lacks explanatory power in explaining large and small stock returns and corporate bond returns. The default spread DEF continues to have explanatory power for large and small stock returns, but not corporate bonds, while the interaction of DEF*DIR is insignificant in forecasting any of the return series. DIR, the proxy for a restrictive monetary environment, continues to have explanatory power in many of the

Table 4 Results of Regressions of Business Conditions and Monetary Policy Proxies on the Expected Returns of Stocks and Bonds: August 1954 through December 1992

Dependent variable	Constant	Term	D/P	DEF	FFRATE	DIR	Adj. R^2	F_{test}
Monthly returns								
(1) Large stock	–0.008	0.242	0.616	0.383	–0.193		0.04	5.93
portfolio	(–0.840)	(2.806)**	(2.577)**	(2.149)*	(–3.210)**			[0.00]
(2) Small stock	–0.012	0.383	0.854	0.526	–0.222		0.04	5.78
portfolio	(–0.947)	(3.219)**	(2.585)**	(2.137)*	(–2.662)**			[0.00]
(3) Bond portfolio	0.004	0.113	–0.039	0.082	–0.017			1.62
	(0.678)	(2.364)*	(–0.297)	(0.829)	(–0.520)		0.01	[0.17]
(4) Large stock	–0.001	0.216	0.294	0.358		–0.011		5.22
portfolio	(–0.074)	(2.456)*	(1.301)	(1.991)*		(–2.749)**	0.04	[0.00]
(5) Small stock	–0.003	0.346	0.476	0.488		–0.015	0.04	5.68
portfolio	(–0.202)	(2.853)**	(1.529)	(1.969)*		(–2.587)**		[0.00]
(6) Bond portfolio	0.007	0.092	–0.090	0.056		–0.005	0.02	2.90
	(1.357)	(1.888)	(–0.724)	(0.570)		(–2.307)*		[0.02]
(7) Large stock	–0.001	0.206	0.529	0.341	–0.163	–0.009	0.05	5.62
portfolio	(–0.151)	(2.356)*	(2.190)*	(1.908)	(–2.634)**	(–2.054)*		[0.00]
(8) Small stock	–0.004	0.335	0.737	0.469	–0.180	–0.011	0.05	5.47
portfolio	(–0.264)	(2.770)**	(2.206)*	(1.900)	(–2.109)*	(–2.015)*		[0.00]
(9) Bond portfolio	0.007	0.092	–0.092	0.057	0.001	–0.005	0.01	2.31
	(1.355)	(1.886)	(–0.684)	(0.570)	(0.032)	(–2.244)*		[0.04]
Quarterly returns								
(1) Large stock	–0.007	0.481	1.277	1.001	–0.527		0.12	5.62
portfolio	(–0.211)	(3.252)**	(1.595)	(2.411)*	(–2.582)**			[0.00]
(2) Small stock	–0.015	0.826	1.754	1.613	–0.527		0.11	5.21
portfolio	(–0.307)	(3.609)**	(1.416)	(2.510)*	(–1.670)			[0.00]
(3) Bond portfolio	0.017	0.161	–0.179	0.357	–0.180		0.03	1.87
	(0.923)	(1.839)	(–0.379)	(1.463)	(–1.498)			[0.12]
(4) Large stock	0.010	0.355	0.464	0.971		–0.029	0.11	4.81
portfolio	(0.292)	(2.145)*	(0.610)	(2.279)*		(–1.960)*		[0.00]
(5) Small stock	–0.005	0.745	0.971	1.642		–0.020	0.10	4.63
portfolio	(–0.099)	(2.915)**	(0.828)	(2.500)*		(–0.864)		[0.00]
(6) Bond portfolio	0.032	0.052	–0.500	0.261		–0.023	0.07	3.31
	(1.693)	(0.548)	(–1.155)	(1.075)		(–2.782)		[0.01]
(7) Large stock	0.010	0.368	1.130	0.865	–0.477	–0.024	0.14	5.07
portfolio	(0.298)	(2.257)*	(1.412)	(2.053)*	(–2.324)*	(–1.620)		[0.00]
(8) Small stock	–0.005	0.758	1.666	1.531	–0.497	–0.014	0.11	4.23
portfolio	(–0.099)	(2.981)**	(1.333)	(2.329)*	(–1.552)	(–0.624)		[0.00]
(9) Bond portfolio	0.032	0.055	–0.314	0.231	–0.133	–0.022	0.07	2.91
	(1.695)	(0.585)	(–0.677)	(0.948)	(–1.123)	(–2.583)**		[0.02]

* Statistically significant at the 0.05 level.

** Statistically significant at the 0.01 level.

Notes: *t* statistics in parentheses; *p* values in brackets.

Table 5 Results on the Stability of the Slope Parameters across Monetary Policy Environments

Monthly returns

Dependent variable	Constant	Term	D/P	DEF	FFRATE	DIR	TERM* DIR	D/P* DIR	DEF* DIR	Adj. R^2	F_{test}
(1) Large stock	0.016	0.106	−0.106	0.438		−0.044	0.286	0.806	−0.299	0.05	4.26
portfolio	(1.134)	(0.844)	(−0.313)	(1.989)*		(−2.314)*	(1.616)	(1.767)	(−0.929)		[0.00]
(2) Small stock	0.012	0.297	0.126	0.810		−0.042	0.163	0.677	−0.557	0.04	3.83
porfolio	(0.588)	(1.715)	(0.268)	(2.164)*		(−1.593)	(0.663)	(1.073)	(−1.113)		[0.00]
(3) Bond portfolio	0.003	0.074	0.025	0.088		0.003	0.030	−0.199	−0.067	0.01	1.79
	(0.335)	(1.055)	(0.132)	(0.585)		(0.267)	(0.312)	(−0.782)	(−0.332)		[0.09]
(4) Large stock	0.014	0.123	0.136	0.558	−0.149	−0.039	0.231	0.740	−0.364	0.06	4.48
porfolio	(0.998)	(0.984)	(0.385)	(2.075)*	(−2.392)*	(−2.037)*	(1.301)	(1.629)	(−1.012)		[0.00]
(5) Small stock	0.009	0.317	0.409	0.834	−0.174	−0.036	0.098	0.600	−0.633	0.05	3.88
porfolio	(0.471)	(1.833)	(0.835)	(2.234)*	(−2.013)*	(−1.357)	(0.399)	(0.951)	(−1.266)		[0.00]
(6) Bond portfolio	0.003	0.073	0.024	0.088	0.001	0.003	0.031	−0.198	−0.066	0.01	1.57
	(0.335)	(1.051)	(0.121)	(0.583)	(0.015)	(0.263)	(0.311)	(−0.779)	(−0.330)		[0.13]

Quarterly returns

Dependent variable	Constant	Term	D/P	DEF	FFRATE	DIR	TERM* DIR	D/P* DIR	DEF* DIR	Adj. R^2	F_{test}
(1) Large stock	0.062	0.324	−0.748	0.747		−0.126	0.209	2.398	0.609	0.11	3.21
porfolio	(1.293)	(1.534)	(−0.654)	(1.296)		(−1.978)*	(0.583)	(1.548)	(0.708)		[0.00]
(2) Small stock	0.045	1.020	−0.424	2.368		−0.103	−0.588	2.154	−1.459	0.10	2.97
porfolio	(0.608)	(3.121)**	(−0.239)	(2.655)**		(−1.094)	(−1.059)	(0.898)	(−1.095)		[0.00]
(3) Bond portfolio	0.003	−0.125	0.307	−0.099		0.023	0.412	−1.205	0.698	0.09	2.83
	(0.125)	(−1.053)	(0.477)	(−0.306)		(0.663)	(2.041)*	(−1.382)	(1.441)		[0.01]
(4) Large stock	0.057	0.395	−0.065	0.782	−0.440	−0.110	0.029	2.144	0.319	0.13	3.41
porfolio	(1.210)	(1.870)	(−0.055)	(1.374)	(−2.057)*	(−1.748)	(0.080)	(1.398)	(0.370)		[0.00]
(5) Small stock	0.038	1.120	0.535	2.417	−0.617	−0.081	−0.840	1.798	−1.866	0.11	3.08
porfolio	(0.524)	(3.414)**	(0.293)	(2.735)**	(−1.859)	(−0.834)	(−1.484)	(0.755)	(−1.396)		[0.00]
(6) Bond Portfolio	0.002	−0.111	0.447	−0.092	−0.089	0.027	0.375	−1.257	0.638	0.09	2.53
	(0.090)	(−0.916)	(0.664)	(−0.283)	(−0.735)	(0.743)	(1.801)	(−1.434)	(1.298)		[0.01]

* Statistically significant at the 0.05 level.
** Statistically significant at the 0.01 level.
Notes: *t* statistics in parentheses; *p* values in brackets.

monthly regressions, particularly for the large stock portfolio. For both small stocks and corporate bonds, we find that DIR is not significant whether FFRATE is included or not.

In the quarterly regressions, we find that only one of the interaction terms (TERM*DIR) has explanatory power in forecasting bond return series. The coefficient on DIR is significant at the 0.05 level in forecasting returns on the large stock portfolio. For the small stock portfolio, we find that both the term spread (TERM) and the default spread (DEF) are significant in explaining quarterly returns.

Overall, from these results, we conclude that monetary policy has explanatory power in forecasting large and small stock portfolio returns, as well as returns on high grade corporate bonds. This is supported by both measures of the stance of monetary policy: the index of change in the discount rate and the federal funds rate. Tests of the stability of the slope parameters across the monetary regimes indicate that the slopes do not change in the restrictive monetary policy environments and that monetary policy continues to forecast large stock returns in most regressions. These results differ from those of Jensen, et al. (1996) in which they cannot determine that monetary policy explains unique variations in security returns beyond that explained by the business conditions proxies. We find that monetary policy has unique explanatory power in forecasting large and small stock monthly portfolio returns, even after controlling for its potential effect through the business conditions proxies. We find that the discount rate change proxy is important in forecasting excess bond and stock returns. After controlling for interaction of this measure and the business conditions proxies, we find it only predicts large stock returns.

SUMMARY AND CONCLUSIONS

We present evidence that the stance of monetary policy has explanatory power for large stocks, small stocks, and corporate bonds. These results confirm earlier findings by Jensen, et al. (1996). Using two measures of monetary policy actions, the federal funds rate and an index based on the change in the discount rate, we show that monetary conditions have explanatory power beyond business conditions proxies. In particular, we find that a restrictive monetary policy stance lowers monthly returns of large and small stock portfolios, and in some cases, corporate bonds.

These results differ from those of Jensen, et al. (1996) in that our business conditions proxies play substantially different roles in explaining variations in expected stock and bond returns, depending on monetary stringency. We do not confirm their findings that only during restrictive monetary policy environments do the business conditions proxies contain significant explanatory power for stocks and bonds. The difference in the findings can possibly be explained by differences in the definitions of the business conditions proxies or by differences in the stock and bond portfolios we examine. If this is the case, it suggests that earlier findings may not be robust to slightly different ways of measuring the business conditions proxies, or they may be sensitive to the particular stock and bond portfolios considered.

Overall, these results indicate that monetary policy actions contain significant information that may be used to forecast expected stock and bond portfolio returns. In addition, we find that information is reflected in the federal funds rate, beyond that indicated by the discount rate changes. This information can be used to forecast stock and bond returns beyond that contained in proxies for the business cycle.

REFERENCES

Banz, R. W. 1981. The Relationship between Return and the Market Value of Common Stock. *Journal of Financial Economics* 9, pp. 3–18.

Bernanke, B. S., and Blinder, A. S. 1992. The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review* 82, pp. 901–921.

Campbell, J. Y. 1987. Stock Returns and the Term Structure. *Journal of Financial Economics,* 18, pp. 373–399.

Campbell, J. Y., and Shiller, R. 1987. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. Unpublished manuscript. Princeton University.

Chen, N.-F. 1991. Financial Investment Opportunities and the Macroeconomy. *Journal of Finance* 46, pp. 529–554.

Chen, N.-F. Roll, R., and Ross, S. 1986. Economic Forces and the Stock Market. *Journal of Business* 59, pp. 383–404.

Dow, C. H. 1920. Scientific Stock Speculation. *The Magazine of Wall Street.*

Fama, E. F. 1981. Stock Returns, Real Activity, Inflation, and Money. *American Economic Review* 71, pp. 545–565.

Fama, E. F., 1990. Stock Returns, Expected Returns, and Real Activity. *Journal of Financial Economics* 45, pp. 1089–1108.

Fama, E. F., and French, K. R. 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics* 25, pp. 23–49.

Fama, E. F., and French, K. R. 1988. Dividend Yields and Expected Stock Returns. *Journal of Financial Economics* 22, pp. 3–25.

Geske, R., and Roll, R. 1983. The Fiscal and Monetary Linkage between Stock Returns and Inflation. *Journal of Finance* 38, pp. 1–33.

Jensen, G. R., Mercer, J. M. and Johnson, R. R. 1996. Business Conditions, Monetary Policy, and Expected Security Returns. *Journal of Financial Economics* 40, pp. 213–237.

Jensen, G. R. and Johnson, R. R. 1995. Discount Rate Changes and Security Returns in the U.S., 1962–1991. *Journal of Banking and Finance* 19, pp. 79–95.

Kaul, G. 1987. Stock Returns and Inflation: The Role of the Monetary Sector. *Journal of Financial Economics* 18, pp. 253–276.

Kiem, D. B., and Stambaugh, R. F. 1986. Predicting Returns in the Stock and Bond Markets. *Journal of Financial Economics* 17, pp. 357–390.

Laurent, R. D. 1988. Testing the 'Spread'. Federal Reserve Bank of Chicago *Economic Perspectives* 12, pp. 3–14.

Romer, C., and Romer, D. 1989. Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz, pp. 121–169. In *NBER Macroeconomics Annual.* Cambridge, MA: MIT Press.

Rozeff, M. 1984. Dividend Yields Are Equity Risk Premiums. *Journal of Portfolio Management* 10, pp. 68–75.

Schwert, G. W. 1990. Stock Returns and Real Activity: A Century of Evidence. *Journal of Finance* 45, pp. 1237–1257.

Shiller, R. 1984. Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity* 2, pp. 457–498.

Waud, R. N. 1970. Public Interpretation of Federal Reserve Discount Rate Changes: Evidence on the Announcement Effect. *Econometrica* 38, pp. 231–250.

This Page Intentionally Left Blank

62

Monetary Policy and Financial Market Expectations
What Did They Know and When Did They Know It?

Michael R. Pakko
Economist, Federal Reserve Bank of St. Louis, St. Louis, Missouri

David C. Wheelock
Research Officer, Federal Reserve Bank of St. Louis, St. Louis, Missouri

On January 31, 1996, the Federal Open Market Committee (FOMC) voted to ease monetary policy, which was widely reported as a lowering of interest rates. Although some interest rates fell with the Fed's action, the declines were generally small, and over succeeding months market interest rates tended to rise. The yield on the Treasury's 10-year note, for example, which had been 5.63 percent on January 30, and which closed at 5.60 percent on January 31, stood at 6.34 percent on March 29, and reached 7.03 percent by June 12. Other rates behaved similarly over this period.

Such seemingly perverse moves in market interest rates have also followed other monetary policy actions, sometimes even on the day those actions were taken. Commonly, Federal Reserve moves to raise or to lower interest rates are followed by changes in market interest rates in the same direction. On May 17, 1994, however, the Fed announced a tightening of monetary policy, which some might expect would cause market interest rates to rise. Instead, many market rates immediately *declined.*

Clearly, the statement that the Fed controls interest rates is, at best, an oversimplification. This article attempts to demystify the relationship between Federal Reserve monetary policy actions and interest rate behavior. Interest rates are set in competitive markets by factors affecting the supply of and demand for individual securities. Monetary policy actions can affect both the supply of and the demand for financial assets, and their effects depend not only on current actions but also on the public's expectations of future policy moves.

We describe in some detail the near-term behavior of government security yields following three recent Federal Reserve policy actions. On the most recent occasion, the Fed's easing action on January 31, 1996, market yields changed little immediately following the

Reprinted from: *Federal Reserve Bank of St. Louis Review* (July/August 1996) 19–31.

policy move, but then yields rose over succeeding months. We contrast this experience with two other events. In early 1994, Fed policy moves to raise interest rates were associated with increases in market interest rates that might be considered greater than justified by the extent of Fed actions. Then, in May 1994, market yields declined following a Fed policy action that was widely interpreted as an effort to raise interest rates. Our review of these episodes reveals how expectations of future monetary policy actions, expectations of the effect of policy on future inflation, as well as nonmonetary influences can cause market interest rates to behave in diverse ways after apparently similar Fed actions.

We begin with a brief description of how the Fed carries out open market policy and the channels through which policy might affect market interest rates. Next, we examine some recent episodes in which market interest rates responded in different ways to Federal Reserve policy moves. Finally, we conclude with a summary of how perceptions of future monetary policy actions affect the behavior of market interest rates in response to current policy moves and hence complicate the assessment of the Fed's credibility as an inflation fighter.

MONETARY POLICY, EXPECTATIONS AND MARKET INTEREST RATES

Open Market Operations and Short-Term Interest Rates

Although Federal Reserve monetary policy is often described in press accounts as the manipulation of interest rates, in fact, monetary policy is carried out mainly by varying the supply of reserves available to the banking system.[1] Open market purchases of Treasury securities by the Fed supply additional reserves, whereas open market sales withdraw reserves.

Banks hold reserves to meet statutory requirements, as well as to meet the payment demands of their customers. A bank with a reserve deficiency might borrow reserves from the Fed, sell securities from its portfolio, or borrow reserves by purchasing federal funds in the interbank reserves market. Similarly, banks with surplus reserves may choose to convert their surpluses into earning assets by acquiring securities or other assets or by selling federal funds. The interest rate that clears the market for federal funds is known as the federal funds rate. The Fed can have a considerable effect on the federal funds rate because its open market operations affect the aggregate supply of bank reserves.

It is generally acknowledged that the Fed has considerable influence on the equilibrium federal funds rate, at least for relatively short periods. But do Federal Reserve operations affect other market interest rates?

The Expectations Hypothesis

The *expectations hypothesis* of interest rate determination states that long-term interest rates will reflect current and expected future yields on short-term securities. For example, the yield on two-year Treasury notes should be the average of the current yield on one-year Treasury bills and the expected yield on one-year bills whose holding period begins one year from now. Interest rate arbitrage ensures that this will occur. If, for example, the interest rate on one year securities that is expected to prevail one year from now would suddenly decrease, arbitrage would cause the current demand for two-year securities to rise. This would tend to lower the market yield on two-year securities to an average of the cur-

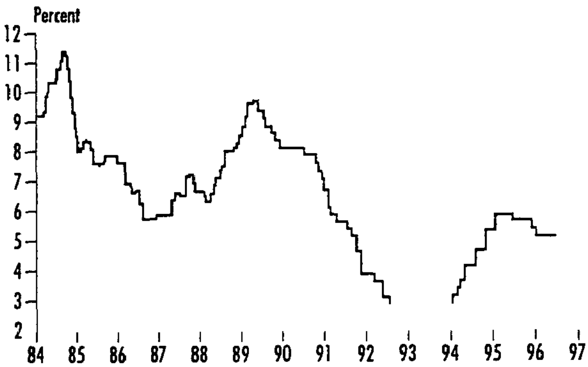


Figure 1 Federal funds rate target (1984–1996).

rent one-year yield and the (now lower) one-year yield expected to prevail one year from now. Similarly, the yield on three-month Treasury bills should reflect the current and expected future path of the federal funds rate over the next three months. As a result, changes in current or expected future short-term interest rates will tend to cause similar movement all along the yield curve.[2]

Because long-term rates are linked to the current and expected future path of short-term interest rates, expectations of future Fed policy moves are important to the movements of interest rates today. It is significant therefore that changes in the Fed's target for the federal funds rate tend to be persistent, with a series of changes accumulating over time. This tendency is clearly illustrated in Figure 1, which shows how the Fed's target has evolved over the past several years.

Financial market participants are well aware of this pattern. For example, after an increase in the federal funds rate target on February 4, 1994, the *Wall Street Journal* reported, "There is little disagreement on where short-term interest rates will be going over the next year: up. The only question is how far they will rise and how fast."[3]

The persistence in federal funds rate changes causes current movements of the funds rate to provide information about future changes. When evaluating the course of short-term interest rates over several months, a current increase (decrease) can be expected to result in further increases (decreases). Because longer-term interest rates are affected by anticipated changes in short-term rates, the yield on a given security might respond to a particular change in the federal funds rate by more than the amount of the funds rate change because the security yield will incorporate the expectation of future changes in the funds rate in the same direction.

Monetary Policy, Inflation Expectations, and the Fisher Relationship

Interest rate arbitrage can explain why market interest rates often move upward when the Fed raises its federal funds target, and downward when the Fed lowers its target. Sometimes, however, market rates fall when the Fed raises its target and rise when the Fed lowers its target. Such apparently perverse changes in market rates can occur because Fed operations are not the sole influences on the supply of and demand for securities. Such changes can also happen because monetary policy is the principal determinant of the long-run rate of inflation—and inflation can have a pronounced effect on interest rates.

Because inflation erodes the purchasing power of money, an increase in inflation causes lenders to require higher interest rates as compensation for receiving future payments in money that has declined in value. Borrowers are willing to pay this *inflation premium* for the same reason. As a result, a fundamental relationship between inflation and interest rates is given by the *Fisher relationship,*

$$i = r + \pi^e, \tag{1}$$

which states that the nominal interest rate (in dollar terms) consists of the following two components: the *real* interest rate (r) and a component that equals expected inflation (π^e).

Thus if market participants interpret a monetary policy action as providing new information about the outlook for inflation, interest rates should change accordingly. This is referred to as the *expected inflation effect* of monetary policy on interest rates. Financial market participants who are interested in the future course of inflation watch Federal Reserve actions closely. If the Fed is viewed as likely to pursue a policy that will prevent significant inflation over time, market yields will be lower. On the other hand, if the public doubts that the Fed is committed to low inflation, then financial markets will reflect fears of future inflation by incorporating an inflation premium in interest rates.

When investors are uncertain about the future course of monetary policy, and hence are uncertain about the future course of inflation, market yields might also be higher than they otherwise would be. For example, although inflation fell substantially in the early 1980s, interest rates remained high, and the difference between the level of market interest rates and the concurrent inflation rate has only recently declined to approximate the difference of the early 1960s. In other words, the *ex post* real interest rate—the difference between the market, or nominal, interest rate and the rate of inflation—was unusually high (see Figure 2).

One interpretation of the high *ex post* real interest rates of the 1980s is that, after experiencing rising inflation from 1965 to 1979, investors feared a return of high inflation and thereby demanded high nominal returns on fixed-income assets. Alternately, if investors viewed the prospects for economic growth favorably, they may have simply demanded higher real returns on fixed-income investments. Still, because disinflations are often accompanied by high *ex post* real rates, both in the United States and in other countries [see, for example, Dueker (1996)], a fear of renewed inflation is a plausible explanation for high real rates in the 1980s.[4]

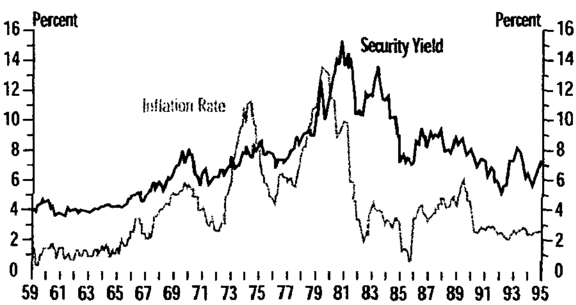


Figure 2 Ten year government security yield and year-over-year CPI inflation (January 1959–June 1996).

Carlstrom (1995) has aptly referred to this effect of Federal Reserve policy on interest rates as a monetary policy paradox. Short-term interest rates can be lowered only by increasing monetary growth, which tends to kindle inflationary expectations and higher interest rates. Lowering interest rates in the long run may require raising them in the short run.

MONETARY POLICY AND INTEREST RATES IN THE SHORT RUN

To evaluate the effect of Federal Reserve policy actions, we focus on the behavior of market interest rates on dates immediately preceding and immediately following recent actions. The Fed made no changes in its target for the federal funds rate during 1993, but on February 4, 1994, the FOMC announced that it had voted to "increase slightly the degree of pressure on commercial bank reserve positions," which it anticipated would increase market interest rates (specifically, the Fed had increased its objective for the federal funds rate by 25 basis points to 3.25 percent).

The official announcement of such a move was unprecedented, and the FOMC stated that it had made the announcement in part because this was the first tightening of monetary policy since 1989.[5] Although it was noted that such a public announcement should not be interpreted as precedent setting, after its meeting on February 2, 1995, the FOMC announced that after each future meeting it would issue a statement indicating whether there would be any change in policy.

By publicly announcing specific policy moves, the FOMC has eliminated uncertainty about its current operational stance.[6] But because the future course of policy remains uncertain, market participants continue to expend considerable effort attempting to forecast upcoming policy actions. Speculation about possible near-term actions often seems to affect the market prices and trading volumes of financial assets as much as actual moves do.

Expectations and Treasury Security Yields

Figure 3 plots the market yields on three U.S. Treasury securities on the date of each announced change in open market policy, that is, change in intended federal funds rate, and each meeting of the FOMC during 1994, 1995, and January 1996. The Fed increased its federal funds target six times in 1994 and once in 1995; the Fed reduced its target twice in 1995 and again on January 31, 1996. The change in basis points, if any, in the Fed's target is noted near the top of each vertical line corresponding to the date of a policy change or FOMC meeting. The market yields on three-month Treasury bills, one-year Treasury bills, and 10-year Treasury notes on each date are plotted, as are the yields five business days before and five business days after the central dates.

Market yields tended to rise during 1994, coincident with the Fed's target rate increases. Yields generally fell in 1995, and the differences in yields of securities with different maturities narrowed. The term structure of yields is often interpreted as revealing market expectations about the future paths of real returns and inflation. Researchers—including Fama (1990), Mishkin (1990), and Estrella and Mishkin (1995)—conclude that yield spreads contain both types of information. Long-term rates tend to be sensitive to inflation expectations, whereas short-term rates follow current and expected real short-term rates more closely. Hence the substantial narrowing in the yield spread across securities of different maturities during 1995 could reflect diminished expectations for real returns, inflation, or both.

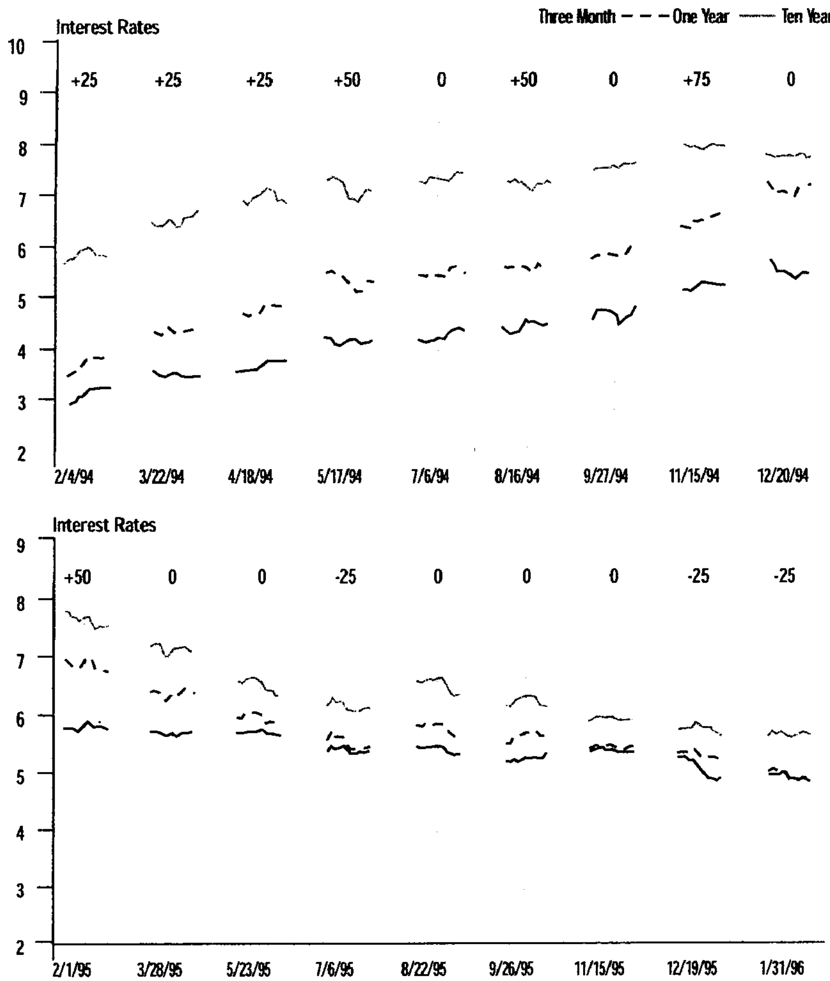


Figure 3 The market response to changes in the Fed funds target and FOMC meetings.

On February 1, 1995, the Fed made the last in a series of federal funds target increases. Although market interest rates rose that day, on subsequent days they resumed a decline that had begun in late 1994. Security yields continued to decline throughout 1995, with the Fed lowering its funds rate target in July and December and again in January 1996.

It is apparent from Figure 3 that when the Fed changes its federal funds target, market rates sometimes, but not always, move in the same direction as the Fed's adjustment. Even when market rates do move in the same direction, they do not move by the same amount as the change in the federal funds rate. A change in expected inflation accompanying a monetary policy action could explain otherwise counterintuitive changes in market interest rates, such as a decline in market rates following a tightening of monetary policy or an increase in market rates following an easing of policy.

In the next sections we examine in more detail the behavior of market rates around three recent episodes of changes in the Fed's target federal funds rate. Knowledge of the extent to which financial market participants anticipated a policy move is important for interpreting each event. Monetary policy actions that are widely anticipated will not convey new information about future inflation, but actions that take markets by surprise may alter forecasts of future inflation. The effect of a policy move on interest rates thus depends on whether the move was expected. One source of information about market expectations of Fed policy moves is the federal funds futures market.

Information from the Federal Funds Futures Market

Since 1988, the Chicago Board of Trade has offered a market in futures contracts based on the federal funds rate. (See the box, *The Federal Funds Futures Market.*) Contracts in this market are based on the monthly average federal funds rate, as reported by the Federal Reserve Bank of New York. The market is used both by financial institutions to hedge their federal funds market positions against changes in the funds rate and by speculators attempting to predict Federal Reserve monetary policy. Because the contracts are based on future monthly averages of the federal funds rate, price movements directly reflect market participants' expectations of policy actions.

Figure 4 shows the accuracy with which the federal funds futures market has predicted actual movements in the funds rate. Both the one-month ahead and two-month ahead rates track the actual funds rate closely, although the two-month ahead forecast fails to predict turning points as accurately as the one-month ahead forecast, lagging behind actual funds rate movements. Nevertheless, Krueger and Kuttner (1995) and Rudebusch (1996) find that one-month, two-month, and three-month future rates are all accurate predictors of subsequent federal funds rate movements.

Information from the federal funds futures market is used in Figure 5 to show expectations of movements in the funds rate implied by futures prices in the days leading up to and following FOMC meetings and policy changes in 1994 and 1995. The figure shows two series of futures yields. One series is the funds rate the market predicts will prevail after the meeting (see the appendix for details of the calculations). The second series is the funds rate derived from a three-month forward contract, indicating market expectations for future levels of the federal funds rate.

Figure 5 illustrates several notable points. First, the three-month ahead futures rate was above the one-month futures rate throughout 1994 and into early 1995. But when the Fed lowered the funds rate in July 1995, its first such move since 1992, the three-month fu-

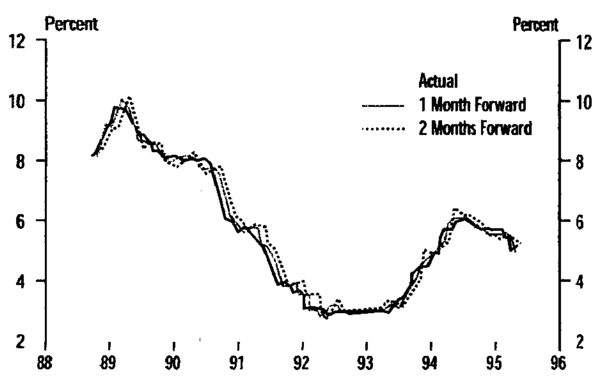


Figure 4 Federal funds futures (Actual, 1 month forward and 2 months forward through 3/96).

tures rate was below both the spot rate and the current month's predicted funds rate. The market had thus correctly forecast the directional change in Fed policy.

The data in Figure 5 also show that many of the Fed's policy actions during 1994 were at least partly anticipated. That is, futures contracts were priced to reflect changes in the federal funds rate before the Fed altered its target. On occasions when it appears that

The Federal Funds Futures Market†

Federal funds futures (formally known as 30-Day Interest Rate futures) have been actively traded at the Chicago Board of Trade since October 1988. The federal funds futures contract is based on the monthly average federal funds rate as reported by the Federal Reserve Bank of New York.

The contract itself calls for delivery of the interest paid on a principal amount of $5 million in overnight federal funds held for 30 days. Contracts are priced in units of 100, with the federal funds rate being 100 minus the price (for example, a price of 92.75 implies a 7.25 percent funds rate). Contracts are settled daily, with the purchaser of a contract paying the seller $41.67 (per $5 million contract) for each basis point increase in the implied federal funds rate (or each 1/100 of a point decline in the contract price) at the close of business. This tick size has been set by using a 30-day month: $5 million × 30/360 × 0.0001 = $41.67.

The following example helps explain the potential hedging use of federal funds futures. Consider a bank that is a consistent buyer of $75 million in federal funds at a current rate of 7 percent. The bank is worried that the federal funds rate will rise in the current month, raising its cost of funds. By selling 15 futures contracts (15 × $5 million = $75 million), the bank stands to profit from the futures transactions in the event that it suffers a loss from a higher cost of funds. For instance, suppose that on the first day of the month, the bank purchases the contracts at 93.00—implying a federal funds rate of 7 percent. If the funds rate immediately rises to 7.2 percent, the bank ends up paying $450,000 in interest on its federal funds purchases over the course of the month [$75 million × .0720 × (30/360)]. However, the buyer of the federal funds futures contract pays the bank $12,501 [15 contracts × 20 ticks × $41.67]. The net cost to the bank is $437,499. The bank's effective cost of funds has been locked in at 7 percent [($437,499/$75 million) × (360/30)].

In addition to banks like the one described in the preceding example—seeking to hedge positions in the federal funds market—futures trade is also carried out by speculators who are betting on a particular course of monetary policy. Each type of trader has an incentive to consider the most likely outcome of monetary policy when deciding whether to participate in a transaction, so the price of federal funds futures represents the market's best estimate of the federal funds rate over the course of the contract month.

† A more complete description of the federal funds futures market can be found in Chicago Board of Trade (1995).

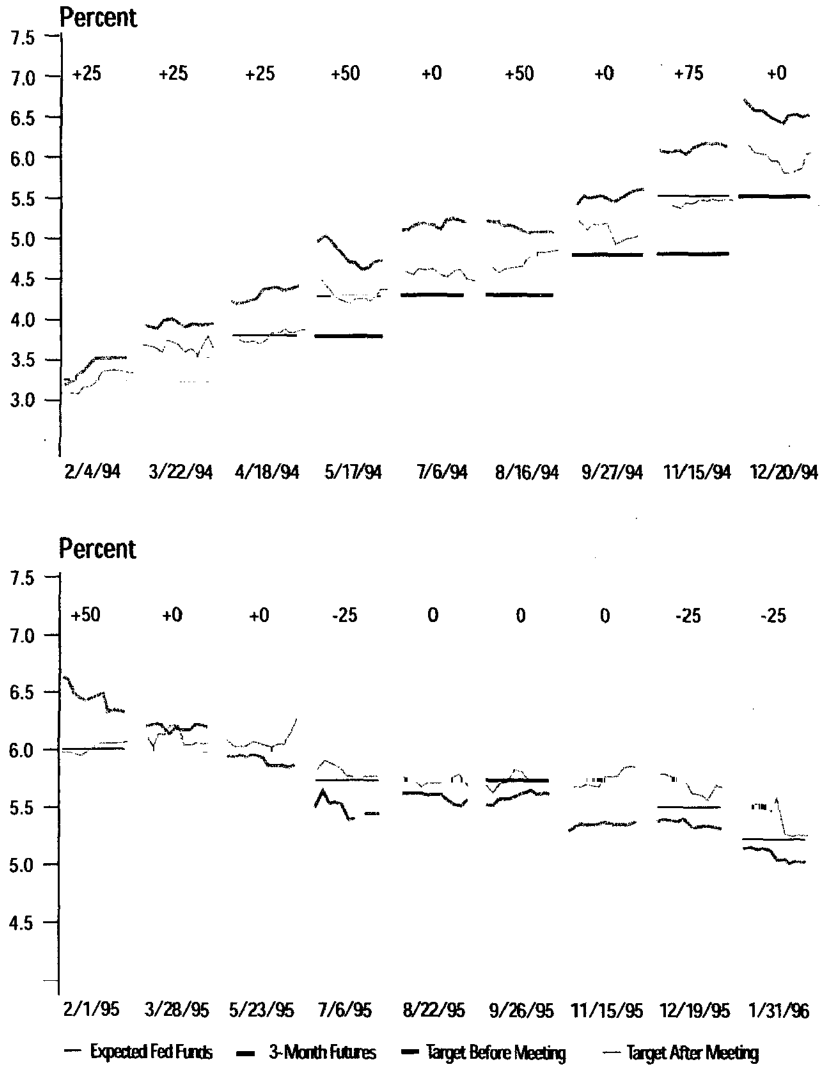


Figure 5 Fed funds expectations derived from the Fed funds futures market.

funds rate changes were not fully anticipated, the three-month forward forecast moved in the same direction as the forecast funds rate for the remainder of the current month. In other words, unexpected changes in the Fed's target led market participants to expect further adjustments to the rate in the same direction as the initial move. The evidence thus indicates that, at least since 1994, the federal funds futures market has forecast specific Fed policy actions fairly well and that futures prices reflect the Fed's tendency to make multiple moves in one direction before reversing course.

Evaluating Market Responses to Specific Monetary Policy Actions

For both policymakers and market participants, the information about expected monetary policy and inflation embedded in interest rates would be useful. As our analysis suggests, however, the effects of monetary policy moves on interest rates can be difficult to disentangle. This difficulty is illustrated by a look at three specific episodes of Federal Reserve policy moves.

February 1994

On February 4, 1994, the FOMC voted to "increase slightly the degree of pressure on reserve positions," resulting in an increase of 25 basis points in the federal funds rate. At the time, some financial analysts claimed that the move took markets by surprise. The move, however, was foreshadowed by Federal Reserve Chairman Alan Greenspan only days earlier. On January 31, 1994, the chairman stated that, "at some point . . . we will need to move [short-term interest rates] to a more neutral stance."[7] This comment was interpreted by some analysts as indicating, "It's a question of when, not whether, they will tighten."[8]

The path of federal funds rate expectations illustrated in Figure 5 makes it clear that the move was anticipated. Beginning on January 31, the expected funds rate rose gradually to the point where the 25 basis point move was almost fully anticipated on the day it occurred. Figure 3 shows that long-term interest rates rose along with the expected federal funds rate. However, bond rates tended to rise by more than the expected funds rate. From January 28 through February 4, the expected federal funds rate rose by 22 basis points, whereas the three-month, one-year, and 10-year Treasury security yields rose by 30, 35, and 26 basis points, respectively.

There are many potential explanations for the larger increases in Treasury security yields. One explanation is rather unique to this particular occasion. It holds that the Fed's policy adjustment was a preemptive move to head off a possible rise in inflation rather than a response to an already-observed increase in inflation. Yet many observers had not seen the emergence of inflation as imminent, so the move was interpreted by some as indicating that the FOMC had information or insight about inflation that was not generally available to the public. Hence inflation expectations were revised upward, and market yields rose.

A related explanation for the large increases in security yields is that the public viewed the relatively small policy move as inadequate to have much effect on incipient inflationary pressures. The market expected a more forceful move from the Fed and in the absence of such a definitive move, revised inflation expectations upward. Either explanation is consistent with the increase in market interest rates that accompanied the Fed's tightening move.

A third explanation—which does not involve any revision to expectations of inflation—seems more plausible, however. Because the FOMC tends to move the federal funds rate in a series of increments, the increase on February 4, 1994 led market participants to anticipate further increases. As a result, long-term rates, which reflect current and expected short-term rates, increased by more than the federal funds rate.

Figure 5 supports the notion that the 25 basis point increase on February 4 led market participants to expect further increases. At the same time that the expected funds rate for February rose in anticipation of the move on February 4, the implied three-month future yield also rose. By the time the February increase in the federal funds rate was announced, the futures market was already predicting another 25 basis point increase within the next three months. This expectation was mirrored in the comments of market analysts at the time: for example, one market observer interpreted the funds rate increase as "the first step on a journey that is going to last some time."[9]

So the behavior of market rates at the time of the Fed's first move to tighten policy could have been caused by an awakening of inflation fears, by the arbitrage effect of current and prospective increases in the federal funds rate, or conceivably by some combination of these effects.

May 1994

After two more increases of 25 basis points each in March and April, the FOMC raised its objective for the federal funds rate by 50 basis points on May 17, 1994. The response in the bond market was the reverse of previous funds rate increases. As the May FOMC meeting approached, long-term bond yields *declined.* After the funds rate increase was announced, bond yields continued to decline. On the day of the funds rate change, the yield on 10-year Treasury notes, for example, fell by 21 basis points.

The decline in bond yields appears to have been directly related to the Fed's move. Reports in the financial press suggest that there was a great deal of uncertainty about the timing and magnitude of the policy move. On the morning of the meeting, a *Wall Street Journal* reporter noted that "several interest-rate watchers expect an increase in rates. The only question is how much?"[10] Figure 5 shows that the federal funds futures market was predicting a high probability of a 50-basis-point increase.

Did the magnitude of the funds rate increase convince market participants that the Fed's anti-inflation strategy would be successful? That is one explanation of the decline in bond yields. That conclusion, however, cannot be drawn with certainty. Once again, the expectations hypothesis suggests an alternate, though not mutually exclusive, interpretation. After the 50-basis-point increase, there was speculation that the FOMC would not have cause to raise the funds rate again in the near future. The official statement released by the FOMC following its meeting contributed to this sentiment: "These actions . . . substantially remove the degree of monetary accommodation which prevailed throughout 1993."[11] A *Wall Street Journal* writer interpreted this statement as being quite clear: "Yesterday's declaration means that the Fed now believes it is very close to neutral and doesn't expect any further rate increases soon."[12] To the extent that bond market participants lowered their expectations of further increases in the funds rate, the expectations theory of interest rates would predict a decline in bond yields, even if inflation expectations remained unchanged.

The reaction of the federal funds futures markets gives some credence to this view. As shown in Figure 5, the implied rate on three-month futures was falling for a period both before and after the meeting. Nevertheless, it continued to indicate that at least one more increase of 25 basis points was likely within the next three months. Hence it is unclear whether the bond market's reaction to the policy move on May 17, 1994, indicated a reduction in expected inflation, a change in the short-term outlook for Fed policy, or both.

January 1996

A third example serves to show the dynamic nature of market expectations and their responses to Federal Reserve policy. On January 31, 1996, the Fed voted, in effect, to reduce its target for the federal funds rate by 25 basis points, from 5.50 percent to 5.25 percent. (At the same time, the Fed lowered the discount rate from 5.25 percent to 5.00 percent.) According to the financial press, the Fed's action was widely expected and the rise in short-term security prices in preceding days reflected anticipation of the move.[13] Between January 1 and January 30, 1996, market yields on short-term Treasury securities fell some 20 to 30 basis points. The yields on government securities with maturities of seven years or more, however, did not fall over the period.

Government security yields did decline, but only modestly, after the Fed's cut in its funds rate target on January 31. Although the Fed reduced its target by 25 basis points, market yield declines ranged from eight basis points on three-month bills to just one basis point on 30-year bonds. Yields on short- and medium-term securities continued to decline through mid-February, however, but those on long-term government securities changed little—some even increased. Then, from mid-February through March, yields on all securities rose. For illustration, the daily yields on three-month, one-year, three-year and 10-year Treasury securities are plotted in Figure 6.

How might we interpret the behavior of interest rates both before and after the Fed's reduction in its funds rate target on January 31, 1996?

The modest changes in interest rates that occurred on January 31, support the press's view that the Fed's action had been widely anticipated. Further evidence of this can be seen in Figure 7, which plots the expected average federal funds rate in different months using data from the federal funds futures market. On January 30, the funds rate the market expected to prevail during February lay between the prevailing Fed target of 5.50 percent and the new target of 5.25 percent established on January 31. That the expected rate lay closer to the new target indicates that on January 30 the market believed that the Fed was more likely than not to reduce its target on January 31. When the Fed validated these expectations, the expected funds rate for February fell immediately to 5.25 percent.

The data charted in Figure 7 also illustrate that on January 30 the futures market expected not only the funds rate cut on January 31, but also further cuts from March through July. After the Fed reduced its target, these expectations only hardened.

Further evidence that the Fed's action on January 31 was widely anticipated is reflected in the lack of change in intermediate- and long-term Treasury security yields on that date. The failure of long-term yields to change significantly on the Fed's easing move is thus consistent with the behavior of short-term rates, the federal funds futures market, and the financial press, all of which suggest that the Fed's move was widely anticipated.

Between mid-February and March 31, 1996, market interest rates generally rose. As illustrated in Figure 6, rates made two especially large jumps in mid- and late February and one more in early March. Throughout the period, new data suggested that the economy was growing more quickly than some previously released indicators had suggested. Moreover, in mid-February, rising commodity prices suggested to some market participants that in-

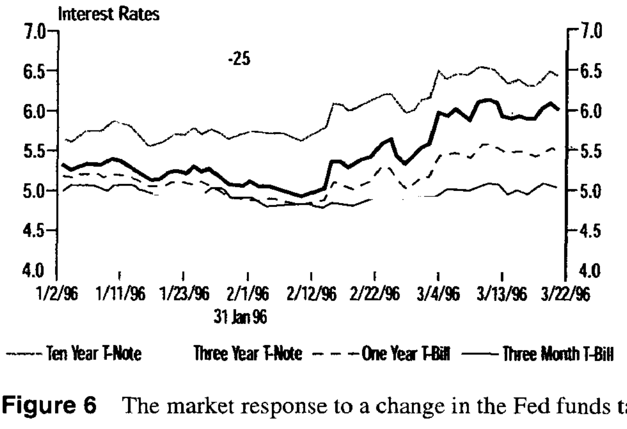


Figure 6 The market response to a change in the Fed funds target.

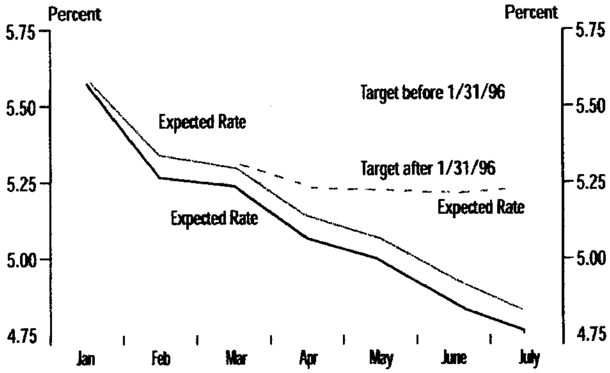


Figure 7 Fed funds futures market implied expected funds rate.

flation was likely to rise, causing market security yields to rise.[14] Although yields rose across the spectrum of maturities, long-term security yields rose most. This pattern of rate changes suggests that the new information caused market participants to revise their expectations of the Fed's target for the federal funds rate upward over ensuing months, and possibly expectations of inflation as well.

Market interest rates again rose when Federal Reserve Chairman Greenspan testified before Congress about monetary policy and the state of the economy on February 20, 1996, which many analysts interpreted as confirmation that additional funds rate reductions over the near term were unlikely. Finally, the release of new employment data on March 8, 1996, revealing an unexpectedly large increase in employment during February is widely cited for a sharp increase in bond yields on that date. According to one report, "The carnage [in the bond market] began immediately after a stronger-than-expected employment report snuffed out hope that Federal Reserve policymakers would lower short-term interest rates anytime soon."[15]

The evolution of expectations about the course of Fed policy was reflected in the federal funds futures market. In addition to the expected future funds rate path implied by market pricing on January 30 and January 31, Figure 7 plots the implied path based on futures market data from March 8. In contrast to the earlier dates, when further funds rate cuts were expected, on March 8 the market expected the funds rate to remain at 5.25 percent through July 1996.

According to the expectations hypothesis, the rise in long-term interest rates on March 8 reflected the expectation that short-term rates would rise in the future. The increase in long-term rates could also reflect a revised anticipation of higher inflation in the future, though other explanations, such as an increase in the real interest rate, could also explain the rise. Inevitably, because many factors affect the supply of and demand for securities, any one move in market yields can have several non–mutually exclusive explanations. Nevertheless, the behavior of market rates after January 31, 1996, is consistent with, first, a period of relative calm in which markets anticipated further reductions in the Fed's interest rate target, with little apparent change in inflation expectations. Then, following new information about the health of the economy and new speculation about Fed behavior, markets changed their expectations about the near-term course of monetary policy and perhaps revised their expectations of future inflation upward.

CONCLUSION

Evaluating the credibility of monetary policy by observing bond market reactions can be difficult. Sometimes market rates rise when the Fed's target is raised, and sometimes they fall. Sometimes rates move by more than the change in the funds rate and sometimes by less. These responses can be interpreted as an amalgam of inflation expectations, anticipated future monetary policy actions, and changes in real rates of return.

Although these influences are difficult to disentangle, the information from the federal funds futures market can help identify the role of expectations in the determination of market interest rates. Specifically, with an understanding of the extent to which a Fed policy action is anticipated in financial markets, we can better interpret subsequent changes in market interest rates.

Throughout 1994 and 1995, however, the behavior of the federal funds futures market suggests that most Fed actions were at least partly anticipated. Moreover, the Fed's tendency to move its target for the federal funds rate incrementally in one direction before reversing course is built into market expectations of future policy actions, as revealed in both the spot markets for Treasury securities and the federal funds futures market. The incremental nature of Fed policy moves, along with interest rate arbitrage, likely also explains why market interest rates typically moved in the same direction as changes in the federal funds rate during 1994–95. When a policy move is widely anticipated, and particularly if it is expected to be one of many in a series of moves in the same direction, market expectations about inflation are not altered. Only surprise moves, or moves that are widely taken as turning points, will typically alter expectations about inflation.

NOTES

1. The Fed also sets the discount rate, which is the rate charged banks when they borrow reserves from the Fed, and required reserve ratios, that is, the percentage of their deposit liabilities that banks are required to hold in the form of vault cash or deposits at Federal Reserve Banks. Neither is changed frequently, however, and open market policy is the principal mechanism by which the Fed conducts monetary policy.
2. See Campbell (1995) for more detail about the term structure of interest rates and empirical evidence on the expectations hypothesis.
3. Thomas T. Vogel, *Wall Street Journal,* February 7, 1994, p. C1.
4. See Dotsey and DeVaro (1995) for empirical evidence suggesting that much of the disinflation of the early 1980s was unanticipated by the public.
5. See Pakko (1995) for a detailed description of FOMC policy moves during 1994 and Gavin (1996) for a discussion of policy moves during 1995.
6. Thornton (1996) finds that financial market volatility has been lower around the time of FOMC meeting dates since the policy of announcing federal funds rate changes was implemented.
7. Statement before the Joint Economic Committee, United States Congress, January 31, 1994. *Federal Reserve Bulletin* (March 1994, p. 233).
8. Joseph Liro, chief economist at S. G. Warburg, quoted by Thomas D. Laurencella and Laura Young, *Wall Street Journal,* February 1, 1994, p. C23.
9. John Lipsky, chief economist at Salomon Brothers, quoted by Thomas T. Vogel, *Wall Street Journal,* February 7, 1994, p. C19.
10. Dave Kansas, *Wall Street Journal,* May 17, 1994, p. C2.
11. *Federal Reserve Bulletin,* July 1994, p. 610.
12. David Wessel, *Wall Street Journal,* May 18, 1994, p. A3.

13. See, for example, Dave Kansas, *Wall Street Journal,* January 31, 1996, p. C1.
14. For example, see Dave Kansas, *Wall Street Journal,* February 15, 1996, p. C1.
15. Vogelstein and Jereski, *Wall Street Journal,* March 11, 1996, p. C1.

REFERENCES

Campbell, J. Y. Some Lessons From the Yield Curve. *Journal of Economic Perspectives* (Summer 1995), pp. 129–52.

Carlson, J. B., McIntire, J. M. and Thomson, J. B. Federal Funds Futures as an Indicator of Future Monetary Policy: A Primer Federal Reserve Bank of Cleveland *Economic Review,* vol. 31, no. 1 (1995 Quarter 1), pp. 20–30.

Carlstrom, C. T. A Monetary Policy Paradox. *Economic Commentary,* Federal Reserve Bank of Cleveland (August 15, 1995).

Chicago Board of Trade. Flexible Futures For Managing Interest Rate Risk. 1995.

Dotsey, M. and Devaro, J. L. Was the Disinflation of the Early 1980s Anticipated? Federal Reserve Bank of Richmond *Quarterly Review* (Fall 1995), pp. 41–59.

Dueker, M. When Are Low-Inflation Policies Credible? Federal Reserve Bank of St. Louis *Monetary Trends,* January 1996.

Estrella, A., and Mishkin, F. S. The Term Structure of Interest Rates and Its Role in Monetary Policy for the European Central Bank. NBER Working Paper #5279, 1995.

Fama, E. F. Term-Structure Forecasts of Interest Rates, Inflation, and Real Returns. *Journal of Monetary Economics* (January 1990), pp. 59–76.

Federal Reserve Bulletin, various issues.

Gavin, W. T. The FOMC in 1995: A Step Closer to Inflation Targeting? this *Review* (forthcoming).

Krueger, J. T., and Kuttner, K. N. The Fed Funds Futures Rate as a Predictor of Federal Reserve Policy. Federal Reserve Bank of Chicago, working paper WP-95-4, March 1995.

Mishkin, F. S. What Does the Term Structure Tell Us About Future Inflation? *Journal of Monetary Economics* (January 1990), pp. 77–95.

Pakko, M. R. The FOMC in 1993 and 1994: Monetary Policy in Transition. Federal Reserve Bank of St. Louis *Review* (March/April 1995), pp. 3–25.

Rudebusch, G. D. Do Measures of Monetary Policy in a VAR Make Sense? Working paper, Federal Reserve Bank of San Francisco (March 1996), #96-05.

The Wall Street Journal, various issues.

Thornton, D. L. Does the Fed's New Policy of Immediate Disclosure Affect the Market? this *Review* (forthcoming).

This Page Intentionally Left Blank

63

The October 1987 Crash Ten Years Later

Robert T. Parry
President and Chief Executive Officer, Federal Reserve Bank of San Francisco, San Francisco, California

This chapter is adapted from remarks delivered by Robert T. Parry, President and Chief Executive Officer of the Federal Reserve Bank of San Francisco, at a conference sponsored by the Graduate School of Management at the University of California, Davis, on October 17, 1997, entitled "The October '87 Crash: What Have We Learned about the Causes and Consequences of Large Market Movements?"

On October 19, 1987, "Black Monday," the Dow Jones Industrial Average plunged 508 points—the largest one-day drop in history. The next day, Tuesday, raised the spectre of a disintegration of the market and with it a very serious threat to the functioning of the entire U.S. financial system. While the "whys and wherefores" of this event remain the subject of research and debate, this chapter focuses instead on the Fed's role in responding to the crisis. Specifically, it discusses why the Fed intervened after the crash, what we did to help stabilize the market, and what lessons we learned.

THE FED'S ROLE IN FINANCIAL CRISES

A stock market crash raises a couple of interrelated policy issues for a central bank: one has to do with our role as monetary policymaker, and the other with our role in ensuring the safety and soundness of financial markets and the payments system—that is, as "lender of last resort."

From the monetary policy point of view, we were concerned that the loss of wealth due to the crash might cut consumer spending and lead to an economic contraction. When the market bottomed out, the loss of wealth owned by individuals amounted to nearly eight hundred billion dollars, and in theory, this could have reduced people's spending on consumer goods. It's true that the response of consumption to a given change in wealth has always been estimated to be relatively small. But the size of *this* crash meant that the wealth effect might have had an important impact on overall economic activity.

Reprinted from: *Federal Reserve Bank of San Francisco Economic Letter,* No. 97-32 (Oct. 31, 1997).

As it turned out, the "wealth effect" *wasn't* a key threat. Looking back, it appears to have had nearly imperceptible effects on spending. One reason may have been that stock prices right after the crash were actually above the levels of the previous year.

Fortunately, we didn't need to wait for the published data—which come in with a substantial lag—to know that we weren't dealing with an overall slowdown in economic activity. Through the District Banks and their branches, the Fed has a nationwide network of contacts that includes current and former Directors on our Boards as well as a number of Advisory Council members. And the message from virtually all of them was that the crash wasn't having a big effect on regional economic activity. Indeed, in the year after the crash, GDP growth was quite strong.

Instead, the greater threat was to the viability of the exchanges and brokerage firms, and, by extension, to the perfectly sound businesses that might have failed if the exchanges had collapsed. Addressing this threat has a lot in common with the Fed's historic role as lender of last resort in preventing banking panics. Like banks, the various intermediaries in the stock and bond markets held relatively small amounts of capital. This made them vulnerable to sudden withdrawals by lenders.

When firms have problems because of their own business decisions, it's clearly *not* the Fed's role to step in and help out. In fact, doing so would only create a moral hazard, inducing firms to take excessive risks and leading to instability in the financial system as a whole. But when the effects of a bank run or a break in the stock market spread to fundamentally sound firms and threaten the stability of the financial system—that is, when there's *systemic* risk—then the Fed has a clear and important role to play.

HOW SYSTEMIC RISK AROSE IN 1987

We usually think of stock exchanges as highly liquid markets, largely because of the financial intermediaries in these markets who stand between buyers and sellers by guaranteeing the execution of transactions. For example, stock exchange specialists *must* buy into falling markets in order to serve as a shock absorber. The system works very well to maintain liquidity when buy and sell orders aren't too far out of balance. Ordinarily, a small stock of inventories relative to gross trade flows is enough to bridge the gap between buy and sell orders. Furthermore, clearinghouses can guarantee the execution of trades in the face of any individual's default risk.

But on October 19, order flows were grossly out of balance—there were virtually no other buyers. That left the specialists at the end of the day with much larger inventories than usual. They had to pay for those purchases within five business days and needed credit to do so. In addition, investors had to meet margin calls as prices fell, and brokerage firms extended credit to many of their customers to enable them to do so. Finally, the solvency of the clearinghouses was in doubt, because the default risk they were insuring was systemic. Although the clearinghouses were well capitalized, they weren't able to bear this kind of risk, and on the morning of October 20 there was a real possibility that they might fail.

At this point, then, all these players needed *additional* credit to continue their functions. But banks were growing nervous and reluctant to lend. And who can blame them? The drop in asset prices cast doubt on the creditworthiness of all parties: investors, specialists, brokerages, and clearinghouses. Furthermore, the reluctance of banks and other creditors to issue further loans itself increased the risk of default. So there was a genuine risk that expectations of a market melt-down would become self-fulfilling.

HOW THE FED INTERVENED

Before the market opened on October 20, the Fed issued the following announcement: "The Federal Reserve System, consistent with its responsibilities as the nation's central bank, affirmed today its readiness to serve as a source of liquidity to support the financial and economic system." This affirmation of the Fed's responsibilities to serve as lender of last resort was intended to reverse the crisis psychology and to guarantee the safety and soundness of the banking system.

The Fed backed up this announcement with a number of critical actions, and I'll highlight four of the most important.

First, we added substantially to reserves through open market operations. The funds rate fell from 7-1/2 percent just before the crash to 6-1/2 percent in early November. This added liquidity helped prevent the crash from spreading to bond prices.

Second, we liberalized the rules governing the lending of securities from our own portfolio to make more collateral available.

Third, we used all of our contacts in the financial system to keep the lines of communication clear and open. In talking with banks, for example, we stressed the importance of ensuring adequate liquidity to their customers, especially securities dealers, and at the same time affirmed that they were responsible for making their own independent credit judgments. We also were in close touch with participants and regulators in the government securities market, officials at the various exchanges and their regulators, and our colleagues at central banks in other countries.

Finally, as a means of gathering real-time information, we placed examiners in major banking institutions to monitor developments—such as currency shipments—to identify the potential for bank runs.

To sum up, performing the lender-of-last-resort activity, and backing it up with close monitoring and close communication, did what it was supposed to do: it transferred the systemic risk from the market to the banks and ultimately to the Fed, which is the only financial institution with pockets deep enough to bear this risk. This allowed market intermediaries to perform their usual functions and helped keep the market open.

WHAT DID WE LEARN?

As our mothers sometimes tell us, we learn more from our mistakes than from our successes, and the policy lesson here really comes from the Great Depression, rather than from the October 1987 crash. During the Depression, the Fed failed to avert the collapse of the banking system, and the loss of intermediary services was one reason why the Depression was so deep and so prolonged.

So the crash of 1987 and the Fed's swift and decisive response serves to reaffirm our understanding of what we need to do. While this should give confidence to the markets, I think it's worth repeating that it should be used sparingly. Such Fed actions must be limited to crises marked by systemic risk. Bailing out individual firms is *not* the job of the Fed, nor is it in the public interest since it would induce excessive risk-taking in the private sector.

Let me conclude with one last caveat that brings me back to monetary policy concerns: once the critical stages have passed, the Fed needs to be especially careful not to generate *another* set of problems—that is, we must be careful not to *overplay* an easier policy stance. That could create inflationary pressures that would menace the process of healing in the financial system and possibly create future economic crises.

I think it's fair to say that the Fed did not make this mistake following the '87 crash. In the face of a strong economy, tight labor markets, and an upward creep in inflation, monetary policy was tightened noticeably between early 1988 and early 1989. The funds rate rose from a low of around 6-1/2 percent in late 1987 to just under 10 percent in early 1989. While some people have debated about whether this action contributed to the 1990–1991 recession, I think it is clear that it helped set the stage for the decline in inflation that has occurred in the 1990s.

64

Bank Deposits and Credit as Sources of Systemic Risk

Robert A. Eisenbeis
Senior Vice President and Director of Research, Federal Reserve Bank of Atlanta, Atlanta, Georgia

There are many different ways to define a financial crisis. Indeed, the economics and finance literature is filled with terms like panic, financial crisis, runs, systemic crisis, or contagion.[1] There is in fact little agreement on even the rudimentary definitions of a financial crisis, the sequence of events constituting a crisis, or the causes of these events.

The professional discussion divides itself into two broad categories. Macroeconomists typically are concerned with explaining business cycle fluctuations and determining when a recession will degenerate into a depression.[2] They are equally interested in the financial system's role as a propagator of this process because most depressions have been accompanied by serious disruptions in the financial system, including banking failures and panics. Eichengreen and Portes, for example, define a financial crisis as "a disturbance to financial markets, associated typically with falling asset prices and insolvency among debtors and intermediaries, which ramifies through the financial system, disrupting the market's capacity to allocate capital within the economy. . . . Our definition implies a distinction between generalized financial crisis on the one hand and bank failures, debt defaults and foreign-exchange market disturbances on the other" (1991, 10).

Financial economists examine the micro behavior of market participants to explain disruptions in financial markets (see Diamond and Dybvig 1983; Chari and Jagannathan 1984). They have tended to focus on banking panics and runs and the reasons depositors withdraw funds rather than on the macro consequences for employment and output in the real economy per se.

While differing in their emphases, the micro and macro approaches to analyzing financial stability share several themes. The first focuses on alternative explanations for why a crisis occurs. One prominent thesis argues that the financial system is inherently unstable and is therefore vulnerable to random shocks. Shocks simultaneously cause market participants to lose confidence in the system and exchange their bank deposits for currency. Others believe that such herd behavior cannot be explained solely by shocks that, like animal spirits, randomly induce depositors to run from bank deposits to currency. They offer more behaviorally oriented explanations and models, the most prevalent being models based on

---

Reprinted from: *Federal Reserve Bank of Atlanta Economic Review* (Third Quarter 1997) 4–19.

the existence of information asymmetries between borrowers and lenders. These models attempt to show that it is sometimes rational for depositors to attempt to withdraw their funds in such a way that it creates a run on the banking system.

Most of the analysis in the random shock and information asymmetries models concentrates on aggregate behavior, assuming essentially that all market actors—both depositors and institutions—are identical. It does not admit differences among depositors and institutions or even the presence of more than one institution in the financial system. When the analysis recognizes more realistic features of market and financial structure, researchers are better able to examine the process by which a shock or problem in one part of the country or sector of the economy is transmitted to other sectors or the system as a whole. These transmission models, representing the second main theme in the literature, have not been the focus of much empirical work and tend to be relatively undeveloped.

The third area investigates the causes of financial crises and their impact on the real economy. For example, do financial crises cause declines in real economic output, or are they instead manifestations of deeper problems in the real economy? What are the channels of transmission? Do deposit runs cause liquidity problems, which in turn induce contractions in lending, thereby affecting real output and production? The final area of analysis examines the role of government policies—both macro and micro—in generating financial crises as well as lessening their potential severity.

The remainder of this article explores these issues in more depth. The discussion gives particular attention to the possible linkages between deposits and credit availability as the transmission mechanism for crises since runs on deposits and payments system disruptions are believed to be transmitted to the real economy through a credit channel.

RANDOM SHOCKS AND INHERENT FINANCIAL INSTABILITY

At the macroeconomic level, models such as those proposed by Minsky (1982) and Kindleberger (1978) embody the claim that the banking system is inherently unstable. Minsky argued that a capitalist economy, and especially its banking system, is inherently unstable. Furthermore, this instability is endogenous, originating within the system itself. He defined instability as "a process in which rapid and accelerating changes in the prices of assets (both financial and capital) take place relative to the prices of current output" (1982, 13).

Simply stated, Minsky assumed that during relatively stable times firms engage in balanced financing, by which he meant that cash flows are sufficient to cover principal and interest payments. However, as the economy grows and enters the expansion phase of the business cycle, firms begin to reach for profits, presumably because of management's preference for short-term gains. Firms start to leverage up, and banks, in particular, begin to shorten the maturity structure of their liabilities relative to their assets. Expanding returns by funding long-term investments with short-term borrowing is driven by the desire to take advantage of an upward-sloping term structure with long-term interest rates exceeding short-term rates.[3] This period of leveraging, which Minsky labels a period of speculative finance, is still one of relative stability.

Cash flows from investment are still sufficient to cover principal payments as debts. This speculation ultimately degenerates into what Minsky calls a period of Ponzi finance, in which cash flows cover neither principal nor interest payments. Debt refunding requires new debt issuance, the proceeds of which are used to cover required interest and principal debt payments. During this period, an exogenous shock will result in a collapse of both the financial system and the real economy. The shock, which can come from many different sources, serves as the trigger for collapse. Minsky was silent on the exact mechanisms by which this happens.

Commenting on Kindleberger's (1978) similar view, Schwartz observes that "those who regard banks as inherently unstable assume no connection between monetary policy and the price conditions under which economic agents make decisions. Proponents of inherent instability see a recurring historical pattern in which many bankers abandon conservative standards of asset management during business expansions only to be caught short when booms collapse. For them instability resides in economic agents. Benevolent government then comes to the rescue. This is the central thesis offered by Charles P. Kindleberger in his 1978 book" (1986, 11).

Minsky puts forth certain stylized facts that would be observed, although they are not the outcome from any specified model.[4] The first is that, during an expansion, credit expands at rates that exceed the growth of income or the capital stock. Second, interest rates and nominal asset prices must be increasing at a rapid rate. Third, debt maturities must become shorter, and, fourth, some exogenous shock must occur to cause a change in expectations. Finally, governments must fail to intervene in ways that cushion any asset reevaluations accompanying any changes in expectations.[5]

Bernanke and James (1991) suggest a different view of the causal relationship. For them, a precipitating force that could lead to a financial collapse is a deflation. Deflation adversely affects credit quality by reducing borrower equity cushions. When companies finally default, intermediaries become owners of illiquid, real assets. To reliquefy their balance sheets, banks are induced to reduce lending and call in loans. Those banks that are unable to reliquefy fail, and, by implication, deposits will be destroyed. In this scenario, credit problems lead to a reduction in bank deposits, contracting the money supply.

The chief distinction from the picture Minsky and Kindleberger paint is that Bernanke and James see banks as passive bystanders in the process. They are not required to take on more risk, nor do they have to misprice risk or adjust their balance sheets to take on more interest rate or maturity risk. The model also suggests that crises occur only during and after an exogenous shock has induced a deflation. Bernanke and James are careful to argue, however, that while deflation is a necessary condition for a crisis to occur, it is not sufficient. They highlight several aspects of banking structure that, if present, also help increase the likelihood that financial institutions would experience a crisis. These include (a) lack of branch banking, (b) universal banking and the commingling of banking and commerce, and (c) funding though short-term, foreign deposits. Thus, banking and financial structure can either mitigate or accentuate the likelihood that a financial crisis will result during a deflationary period.

Unlike the macroeconomists' models discussed, the random shock models of the financial economists, most closely associated with Diamond and Dybvig (1983), look more deeply at the structure of the deposit contract and the process by which it is redeemed.[6] Because deposits are payable upon demand at par, they offer depositors nearly costless liquidity, provided that not all depositors wish to withdraw their funds at the same time. With sequential servicing, in which depositors are treated on a first-come, first-served basis, depositors, especially if they are geographically dispersed rationally know that not everyone can withdraw simultaneously. If bank loans are inherently not marketable, or cannot be easily liquefied, then at the first hint of potential trouble, it is rational for depositors to step to the head of the line rather than incur costs to determine whether and exactly when deposits will be paid.

These micro random shock models pay no particular attention to the source, or nature, of the random shock that causes depositors to line up. Depositors just decide to run, and once they do, all depositors run. These models also do not consider the credit side of the balance sheet as a factor in crises, other than the fact that loans are less liquid than deposits so that bank cannot pay all claims in currency. Nonetheless, they make it easy to see that shocks affect depositors' willingness to hold bank deposits, and, when that willingness is reduced, a contraction in credit follows as loans must be liquidated to meet the deposit-redemption demand.

The Diamond and Dybvig model approximates the situation that prevailed in early U.S. history. Individual banks issued their own bank notes to the public promising to redeem these notes at par for specie. Since note issues typically were not backed 100 percent by specie, periodic liquidity problems arose whenever noteholders became concerned that a bank might not be able to honor its redemption commitment and suspend convertibility of deposits into specie. Runs on individual banks and the system sometimes occurred, and these resulted, albeit infrequently, in cumulative contractions in the money stock.[8] Suspension of convertibility of deposits into specie was a common way for early banks to deal with temporary liquidity problems. It often resulted, however, in a decline in purchasing power since the value of deposits declined. By shifting the cost of nonconvertability at least temporarily to the creditors (depositors) of the bank, they gave all liability holders an important incentive to worry about bank solvency.

Diamond and Dybvig (1983) investigate the suspension of convertibility as one equilibrium solution to the problem of runs, but they do not consider the price level effects or how the costs of suspension of convertibility are distributed because their model has only consumption good and no currency. Another weakness of their model is that there is only one bank in the system, and hence runs are on the banking system as a whole and involve flights to the currency rather than runs on one of many banks in the system.[9]

For these early banks, avoidance of runs mean maintaining public confidence. Depositors needed believe that the institution could convert notes into specie in sufficient amounts and would not need to suspend convertibility.[10] Indeed, the first forms of public regulation designed to deal with the problems involving suspension of convertibility imposed reserve requirements specifying permissible ratios of notes to specie. The regulations sought to assure public confidence by requiring banks to engage in minimal maturity intermediation, maintain sufficient specie reserves, and have adequate capital and liquidity.[11]

INFORMATION ASYMMETRIES MODELS

The micro random shock models have been less than satisfying, both because they appear generally inconsistent with economic events, as will be discussed in the next section, and because many economists find it hard to believe that people randomly decide to run without some just cause rooted in economics. Recent modeling efforts have applied concepts of information asymmetries to derive conditions that might make it rational for depositors to engage in runs on banks. Under the information asymmetries models, banks are viewed as being "opaque" to depositors and thus costly for depositors to monitor. With imperfect and costly information, a type of Akerlof (1979) lemons model applies in which depositors have a great deal of difficulty distinguishing between healthy and unhealthy banks. Any shock or news event that might induce depositors to reassess their bank's riskiness (in combination with the sequential servicing constraint) will cause depositors to assume that all banks are riskier than previously believed. Under these circumstances, it is more rational for depositors to withdraw funds than to seek out and evaluate costly information or risk losing their funds by not withdrawing. In these models, as in the micro random shock models, the source of the shock is not specified, in that no particular cause is suggested for a failure. But usually it is hypothesized that the shock originates in credit markets and in releases of relevant news about bank asset quality. The model's predictions are consistent with the view that shocks are more likely to result from disturbances in the real sector than from the default of a single borrower.

Macroeconomists have articulated a form of this same asymmetric information hypothesis in attempting to counter the inherent instability arguments. As Schwartz describes it, "a widely held belief in the United States and the world financial community is that the default of major debtors—whether companies or municipalities or sovereign countries—could lead to bank failures that would precipitate a financial crisis. . . . A financial crisis is fuelled by fears that means of payment will be unobtainable at any price and, in a fractional-reserve banking system, leads to a scramble for high-powered money. It is precipitated by actions of the public that suddenly squeeze the reserves of the banking system. In a futile attempt to restore reserves, the banks may call loans, refuse to roll over existing loans, or resort to selling assets. . . . The essence of a financial crisis is that it is short lived, ending with a slackening of the public's demand for additional currency" (1986, 11).[12]

Under this scenario, a banking crisis is precipitated by the failure of a major debtor, which induces a sudden shift in the public's demand for currency. In turn, banks scramble for reserve assets by curtailing lending and selling assets. By implication the decline in lending and refusal to roll over existing credits leads to a decline in economic output. The process becomes systemic in that problems in one or several major creditors raise questions about the quality of bank assets in general and induce the public to switch to holding currency.

The hypothesis implies a direct linkage between increased demand for money and the availability of credit—and hence the ability to finance and maintain the real economy. The information the public perceives is not assumed to be bank specific; instead, it is the fact that the information concerns the quality of banking assets in the aggregate that increases the private sector's demand for currency relative to deposits. The channel envisioned in this scenario results in banks calling in loans and building up liquidity to meet the public's desire for currency.

Empirically, three elements are necessary for this view to hold. First, there must be a credit-related shock that affects the public's desire to hold currency relative to deposits. Second, this shock must induce a liquidation of deposits for currency by the public. Third, bank credit must contract.

There are several important differences between the various random shock models and asymmetric information models. First, Minsky's random shock model includes leveraging up of both bank and corporate balance sheets across the board, and, furthermore, it does not require an inflationary environment. Second, the collapse that results is not driven by runs forcing institutions to liquefy balance sheets to meet deposit withdrawals. Third, under this type of model financial institutions accommodate the leveraging up of balance sheets by underpricing credit risk. They also take on more interest rate and maturity risk by shortening the maturity structure of liabilities relative to assets. Fourth, no interdependence among either borrowers or lenders is necessary for a collapse to take place. Finally, the direction of causation, in terms of propagators of the crisis, appears to run through credit channels by eroding depository institution real equity values. Only if institutions fail is the money supply affected.

In the random shock model of Diamond and Dybvig (1983), the crisis does not result from asset mispricing or from rational economic behavior but rather from an exogenous event. Since there is only one bank in the economy in this model, runs take the form of flights to currency (or more precisely, the consumption good) and not to other healthy banks. The panic is due solely to the existence of the sequential servicing requirement discussed above and the fact that bank assets are not perfectly liquefiable.

Like the micro random shock models, the asymmetric information models do not rest upon systematic ex ante asset mispricing or other problems of bank behavior. Changes in expectations and market assessment of bank asset quality, combined with the opaqueness of bank balance sheets and sequential servicing make runs a rational customer response.

EMPIRICAL EVIDENCE ON SYSTEMIC RISK

When examined carefully, many of these alternative explanations of panics and financial crises appear to overlap, differing only slightly in their details. Separating them empirically can therefore be very difficult. Empirical tests of various hypotheses about financial crises and panics have generally focused on the National Banking Era and the period of the Great Depression. The reason for studying these periods is that no broad-based panics have occurred since (in part because of the existence of federal deposit insurance and lender-of-last-resort actions followed by the Federal Reserve). In this section, the empirical evidence is examined to determine which of the models appear to be more consistent with the data.

The question of whether this empirical work provides useful insights or is relevant today is a legitimate one, given the changes in financial structure and markets, the rise of technology, the proliferation of information, and the globalization of markets. This issue will be addressed in the next section.

The Random Shock and Financial Fragility Hypothesis.

Given the lack of precision in specifying the models, does the evidence suggest that one or more of the models may be correct? With respect to the macro models, critics of the Minsky financial fragility hypothesis argue that it does not yield testable hypotheses and is inconsistent with the data (see Sinai 1977; Lintne 1977; Mishkin 1991; Schwartz 1986). As mentioned previously, for the hypothesis to hold, a sequence of several factors must be present: debt burdens increasing faster than the growth of income or capital stock, interest rates and nominal asset prices increasing rapidly debt maturities at depository institutions becoming shorter, an exogenous shock occurring to cause change in expectations, and, finally, governments failing to intervene in ways that would provide a soft landing to any asset revaluation that must accompany the change in expectations.

Unfortunately, data do not readily exist for examining a number of the conditions Minsky sets forward. As an alternative Table 1 lists the periods of economic recession with information on when panics took place and, where possible, what possible shocks may have existed. Looking first at the timing of the panics relative to the peaks and troughs of the business cycles shows that in only one instance was there a panic before the peak of the business cycle. In most cases, the panic occurred anywhere from three to six months after the business cycle had peaked. Such long lags would seem to be logically inconsistent with Minsky's view.

Mishkin (1991) devotes considerable attention to the rate pattern and to risk premiums and their relationship to the onset of panics. In general, the spread between rates on high- and low-quality bonds rose before the panic began. However, these spreads generally widened after the recession started rather than before as the Minsky hypothesis would require.[13]

The Asymmetric Information Hypothesis versus Micro Random Shock Models

Gorton (1988), Mishkin (1991), and Donaldson (1992) specifically investigate the information asymmetric hypothesis in detail. Examining the National Banking Era and the post–Federal Reserve Era through 1933, Gorton models depositor behavior in terms of the currency/deposit ratio. He poses two questions. First, if panics are random events, then is the model predicting a different currency/deposit ratio during panic periods than exists in other times? Second, are panics predictable in terms of movements of perceived risk? From these two questions Gorton suggests the following testable hypothesis: "Movements in variables predicting deposit riskiness cause panics just as such movements would be used to price such risk at all other times. This hypothesis links panics to occurrences of a threshold value of some variable predicting the riskiness of bank deposits" (1988, 751). Such p redictive variables might be extreme seasonal fluctuations, unexpected failure of a large corporation (usually a financial corporation), or a major recession.

A third question Gorton asks is whether certain predictors of risk stand out as important predictors of panics. Finding no evidence that panics are random events, he concludes that there is strong support for the asymmetric information hypothesis. Furthermore, panics appear to be predictable ex ante. Evidence also suggests that recession, and not a triggering bank failure, is the critical factor in determining whether a panic will occur. Gorton explains: "the recession hypothesis best explains what prior information is used by agents in forming conditional expectations. Banks hold claims or firms and when firms begin to fail (a leading indicator of recession), depositors will reassess the riskiness of deposits" (1988, 778). In short, causation seems directed from the real sector to the financial sector rather than vice versa.

Donaldson (1992) extends Gorton's analysis using a somewhat different specification of the model and weekly data between 1867 and 1907 to determine whether panics are systematic and predictable events. Unlike Gorton, Donaldson rejects the conclusion that panics are systematic events and argues that the data are more consistent with the random shock model than the asymmetric information model.[14] However, for the panics of 1914 and 1933 (which required expansion of the money supply during crisis periods), he finds behavioral patterns of earlier panics had been dampened. Given that the later panics followed the creation of the Federal Reserve in 1913 and passage of the Aldrich-Vreeland Act of 1912, this finding suggests that government involvement to increase liquidity can truncate panic situations. He concludes that panics are therefore special events. But he also finds evidence that panics are more likely to occur when seasonal and cyclical factors are present.

Mishkin (1991) formulates the asymmetric information hypothesis somewhat differently. He argues that key variables help to capture differences in depositor assessment of bank risk. In particular, during periods of financial distress high-quality firms will be less affected and lenders will have less uncertainty about the riskiness of such firms than they will have for low-quality firms. To the extent that these risks are priced, an impor-

Table 1 *Panic Dates and Causes*

Panic date	Business cycle peak and trough	Percentage (and number) of national bank failures	Prepanic interest rate movement	Percent change in currency-to-deposits ratio from previous year's average	Possible prepanic exogeneous shock
October 1857	Fall 1853–July 1857	NA	Rates fell until the recession began. Spread on bonds did not widen until after the onset of the recession.	NA	With failure of Ohio Life and Trust, reserves were pulled from New York City banks. First bank failures occurred in September. Several railroads went bankrupt in September, and major runs on New York banks in October culminated in specie suspensions in mid-October.
September 1873	October 1873–March 1879	2.8 (56)	Rates rose about 5 percentage points in August, five months before the beginning of the recession. Spread on bonds did not rise until the month of the panic and did not rise prior to that.	14.5	Crisis began when New York Warehouse and Security Co. failed on September 8. Other failures and suspensions followed: Kenyon, Cox & Co., Jay Cooke & Co., and Fisk & Hatch. Panic-selling on the New York Stock Exchange led to closing of the market for ten days. The initial failures appeared related to debt problems and problems with railroad bonds.[a]
June 1884	March 1882–May 1885	0.9 (19)	No obvious pattern preceded the panic.	8.8	On May 6 Marine National Bank failed. The Wall Street

			With the exception of a three-month period, spreads on bonds declined steadily for two years prior to the panic.		brokerage firm Grant and Ward was linked to a bank that failed on May 8. That failure was followed by a run on Metropolitan National Bank and suspension of several other banks. However, an inflow of foreign capital and the issuance of clearinghouse notes moderated the panic. It appeared the clearinghouse notes provided a signal to the market of bank solvency.
No panic	March 1887–April 1888	0.4 (12)	—	3.0	—
November 1890	July 1890–May 1891	0.4 (14)	Rates did not rise appreciably until after the recession had begun. Spread on bonds was essentially flat for a year preceding the panic.	9.0	New York Stock Exchange prices began falling in early November. On November 11 Decker, Howell & Co. failed, involving the Bank of North America. On November 12 a stock broker failed, and on November 15 Baring Brothers failed in London.
May 1893	January 1893–June 1894	1.9 (74)	Rates rose beginning in January 1892 approximately 2.5 percentage points in the seven months preceding the recession. Spread was essentially flat for more than one year preceding the recession.	16.0	On February 26 the Philadelphia & Reading Railroad went into receivership, and on May 4 National Cordage Co. failed and a stock market crash followed. New York banks weathered the situation until banks in the West and

continues

Table 1 Continued

Panic date	Business cycle peak and trough	Percentage (and number) of national bank failures	Prepanic interest rate movement	Percent change in currency-to-deposits ratio from previous year's average	Possible prepanic exogeneous shock
					South experienced runs and began withdrawing reserves from New York City banks to meet liquidity needs. In August there was a general suspension of specie payments. National banks were reopned after examination and certification by the Comptroller of the Currency.
October 1896	December 1895– June 1897	1.6 (60)	Rates rose only about 75 basis points in the three months preceding the peak of the expansion but rose substantially in the three months before the panic. Spread on bonds did not widen until after the beginning of the recession and peaked just prior to the panic.	14.3	The period of 1895–96 was a mild, paniclike period. Although the New York Clearing House Association made emergency credits available in the form of loan certificates, none were used.
No panic	June 1899–December 1900	0.3 (12)	—	2.8	—
No panic	September 1902–August 1904	0.6 (28)	—	–4.1	—

October 1907	May 1907–June 1908	0.3 (20)	Rates were flat preceding the recession and rose only slightly thereafter. Spread on bonds was essentially flat prior to the beginning of the recession.	11.5	Stock market declined in October. During the week of October 14, five New York members of the New York Clearing House Association and three outside banks required assistance. These banks had been used to finance speculation in copper-mining stocks. On October 12 Knickerbocker Trust Co. (third largest in New York City) began to experience clearing problems, and it suspended operations on October 22.
No panic	January 1910–January 1912	0.1 (10)	—	–2.6	—
August 1914	January 1913–December 1914	0.4 (28)	—	10.4	This crisis was linked to problems in London and disruptions to payments on discount bills by foreign (European) borrowers. London stopped discounting foreign bills, and the effect was to disrupt New York banks, who were in a net debt position during the summer (as apparently was usual). New York banks were forced to remit gold, draining reserves. Both London and New York Stock Exchanges closed on July 31, and a panic threatened.

[a] Sprague ([1910] 1968) notes that all the failure swere due to criminal mismanagement or to neglect or violation of the National Banking Act and not to questions about bank solvency. Source: Gorton (1988) and Schwartz (1986), except for Panics of 1857 and 1893 (Mishkin 1991).

tant index of asymmetric information uncertainty should be captured by the spread between the rates on high- and low-quality bonds, by stock prices (as a measure of net worth and collateral value), and by interest rates (as a measure of agency costs and adverse selection). His analysis, like that of Gorton (1988), supports the information asymmetries hypothesis to the extent that the proxy variables are in fact good proxies. He concludes that most financial crisis periods begin with an increase in interest rates and a widening of the spread between high- and low-quality bonds and a decline in stock prices, rather than with a panic. "Furthermore," Mishkin observes, "a financial panic was frequently immediately preceded by a major failure of a financial firm, which increased uncertainty in the marketplace" (1991, 97). He also asserts that the information hypothesis offers a better explanation than the macro theories of financial fragility for the pattern of rate spreads and stock market movements both before and after a panic as well as the panic's likely occurrence.

Finally, Park (1991) argues that the provision of bank-specific information overcame the information asymmetries that played a role in runs on banks. In particular, by analyzing the panics of 1873, 1884, 1893, 1907, and 1933 he concludes that clearinghouse and government intervention were effective devices in settling panics but only when they provided information on bank-specific solvency.[15] In the panic of 1884, a run was abated following certification of solvency by the Comptroller of the Currency and by subsequent extensions of clearinghouse certificates to Metropolitan National Bank, which was the bank suffering the greatest withdrawals.

The panic of 1893 followed a long period of depression during which banks suffered prolonged periods of withdrawals of gold and uncertainty about U.S. adherence to the gold standard. Gold hoarding culminated in suspension of convertibility, and repeal of the Sherman Silver Act was promised by the president. Banks lifted the suspension of convertibility, and the runs stopped. Because no systematic attempt was made to release information on individual banks, public confidence in all banks remained low until the source of uncertainty—lack of confidence in U.S. maintenance of the gold standard—was removed. Park (1991) interprets the Comptroller of the Currency's certification of individual bank solvency before their reopening following the panic of 1893 as the major information factor that quelled depositor uncertainty.

In the panic of 1907, the problem began with runs on individual New York banks and trust companies that had been directly or indirectly associated with a failed attempt to corner the market in copper stocks. Only intervention by the New York Clearing House Association, which attested to the solvency of banks experiencing runs and provided financial assistance, resolved the situation. Again, release of firm-specific information appeared to have addressed the information asymmetries and helped stabilize the crisis.[16] Unlike other cases, in the panic of 1907 runs did not affect all banks, and, indeed, some New York Clearing House member banks experienced reserve inflows (Park 1991).

Transmission Models

Neither the basic random shock models nor the information asymmetry models specifically address the issue of which mechanisms transmit panics or financial crises through the economy. In fact, no models admit more than one institution, a condition that would be necessary to model customers simply transferring funds from an unhealthy to a healthy institution as distinct from retreating to currency.[17] The models provide no information on what, if any, real impacts such funds transfers among banks have. Nor have the models addressed when depositors will run on one bank and when they will run on the entire system.

Researchers have addressed the question of transmission mechanism more indirectly by attempting to generalize from the basic models. For example, Calomiris and Gorton (1991) maintain that it is the sequential servicing constraint imposed in the Diamond and Dybvig–type models that can induce banks to run on other banks. Such runs are especially likely when banks are geographically dispersed but are permitted to count interbank deposits as legal reserves, as under the National Banking system. Two other regulatory constraints—restrictions on branching and on the payment of interest on interbank deposits—have also been regarded as important.[18]

The structure of the National Banking system prior to creation of the Federal Reserve in 1913 added a further source of instability to the economy. Under that system, legal reserves for National Banks included not only cash in vault but also deposits in Reserve City and Central Reserve City banks. In such a fractional-reserve banking system that has pyramiding of reserves, a run on an individual bank could more easily have systemic, systemwide effects. Shocks originating in the countryside, for example, could induce country banks to improve their liquidity positions by recalling interbank deposits from the Reserve City and Central Reserve City banks. Hence, panics were also endemic to the structure of the system as a whole, and it is clear how a panic or run in a rural region could blossom into a systemic crisis for healthy banks in Reserve Cities and Central Reserve Cities. Chari (1989) addressed this issue directly in considering a model of spatially separated banks. He argued that the most likely source of a shock that would cause country banks to withdraw reserve funds was seasonally related, with differences in currency demands rising significantly during planting and harvest times.

Calomiris and Gorton (1991) attempted to determine specifically whether panics were transmitted from rural areas through the National Banking system, as the analysis suggested, and also whether the patterns were more consistent with the random shock or information asymmetries models. They found that three important differences between the models have empirical implications.

The first concerns the origin of problems. The random shock model suggests that shocks would occur in rural areas because of seasonal demands for currency. In contrast, the asymmetric information model implies that adverse economic news related to asset-quality problems would precede a panic.

Second, the two theories would seem to predict different patterns of failures during a crisis. The asymmetric information model suggests that banks whose asset portfolios were closely linked to the specific shock would be more prone to failure whereas the random shock model would predict that failures would be experienced in the areas suffering currency withdrawals.

Finally, the models differ in the conditions required to resolve a crisis. In the random shock model, the key to resolving a panic is liquefication of assets. In the asymmetric information model, it is the effectiveness of mechanisms initiated to resolve depositor uncertainty about bank solvency.

Calomiris and Gorton's exhaustive investigation of the sources of panics between 1873 and 1907 led them to reject the idea that seasonal money-demand shocks were the cause of banking panics. Rather, their analysis suggests that panics originated in bad economic news and bank vulnerability to that news. Moreover, their inspection of failure patterns shows virtually no support for the random shock model. Finally, they conclude that in terms of resolving crises, the mere availability of currency, which would provide the ability to liquefy assets, was not sufficient to stop panics during the periods studied. Again, this conclusion suggests that the asymmetric information model was more consistent with the data than was the random shock model.

Smith (1991) provides some specific evidence on country banks' behavior vis-à-vis their holdings of cash reserves as compared with reserves held in the form of interbank deposits when panics occurred. He provides analysis of some anecdotal and other evidence, derived mostly from Sprague ([1910] 1968), about the behavior of Reserve City banks during the crises of 1873, 1893, 1907, and 1930–33.

Smith describes the situation leading up to the panic of 1873, indicating that interbank deposits were concentrated in seven of the New York City banks. These interbank deposits constituted about 45 percent of the sources of funds for the New York banks and were the base upon which their bond holdings and loans were built. These banks were clearly vulnerable to demands by country banks for withdrawal of reserves, and funds were especially tight in the few months before the crisis. When the key triggering events occurred (see Table 1), a combination of circumstances made the crisis severe. In addition to having virtually no excess reserves, several of the banks were in weak financial condition. As subsequent events would prove, several had been the victims of fraud and defalcations, probably accounting in part for their financial weakness. Clearly, however, the institutions' problems stemmed primarily from reserve withdrawals and their inability to call in loans in that economic environment rather than from major credit problems in the New York Central Reserve City banks.

The Reserve City banks experienced similar problems caused by currency outflows during the panics of 1893 and 1907. Thus, it seems clear that reserve outflows, coupled with the lack of excess currency reserves at the Central Reserve City banks in New York and Chicago, forced contractions in loans and finally resulted in the suspension of currency payments. Smith notes that currency suspension was the prime transmission mechanism of panics once a triggering mechanism occurred. He also concludes that the problems during the 1930–33 period originated in the rural agricultural areas as well and were intimately intertwined with the correspondent banking system.

Despite a fairly clear pattern in the transmission mechanism of panics emanating from large reserve deposit withdrawals (rather than from uncertainties about credit quality in Reserve City banks, as the Minsky hypothesis would imply), a number of questions remain. For example, Tallman (1988) indicates that looking at the data over longer time periods does not suggest a clear linkage between the incidence of panics and either increases in currency demand relative to deposits or contractions in loans. He presents evidence that loan contractions occurred at several intervals during the period between 1893 and 1907, for example, that exceeded the declines during periods when panics occurred. Similarly, during some periods of time between 1873 and 1930 the number of bank failures far exceeded those observed during panic periods. Finally, Tallman provides aggregate data on the growth in loans relative to high-powered money and on the growth of manufacturing output between 1873 and 1914. Two observations are important. First, loans do increase in the years prior to panic periods, but the panics occur after loan growth has fallen significantly. Second, numerous periods during the interval show the same patterns in loan and output growth and decline but are not accompanied by a panic. These aggregate data do not reveal whether there are differences in the loan-contraction periods in terms of their concentration in particular parts of the country during episodes of panic and nonpanic periods.

Causal Direction

The research evidence seems to indicate fairly consistently that the dynamics between financial panics and changes in real economic output begin in the real sector and move to the financial sector rather than starting in the financial sector. There are no examples in U.S. history of the economy operating at high levels of output when a financial crisis occurred that resulted in a contraction in the real economy. As the discussion in the previous sections suggests, however, banks were sometimes under pressure and were forced to call in loans. It seems reasonable to assume that once problems in the financial sector become severe, there could be negative feedback effects to the real sector.

Indeed, Bernanke (1983) has made precisely this point. Financial crises can have real effects outside the normal reserve/loan transmission mechanism because of the disruptions to the intermediation process. Bank failures disrupt borrower/lender relationships and make attaining financing more difficult and costly. But this observation should not obscure the fact that financial crises are better viewed as creatures of recession and economic downturns rather than primary causal agents precipitating the downturns.

The Role of Government

A substantial body of evidence indicates that government actions have played significant roles in contributing to crises as well as in mitigating them. For example, Sprague ([1910] 1968) notes that lack of access to a reliable lender of last resort to provide short-term liquidity can help escalate a period of financial tightness into one of crisis. Friedman and Schwartz (1963) argue that several Federal Reserve actions during the Great Depression contributed to both its duration and magnitude. For instance, they observe that the Federal Reserve's failure to liquefy the assets of many small nonmember institutions (the Fed was not obligated to lend to nonmember banks), together with its insistence that it would lend only upon sound collateral, added to the number of bank failures. This policy, in conjunction with the Fed's attempt to adhere to the rules of the gold standard, contributed to a 33 percent decline in the money supply and clearly exacerbated the severity of the recession.

While it has become fashionable to criticize the Fed for its policy failures during the Great Depression, it is also the case that government interference affected financial soundness long before the Fed was created. For example, during the National Banking Era the monetary base was tied, except for a period of suspension, to gold and silver specie monies through the Treasury.[19] When the United States adhered to the gold standard, fluctuations in the gold supply expanded and contracted the monetary base, directly affecting banks' lending behavior. Decisions about how much in the way of international gold flows would be permitted before conversion could be suspended was a matter of Treasury and government policy. European central banks, and to a lesser extent the U.S. Treasury, often intervened to prevent loss of gold reserves by raising short-term interest rates. Government policies frequently exacerbated gold flows and, by implication, induced fluctuations in the monetary base. For example, following passage of the Sherman Silver Purchase Act in 1890, foreigners' concern that the United States would remain on the standard precipitated gold outflows and contributed to the panic of 1890.[20]

Tallman and Moen note that each panic after 1897 was preceded by unusual gold flows. They conclude that political uncertainties concerning the U.S. commitment to the gold standard were important influences on gold flows and, hence, the U.S. monetary base.

Political conditions outside the United States also affected gold flows. For example, in 1907 the Bank of England responded to problems in the London money markets by raising its discount rate to stem potential speculative outflows of gold to the United States. At that time London was the most important market for discounting U.S. trade bills. The increase in the discount rate not only disrupted the flow of gold to the United States but also discouraged the discounting of trade bills and caused a liquidity crisis in the United States.

Since the debacle of the Great Depression, U.S. intervention in markets has often had as its objective providing liquidity to avoid a crisis. Numerous examples exist of emergency liquidity having been provided through the efforts of the Federal Reserve either directly or indirectly, such as during the Penn Central scare, the Chrysler problem, the collapse of Drexel-Burnham, and the failure of Continental Illinois Bank, to name just a few. The Federal Reserve has on occasion attempted to provide liquidity not only to cushion problems in interbank markets but also to prevent disruptions in other markets.

RELEVANCE IN TODAY'S WORLD

It can be argued that inferences about financial crises and systemic risk drawn from study of the banking situation during the National Banking Era and early 1900s are not particularly meaningful or relevant in today's economic environment. Pyramiding of legal reserves in private banks is not a structural feature of the present reserve requirement regime. Markets are no longer isolated, and information-availability problems that might have resulted in the past in information asymmetries have been reduced significantly. Communications technology and new instruments have increased the liquidity of all banking assets and have given rise to new markets that make the kinds of liquidity crises that occurred in the National Banking Era unlikely today. Furthermore, the United States has abandoned the gold standard, and thus the domestic money supply is not subject to the random fluctuations and shocks that it was vulnerable to under strict adherence to the gold standard rules. Deposit rate ceilings of the 1930s have been phased out, and branching restrictions, which essentially prevented institutions from achieving geographical diversification, are a thing of the past.

Certainly the focus on protecting the money supply from sudden shocks is no longer of prime policy concern for three reasons. First, it seems unlikely that significant runs to currency will occur (see Kaufman 1988). Deposits still have large advantages over currency for a variety of purposes, and there are many banks to chose from. Runs on individual banks would simply transfer reserves from one institution to another. Second Federal Reserve policy is likely to provide emergency liquidity to prevent such runs from disrupting other institutions. Finally, while still accounting for the bulk of payment items, checks and currency are no longer the dominant forms in terms of dollar volume of transactions in the economy. The concerns and risks have shifted to other sectors of the payments system that did not exist during the National Banking Era.

Today, the payments system is larger, has many more components (both private and public), and is subject to different risks than in the past. The checks demand deposit system, which accounts for the bulk of individual payments (except for currency), and the one that the present regulatory structure was primarily designed to protect, is small in terms of the dollar volume of payments. The rest are made in the form of computerized transfers of reserve balances on the Federal Reserve's Fedwire system and the privately owned Clearing House Interbank Payments System (CHIPS) and in the form of automated clearinghouse (ACH) transactions. Payments on the former two systems account for about 85 percent of the dollar value of transactions. Closely related to these systems are the automated transfers of book-entry Treasury securities which also take place on Fedwire and involve substantial volumes of transactions.

Finally, as markets have become increasingly global, timing differences and differences in clearing and settlement conventions can add temporal and other dimensions to credit risks not always found in the domestic markets that characterized earlier times. Many other significant sources of uncertainty can also be identified in the clearing and settlement processes in modern financial markets (see, for example, Eisenbeis 1997 and McAndrews 1997).

Maintaining the integrity of payment flows is a substantially more complicated and difficult problem today than protecting the stock of demand deposits for a number of reasons. First, given the large size of transactions in the system and the size of the system itself, the resources required to support unwinding even a short-run problem may be enormous and could exceed the capacity of private participants to self-insure. Second because the transactions are electronic and occur instantaneously, monitoring them and the net position of each participant is critical to controlling participants' credit risk exposure. Third, when the international activities of U.S. banks and the links between the U.S. domestic payments system and foreign banking organizations are recognized, it becomes difficult to conceive of ensuring domestic financial stability without also ensuring international financial stability.

Clearly, different types of uncertainties exist with respect to systemic risk exposures today than existed in the past. There is also reason to believe that liquidity problems for borrowers may be significantly different than they were for borrowers in the 1800s. The growth of new mortgage lending instruments and, particularly, the development of home equity lines of credit provide ways for borrowers instantaneously to liquefy previously illiquid assets during tight times. While this ability to liquefy assets more easily may enable borrowers to maintain payments on outstanding debts and lessen the severity of the credit component of a recession, it also suggests introduction of a new discontinuity that might systematically transfer risks to the banking system at a critical trigger point. If during times of financial distress borrowers draw down lines on home equity and similar lines of credit and are then forced into default, the burdens of these defaults will be shifted to the providers of the home equity lines. Should these losses be large, capital might be wiped out, with few options available to lenders to avoid the costs of those defaults. Examples of similar impacts in commercial and real estate lending markets occurred when commercial paper borrowers drew down banks' back-up commitments during the Penn Central and Real Estate Investment Trust (REIT) crises.

SUMMARY AND CONCLUSIONS

This article has investigated the various theories of financial panics and crises with particular emphasis on the links between credit and deposits. The survey suggests that panics are not random events, as some of the theories may suggest, but neither are they perfectly predictable. Nevertheless, it does appear that information asymmetries about the ability to liquefy deposits were a major contributing factor to banking panics in the past. Moreover, while panics do appear to be associated with recessions and deflationary periods, the direction of causation seems to run from the real sector to the financial sector rather than the other way around. It is not that financial crises cannot exacerbate economic declines; rather, they are not primary causal agents of recessions.

The analysis also suggests that government policies can affect the likelihood of a financial crisis as well as play a role in its solution. These considerations are as relevant today as they have been historically. At the same time, the article raises a cautionary note that the dynamics of crises and how they might play out may be significantly different in the future given recent, rapidly developing changes in the U.S. and world financial system.

NOTES

1. For representative examples see Smith (1991), Kaufman (1995), Donaldson (1992), Bartholomew, Moe, and Whalen (1995), and Eichengreen and Portes (1991). See Benston and Kaufman (1995) for a review of the evidence on fragility.
2. Eichengreen and Portes (1991) require declines in real output for a true financial crisis to occur.
3. Before 1910, however, the most common yield curve in the United States was downward-sloping.
4. It is generally argued that the theory as put forth by Minsky is not a unified theory that yields testable hypotheses. See, for example, Sinai (1977), Lintner (1977), Mishkin (1991), and Schwartz (1986).
5. Minsky argues that the ability to intervene is directly correlated with the size of government: and big government, with its revenue capacity, has the resources to support, through fiscal and monetary policies, a longer run-up of leverage. Also, through its lender-of-last-resort capabilities, it can soften the landing during an exogenous shock period by supporting a gradual rather than precipitous liquidation of assets. It thereby avoids the corresponding collapse of credit, bank failures, and destruction of the money supply.
6. For other examples of models in this mode see Haubrich and King (1984), Cone (1983), Jacklin (1987), Wallace (1988), Bhattacharya and Gale (1987), Smith (1991), and Chari (1989).
7. In the Diamond and Dybvig (1983) model there is really no nonbank money in circulation. Individuals deposit a real consumption good in the bank in exchange for a deposit or warehouse receipt. This consumption good is close, but not identical, to specie.

   In early U.S. banking, it was not uncommon for notes issued by out-of-area banks to trade at discounts, which reflected several factors, including transportation and transaction costs, lack of information on the issuing bank, and uncertainties about the creditworthiness of the issuing bank. This lack of par clearance in no way affected the ability of state bank notes to function as money.
8. For discussions of the evidence on runs see Kaufman (1988) and Gorton (1987).
9. Because of the way the model is constructed, runs necessarily have an adverse impact on the real economy.
10. For a discussion of these early bank runs see Kaufman (1988) or Bryant (1980).
11. Clearinghouses and other banks in the region often provided temporary credit to institutions experiencing liquidity problems (see Kaufman 1988). Kaufman (1994) notes that bank capital ratios were substantially higher during this period than they were after deposit insurance was introduced.
12. Although Schwartz articulates this view, she clearly does not believe it is correct or that the policies designed to protect against the events are appropriate.
13. The exception is the panic of 1873.
14. As a robustness test, he also reruns the analysis using monthly data as Gorton does and gets similar results to those found by Gorton. He concludes that monthly data are too spaced out to provide a sharp test of the hypothesis.

15. A more complete test of the Gorton-Mishkin-Park hypothesis about information asymmetries would be provided by examining fund flows from individual solvent and insolvent institutions. Relying upon aggregate statistics can be only circumstantial, not conclusive.
16. The Roosevelt administration, following the declaration of the bank holiday on March 6, 1933, employed this same policy.
17. Smith (1991) does provide a model in which banks are permitted to hold funds at a Reserve City bank. Bhattacharya and Gale (1987) provide a model with geographically dispersed depositors and banks. Again, however, these models only look at the interdependence among banks through the interbank deposit markets.
18. See also the discussions in Haubrich (1990), Bordo (1986), and Williamson (1989). All emphasize the advantages over U.S. banks that banks in Canada and other countries that permitted branching had in weathering panics. Calomiris and Schweikart (1991) have explored in detail for the United States the effects that structure had on failure rates in different states with different branching statutes. They show that branch banks had both lower failure rates and in general paid lower premiums on their notes during the crisis of 1857 than banks in other parts of the country.
19. Specifically, the monetary base included gold coin, gold certificates backed 100 percent by gold, silver dollars, silver certificates, other small silver coins. U.S. notes and other Treasury fiat, and national bank notes. See Tallman and Moen (1993).
20. Tallman and Moen (1993) indicate that this uncertainty was greatly reduced with the discovery of large gold supplies in the late 1890s.

REFERENCES

Akerlof, G. 1979. The Market for Lemons: Qualitative Uncertainty and the Market Mechanism. *Quarterly Journal of Economics* 84:488–500.

Bartholomew, P. F., Moe, L. R. and Whalen, G. 1995. The Definition of Systemic Risk. Office of the Comptroller of the Currency. Presented at the seventieth annual Western Economic Association International Conference, San Diego, California, July.

Benston, G. J., and Kaufman, G. G. 1995. Is the Banking and Payments System Fragile? *Journal of Financial Services Research* 9 (December): 209–40.

Bernanke, B. S. 1983. Non-Monetary Effects of the Financial Crisis in the Propagation of the Great Depression. *American Economic Review* 73 (June): 14–31.

Bernanke, B., and James, H. 1991. The Gold Standard, Deflation, and Financial Crisis in the Great Depression: An International Comparison. In *Financial Markets and Financial Crises,* edited by R. Glenn Hubbard, 33–68. Chicago: University of Chicago Press.

Bhattacharya, S., and Gale, D. 1987. Preference Shocks, Liquidity, and Central Bank Policy. In *New Approaches to Monetary Economics,* edited by William Barnett and Kenneth Singleton, 69–88. New York: Cambridge University Press.

Bordo, M. 1986. Financial Crises, Banking Crises, Stock Market Crashes, and the Money Supply: Some International Evidence. In *Financial Crises and the World Banking System,* edited by Forest Capie and Geoffrey E. Wood, 190–248. London: Macmillan.

Bryant, J. 1980. A Model of Reserves, Bank Runs, and Deposit Insurance. *Journal of Banking and Finance* 4:335–44.

Calomiris, C. W., and Gorton, G. 1991. The Origins of Banking Panics: Models, Facts, and Bank Regulation. In *Financial Markets and Financial Crises,* edited by R. Glenn Hubbard, 109–73. Chicago: University of Chicago Press.

Calomiris, C. W., and Schweikart, L. 1991. The Panic of 1857: Origins, Transmission, and Containment. *Journal of Economic History* 51:807–34.

Chari, V. V. 1989. Banking without Deposit Insurance or Bank Panics: Lessons from a Model of the U.S. National Banking System. Federal Reserve Bank of Minneapolis *Quarterly Review* 13 (Summer):3–19.

Chari, V. V., and Jagannathan, R. 1984. Banking Panics, Information, and Rational Expectations Equilibrium. Northwestern University. Photocopy.

Cone, K. 1983. Regulation of Depository Institutions. Ph.D. diss., Stanford University.

Diamond, D., and Dybvig, P. 1983. Bank Runs, Liquidity, and Deposit Insurance. *Journal of Political Economy* 91 (June): 401–19.

Donaldson, R. G. 1992. Sources of Panics: Evidence from Weekly Data. *Journal of Monetary Economics* 30 (November):277–305.

Eichengreen, B., and Portes, R. 1991. The Anatomy of Financial Crises. In *Financial Markets and Financial Crises,* edited by R. Glenn Hubbard, 10–58. Chicago: University of Chicago Press.

Eisenbeis, R. A. 1997. International Settlements: A New Source of Systemic Risk? Federal Reserve Bank of Atlanta *Economic Review* 82 (Second Quarter): 44–50.

Friedman, M., and Schwartz, A. J. 1963. *A Monetary History of the United States, 1867–1960.* Princeton, N.J.: Princeton University Press.

Gorton, G. 1987. Banking Panics and Business Cycles: Data Sources, Data Construction, and Further Results. The Wharton School, University of Pennsylvania. Photocopy.

Gorton, G. 1988. Banking Panics and Business Cycles. *Oxford Economic Papers* 40:751–81.

Haubrich, J. 1990. Non-Monetary Effects of Financial Crises: Lessons from the Great Depression in Canada. *Journal of Monetary Economics* 25 (March): 223–52.

Haubrich, J., and King, R. 1984. Banking and Insurance. NBER Research Working Paper No. 1312.

Jacklin, C. 1987. Demand Deposits, Trading Restrictions, and Risk Sharing. In *Contractual Arrangements for Intertemporal Trade,* edited by Edward D. Prescott and Neil Wallace, 26–47. Minneapolis: University of Minnesota Press.

Kaufman, G. G. 1988. Bank Runs: Causes, Benefits, and Costs. *Cato Journal* 7 (Winter): 559–87.

Kaufman, G. G. 1994. Bank Contagion: A Review of the Theory and Evidence. *Journal of Financial Services Research* 8:123–50.

Kaufman, G. G. 1995. Comment on Systemic Risk. In *Research in Financial Services,* vol. 7, edited by George G. Kaufman. Greenwich, Conn.: JAI Press.

Kindleberger, C. P. 1978. *Manias, Panics, and Crashe,* New York: Basic Books.

Lintner, J. 1977. Discussion. In *Financial Crises: Institutions and Markets in a Fragile Environment,* edited by Edward I. Altman and Arnold W. Sametz, 204–207. New York: John Wiley and Sons.

McAndrews, J. J. 1997. Banking and Payment System Stability in an Electronic Money World. Federal Reserve Bank of Philadelphia Working Paper No. 97-9.

Minsky, H. P. 1982. The Financial-Instability Hypothesis Capitalist Processes and the Behavior of the Economy. In *Financial Crises: Theory, History, and Policy,* edited by Charles P. Kindleberger and Jean-Pierre Laffargue, 13–38. Cambridge: Cambridge University Press.

Mishkin, F. S. 1991. Asymmetric Information and Financial Crises: A Historical Perspective. In *Financial Markets and Financial Crises,* edited by R. Glenn Hubbard, 69–108. Chicago: University of Chicago Press.

Park, S. 1991. Bank Failure Contagion in Historical Perspective, *Journal of Monetary Economics* 28 (October): 271–86.

Schwartz, A. 1986. Real and Pseudo Financial Crises. In *Financial Crises and the World Banking System,* edited by Forest Capie and Geoffrey E. Wood, 11–31. London: Macmillan.

Sinai. A. 1977. Discussion. In *Financial Crises: Institutions and Markets in a Fragile Environment,* edited by Edward I. Altman and Arnold W. Sametz, 187–203. New York: John Wiley and Sons.

Smith, B. D. 1991. Bank Panics, Suspensions, and Geography: Some Notes on the 'Contagion of Fear' in Banking. *Economic Inquiry* 29 (April): 230–48.

Sprague, O.M.W. [1910] 1968. *History of Crises under the National Banking System.* National Monetary Commission. Reprint. New York: Augustus M. Kelley.

Tallman, E. W. 1988. Some Unanswered Questions about Bank Panics. Federal Reserve Bank of Atlanta *Economic Review* 73 (November/December): 2–21.

Tallman, E. W., and Moen, J. 1993. Liquidity Shocks and Financial Crises during the National Banking Era. Federal Reserve Bank of Atlanta Working Paper 93-10, August.

Wallace, N. 1988. Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig Model with Sequential Service Taken Seriously. Federal Reserve Bank of Minneapolis *Quarterly Review* 12 (Fall): 3–16.

Williamson, S. 1989. Bank Failures, Financial Restrictions, and Aggregate Fluctuations: Canada and the United States, 1870–1913. Federal Reserve Bank of Minneapolis *Quarterly Review* 13 (Summer): 20–40.

This Page Intentionally Left Blank

65
Financial Crises and Market Regulation

Jerry L. Jordan
President and Chief Executive Officer, Federal Reserve Bank of Cleveland, Cleveland, Ohio

Financial crises are inevitable. Both government intervention and market innovations can influence the frequency and severity of these episodes, but they cannot eliminate them. Evolution toward stronger political and economic institutions is a discovery process, and the sometimes dramatic financial market adjustments labeled "crises" are an unavoidable part of that process.

Government intrusions into financial markets typically make financial crises more serious. For the most part, official programs seem designed to act as sponges for absorbing risk exposures from particular groups of economic agents. This can lead market participants astray. Unless the resulting incentive to overinvest in risky projects is offset by an effective program of supervision, agents are likely to misallocate resources. Moreover, especially before a crisis, a government may act as though the capacity of its risk sponge is unlimited. Only when that capacity is tested—by calls on foreign exchange reserves, by demands on taxpayers through the budget—does information about limits emerge. Crisis ensues.

Financial crises are not predictable. If they were, actions would be taken to alter the predicted event, and the crisis would be avoided. It is the surprise contained in new, unexpected information that sets off these episodes. If there were no surprise, there would be no basis for the sudden and substantial changes in market prices of financial instruments that are characteristic of crises.

THE ELEMENT OF RISK AND THE LESSONS OF FAILURE

Financial outcomes always have an element of uncertainty as well as an element of risk. I choose the words uncertainty and risk deliberately.[1] They describe two different kinds of exposure to chance events. Uncertainty is impossible to describe probabilistically because the distribution of possible outcomes essentially is unknown. Therefore, exposure to uncertainty cannot be hedged or insured against. Gains and losses from uncertain events are

Reprinted from: *Federal Reserve Bank of Cleveland Economic Commentary,* Oct. 1, 1999.

pure economic rents. They are either borne by those exposed to them or socialized in an act of community greed—such as a steeply progressive tax—or in an act of community compassion—such as Red Cross, Federal Emergency Management Agency, or Marshall Plan assistance. We cannot help but be surprised by the occurrence of uncertain events. This is why crises are unpredictable.[2]

Risk, on the other hand, is used for exposure to the other kind of outcome. This exposure can be described probabilistically, typically by using a model of some sort to filter past experience. Exposure to losses from risky events can be bought and sold in the marketplace. An insurance company will assume your exposure to loss from the death of a partner in return for an insurance premium. Financial institutions accept exposure to the bankruptcy of debtors in return for a sufficiently large risk premium.

The existence of risky situations is not a source of crises, for rational economic agents will have incorporated calculations about potential outcomes into their decision-making. Instead, crises typically result from risk mismanagement.

Risks can be transferred from one party to another, but they cannot be eliminated. Exposure to loss from a partner's death can be shifted through an insurance company to a diversified set of policyholders. Exposure to loss from lending can be shifted through a bank and the FDIC to all insured depositors or general taxpayers. Exposure to exchange-rate changes can be shifted to someone else through the organized exchanges or an over-the-counter transaction or by government to general taxpayers.

While individual parties can take actions to reduce, minimize, or avoid their own risk exposure, they can do so only when another party is willing and able to bear such risks. This is where government behavior can become a problem. Usually, with the most sincere and honorable of intentions, governments seek to reduce, minimize, or eliminate the exposure to risk of some constituent.

But allowing borrowers and investors to suffer the losses incurred in a crisis is a necessary and useful way to bring about adjustments in the technology of investment. For example, adopting regulations now that would proscribe practices that appear to have made the 1997 Asian crisis so severe should be unnecessary because borrowers and lenders now understand what happened and will not be surprised again. If the parties were to repeat the same behavior, markets would discount the now-expected result, leading to substantial risk premiums in loan and currency markets. There would be no crisis because there would be no surprise. More likely, however, borrowers and investors would not repeat the behavior that brought them grief the last time. Instead, faced with markets that extract more accurate premiums, they would adopt new practices—that is, innovations—some of which would work and some of which would not.

Lessons learned in a crisis lead to changes in behavior that prevent a repeat of the conditions that led to the crisis. The discipline exerted by global financial markets is beneficial in that it erodes local resistance to more efficient domestic markets. This is what the president of Korea had in mind when he said recently that there is a "silver lining" to the Asian currency crisis. The restructuring and reforming of the banking institutions now occurring in Asia will leave them better off. It would have taken much longer to implement these reforms without the "crisis atmosphere." As a result, these nations may soon have better credit risk analysis, have better asset and liability management techniques, be less subject to politically connected bank lending, and develop both internal audit and external supervision that is essential to sound banking.[3]

Crises are not desirable, but given that they are inevitable, we would do well to recognize that they lead to more efficient financial systems—so long as the consequences of risk taking are borne by those responsible.

GOVERNMENT INTRUSIONS AND ALTERED INCENTIVES

A myriad of government intrusions into the marketplace—often in the form of controls or guarantees—have altered the incentives of participants to accept risk exposures in affected markets. The stated justification for regulation often is a contention that a market failure or imperfection is present. It is important to understand that the perceived failure or imperfection often was introduced by another government intrusion. In that circumstance, then, regulation or supervision may represent an effort to recreate the competitive situation that would have been present without the initial government involvement in the market.

Banking supervision is a prime example of this effort. One of the most familiar government intrusions into financial markets of this century has been a government-supplied guarantee called deposit insurance. Much of the regulation governing firms covered by deposit insurance is rationalized by the need to neutralize the moral hazard introduced by the insurance (that is, to reduce the bank's incentive to make risky investments because depositors, knowing that insurance will bail them out in the event of bank failure, do not demand a risk premium). The challenge always has been to construct government regulations that do not undermine the effectiveness of market regulation—the discipline that comes from having to compete.

The prior presence of guarantees by government has been a common element underlying many episodes labeled crises. Most often, such guarantees are introduced as a well-intentioned effort to shelter someone from exposure to the costs of some type of risk. But those costs and risks are real; as I mentioned earlier, they are not eliminated by the presence of government-supplied guarantees. They merely have been shifted to someone else, usually in a very diffused or opaque way.

When governments offer guarantees, one of two possible problems is created. Because market participants engage in behavior that is conditioned on the existence of the guarantee, government has become a third contractual partner to any transaction or agreement. In so doing, government has caused someone else—often the general taxpayer—to bear a risk or cost, usually unknowingly. Then, when unforecastable events force government to make good on the guarantee, general taxpayers belatedly are informed that they must incur a wealth loss as a result of risk having been shifted to them without their prior knowledge and agreement. As long as economic agents are unable to fully internalize the potential costs of government promises, financial markets are not complete and thus are not efficient.

The alternative problem is that withdrawal of a guarantee has the effect of breaking a contract, thereby imposing losses on someone. Abandoning an exchange-rate peg is a good example of this breach of contract. Many of the events that have come to be labeled financial or banking crises involve the breaking of an explicit or implicit contract or the withdrawal by government of the offer to make new contracts—that is, to provide guarantees—for the future. Because past behavior was influenced by past guarantees, the anticipated withdrawal of the guarantees must be reflected in surprise adjustments of relative prices. If these price changes are large and occur within a short time interval, the event is labeled a crisis.

For example, exchange-rate crises of the early 1970s involved the withdrawal of the U.S. guarantee that foreign governments could exchange dollars for gold at a ratio of 35 to 1. Most exchange-rate crises since then have simply been reflections of unsustainable government promises to redeem their own fiat currency for a foreign currency at preannounced rates.

A related aspect of crises concerns the quality of a nation's legal and financial infrastructure. The surprise that triggers a crisis is some new piece of information about the inability of debtors to meet the terms of financial contracts. Contracts are drawn up according to a nation's rules and regulations.[4] If those rules are lacking, unclear, or capable of manipulation, then the opportunity for surprise is greater than if the rules are clear and stable.

Of course, I don't mean to imply that all crises are purely domestic, for there is little point in trying to distinguish between an international crisis and a domestic crisis. In the context of late twentieth century global capital markets, all financial markets have an international element. In fact, global financial asset portfolio managers can be thought of as a new class of voters—stateless citizens of the world, empowered to roam the globe casting votes for and against economic policies of the 200 or so nation-states. As a result, a common, and ultimately hopeful, element of crises has begun to emerge. Various types of government guarantees and promises—on deposits, on foreign investments, on domestic pensions—may be on the "endangered species list" because so many are being revealed as impossible to honor. That is the lesson of Mexico in 1994–95 and Asia in 1997. That is what the current political turmoil in Germany may be all about. To the extent that global capital market vigilantes are becoming more effective in evaluating the true costs of national governments' promises, there should be less opportunity for big surprises to bring on financial crises.

THE ROLE OF MARKET INNOVATIONS

Crises also originate in failed market innovations. For example, the collapse of Long Term Capital Management last year raised questions about the relationship between financial innovations and crises. Innovations are tests of ways of doing things. Some succeed, but, just as surely, others fail. Innovation is bound to bring profits to some and losses to others. That's the way the market system works.

In the case of hedge funds, massive doses of computer technology are combined with finance theory and online, real-time global information gathering to exploit perceived anomalies in the prices of related financial instruments. This is a valuable role to play in perfecting markets, and it can be a very profitable role if all goes according to plan. However, as we have seen in a few cases, it can be a very unprofitable role if market prices do not move toward the assumed past relationships. The resulting losses can be multiplied manyfold to the extent such funds are leveraged. As long as there is uncertainty about market relationships, surprises are capable of inflicting big losses—perhaps big enough to be called crises.

Often, what gets missed in considering new vehicles for risk bearing is that they are methods of redistributing risk. Far from increasing risk in the financial system, as some have thought, they in fact serve the primary purpose of unbundling risk into its components and, thus, allowing each type of risk to be managed separately. This allows for specialization in risk bearing, as particular types of risk can be transferred from one party, who is averse to that risk, to another party, who is less averse to that risk. In this way, financial innovations lower the cost of bearing risk by transferring it to those most willing, and presumably able, to bear it. However, some financial innovations don't work, and investors lose money.

MORAL HAZARD

"All that's very well," you will say, "but haven't many market participants come away from the 1997 Asian crisis with an entirely different lesson?" Namely, "Don't worry, 'cause the IMF or someone else will pull our chestnuts out of the fire next time—just as they did last time." Indeed, moral hazard is a real and serious problem. Not only are financial rescues likely to impede effective reforms, but they may also induce back-sliding by encouraging an inattention to risk that will make the next crisis worse than the last one and thus induce more pernicious moral hazard through even stronger rescue efforts. Notice that moral hazard involves intervention in the market by a government-type authority that relieves certain market participants of a risk exposure by assuming that exposure itself—that is, on behalf of the general taxpayer. Especially in cases where the risk transfer is only implicit, a big danger is that the authority fails to create a supervisory mechanism to manage or control the risk it has assumed. At least the current U.S. deposit insurance and lender-of-last-resort institutions maintain an active supervisory presence in the banking system. Without a supervisory presence, rescue efforts may promise relief in the short run, but will be incapable of keeping that promise in the long run.

WHAT CAN BE DONE

One approach to reducing the perceived need for government agencies to mount rescue operations is to strengthen private market facilities for monitoring and dealing constructively with potential crisis situations. I'll mention just two examples of proposals that have been getting attention. One would improve the transparency of the financial positions of debtors through more consistent and thorough financial disclosure. Creditors could then derive earlier warnings of trouble and force actions to limit losses. Greater transparency is a desirable feature of efficient financial markets in its own right. It is not clear that it would do much good in preventing the kinds of surprises that trigger financial crises; however, it might aid in crisis resolution.

Another approach would include "collective action clauses" or other such provisions in debt contracts that would facilitate action by private creditors to reschedule, restate, or otherwise resolve situations in which debtors need relief. Especially with a growing share of credit going directly to private firms rather than through sovereign borrowers, creditors should be better equipped to protect their own interests without the intervention of a government or international agency.

Of course, arming private investors will not provide an incentive for them to use new provisions as long as government agencies appear willing to intervene and achieve settlements on more favorable terms than produced by private negotiations. Given that a "just say no" intervention policy is politically difficult, judgment must be exercised in drawing the line between productive and unproductive intervention.

Critical to any successful reform is a reduction or elimination of uncertainty arising from official responses to international market disruptions. Governments and international agencies might clarify for markets what actions they would be prepared to take, and under what circumstances. This could reduce incentives for creditor runs on sovereign borrowers. Moreover, it would increase the accountability of policymakers for their actions and, in turn, reduce the attractiveness of ex post creditor bailouts. This is by no means as easy as it might sound. More fruitfully, perhaps, agencies might explore a structured format for pre-commitment, comparable to that made familiar in bank capital regulation.

In summary, financial crises are inevitable, I'm afraid. They are not predictable. Typically, they arise from financial mismanagement by governments, sometimes in response to failed financial innovations. Moral hazard compounds the inevitability of crises by encouraging private inattention to risk exposures without any assured supervisory offset.

NOTES

1. This taxonomy was introduced by Frank Knight. See Risk, Uncertainty, and Profit. Chicago: University of Chicago Press, 1971 (1921).
2. Others describe the origins of crises by suggesting that an objective distribution of possible outcomes of investments includes a significant region of bad results. This region (sometimes referred to as the "fat tail" of the distribution) reflects the slight, but not negligible chance of simultaneous losses to many investors in a crises. Investors are shortsighted, however, in ignoring this region of possibilities when constructing portfolios and, therefore, surprised when it actually materializes.
3. Or they may not. Suppose, instead, that these governments heed the advice of those who would throw more sand in the gears of cross-border financing to discourage susceptibility to capital flight. A turnaround tax on capital withdrawn within a few months or a year of entering a country is a popular suggestion. The approach might work, just as capital controls seem to have worked for most nations during and after World War II, but at the cost of a less efficient global allocation of capital. If regulation were to be part of the reform process, it would make more sense to remove regulations that discourage long-term capital flows than to add regulations that discourage short-term capital flows.
4. An interesting exception has been proposed by Howell E. Jackson in "The Selective Incorporation of Foreign Legal Systems to Promote Nepal as an International Financial Services Center," forthcoming in an Oxford University Press symposium volume on regulatory reform. His suggestion is that firms be permitted to establish operations in Nepal if they agree to abide by their home countries' regulations and submit their Nepalese operations to their home countries' supervision.

66

Disruptions in Global Financial Markets

The Role of Public Policy

Michael H. Moskow
President and Chief Executive Officer, Federal Reserve Bank of Chicago, Chicago, Illinois

When we think of the state of the world economy over the last 20 years, what we see is a mixture of the good, the bad, and the puzzling. On the one hand, we in the U.S. have experienced the longest period of uninterrupted growth in our history. We've actually had an extraordinary 17-year run, interrupted only by the rather short recession of 1990–91. On the other hand, this period has also been characterized by financial turmoil.

Worldwide, this period has seen the greatest concentration of financial crises since the 1930s. In the U.S., the cost of resolving the savings and loan crisis amounted to around 3 percent of our gross domestic product (GDP).[1] Yet this cost was dwarfed by crises in Scandinavia, Latin America, and, most recently, East Asia. Estimates of resolution costs for the Asian crisis countries are between 20 percent and 65 percent of GDP.[2] And there are still rumblings of concern in certain markets. For example, the Japanese banking system is generally regarded as undercapitalized, with official reported bad loans amounting to over 6 percent of total loans. The true volume of bad loans may be quite a bit higher.[3] Most observers would agree that financial disruptions of these magnitudes have substantial welfare costs.[4]

I'd like to focus on a particular manifestation of this problem: the so-called twin-crisis phenomenon, where banking crises go hand-in-hand with currency crises in emerging economies. We saw this in Mexico in 1995, in East Asia in 1997, and in Russia in 1998.

This new development is part of a broader phenomenon that creates both opportunities and dangers: the rapid globalization of financial markets. This explosion in cross-border financial transactions resulted from a confluence of economic, political, and technological factors. The rapid export-led growth of developing free market economies, notably the Asian tigers, especially by comparison to the relative lackluster performance of many state-controlled economies, has dramatized the potential gains from decentralization, deregulation, and reduction in restrictions on free movement of goods and capital. Tech-

Reprinted from: *Federal Reserve Bank of Chicago Economic Perspectives* (Third Quarter 2000) 2–8.

nological developments have increased the ease and speed with which large volume cross-border transactions can be executed.

The great opportunity from globalization is that standards of living worldwide can grow as more and more countries exploit the gains from trade, and as capital flows to its most productive uses. The great danger is that globalization may carry with it new sources of financial instability and may exacerbate financial disruptions when they do occur. So I've briefly discussed the good and the bad—now the puzzle: "How should we respond to these challenges?"

This issue is of particular concern to the Federal Reserve System. The long-run goal of the Federal Reserve is to promote maximum sustainable growth through price stability. However, the Federal Reserve is also committed to safeguarding the safety and soundness of the financial system. The approaches we have used in the past are designed mainly for national financial systems. Now, I believe, is the appropriate time to consider whether these approaches are adequate in an environment where national borders are less and less important. In light of recent developments, how should we proceed?

A good place to start is with a discussion of the twin-crisis phenomenon: where a banking crisis and a currency crisis occur simultaneously and feed on each other. Perhaps the most dramatic example of this phenomenon is the Asian crisis of 1997. What happened? Why did these countries get hit by a sudden crisis so strong that it engendered output declines on the order of the Great Depression?[5]

First, let's rule out one candidate explanation. The crisis was not the result of poor macroeconomic policies. In fact, the crisis countries pursued rather conservative policies. Their economies were characterized by low inflation, budgets generally in surplus, and declining government foreign debt. They engaged in responsible credit creation and monetary expansion. In short, these countries seemed to be following the usual prescription for responsible economic governance.[6]

Then something else must be going on. A number of observers are giving significant emphasis to this twin-crisis argument.[7] A twin crisis can occur when two factors are present. First, an emerging economy must provide its banks *with implicit or explicit government guarantees.* There may be valid reasons for such guarantees. They may reflect efforts to increase the flow of investable funds. I'll return to this point later. Alternatively, the guarantees may reflect "crony capitalism" or other forms of politically directed interference in the economy. Either way, they imply for the government a huge potential liability in the event of widespread bank failure. The second factor in a twin crisis is that banks in this emerging economy must rely on *short-term loans from abroad* denominated in *dollars* or other hard currencies.

How might these two factors interact in potentially malign ways? First, the government guarantees lead to moral hazard problems. They reduce the monitoring of banks by investors, so banks are less likely to make prudent investment choices.[8] Of course, moral hazard isn't a problem just for emerging countries. We don't have to go back far in U.S. history to find examples of moral hazard induced distortions. But moral hazard doesn't seem to be the whole story. Eventually, it is likely that these poor investments go bad, at which point the government may feel compelled to bail out the failing banks. The problem is that this bailout is costly. It acts as a fiscal drag on the government. There seems to be a connection between the fiscal burden to resolve these crises and the resulting currency attacks. We've known since the early work of Paul Krugman that fiscal shocks tend to foreshadow speculative attacks on a country's currency. If the government finances the bank bailouts by borrowing or increasing taxes, its ability to defend its currency against speculative attack is reduced. If the bailout is financed by monetary expansion, the resulting inflation directly weakens its currency. Either way, the banking crisis is likely to lead to a fall in the country's exchange rate.

Where does globalization fit in here? Recall that the banks in this emerging economy have a large volume of dollar-denominated liabilities outstanding. When the currency depreciates, these liabilities become harder to repay. As a result, more banks fail, requiring an even bigger government bailout, which in turn places even more stress on the currency. The banking crisis generates a currency crisis, which deepens the banking crisis, which exacerbates the currency crisis, and the vicious cycle continues.

Notice the fiendish way this twin-crisis phenomenon renders national banks virtually powerless. The usual weapons central banks have in their arsenal are *general liquidity provision,* usually through open-market operations, and *directed liquidity provision* through their role as lender of last resort. In a twin-crisis event, *neither* of these weapons can be effective. Open market operations, by increasing the relative supply of the national currency, act to drive down the exchange rate. Loans to assist banks in paying off their dollar-denominated debt simply strip away foreign reserves that are needed to defend the currency. If a central bank is seen to be depleting its hard currency reserves, a speculative attack is almost inevitable. The government has little choice but to go to the International Monetary Fund (IMF) or the U.S. Treasury for relief.

Now, this twin-crisis phenomenon seems mainly to be a problem for emerging markets. Why should we, in the developed world, care? I believe that we must be concerned. In the modern global economy, there are numerous pathways whereby weaknesses in developing countries can be harmful to our own well-being. Take, for example, the recent concern about our current account deficit. In the year that just ended, the U.S. experienced a current account deficit exceeding $300 billion. At over 3.5 percent of GDP, this is the largest current account deficit in U.S. history.[9]

The main reason for this deficit clearly is the strength of the U.S. economy relative to our trading partners. However, in recent years this deficit seems to have been exacerbated by the changing capital flows due to the financial crises in developing countries. As these economies weaken, investors who had exported capital to these countries now look for a safe haven for their money. The safest economy in the world is the U.S. This inflow of capital produces a larger capital account surplus, which, as a balance-sheet identity, implies a larger current account deficit than would have been produced in the absence of these capital flow distortions.

The way this process works itself out is that these foreign investors bid up the prices of U.S. assets and drive down U.S. interest rates. This induces a *wealth effect:* Americans see their wealth increasing and their relative borrowing costs decreasing, so they tend to save less and consume more. Since this increased consumption can't be satisfied by domestic production alone, we buy more from abroad, increasing the trade deficit.[10]

While this flow of funds into the U.S. has immediate benefits to us, it carries with it potential problems. Sudden capital inflows can be reversed. If American consumers increase their indebtedness in response to temporary capital inflows, this indebtedness remains even after the capital inflow has been reversed. Thus, our record current account deficit could actually trigger a period of consumption volatility for the American consumer.

A further reason why we must be concerned about financial turmoil in emerging economies is that financial turmoil knows no borders. The danger to U.S. financial markets in late September and early October of 1998 was very real. Triggered by the Russian default and devaluation in August 1998, the resulting uncertainty about who was affected by the default, who was creditworthy, and who was overextended induced a widespread drying-up of liquidity. The turmoil that followed in the U.S. exemplifies how a disturbance in an emerging economy can be propagated through U.S. markets in rather unpredictable ways. Additionally, as a result of this event, the Federal Open Market Committee (FOMC) decreased the fed funds rate by 75 basis points to ensure a sufficient supply of liquidity in the economy. In doing so, the committee took a conscious risk that the monetary expansion would not exacerbate inflation pressures, with the associated costs. We believe things worked out well in this incident, but foreign turmoil of this type makes the FOMC's job more difficult.

So it is clearly in our interest to consider ways to respond to challenges of globalization such as the twin-crisis phenomenon. To start thinking about potential responses, let's first remember that twin crises arise out of a confluence of *globalization* (in the form of short-term dollar-denominated loans to emerging economies) and *government action* (in the form of government guarantees of bank liabilities). One could address this problem by restricting globalization—say, by imposing capital controls. Indeed, this approach has been recommended by some, and has been implemented in Malaysia.[11] While there are arguments in favor of capital controls as a short-term fix, I don't think this is the place to look for a long-term solution. The gains from international capital mobility are just too great, and the costs in economic growth of restricting this mobility too large, to consider capital controls as a permanent solution to the troubles associated with globalization.

Rather, we should focus on the second factor: *government action.* A useful set of principles for appropriate government action in the economic arena are as follows:

governments should have a clear policy objective;
they should be minimally intrusive in achieving that objective;
they should rely to the greatest extent possible on market mechanisms and market incentives; and
they should seek to influence ex ante behavior, rather than focusing exclusively on ex post crisis management.

Most important, government policies should not actively encourage poor choices in the private sector. The Hippocratic maxim, "First, do no harm," applies to financial regulation as well as medicine.

Since the twin-crisis phenomenon starts with government guarantees of banking sector liabilities, let's look at the role of a government-provided safety net in light of these principles. Governments everywhere tend to provide some type of safety net for their national banking systems. Presumably, the intention is to promote confidence in the financial system and to reduce the possibility of financial panics and bank runs. Indeed, the Federal Reserve System was first proposed in the aftermath of the Financial Panic of 1907. There are a number of reasons why emerging countries have special pressures to provide guarantees. These countries often do not have financial structures encouraging to investors. Accounting practices are not fully transparent, disclosure is inadequate, assets are opaque, and property rights are ambiguous. For example, at the time of the 1997 crisis, Thailand did not even have an effective bankruptcy code.[12] Similarly, I once met a delegation of central bankers from a foreign country who discussed in detail the process they used to impose a haircut on collateral used for the equivalent of our discount window loans. This sounded impressive until they acknowledged that there was no legal basis for perfecting the collateral in case the borrowing bank failed. So what we see in certain developing countries is the use of an implicit government guarantee in effect being substituted for the legal and accounting infrastructure necessary to create a credible investment environment: the kind of environment investors in the U.S. can take for granted.

While these reasons for a bank safety net may be understandable, it is clear that the safety net does not banks to change their poor lending practices or would induce the regulatory authorities to take corrective action. As it happened, the western creditors of these banks were not induced to walk from the banks, since they believed that the governments in these countries would never let the banks fail. Indeed, the crisis started when the largest Thai finance company failed and the government bailout was *not* forthcoming. That is, western creditors did not move until they became convinced that these governments simply did not have the wherewithal to engage in a large-scale bailout. At that point, the creditors did not walk away from the banks. They ran.

Is it feasible for incentive-based regulation to be implemented as a global standard, including emerging markets? The evidence is looking more and more positive. One bright light shines from Argentina. The Argentine government has imposed strong market discipline on its banking system. Deposit insurance has been scaled back, banks are required to hold substantial dollar-denominated reserves, there is *significant* market disclosure,[18] and, since 1996, there has been something very similar to a subordinated debt requirement equal to 2 percent of assets. These steps were not taken in response to prodding from western governments. On the contrary, they were taken in response to *market pressures.* Specifically, Argentina wanted to avoid the high interest rates it was forced to pay in the wake of the Mexican "Tequila" crisis of 1995.[19] As the results of the Argentine experiment become known, other emerging economies may take similar actions.

There also appears to be a role for international coordination. In setting required levels of bank capital to cushion against losses, international standards rely on somewhat arbitrary criteria and place bank assets into "risk buckets." An alternative to this would be the use of market evaluations of the bank's risk profile. Indeed the Basel Committee has recently moved in this direction with a proposed new capital adequacy framework that relies on external risk evaluations. Additionally, current international standards significantly limit the use of certain capital instruments, such as sub-debt, by banks. Yet, as just discussed, such instruments could have characteristics that make them attractive as a disciplining force and as a cushion against losses. The Basel Committee has received comments suggesting that they relax current restrictions to allow countries to more fully utilize, and benefit from, alternative capital instruments. Finally, the recent Meltzer report on the future role of the IMF strongly encourages pre-certification before countries could borrow from the IMF. Part of this pre-certification procedure could be requirements that national bank regulations incorporate adequate market incentives into their regulatory policies.

Now, I must stress that market discipline is not a panacea. It is a guiding principle that directs us towards steps that need to be taken. It is a direction in which we should move, not a magic bullet for all problems. I believe, however, that market discipline will be an essential tool for managing the changes that will occur from increased globalization of the economy.

Earlier, I posed the rhetorical question, "Why should we care?" If there's a message I want to leave you with, it's that everyone has an important stake in how we resolve these international issues. I spoke earlier about the financial crisis in Asia influencing our current account and perhaps creating wealth effects that may lead to problems in the future. Recently, a banker from a small town in Iowa wrote to me when he read about a session on implicit government guarantees at our May 2000 Conference on Bank Structure and Competition. Now we typically think of implicit government guarantees as a large bank issue, just as we think of international disruptions as being associated with money center banks. But this small-town banker understood well how these implicit guarantees affected him and his customers. He argued that as larger banks encounter problems, their customers would realize that they were likely to be protected from losses. Customers would then flock to the protected bank, causing his funding sources to become more expensive or disappear. As a result, he would be less able to fund his customers' credit needs. This is an example of how inappropriate government guarantees may have unintended consequences, affecting not only Wall Street, but Main Street as well. We need to ensure that we do what we can to promote the unfettered flow of capital in international markets as well as in that small Iowa community.

To conclude, I think it's useful to consider how far we've come. Back in the 1930s, policymakers believed that financial markets were too important to be left to the marketplace. The renowned economist Abba Lerner expressed this view in a remarkable metaphor. He suggested that the Great Depression was like the scene of a multi-car pileup, with bodies strewn all over. A passer-by might wonder what caused the disaster, until he looked inside the wrecked cars and noticed that *there were no steering wheels*! Lerner's implication was that private market participants simply did not have the tools to avoid financial crises. Only the government could provide the controls to keep the cars smoothly riding towards their destinations. More recently, numerous scholars, including Milton Friedman and Anna Schwartz, offered a very different interpretation of the Depression. In this view, always work as intended. For example, the greatest bank losses in U.S. history came *after* the establishment of safety net institutions.[13] The twin-crisis phenomenon shows how the safety net can result in more financial disruption for emerging economies, not less. Why might this be so? The answer lies in the negative effect of an excessive safety net: By insuring banks' creditors, it makes them less aggressive in monitoring the business practices of banks. For example, banks in the U.S. held more capital and more cash reserves prior to the 1930s than they do currently.[14] Why? Because the market demanded that they do so. In the absence of a safety net, investors would not provide banks with funding unless they were adequately capitalized.

The insight from this example is that *well-functioning markets can go a long way to induce firms to make socially optimal decisions.* There is a role for government, but the best way to fulfill that role is to encourage markets to do as much of the work as possible. Ideally, we want to direct market incentives to achieve the regulatory goal of safe and stable financial markets, which foster maximum sustainable growth.

This basic principle is the driving motive for regulatory reform along several fronts. The Basel Committee on Bank Regulation is currently reevaluating bank capital standards to reduce distortions induced by the 1988 Capital Accord. That agreement introduced asset categories that carried specific risk weights for use in determining required levels of capital. It is generally recognized that the weights are not closely associated with risk: They favor bank-to-bank lending and place much sovereign debt in the same risk category. Furthermore, the risk weights favor short-term lending of foreign currencies that can have profound effects on lending patterns to developing countries. Some observers argue that these distortions may have been important causal factors in the Asian crisis. They create a regulatory environment where Korean sovereign debt has the same capital charge as U.S. Treasury securities, and where short-term loans to banks in developing countries can carry lower capital charges than loans to American AAA-rated nonbanks. In this environment, there is a clear incentive for western banks to channel money to risky emerging markets.

Similarly, there is a clear incentive for these markets to take the loans that are offered. By the standard, "First, do no harm," the current international capital standards appear to be wanting.

The Basel Committee recognizes that problems exist with the current accord and a public comment period is currently underway to reform the international capital standards. In its comment letter, the Federal Reserve Bank of Chicago emphasized the need for incentive-compatible regulation, disclosure, transparency, and market-driven risk assessment. Banks should be required to pass the test of the marketplace.

One example of how regulation could be used to promote, rather than suppress, market discipline is a recent proposal to require larger banks to issue medium- to long-term subordinated debt at regular intervals to satisfy a portion of their capital requirement. This proposal, which has been advocated by the Chicago Fed since the late 1980s, has recently gained increased support. How would it work?

Without getting into specifics, most sub-debt proposals would have the capital requirement be modified to have a sub-debt requirement in addition to an equity requirement. Why is this more in line with incentive-compatible regulation? Because the risk preference of these debtholders would closely approximate that of bank supervisors. They would be much more concerned with downside risk than they would with potential upside gains from bank portfolio choices. Because their debt holdings are subordinated to other liabilities, the sub-debt holders would be risk-sensitive and would monitor and discipline bank behavior. They would demand a higher interest rate from riskier banks. The debtholders would also have strong incentives to quickly resolve problems and to avoid forbearance and its associated costs. Most sub-debt proposals would impose certain characteristics on the sub-debt issues, such as minimum maturities, and would require the bank to stagger the debt issues to force the bank to "go to the market" on a regular ongoing basis, perhaps semiannually.[15] The purpose, again, is to ensure that the bank can pass the test of the market.

Suppose a regulatory framework like the subordinated debt proposal was accepted as a worldwide standard. How could this potentially mitigate the problem of twin crises? Holders of subordinated debt of banks in emerging markets would act as the equivalent of "mine-shaft canaries." At the first sign that bank loan quality was poor, they would refuse to roll over their debt or require a much higher yield. Most studies of U.S. bank markets suggest that debtholders can distinguish between the asset quality of banks and price the debt accordingly. In the Asian crisis countries, evidence suggests that investors were aware that problems were brewing well before the onset of the crisis.[16] The higher sub-debt yields would send a clear signal to both the markets and to regulators that potential problems existed. Seeing this, the subordinated debtholders would start their "walk" away from the bank. This walk would be more methodical and less disruptive than a run from troubled banks by uninsured depositors.[17] This would either induce the the cataclysm was not a result of insufficient government action but, of inappropriate government action.

Now, at the turn of the millennium, in the midst of the greatest prosperity our country has ever known, we have the opportunity to combine Lerner's profound concern for the costs of financial instability with a realistic appreciation of both the power of market incentives and the limits of government action. I think we can navigate the uncertainties of globalization with creativity and courage to maximize the chances for continued prosperity both here and abroad.

ACKNOWLEDGMENTS

The author would like to thank Douglas Evanoff and David Marshall of the Economic Research Department of the Federal Reserve Bank of Chicago for their research support and input.

NOTES

1. Based on "cleanup" costs of $165 billion. Hellmann, Murdock, and Stiglitz (2000) recently put the figure at $180 billion or 3.2 percent.
2. Drawing upon information from *Standard and Poor's Sovereign Ratings Service* (issues June 1999, November 1999, and December 1999) and the World Bank, the costs (as a percent of GDP) of bank recapitalization were estimated to be Korea, 24 percent; Malaysia, 22 percent; Thailand, 35 percent; and Indonesia, 65 percent. Kaufman (1999, table 2) puts the range at 45 percent to 80 percent.
3. The 6 percent figure is from a January 27, 2000, press release of the Financial Supervisory Agency of Japan. Others place the figure even higher: Hellmann, Murdock, and Stiglitz (2000) say nonperforming loans may approach 25 percent of GDP.
4. The previous cost figures (for example, 20 percent to 65 percent resolution costs) are not measures of welfare loss; they are simply transfers between agents in the economy. But welfare losses are typically associated with these transfers as a result of market distortions and the resulting inefficiencies. For example, the International Monetary Fund estimates that crises-induced output losses (actual versus trend growth) have been in the range of 17–18 percent (see Kaufman, 1999, table 1). Additionally, there could be welfare losses on a regular ongoing basis (not just during the crisis) as inefficient investments are undertaken. Thus, losses associated exclusively with crises probably understate the true welfare costs.
5. For the two worst quarters of the Asian crisis the annualized rates of GDP decline were over 25 percent for Indonesia, Korea, Thailand, and Malaysia. In the U.S., over the 1929–33 period GNP decreased by nearly 50 percent—about 15 percent per year. While the declines in the Asian crisis were less persistent than the Great Depression, the rate of decline was comparable.
6. This is further discussed in Marshall (1998) and Dooley (1999).
7. The twin-crisis hypothesis is associated with Kaminsky and Reinhart (1999) and Burnside et al. (1999). It should be emphasized that this may not be the sole explanation of what occurred in the Asian crisis.
8. This is consistent with arguments that suboptimal investments were undertaken in the Asian crisis countries. For example, a standard measure of investment efficiency is the "incremental capital output ratio" (ICOR), defined as the ratio between the investment rate and the rate of output growth. Higher ICOR implies less efficient investment. Corsetti, Pesenti, and Roubini (1999) show that for most East Asian countries, ICOR increased sharply in 1993–96, relative to 1987–92. They also claim that a substantial fraction of the new investment was directed toward real estate, as opposed to increased manufacturing capacity. Indirect evidence that this real estate investment was inefficient comes from data on rental yields for commercial office buildings. The yields were quite low (and vacancy rates quite high) before the onset of the crisis. Also, the rate of nonperforming loans before the crisis was above 15 percent in Thailand, Korea, Indonesia, and Malaysia. In 1996, 20 of the 30 largest Korean conglomerates showed a rate of return on invested capital below the cost of capital. While certainly not definitive, these patterns are consistent with less prudence on the part of investors.
9. This figure (3.66 percent of GDP in 1999) and the general discussion that follows are based on Hervey and Kouparitsas (2000).
10. This increase in the current account could be associated with either the "wealth effect" from the financial side or from demand side effects resulting from the strong U.S. economy. Again, while

both are operative, the data appear consistent with the wealth effect having an impact as described here. While the U.S. current account has been in deficit during most of the 1990s, the biggest increase occurred during 1998 and 1999; a timing that is consistent with the Asian crisis. Additionally, if driven by increased demand for foreign goods by U.S. consumers, it would result in increases in domestic interest rates. If the deficit was driven by an increased supply of foreign capital to the U.S., it would result in decreases in rates. In fact, the recent acceleration in the current account deficit was associated with a fall in medium- and long-term U.S. interest rates. This is consistent with the wealth effect argument.

11. Controls were imposed in August–September 1998.
12. See Renaud, Zhang, and Koeberle (1998).
13. Baer and Mote (1992) present evidence showing the rate of loss per dollar of deposits in the 1980s exceeded that experienced during the Depression years.
14. Boyd and Rolnick (1998) argue that "... before 1933, banks held much more capital than they now do. In fact, from 1844 to 1900, average capital ratios exceeded 20 percent of assets. In recent years, the average has been around 6 percent." Indeed, steady bank capital declines after 1930 were a major reason for the introduction of explicit bank capital requirements in the early 1980s. For a discussion of bank capital trends, and a discussion of why reported capital levels after 1930 are actually overstated, see Kaufman (1992).
15. Examples of specific proposals include U.S. Shadow Financial Regulatory Committee (2000) and Evanoff and Wall (2000).
16. Examples of the evidence in Asia include 1) the fact that the relative stock market valuation of the banking sector in Thailand started to decline in mid-1994, and 2) the IMF was "warning" the Thai government of economic misalignments for over two years prior to the onslaught of the crisis; see Lissakers (1999). The U.S. empirical evidence on the market's ability to distinguish between banks is summarized in Kwast et al. (1999).
17. The difference between the behavior of uninsured depositors and subordinated debtholders is driven by their different maturity structures.
18. Detailed information for individual banks concerning loan customers, status of loans, past dues, and so forth, is accessible via the Central Bank of Argentina's website.
19. See, for example, Calomiris and Powell (2000). They also provide evidence that banks are responding to this increased market discipline. They find that banks that were in compliance with the requirement before the Asian crisis were stronger (lower default risk), had faster deposit growth, and paid lower deposit interest rates than the banks not in compliance.

REFERENCES

Baer, H. L., and Mote, L. R. 1992. The United States financial system. In *Banking Structures in Major Countries,* George G. Kaufman (ed.). Boston: Kluwer Academic Publishing.

Boyd, J. H., and Rolnick, A. J. 1988. A case for reforming federal deposit insurance. *Annual Report,* Federal Reserve Bank of Minneapolis, available on the Internet at http:/-woodrow.mpls.frb.fed.us/pubs/ar/ar1988.html.

Burnside, C., Eichenbaum, M., and Rebelo, S. 1999. What caused the recent Asian currency crisis? In *The Asian Financial Crisis: Origins, Implications, and Solutions,* William C. Hunter, George G. Kaufman, and Thomas H. Krueger (eds.). Boston: Kluwer Academic Publishing, 1999.

Calomiris, C., and Powell, A. 2000. Can emerging market bank regulators establish credible discipline? The case of Argentina, 1992–1999. Paper presented at the National Bureau of Economic Research Conference on Prudential Supervision, January 15.

Corsetti, G., Pesenti, P. and Roubini, N. 1999. What caused the Asian currency and financial crisis. *Japan and the World Economy,* Vol. 11.

Dooley, M. P. 1999. Origins of the crisis in Asia. In *The Asian Financial Crisis: Origins, implications, and solutions.* William C. Hunter, George G. Kaufman, and Thomas H. Krueger (eds.). Boston: Kluwer Academic Publishing, pp. 27–31.

Evanoff, D. D., and Wall, L. D. 2000. Subordinated debt as bank capital: A proposal for regulatory reform. *Economic Perspectives,* Federal Reserve Bank of Chicago, Second Quarter, pp. 40–53.
Hellmann, T. F., Murdock, K. C. and Stiglitz, J. E. 2000. Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American Economic Review,* Vol. 90, March.
Hervey, J., and Kouparitsas, M. A. 2000. Should we be concerned about the current account? *Chicago Fed Letter,* Federal Reserve Bank of Chicago, April, No. 152.
Kaminsky, G. L., and Reinhart, C. M. 1999. The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review,* Vol. 89, June.
Kaufman, G. G. 1999. Banking and currency crises and systemic risk: A taxonomy and review. Federal Reserve Bank of Chicago, working paper, No. WP-99-12.
Kaufman, G. G. 1992. Capital in banking: Past, present, and future. *Journal of Financial Services Research,* Vol. 5, April.
Kwast, M. L., Covitz, D. M., Hancock, D., Houpt, J. V., Adkins, D. P., Barger, N., Bouchard, B., Connolly, J. F., Brady, T. F., English, W. B., Evanoff, D. D., and Wall, L. D. 1999. Using subordinated debt as an instrument of market discipline. Report of a study group on subordinated notes and debentures, M. Kwast (chair), Board of Governors of the Federal Reserve System, staff study, No. 172, December.
Lissakers, K. 1999. The IMF and the Asian crisis: A view from the executive board. In *The Asian Financial Crisis: Origins, implications, and solutions,* William C. Hunter, George G. Kaufman, and Thomas H. Krueger (eds.). Boston: Kluwer Academic Publishing, pp. 3–10.
Marshall, D. 1998. Understanding the Asian crisis: Systemic risk as coordination failure. *Economic Perspectives,* Federal Reserve Bank of Chicago, Third Quarter, pp. 13–28.
Renaud, B., Zhang, M., and Koeberle, S. 1998. How the Thai real estate boom undid financial institutions. What can be done now? paper presented at the NESDB seminar on Thailand's Economic Recovery and Competitiveness, Bangkok, May 20.
U.S. Shadow Financial Regulatory Committee. 2000. *Reforming Bank Capital Regulation,* Washington, DC: AEI Press, policy statement, No. 160, March 2.

67

How Powerful Is Monetary Policy in the Long Run?

Marco A. Espinosa-Vega
Senior Economist, Federal Reserve Bank of Atlanta, Atlanta, Georgia

Press reports about the state of the economy often give readers the impression that monetary policy and the people who direct it are quite powerful. For example, an article in the *Washington Post* in March 1997 asserts that "second to the president, Alan Greenspan is arguably the nation's most powerful person. As chairman of the Fed, he guides U.S. monetary policy, adjusting short-run interest rates."[1]

Many prominent academic economists seem to agree that monetary policy is quite powerful. In reviewing the monetary policy experience of the 1970s, Nobel Laureate James Tobin wrote, "In one respect demand-management policies worked as intended in the 1970s. . . . the decade is distinguished by its three recessions, all deliberately induced by policy. Likewise the expansionary policies adopted to reverse the first two recessions, beginning in 1971 and 1975 respectively, promoted recoveries and in 1977 expansion . . . The major turns in direction conformed to the desires and intentions of the managers of aggregate demand" (1980, 20–21).

Monetary policymakers themselves often describe their role as powerful. Consider, for example, Federal Reserve Chairman Alan Greenspan's testimony before Congress in July of last year. Attempting to explain the influence of monetary policy on the current state of the economy, he stated that "the preemptive actions of the Federal Reserve in 1994 contained a potentially destabilizing surge in demand, short-circuiting a boom-bust business cycle in the making" (1997). Without attempting to explain the full meaning of Greenspan's statement here, it is clear from his language that he believes the Federal Reserve System is powerful enough to have a profound influence on the course of economic activity.

Both Greenspan's statement and Tobin's comments focus on the short-run effects of monetary policy. One might suspect that if Greenspan really is the second most powerful person in the United States then the policy tools he controls must have some long-run influence on the U.S. economy. Ironically, however, although many academic economists and most Federal Reserve policymakers believe that monetary policy is quite powerful in the short run, they also believe that it is virtually powerless in the long run.

Reprinted from: *Federal Reserve Bank of Atlanta Economic Review* (Third Quarter 1998) 12–31.

Although opinion on this subject is far from uniform, most economists seem to believe that monetary policy can affect the level of real (inflation-adjusted) economic activity—that is, economic variables such as real interest rates, real gross domestic product (GDP) and the unemployment rate—over periods of one or two years. For example, the Fed can create economic recessions or strengthen cyclical expansions. It can do so, according to the conventional view, by increasing the growth rate of the money supply if it wants the economy to grow faster and reducing it if it wants the pace of economic activity to slow. However, increases in the money supply growth rate eventually cause the inflation rate to rise, and decreases in the money growth rate have the opposite effect. When policymakers discuss the short-run effects of monetary policy they usually describe some version of this trade-off between higher inflation, which almost everyone considers undesirable, and desirable changes in other economic variables: higher inflation vs. lower interest rates, lower unemployment, or faster growth in real GDP.

As indicated, however, most economists believe that the long-run effects of changes in monetary policy are very different from their short-run effects. Federal Reserve Governor Meyer has clearly stated the nature of this difference in beliefs, commenting that "there is, to be sure, no trade-off and hence no inconsistency between full employment and price stability in the long run" (1997, 19).

A pair of simple diagrams illustrates the conventional views about the short- and long-run effects of monetary policy. Figure 1 depicts a negatively sloping curve that describes a short-run trade-off between inflation and unemployment. In contrast, Figure 2, which displays a vertical line at the level of full employment, illustrates a scenario in which there is no trade-off between the level of unemployment and the rate of inflation. A low rate of inflation (price stability) is compatible with full employment, but so is a high rate of inflation. If Chart 2 accurately describes the long-run relationship between unemployment and inflation, then changes in monetary policy that lead to changes in the inflation rate have no effect on the long-run levels of unemployment or real output. In the jargon of economists, this diagram describes a situation in which money is superneutral in the long run.[2]

Although the view that monetary policy has real effects in the short run but is superneutral in the long run is widely accepted by academic economists, business economists,

Figure 1 A short-run inflation-unemployment trade-off.

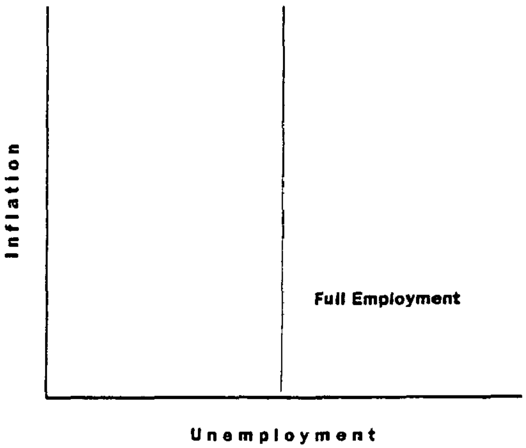


Figure 2 An inflation-unemployment relationship with no trade-off.

and economic policymakers, these groups are not in complete agreement about the ultimate real effects of monetary policy. One source of differences involves the magnitude of the short-run effects. Business economists and policymakers tend to believe that the short-run effects of monetary policy are very large, but most of their academic counterparts see these effects as rather tame and inconsequential. A related difference involves the questions of whether any short-run power that the Fed may have can survive repeated use. Most nonacademics seem to believe that the Fed can use its policy tools as often as it wishes without fear that they will lose their short-run effectiveness. On the other hand, most academic economists believe that repeated, systematic efforts to use the Fed's power to affect real economic activity will grow less and less effective over time.

This article reviews the development of the consensus view that monetary policy can have short-run effects but that it is long-run superneutral. The discussion emphasizes the fact that the basis for this view is the assumption that the only way monetary policy can affect real economic activity is via "money illusion"—that is, by creating changes in the price level that are misunderstood by households and firms and cause them to make bad economic decisions. If monetary policy can affect real economic activity by means other than money illusion then it may be possible for money to be nonsuperneutral in the long run.

This hopes to challenge economists and policymakers to devote more attention to investigating alternative explanations for the real effects of monetary policy—explanations that may imply that money is not long-run superneutral. In order to develop these alternative explanations it is necessary to make very explicit assumptions about the role of money in an economy, how it interacts with real variables and how economic decisionmakers react to the changes in its supply. Different assumptions will turn out to have very important implications for both the nature and the magnitude of the results of policy changes. This point is illustrated in the review of the small but growing branch of the academic literature on the real effects of monetary policy literature that studies models in which money may not be long-run superneutral. In these models the ultimate source of the real effects of monetary policy is the credit markets. By linking monetary policy with the supply of credit these models can analyze an alternative mechanism for evaluating the long-run effects of monetary policy that does not rely on money surprises.

Another important message of this article is that the very idea that monetary policy is powerful in the short run but powerless in the long run may be internally inconsistent.[3] If monetary policy is indeed as powerful as many informed people seem to believe, then theories of its real effects that rely on money illusion may have to be replaced by theories in which money is not superneutral in the long run.

THE PRECURSORS

This section briefly reviews the evolution of two prominent views on the neutrality of money: the Keynesian view and the monetarist view. The discussion begins with a look back at the classical theory that preceded Keynesianism and monetarism. It concludes by describing the clash between the Keynesians and the monetarists and the resulting "unilateral synthesis" of the 1970s.

The Early Quantity Theory

Classical macroeconomic theory, which developed during the late nineteenth and early twentieth centuries, was characterized by its focus on economic fundamentals ("real" economic conditions) such as individuals' propensity to save, the state of technology, and so on. In the classical view monetary policy played no long-run role in determining real economic activity. In particular, it had no long-run effect on the level of real interest rates. Classical theorists acknowledged that monetary policy might have a minor influence over economic activity (particularly interest rates) in the short run. In the long run, however, they viewed money as having a direct influence only on prices.

This early view of the influence of money on prices came to be known as the quantity theory of money. Like many economic concepts, the quantity theory has a rich history of reinterpretations. One of the earliest statements of the theory in its modern form was presented by Fisher (1926). According to Fisher, an economy's general price level is a function of the quantity of money in circulation, the economy's efficiency, or velocity, of circulation (the average number of times a year money is exchanged for goods), and the volume of trade (the quantity of goods purchased with money). Notationally, Fisher expresses the equation of exchange as:

$$M \times V = P \times T.$$

where M is the supply of money, V is the velocity of money, P is the general price level, and T is the total value of transactions or trade. Fisher held that in the long run there was a "natural" level of real economic activity determined by economic fundamentals that could not be affected by increases in the amount of money in the economy. In his words. "An inflation of the currency cannot increase the product of farms and factories . . . The stream of business depends on natural resources and natural conditions, not on the quantity of money" (1926, 155). This hypothesis that there was a natural long-run level of real economic activity, together with the assumption that the only role of money is to serve as a unit of account, formed the basis of Fisher's quantity theory analysis that prices varied proportionately to changes in the quantity of money. According to this equation, if velocity of money and the value of transactions were fairly stable—at least in the long run—as the economy approached its natural level, then changes in the quantity of money would be met with proportional changes in the price level.

Fisher conceded that monetary policy might have some temporary effects on real economic activity, commenting that "the 'quantity theory' will not hold true strictly and absolutely during transition periods" (1926, 161). In his mind, however, these effects were mainly due to temporary changes in velocity. If velocity was fairly stable in the long run, though, as he assumed, then it had to be the case that prices varied proportionately with the supply of money.

As this description indicates, the basic current consensus on the short- and long-run effects of monetary policy can be traced to the early quantity theorists. According to their view, as represented by Fisher, monetary policy could have temporary real effects but it would be superneutral in the long run. However, even the theory's adherents understood that the theory needed further refinement.[4] This task was undertaken a few years later by Milton Friedman (see below). By the time of the Great Depression, moreover, classical theory had lost much of its popularity and a new, completely different macroeconomic theory was appearing.

The Keynesians and Money: The First Time Around

The first nonclassical macroeconomic theory was the creation of John Maynard Keynes and is laid out in his *General Theory* (1936). One of Keynes's principal goals was to identify the causes of the persistently high rates of unemployment that were afflicting virtually the entire world during the Great Depression. He also sought to identify government policies that could help reduce these high levels of unemployment. Although Keynes's theory discussed the long-term implications of government policies, his focus was on the short run. And although monetary factors played a role in determining real economic activity in his theory (unlike in classical theory), Keynes's analysis emphasized fiscal policy. Keynes believed fiscal policy was the most powerful tool a government could use to lift the economy out of a recession or depression. In fact, his theory predicted that under certain conditions increases in the money supply would be unable to drive interest rates down low enough to stimulate economic activity by generating additional demand for credit. This situation was known as the liquidity trap. In liquidity trap situations money was superneutral even in the short run.

Nonneutrality of Money in the Long Run: The Chicago School

For many years after the Depression and the world war that followed it the question of the long-run implications of government policies received very little attention. One of the early assessments of the long-run effects of monetary policy came from, of all places, the University of Chicago. The university's department of economics—which was and remains perhaps the world's most influential collection of academic economists—has always been associated with the economic principles of the classical economists. The Chicago economics department was instrumental in the development of monetarism, which is usually considered to be a direct descendant of classical macroeconomic theory. In 1951, however, Lloyd Metzler, a prominent member of the economics faculty at Chicago, published a paper describing the long-run implications of central bank open market operations in which he asserted that under some circumstances monetary factors could interact with real variables in such a way as to help determine the level of real economic activity in both the short run and the long run. Metzler wrote that "by purchasing or selling securities, the banking authorities can alter not only the temporary interest rate which prevails while the open-market transaction is taking place but also the rate at which the system will return to equilibrium after the bank's transactions in securities have ceased" (1951, 107). He continued. "By purchasing securities, the central bank can . . . [cause] the system to attain a new equilibrium at a permanently lower interest rate and a permanently higher rate of capital accumulation" (112).

It is important to note that Metzler's conclusion that monetary policy—induced changes in the government's portfolio of liabilities could potentially have long-run real effects does not rely on the monetary authority's ability to produce inflation surprises or on workers' or firms' inability to correctly appraise conditions in the labor market. His analysis is therefore very different from most modern analyses, which view monetary surprises and their impact on naive participants in labor markets or short-run market frictions as the main channel by which monetary policy can affect real economic variables.

Although Metzler's analysis is less than fully rigorous by modern standards, it is worth recalling because it represents one of the first careful descriptions of a mechanism through which monetary policy can have long-run real effects. As noted above, Metzler's conclusions ran counter to the classical tradition of the Chicago school. Perhaps for this reason Metzler's ideas failed to stimulate a research program at Chicago (or elsewhere). Instead, Chicago's monetarists pressed ahead with refinements of the quantity theory of money.

The Monetarists and the Modern Quantity Theory

During the 1950s the monetarists attempted to recover the popularity that classical theory had lost as a result of the Great Depression and the development of Keynesian macroeconomics. The most prominent monetarist was (and remains) Milton Friedman, an economist at the University of Chicago. One of Friedman's first major contributions to monetarism was a refinement of Fisher's quantity theory.

As Friedman pointed out, both the size of the money supply and the general level of prices can be considered public knowledge. However, for a theory to be able to make predictions regarding the effects of changes in monetary policy or changes in the price level, it is necessary to establish some assumptions about what differentiates the behavior of money supply from the behavior of money demand. In Friedman's words. "The quantity theory is in the first instance a theory of the demand for money. It is not a theory of output, or of money income, or of the price level. Any statement about these variables requires combining the quantity theory with some specifications about the conditions of the supply of money . . ." (1956, 4).

Friedman's version of the quantity theory is based on the postulate that there is a stable demand for real money balances—that is, for purchasing power in monetary form. He assumes that in the long run the level of money demand depends on economic fundamentals such as real income, the interest rate, and the nature of the technology for conducting transactions. Under this assumption, changes in the nominal supply of money engineered by the Fed have no long-run impact on the real demand for money and consequently lead inevitably and exclusively to changes in the price level. This observation is true both for one-time changes in the money supply and for changes in the rate at which the money supply is growing, which would result in changes in the inflation rate but not in the levels or growth rates of real variables. Thus, one implication of Friedman's restatement of the quantity theory of money is that changes in monetary policy would have no real effects in the long run—that is, money would be long-run superneutral.

Money in Keynesian Analysis: The Second Time Around

As the discussion has shown, early Keynesians focused their attention on fiscal policy. They believed that under normal circumstances changes in the general price level would be both infrequent and relatively inconsequential. As a result, for many years after the Second World War monetarists enjoyed a virtual monopoly over monetary analysis. This situation changed in the mid-1960s, when Keynesians developed a strong interest in the role of monetary policy.

Keynes himself rejected the quantity theory approach to determining the price level. For Keynes, the magnitude of the money supply in the economy was only one of a number of factors affecting the general level of prices. Another important factor was the level of employment. In Keynes's view it was impossible to determine the ultimate impact of a change in the quantity of money on the price level without considering the economy's overall level of employment. More specifically, Keynes believed that "an increase in the quantity of money will have no effect whatever on prices, so long as there is any unemployment" (1964, 295). Since Keynes saw persistent unemployment as the central problem facing industrialized economies, he did not think it would be unusual for economies to go for extended periods of time without observing significant changes in the price level. During the 1950s the general price level was indeed fairly stable. This circumstance lent credence to the Keynesian view that focusing on fiscal policies that might help solve chronic unemployment problems was likely to be more fruitful than devoting a lot of energy to analysis of price level determination.

As the postwar era wore on, inflation began to pick up in both the United States and Western Europe. This development stimulated interest in analyzing the causes of and cures for inflation. In 1958 British economist A.W. Phillips published an empirical analysis of historical data for the U.K. labor market. He hoped to find empirical support for the Keynesian hypothesis that the rate of wage inflation depended on the tightness of the labor market. Phillips found that from 1861 to 1957 the growth rate of nominal wages was negatively related to the rate of unemployment. This "Phillips curve" seemed to link the real side of the economy (the rate of unemployment) to the nominal side (nominal wages). And since wages are the biggest single component of firms' costs, most economists were willing to assume that persistent increases in wage rates would eventually force firms to begin increasing their prices, producing economywide price inflation.[5]

Although Phillips's findings were empirical in nature, they have had a profound and lasting effect on the development of economic theories about the relationship between inflation and real economic variables. As the discussion has shown, Keynesian theory holds that it is possible to use fiscal or monetary policy to increase or decrease the level of aggregate demand and through it the level of employment. The Phillips curve created a link between the level of aggregate demand and the rate of inflation. As a result economic policymakers began to think of demand management policies as involving a trade-off between the unemployment rate (and, more generally, the level of real economic activity) and the inflation rate. And if the Phillips curve was stable over time then this trade-off would exist in both the short run and the long run.

Long-Run Nonsuperneutrality of Money: The Keynesian School

The first attempt to formalize the Keynesian view about the long-run real effects of monetary policy was presented by James Tobin. Unlike the classical economists (but like

Metzler), Tobin saw real economic activity in general, and real interest rates in particular, as being determined jointly by economic fundamentals and by monetary policy—even in the long run. In Tobin's words, "Keynes gave reasons why in the short run monetary factors and portfolio decisions modify, and in some circumstances dominate, the determination of the interest rate and the process of capital accumulation. I have tried to show here that a similar proposition is true for the long run. The equilibrium interest rate and degree of capital intensity are in general affected by monetary supplies and portfolio behavior, as well as by technology and thrift" (1965, 684).

Tobin's analysis resembled Metzler's in abstracting from labor markets and concentrating on portfolio adjustments as the channel by which monetary policy could have long-run real effects. According to Tobin's theory, both money and physical capital were elements of an individual's portfolio of savings. For a given real rate of return on capital, an increase in the rate of inflation would make money less attractive and capital more attractive, inducing individuals to reduce their holdings of money in favor of holdings of physical capital. As a consequence, one would observe additional accumulation of capital, a higher capital stock, and a higher output level in the long run.

Long-Run Superneutrality of Money: The Monetarist School

What was the monetarist reaction to Keynesian's claim about the long-run effects of monetary policy? Monetarists did not address Tobin's arguments directly. Instead, they attempted to provide a theoretical underpinning for empirical work of the type conducted by Phillips (1958) that analyzed the relationship between nominal and real variables. Once the theoretical framework was in place, the monetarists used it to explain why monetary policy–induced changes in real economic activity would be short-lived.

While Phillips's statistical evidence involved nominal wages, standard economic theory assumes that households and firms base their employment decisions on real (inflation-adjusted) variables such as real wages, real interest rates, real profits, and so forth. Thus, additional assumptions were needed to reconcile standard economic theory with Phillips's findings. Ironically, the point of departure for this reconciliation was Keynes's observation that "every trade union will put up some resistance to a cut on money wages, however small, but no trade union would dream of striking on every occasion of a rise in the cost of living" (1964, 14–15). Friedman (1968) and Phelps (1967) used Keynes's observation in an attempt to extract some economic content from the statistical relationship discovered by Phillips. Their explanation for the behavior Keynes described was based on two assumptions—one about the nature of monetary policy, and the other about economic decisionmaker's responses to the effects of monetary policy. The first assumption is that increases in the money supply often cause "monetary surprises"—unexpected increases in the rate of inflation. The second assumption was that economic decisionmakers' reaction to monetary surprises often involves temporary money illusion, which is a temporary failure to recognize that there has been an increase in the price level. The basic idea here is that although monetary surprises increase the prices of all goods and services, economic decisionmakers usually notice the effects of these increases on particular prices in which they have a special interest—their wages or the prices of the goods they produce—well before they notice their effects on the overall price level. Until they discover that the overall price level has increased, they mistakenly believe that the increases in the money (nominal) prices of the goods they care about represent increases in the real prices (relative prices) of those goods. This mistaken belief can lead households or firms to make decisions about saving, consumption, work effort, investment, production, and so forth that are quite different from the decisions they would have made otherwise. As a result, by creating monetary surprises monetary policy can influence the level of real economic activity.

Here is a hypothetical sequence of events that illustrates how temporary money illusion can empower monetary policy: Suppose the economy starts out in its long-run equilibrium at its normal inflation rate and its "natural" real rate of interest. Suppose further that the monetary authority begins to increase the money supply at a faster pace than previously. The most immediate consequence of this move will be a drop in nominal interest rates. Friedman explains. "Let the Fed set out to keep interest rates down. How will it try to do so: By buying securities. This raises their prices and lowers their yields . . . In the process, it also increases . . . the total quantity of money. The initial impact of increasing the quantity of money at a faster rate than it has been increasing is to make interest rates lower for a time than they would otherwise have been" (1968, 5–6).

The next step in Friedman's chain of causation is that lower interest rates will stimulate spending, and this increase in spending will have a multiplier effect on the overall level of economic activity. Friedman writes. "The more rapid rate of monetary growth will stimulate spending . . . one man's spending is another man's income" (1968, 5–6). From this point, Friedman's analysis can be illustrated using the aggregate demand and aggregate supply (AD and AS) diagram that appears in many textbooks in introductory macroeconomics. The economy starts out in a long-run equilibrium at the intersection of the AD and AS curves. The intersection point represents the long-run equilibrium levels of real output and the price level. In the AD-AS framework, a change in a variable like the market interest rate leads to changes in the market environment that determined the location of the AD and AS curves and consequently produces a shift in at least one of these curves. In this case, the increase in spending that results from the decline in interest rates (which was caused by the increase in the money supply growth rate) is represented by a rightward shift in the AD curve. This increase in aggregate demand produces an increase in output and prices along the original aggregate supply curve.

According to Friedman, this change in the equilibrium will be strictly a short-run phenomenon. As soon as households and firms realize that lower interest rates and faster-rising wages and product prices are also associated with a more rapid rate of increase in the overall price level—as soon, that is, as they realize that real wages and prices have not changed—the households will reduce their supply of labor and the firms will cut back their production. On the diagram, this behavior is represented by a leftward shift in the aggregate supply schedule that exactly offsets the effects of the increase in aggregate demand. In the end, the economy will return to the original long-run natural level of economic activity but a higher rate of inflation. Friedman writes. "Rising income will raise the liquidity preference . . . and the demand for loans; it may also raise prices, which will reduce the real quantity of money. These three effects will reverse the initial downward pressure on interest rates in something less than a year. Together they will tend, after . . . a year or two to return [real] interest rates to the level they would otherwise have had" (1968, 5–6).

Friedman's theory of the short-run effects of monetary policy is sometimes described as the liquidity effect theory. In recent years this theory has been the basis for a great deal of recent research, both empirical and theoretical, about the short-run effects of monetary policy.

As the discussion has indicated. Friedman's liquidity effect theory is based on the belief that in the short run the decisions of firms and households are influenced by money illusion. In this theory, an increase in production and employment occurs not because there has been a change in economic fundamentals but because a more rapid rate of monetary growth has produced a higher rate of inflation. In Friedman's words. "The monetary authority can make the market rate less than the natural rate [of interest] only by inflation" (1968, 7).

The monetary surprises/money illusion hypothesis of Friedman and Phelps seemed to reconcile classical economic principles with the existence of Phillips-type relationships (a negative relationship between inflation and the real interest rate, a positive relationship between inflation and the level of real output, and so forth) created by monetary policy. Under this hypothesis the Phillips curve continued to represent a menu of choices involving trade-offs between real and nominal variables that were available to monetary policymakers—but only in the short run.

The Accelerationist Hypothesis

In tandem with this money illusion hypothesis, monetarists held firm to the classical premise that in the long run all real economic variables such as the real interest rate or the real unemployment rate have a natural level that is determined by economic fundamentals and is completely independent of the nature of monetary policy. In their view, temporary money illusion was the only mechanism by means of which monetary policy could affect real economic activity. It followed from these premises that continuous efforts by monetary policymakers to stimulate economic activity would translate mostly into an ever-increasing rate of inflation. While it might be possible for monetary policy to influence the level of interest rates (in particular) and real economic activity (in general) in the short run, once households and firms recognized that the rate of inflation had increased, the aggregate supply would shift back and the real effects of an increased inflation rate would disappear. Further reductions in interest rates and further stimulus to economic activity could be attained only via further increases in the rate of inflation. In Friedman's words, "Let the monetary authority keep the nominal market rate for a time below the natural rate by inflation. That in turn will raise the nominal natural rate itself, once anticipations of inflation become widespread, thus requiring still more rapid inflation to hold down the market rate" (1968, 7–8). The view underlying this "accelerationist" hypothesis is that while economic decisionmakers cannot be fooled permanently by a single increase in the inflation rate, they can be fooled persistently by accelerating inflation—that is, by a price level that increases over time at an increasing rate.

THE MONETARISTS AND THE KEYNESIANS IN PERSPECTIVE

The monetarists' persistent attacks on the Keynesians failed to convince the Keynesians that systematic efforts to use monetary policy to affect economic activity would fail. The monetarist argument that attempts to exploit the short-run inflation-unemployment trade-off would lead to accelerating inflation convinced Keynesians that balancing the competing economic goals of keeping inflation low and keeping real economic activity brisk would be harder than they had thought. However, the argument did not convince them that this balancing act was impossible.

To reiterate, during the 1960s Keynesian theorists came to regard the Phillips curve as a menu of options between inflation and unemployment from which policymakers could choose. They assumed that the Phillips curve was stable, which implied that monetary policy was powerful both in the short run and in the long run (that is, that money was not long-run superneutral). To Keynesians, the job of macroeconomic policymakers was to design demand-management policies that would strike the right balance between the competing problems of sustaining robust economic activity and controlling inflation.

Monetarists, on the other hand, believed the economy would be better off if the Federal Reserve supplied money according to a fixed, publicly announced formula and did not try to influence the level of real economic activity. Monetarists such as Friedman and Phelps disagreed with Keynesians regarding the effectiveness and usefulness of demand management. They viewed the "natural rate of unemployment" (the analog of Friedman's natural rate of interest: see above), together with the quantity theory of money, as solid enough arguments to assert beyond doubt the undesirability of activist monetary policy and the long-run superneutrality of money. Monetarists acknowledged the possibility that monetary policy might have short-run effects on employment, interest rates, and private spending, but they believed that these effects arose exclusively from the public's misperception of the impact of changes in the price level. According to the monetarists, the only way the monetary authority could have persistent real effects was by producing an ever-accelerating rate of inflation.

The debate between the monetarists and the Keynesians sometimes took the form of disagreements about the slope of the Phillips curve. These disagreements reflected differing views about the effectiveness of monetary policy in the short run versus the long run. During the 1970s, the Keynesians attempted to capitalize on the monetarists' tendency to frame the debate about monetary policy in terms of short- and long-run effects. Their strategy involved reinterpreting the Phillips curve in a way that reconciled the Keynesian and monetarist views of the timing of the inflation-unemployment relationship. This strategy forced the Keynesians to acknowledge that there were limits on the exploitability of the Philips curve.

A key element of the "compromise" offered by the Keynesians was the NAIRU, an acronym that stands for "non-accelerating inflation rate of unemployment" (see Espinosa and Russell 1997). In a diagram of the Phillips curve, the NAIRU is the unemployment rate at which the negatively sloping Phillips curve intersects Friedman's natural rate of unemployment. Monetarists believed that the existence of a natural rate implied that there was no useful trade-off between inflation and unemployment. Keynesians, however, interpreted the natural rate as a long-run constraint that policymakers have to face when trying to exploit an inflation-unemployment trade-off that remained both available and helpful in the short run. This revised Keynesian view of the trade-off was accepted by most policy-oriented economists and most economic policymakers. In the words of Tobin, the "consensus view accepted the notion of a nonaccelerating inflation rate of unemployment . . . as a policy constraint on policy" (1980, 24).

In retrospect it is clear that as much as the monetarists tended to dismiss Keynesian views, in many ways the two schools were not very far from each other—particularly in their analyses of the short-run consequences of monetary policy. These similarities become more evident when the monetarist-Keynesian debate is put in historical perspective. The years since the 1970s have witnessed the development of "neoclassical" macroeconomics—a new school of macroeconomic thought that is based on classical principles even more firmly than monetarism. One of the most influential branches of neoclassical macroeconomics is real business cycle theory. According to real business cycle pioneers such as Kydland and Prescott (1982) and Nelson and Plosser (1982), the cyclical pattern of recessions and expansions has little to do with monetary policy and can be explained almost entirely by "real shocks"—technological developments, changes in tax policy, and other unpredictable changes in economic fundamentals. Thus, the real business cycle theorists believe monetary policy has few or no effects even in the short run.

As economist Joseph Stiglitz points out. "Friedman is, in many ways, closer to the Keynesians than to the real business cycle theorists. He believes, for instance, that there are short run rigidities . . . such that any action by the monetary authority cannot immediately and costlessly be offset by changes in the price level" (1991, 48). Stated differently, the short-run predictions of the Keynesians and the monetarists differed in magnitude but not in direction. Both groups believed in a monetary policy transmission mechanism under which an increase in the money supply leads to an increase in economic activity accompanied by an increase in the general price level. The disagreement about magnitudes could, in principle, have been settled by the analysis of the empirical evidence (although in practice this was no easy task). But as long as the monetarists conceded that monetary policy had some short-run real effects it was impossible for them to make an unequivocal case against the exploitability of the Phillips curve.

In summary, the classical school saw the long-run level of economic activity as being determined independently of monetary policy. Metzler (1951) accepted much of the classical analysis but believed that there were situations in which monetary policy could have long-run real effects. The monetarists focused on money illusion as the only mechanism through which monetary policy could have real effects. In their view, economic fundamentals helped determine an individual's demand for money for transaction purposes. In the absence of surprises this money-demand relationship was fairly stable. It followed that in the long run, changes in the rate of money growth would produce proportional changes in the rate of inflation but would not affect real variables. Tobin (1965) sketched out a formal model in which changes in the rate of money growth could have long-run real effects. In his portfolio-based analysis, a permanent increase in the inflation rate led to more capital accumulation and a lower real rate of return on physical capital. Keynesians implicitly accepted the monetarist view of the role of money illusion in empowering monetary policy. They came to view Friedman's natural rates, which could be interpreted as long-run equilibrium values determined exclusively by fundamentals, as long-run constraints on policy strategies that remained effective in the short run. The short-run policy effect predictions of the Keynesians and the monetarists differed in regard to magnitude and persistence but not in regard to direction. Both schools agreed that in the short run a higher rate of money growth was associated with a higher rate of inflation, a lower real interest rate, and a spurt in economic activity. Keynesians did not themselves develop theories in which monetary policy was powerful in the short run but money was superneutral in the long run. Instead, they implicitly accepted the theoretical framework provided by their critics, the monetarists, although the two schools continued to disagree about some of the implications of this framework.

To this day, much of the economics profession continues to regard Keynesians' acceptance of the monetarists' position regarding long-run superneutrality as proof that there has been a rigorous scientific synthesis of the two theories. As discussed below, however, any synthesis of this sort is likely to be internally inconsistent.

THE NEOCLASSICAL SCHOOL

The arguments made by Friedman and Phelps against Keynesian theory were extended by economists such as Lucas (1972) and Sargent and Wallace (1976), who became the founders of the neoclassical school.[6] Lucas's 1972 article set the standards for neoclassical macroeconomics and, to a large extent, for all modern macroeconomics. The two pillars of his analysis were his assumption that economic decisionmakers had rational expectations and his use of a dynamic general equilibrium model. A dynamic general equilibrium model is a model that takes into account the intertemporal nature of many economic decisions and recognizes that economic variables interact with each other. Therefore, to determine the consequences of a postulated policy experiment one has to consider the relevant economic variables simultaneously and through time.

Lucas's article presented a very rigorous description of a situation in which (1) money is superneutral in the long run, and (2) the short-run real effects of monetary policy are bound to be rather limited, even in a scenario involving accelerating prices. A first step toward understanding Lucas's analysis is to recognize a key distinction between his assumptions and those of Friedman and Phelps. A simple way to describe this distinction is to say that the Friedman and Phelps analysis permitted persistent money illusion while Lucas's analysis ruled out persistent money illusion. Stated differently, the Friedman and Phelps analysis was based on the assumption that changes in prices or wages could cause households and firms to make "bad" economic decisions—decisions they would not have made if they had used available economic information more efficiently and had displayed more flexibility in reacting to the changes. Lucas, in contrast, assumes that the public processes economic information as efficiently as possible: in particular, individuals base their current decisions on the best possible forecasts of future events. His description of this decision-making process includes specific assumptions about how people form their economic expectations.

In Lucas's model there are two types of changes in prices: temporary changes in prices in particular industries, which are caused by short-run fluctuations in the demand for the goods produced by those industries, and changes in the price level, which are caused by changes in the growth rate of the money supply. There are also two types of changes in the growth rate of the money supply: systematic, permanent changes in the average (long-run) money growth rate and unsystematic, temporary changes in the current (short-run) money growth rate. The systematic changes result from deliberate changes in policy by the central bank; they produce a permanent increase in the average rate of inflation. The unsystematic changes result from errors in the implementation of the central bank's operating procedures. They do not reflect deliberate policy decisions, and they do not affect the long-run average money growth rate or inflation rates. They do, however, produce temporary changes in the current rate of inflation.

As has been indicated, Friedman and Phelps had assumed implicitly that economic decisionmakers have access to complete economic information but fail to use it efficiently. Lucas, on the other hand, assumes explicitly that decisionmakers use any information available to them in the most efficient way but do not always have access to complete information. The particular aspect of the economy that Lucas assumes decisionmakers do not have complete information about is the relationship between changes in the relative prices of the particular goods they produce and changes in the overall price level. This information gap is important because fully informed decisionmakers will react very differently to changes in the prices of their goods that represent changes in relative prices—that is, to situations in which the prices of their goods change but the general price level remains constant, or situations in which the general price level changes but the prices of their goods change by a larger or smaller proportion—than to changes in the prices of their goods that simply follow along with changes in the overall price level. More specifically, decisionmakers have no incentive to increase their work effort and production in response to increases in the overall price level for the same reason that one would not be any happier if a doubling of salary coincided with a doubling of the price of every good purchased. On the other hand, it makes sense for a person to increase effort and output if the relative price of the good produced has increased. Under Lucas's assumptions any such increases in effort and output will be temporary because the demand fluctuations that induce them are also temporary.

Now suppose that at a given moment in time, and in the absence of any changes in the economy's fundamentals, the overall inflation rate increases because of an unsystematic increase in the money supply. As the overall inflation rate increases, prices in every sector or industry increase. However, individuals are unable to tell, immediately, whether the price increases affecting their sector are relative or absolute changes. The reason is that people are assumed to have better information about prices of the goods and services in their industry than about the many different prices that figure in the overall price level. This lack of complete information about the overall level of prices leads people to assume that at least part of the increase in the price of their product has been caused by an increase in its relative price. As a result, they increase their work effort and production.

The situation just described seems quite consistent with the Keynesian notion that there is a short-run trade-off between rate of inflation and the level of economic activity. But does this trade-off indicate that monetary policy is powerful, in the sense that the central bank can use it to control the level of economic activity? Is this a model of the "tightrope walk" that aggregate demand managers are often described as having to perform? If the central bank in the model can use monetary policy actions to exert continuous and repeated influence over individual decisions concerning work effort and production, then the answer to these questions may be yes.

This situation turns out not to be possible, however. Suppose that the central bank announces a permanent change in the average growth rate of the money supply, Lucas's assumption that people have rational expectations implies that they understand the nature of the relationship between money growth and inflation. As a result, they will not increase their work effort or production in response to the resulting increase in the average rate at which prices change. Thus, permanent increases in the money growth rate have no effect on the level of output or employment, while temporary increases in the money growth rate will produce temporary increases in both output and employment.

Thus, in Lucas's model the rational expectations assumption implies that systematic changes in monetary policy should not have real effects. The rigorous nature of Lucas's analysis made his argument seem very convincing. It is important to note, however, that the argument also depends on Lucas's assumption, which is built into the structure of his model, that the effects of monetary policy on real economic activity are caused only by money illusion.

Lucas's argument can be illustrated further by returning to the context of his model and exploring its implications in a somewhat less rigorous way. Suppose that the central bank in his model attempts to exploit the apparent trade-off between inflation and output by increasing the average money growth rate without making any announcement that it has done so. It is hoping that people will make inflation-forecasting mistakes because they will not recognize that any policy change has occurred. The increase in money growth will, of course, produce a permanent increase in the average inflation rate. Initially, people will mistake this systematic, policy-induced increase in the inflation rate for an unsystematic inflation rate increase caused either by a temporary demand disturbance or by an error in the execution of the original monetary policy rule. Since they will not be sure which of these two types of unsystematic increase has occurred, their work effort and production will rise (see above). Soon, however, people will start to recognize that there is a pattern to the unusually high rates of inflation they are observing. As a result, they will begin to think it less and less likely that the next above-average increase in the inflation rate was caused by a demand disturbance, and they will begin to cut back on their above-normal production and work effort. Ultimately, they will realize that the central bank has changed policy in a way that has caused the average inflation rate to increase. At this point, the increase in the average inflation rate will no longer have any effect on work effort and production.

The scenario just described suggests that systematic changes in monetary policy may have substantial short-run effects but no long-run effects, just as the monetarists argued and just as the Keynesians ultimately conceded. Suppose, however, that the central bank tries to repeat its short-run success by changing the average inflation rate from time to time in response, say, to other changes in the state of economy. In the real world people learn from past mistakes: as a result, each time the central bank engineers another systematic change in the inflation rate people will catch on to the policy change more quickly and the effects of the change will disappear more quickly. At some point, moreover, people will figure out which events motivate the central bank to change policy; they will then be able to detect policy changes as soon as they occur. At this point the policy changes will no longer have any effects, even in the short run.

These modified rational expectations assumptions about the way people obtain and use information seem consistent with one's economic intuition about the behavior of actual households and firms. In real-life economies, most people have a very good "micro" picture of the status of their firm or industry but a fairly fuzzy "macro" picture of the state of the economy at large. However, once they start getting surprised by unexpected price changes that make their decisions work out badly they become more interested in identifying the causes of changing prices. They start to use any information available to them to try to anticipate changes in the price level and distinguish them from changes in relative prices. As a result, future price level increases have less and less surprise effect. This sort of intelligently adaptive behavior is the real-life analogue of Lucas's formal assumption that economic decisionmakers have rational expectations.

Lucas's argument, and the closely related arguments of neoclassical economists such as Sargent and Wallace (1976), left Keynesians with only two intellectually legitimate choices. First, they could have tried to capture their intuition about the effects of monetary policy in a rational expectations general equilibrium model in which money was not long-run superneutral because monetary policy derived its real effects from some source other than monetary surprises. Many economists expected this approach. Sargent, for example, writes that "in the early 70's, I thought that Modigliani. Solow, and Tobin—our heroes in those days—were missing the boat by resisting the intrusion of rational expectations into macroeconomics, instead of commandeering it. Despite the appearances of its early incarnations like Lucas's 72 JET paper, the canons of rational expectations models . . . were evidently wide enough to include Lucas's brand of monetarism or, just as readily, accommodate the completion of Tobin's criticism of monetarism by fully bringing to bear the logic of Modigliani and Miller. Modigliani, Solow and Tobin chose not to commandeer the movement, and left it to Kareken. Wallace, Chamley, Bryant and others to draw out many of the nonmonetarist implications then waiting to be exposed." (1996, 545). In retrospect it seems clear, as this quotation indicates, that an important reason Keynesians did not pursue this strategy was because they mistakenly believed that rational expectations implied long-run superneutrality of money.

Another alternative for Keynesians might have been to concede that monetary policy was long-run superneutral but to argue that frictions of various sorts might allow monetary policy to have real effects in the short run. Taylor's work on staggered contracts (1980), his work on slow adjustment of prices (1994), and the work of Ball and Mankiw (1995) concerning "menu costs" are illustrations of this line of research. This research has faced criticisms on two fronts. First, there is little empirical evidence to support this type of nominal rigidities (see, for example, Wynne 1995). Second, there has not been a clear explanation as to why these frictions could prevent people from changing their behavior so as to offset the effects of systematic changes in monetary policy. For example, what prevents individuals from relying on mechanisms such as indexing of nominal contracts to guard against the potential negative effects of nominal rigidities?

Most Keynesians chose to ignore the neoclassical critique and the potential problems with short-term frictions. They continued to claim that monetary policy had powerful short-run effects, while accepting the monetarist critique that it was powerless in the long run. For the most part, economists outside academia—business economists and economic policymakers—have adopted this "Keynesian consensus" view. To the extent that either group of economists has attempted to justify this view, they have done so by arguing that rational expectations is a sensible assumption only in the long run. In the short run, they argued, people could and often did misread the nature and effects of changes in monetary policy.

What is wrong with the Keynesian consensus? Lucas points out that, while it may seem reasonable on its face, it suffers from serious logical problems. Com-menting on Tobin's description of the Keynesian consensus, Lucas writes, "Here we have Model A, that makes a particular prediction. We have model B, that makes a strikingly different prediction concerning the same event. The event occurs, and Model B proves more accurate. A proponent of model A concludes: 'All right, I "accept" Model B too.' Consensus economics may be a wonderful thing, but there are laws of logic which must be obeyed . . . These models gave different predictions about the same event because their underlying assumptions are mutually inconsistent. If the Friedman-Phelps assumptions are now 'accepted,' which formerly accepted Keynesian assumptions are now viewed as discarded? Tobin does not say" (1981, 560–61). Lucas goes on to spell out the monetarist (model B)-Keynesian (model A) consensus, as viewed through the Keynesian glass. He writes, "Though I refer to Tobin as 'evading' a central issue, I do not think he sees it this way at all. He writes as though he is willing to concede the 'long-run' to Phelps and Friedman [the Monetarists], claiming only the 'short-run' for Keynesians. Where is the conflict?" (561). Lucas goes on to explain that the long run consists of a sequence of short runs. If a policymaker conducts short-run policy by choosing an annual money growth rate based on model A (the Keynesian model) every year, then he or she has implicitly used model A to pick the average rate of money growth for the long run. It is logically inconsistent to pretend that the long-run average money growth rate using model B (the monetarist model) can be a guide. Suppose, for example, that the central bank decides that in the long run the optimal growth rate of the money supply is 5 percent per year. However, it decides on the basis of short-run considerations that it would be a good idea to increase the money growth rate to 6 percent for the coming year. The same thing happens again in the following year, and in the year after, and so on. The end result is a departure from the optimal long-run money growth rate that may have adverse consequences for the economy. Thus, Lucas observes that "if we concede that Model A gives us an inaccurate view of the 'long-run.' then we have conceded that it leads us to bad short-run situations as well" (560–61).

Monetary Policy after Lucas

Starting in the late 1960s, Keynesian economic theory was the victim of a succession of setbacks, including the monetarist critique of Friedman and Phelps, the combination of high inflation and high unemployment that the United States experienced during the 1970s, and the neoclassical critique of Lucas (1972) and Sargent and Wallace (1976). As Keynesian theory lost ground in the academic community, so did belief in the power of monetary policy. In fact, much of the early work by neoclassical economists followed Lucas (1972) by constructing models that made debating points against the Keynesians by demonstrating that systematic changes in monetary policy would have no real effects, even in the short run. Unsystematic policy actions might have a short-lived influence on the level of economic activity, but any attempt to use systematic changes in policy to exploit this influence would be frustrated by changes in the expectations of the public.

The fact that the model described by Lucas (1972) became the "industry standard" in neoclassical theory has encouraged other neoclassical economists to focus on models that display long-run superneutrality of money. In recent years, the most popular vehicle for research on monetary policy by neoclassical economists has been the real business cycle model. In this model money is long-run superneutral, but temporary changes in monetary policy can generate small "liquidity effects" of the sort described in Friedman (1968). (See, for example, Lucas 1990: Christiano and Eichenbaum 1991, 1992: Fuerst 1992: Dow 1995).

During the mid-1970s economist Harry Johnson, reviewing what he labeled the Keynesian revolution and the monetarist counterrevolution, commented that "the monetarist counterrevolution has served a useful purpose, in challenging and disposing of a great deal of the intellectual nonsense that accumulates after a successful ideological revolution . . . If we are lucky, we shall be forced as a result of the counterrevolution to be both more conscious of monetary influences on the economy and more careful in our assessment of their importance" (1975, 106).

In fact, the monetarist counterrevolution had mixed effects on economists' views concerning the importance of monetary policy. On the one hand, monetarist arguments convinced many Keynesians that monetary policy had many of the same powers that they had attributed to fiscal policy. On the other hand the monetarists, as has been pointed out, completely dismissed the possibility that monetary policy might have long-run real effects. To the extent that Keynesians conceded this point they were also conceding that the importance of monetary policy was quite limited.

As shown above, the period of the monetarist counterrevolution was also a period when a few economists began to try to identify explicit mechanisms that would allow monetary policy to have long-run real effects. Metzler (1951) and Tobin (1965) developed theories that allowed the Keynesian, conventional wisdom to be extended to the long run. These theories allowed permanent increases in the money supply growth and inflation rates to be causally associated with permanently lower real interest rates and permanently higher levels of output.

Lucas's (1972) work suggested that macroeconomic theories of all sorts were in need of reevaluation. The theories of Metzler, Tobin, and the ones derived from Phillips's analysis were no exception. Lucas's interpretation of the Phillips curve analysis has been described above. The next section reviews subsequent research that tries to reformulate Metzler's and Tobin's theories using the neoclassical methodology Lucas introduced. This research indicates that departures from long-run superneutrality are possible because monetary policy does not necessarily derive all (or any) of its power from money illusion.

Instead, changes in monetary policy may have lasting effects because it affects the supply of or the demand for credit.

SOME NEOCLASSICAL MODELS THAT DELIVER LONG-RUN NONSUPERNEUTRALITIES

This section looks at the three challenges facing economists who want to develop neoclassical models that deliver results similar to those of Metzler, Phillips, and Tobin. The first challenge is simply to construct a plausible neoclassical model in which money is not long-run superneutral. The second challenge is to identify a mechanism under which the departures from superneutrality work in the "right direction," that is, a mechanism that allows increases in the money supply growth rate to be causally associated with lower real interest rates and higher levels of output. The third challenge is to find a mechanism that has some hope of generating departures from superneutrality that are large enough to have practical importance.

The Tobin Effect

An answer to the first challenge is to rely on the credit market as the ultimate source of the real effects of monetary policy. In this respect one could follow Tobin (1965). Tobin's analysis is based on the idea that the increase in the inflation rate that is induced by an increase in the money supply growth rate increases the supply of credit at any real interest rate. It does so because when the inflation rate rises money becomes a relatively unattractive asset, and the public wishes to cut back on its money balances and increase its holdings of bank accounts, bonds, stock, and other financial assets. Thus, there is a decrease in the demand for money and a matching increase in the supply of credit. The Tobin effect mechanism allows a permanent easing of monetary policy (a higher money growth rate) to lead to a higher inflation rate, a lower real interest rate (due to the increased supply of credit), and a higher level of output (due mostly to an increase in the capital stock).

Many economists believe that financial intermediation is one of the most important channels through which changes in monetary policy affect the economy (see for instance Bernanke and Gertler 1995). The central bank may be able to affect the composition of financial intermediaries' portfolios without relying on monetary surprises. Thus, permanent changes in monetary policy may affect financial intermediaries in a fundamental way and may have long-run real effects. It follows that a natural environment in which to analyze the Tobin effect would be one in which financial intermediaries were explicitly developed.

The starting point for assessing this possibility should be a realistic model of financial intermediation. Bencivenga and Smith (1991) were among the first economists to include financial intermediaries in a dynamic general equilibrium macroeconomic model. The Bencivenga-Smith intermediaries are similar to actual intermediaries in accepting deposits from, and lending to, a large number of individuals. They are also similar to actual intermediaries in making loans that are less liquid than the deposits they accept. As a result, they are forced to hold a liquid asset (money) on reserve to cover sudden deposit withdrawals.

In the Bencivenga-Smith model individuals could, in principle, manage their own asset portfolios (as in Tobin 1965). However, the financial intermediaries have an actuarial advantage over individuals in structuring a portfolio. Consequently, under most circumstances people will prefer to delegate this activity to financial intermediaries. Although Bencivenga and Smith's work contains the elements needed to pursue an analysis of the long-run effects of permanent changes in monetary policy, their analysis concentrates on the long-run implications of financial intermediaries for an economy's long-run performance. Based on the Bencivenga and Smith model. Espinosa and Yip (1998) study the growth-inflation implications of alternative fiscal and monetary policies in the presence of financial intermediaries. Espinosa and Yip can, thus, draw some qualitative lessons on the Tobin effect in a dynamic general equilibrium model that explicitly models financial intermediaries. Before listing their findings, it is useful to briefly review some recent empirical results on the relationship between inflation and growth.

Inflation and Growth

In part because money has been assumed to be long-run superneutral, there has not been much research on the long-run relationship between inflation and growth. Recently, however, interest in theoretical and empirical analysis of this relationship has revived. The empirical findings are not always in agreement. DeGregorio (1992) and Barro (1995) uncover a significant negative correlation between inflation and economic growth. On the other hand, Bullard and Keating (1995) and Bruno and Easterly (1998) do not find strong support for such an inverse relationship. Bullard and Keating find that the direction of the growth-inflation relationship depends crucially on the initial level of the inflation rate. In countries in which the rate of inflation starts out relatively low, a permanent increase in the inflation rate actually increases the long-run level of economic activity. Only for countries with relatively high initial inflation rates do Bullard and Keating find that permanent increases in the rate of inflation negatively affect long-run growth. These findings are partly confirmed by Bruno and Easterly, who are able to find an inverse relationship between inflation and growth only when the rate of inflation exceeds some critical value.

Clearly, these empirical studies do not settle whether monetary policy can have real effects that do not spring from monetary surprises and whether these effects are likely to be of the type described by Tobin, with higher inflation being associated with higher rates of growth. Espinosa and Yip (1998) address these questions in a model based on the model developed by Bencivenga and Smith (1991). Their analysis emphasizes the point (to be made very explicitly below) that fiscal and monetary policy are inevitably linked by the government budget constraint. In their model, monetary policy consists of changes in the growth rate of the money supply that are necessitated by changes in fiscal policy—specifically, by changes in the government budget deficit as a fraction of GDP.

Espinosa and Yip show that their model can produce the positive long-run relationship between inflation and growth that was predicted by Tobin. However, it is also possible for the model to produce situations in which lower rates of inflation result in higher rates of growth. The direction of the inflation-growth relationship depends on, among other things, the initial inflation rate, the degree of risk aversion of the average depositor, and the size of the government budget deficit. Thus, the Espinosa-Yip analysis provides a theoretical framework that helps reconcile the conflicting empirical findings about the direction of the long-run relationship between inflation and growth that were described in the preceding subsection.

Fiscal Policy and Open Market Operations

The Tobin effect has a potential drawback as a theory of the real effects of monetary policy (see, for example, Danthine, Donaldson, and Smith 1987). The shift in the credit supply curve produced by an increase in the inflation rate is essentially equal to the reduction in money demand that the increased inflation induces. Money demand is quite small (a small fraction of total output, or total assets, and so forth) and statistical evidence (for example, Hoffman and Raasche 1991) suggests that it is not very sensitive to changes in the inflation rate. As a result, the Tobin effect of moderate changes in the inflation rate on real interest rates and output is likely to be small.

An alternative mechanism for linking monetary policy and the supply of credit has been developed by Sargent and Wallace (1981). This mechanism is based on the fact that changes in monetary policy affect the government's stream of revenues and thus necessitate changes in fiscal policy. To gain a better understanding of this mechanism, it is useful to review some basic elements of a government's budget constraint.

An important premise of the research described in this section is that both fiscal and monetary policy actions are constrained by the government's need to finance its expenditures. Consequently, these two types of government policy cannot be devised or executed independently from each other. The government of a country must decide on the level of government spending to finance domestically, how much of its domestic financing will rely on current taxes, and how much will take the form of newly issued debt. Stated differently, it is the government budget deficit that determines the need for new issues of government debt. Since government borrowing competes with private borrowing in the credit market, the amount of government borrowing is likely to influence the level of real interest rates.

Government policy concerning taxes, debt, and deficits is usually described as fiscal policy. The analysis just presented suggests that fiscal policy may affect real interest rates. However, monetary policy has a fiscal policy aspect to it because it may play a role in determining the size of the government budget deficit. To the extent that monetary policy has this effect, this analysis suggests that it will also have an impact on real interest rates. Thus, monetary policy may influence the real economy in ways that do not involve inflation surprises. If this influence can persist in the long run then money may not be long-run superneutral.

In practice, monetary policy is carried out via open market operations. Open market operations produce changes in the composition of the government's portfolio of liabilities—debt (bonds and bills) versus money.[8] Given the amount of government bonds currently outstanding, the government must decide what fraction of these bonds (if any) it will "monetize" by purchasing them with newly created currency. This decision, which determines the composition of the government's liability portfolio, also determines the amount of outstanding government debt in the credit markets and consequently has an impact on the market real rate of interest. More specifically, changes in the growth rate of the money supply affect the volume of government revenue from currency seigniorage (the "inflation tax"). Sargent and Wallace (1981) assume that the government's primary (net of interest) budget deficit is fixed by the tax and spending decisions of Congress and is not affected by changes in monetary policy. Consequently, the only way the government can offset the changes in its revenues that are caused by changes in monetary policy is to change the size of the national debt. Thus, this mechanism can be thought of as the neoclassical successor of Metzler (1951) (because of Metzler's emphasis on open market operations as the mechanism through which non-long-run superneutrality results could be attained).

The size of the national debt has substantial effects on the state of the government budget. On the one hand, the government has to pay interest on the debt. On the other hand, as the economy grows the government can allow the national debt to grow at the same rate without increasing the size of the debt relative to the economy.[9] The relationship between these two factors determines whether debt service is a financial burden for the government or whether the existence of the national debt actually increases the amount of government revenue.

To see why the latter situation is possible, suppose the government borrows just enough each year to keep the debt-GDP ratio constant. If the economy is growing, it will increase its borrowing each year by an amount that causes the real national debt to grow at the same rate as real GDP. Although the government will have to use some of the proceeds of this new borrowing to pay the interest on the current debt, if the real (also inflation-adjusted) interest rate on the debt is lower than the real GDP growth rate then the government will have funds left over to use for other purposes. In this case, the national debt actually provides the government with revenue on net. This source of revenue is sometimes referred to as bond seigniorage.[10] The difference between the real growth rate and the real interest rate is the net real amount that each real dollar of debt contributes to the government budget each year.

If the real interest rate on the government debt is higher than the output growth rate then the government's new borrowing will not be enough to cover the interest on the existing debt. As a result, some of this interest will have to be covered by funds from other sources. In this case the national debt is a financial burden for the government. (One can think of this as a case in which bond seigniorage revenue is negative.) The difference between the real interest rate and the real growth rate is the net real amount that each real dollar of debt costs the government each year.[11]

Once it is known whether the national debt is a source or a use of government funds one is in a position to determine how a change in the size of the national debt will affect the government's budget position. Other things being equal, an increase in government borrowing that increases the size of the national debt represents an increase in the quantity of credit demanded at each real rate of interest and will consequently produce an increase in the real interest rate. If the real interest rate is relatively high, so that the debt is a burden on the government budget, then the combination of a larger debt and a larger unit cost of financing the debt means that the debt will definitely become costlier to the government. Conversely, a smaller debt that will result in a lower real interest rate will reduce the government's costs. As a result, when the government cuts the money supply growth and inflation rates and loses money from currency seigniorage, it must compensate by cutting back on its borrowing and driving the real interest rate down. As a result, tighter monetary policy produces lower real interest rates and a higher level of output.

Suppose, on the other hand, that the real interest rate is relatively low, so that the national debt is a source of revenue for the government. In this case, a given change in the size of the debt (say, an increase) can either increase or decrease government revenue from bond seigniorage. An increase in the size of the debt tends to cause bond seigniorage revenue to increase: this is the "tax base effect" of the increase. On the other hand, an increase in the government debt drives the real interest rate closer to the output growth rate and reduces the real seigniorage revenue produced by each real dollar of debt. This is the "tax rate effect" of the increase. If the tax base effect is stronger than the tax rate effect then an increase in the size of the debt will increase the government's bond seigniorage revenue: otherwise, the amount of revenue will fall.

The tax rate effect tends to be largest when the government debt is large, because in this case any change in the real interest rate affects the revenue produced by a large volume of debt. Conversely, the tax base effect tends to be largest when the real interest rate is low because each dollar of debt generates a lot of revenue. A low real interest rate tends to be associated with a small volume of government debt, since when the real interest rate is low private credit demand is high and private debt crowds out government debt. Conversely, a high real interest rate tends to be associated with a large government debt. As a result, when the real interest rate is relatively low—well below the output growth rate—an increase in the size of the national debt tends to increase bond seigniorage revenue while when the real interest rate is higher an increase in the size of the debt tends to decrease the amount of revenue.

One can now put all the pieces of this story together to determine the possibilities for the long-run real effects of monetary policy. If the real interest rate is higher than the output growth rate, or lower than the output growth rate but not too much lower then an increase in the size of the national debt decreases government bond seigniorage revenue and vice-versa. Thus, a decrease in the money growth and inflation rates that reduces government revenue from currency seigniorage will force the government to reduce the size of its debt and will drive the real interest rate down. This is the scenario described by Wallace (1984): it has the implication that monetary tightening will increase the level of real GDP in the long run. On the other hand, if the real interest rate is substantially below the output growth rate then a decrease in the money growth and inflation rates will allow the government to increase the size of its debt and will drive the real interest rate up. This is the scenario described by Espinosa and Russell (1998a, b). It is similar to the Tobin effect in having the Keynesian, or conventional, implication that a monetary tightening will reduce the level of real output.

Historically, the average real interest rate on U.S. government debt has been well below the average U.S. output growth rate. This situation makes Espinosa and Russell's Keynesian scenario seem plausible empirically. An additional reason why the scenario is appealing is that it weakens the link between the size of the money supply and the size of the shift in the credit supply curve that is induced by a change in monetary policy—the link that keeps the Tobin effect small. Although the fact that the money supply is small relative to GDP means that a change in the inflation rate will have a relatively small impact on government revenue (from currency seigniorage), if it takes a relatively large increase in the real interest rate to produce a substantial decrease in government revenue from bond seigniorage then the resulting change in the real interest rate and the level of output could still be large.

Why might it take a large change in the real interest rate to produce a substantial change in the revenue from bond seigniorage? When the real interest rate is low, the tax rate effect and the tax base effect tend to work against each other. As a result, the net change in the amount of revenue that is produced by a change in the real interest rate can be quite small. In fact, there is always a range of real interest rates over which the two effects offset each other almost perfectly. Over this range, the ratio of the change in the real interest rate to the change in the amount of revenue it produces will be extremely large.

The bottom line here is that, at least in principle, the Espinosa-Russell variant of the Sargent-Wallace "unpleasant arithmetic" can give us just what is needed: a theory that explains how a moderate but permanent increase in the money supply growth and inflation rates might result in a fairly large decrease in the real interest rate and a fairly large increase in the level of output.

Before concluding this section it is important to emphasize that the research just reviewed composes a relatively small part of the growing academic literature that studies the long-term effects of monetary policy in neoclassical models. Related work in this area includes Haslag (1998). Bhattacharya and others (1997), Schreft and Smith (1997), and Bullard and Russell (1998a. b). One implication of this line of research is that monetary economists may have spent too much time trying to forge direct links between changes in monetary policy and changes in the unemployment rate and the output growth rate. Instead, they perhaps should be devoting more effort to understanding the relationship between monetary policy and the economic fundamentals that drive saving and production decisions and also to exploring the relationship between monetary policy variables and "real" macroeconomic variables such as the government deficit, real interest rates, reserve requirements, and other variables that link the money market to the credit market.

CONCLUSION

This article has reviewed the history of the view that monetary policy has real effects in the short run but no such effects in the long run (so that money is long-run superneutral). This view grew out of a debate between the adherents of two influential schools of macroeconomic thought, the monetarists and the Keynesians. The final result of this conflict was a unilateral, Keynesian-produced synthesis that developed during the 1970s. Under this synthesis the Keynesians accepted the monetarists' view that money was superneutral in the long run but continued to disagree with them about the magnitude and desirability of the short-run effects of monetary policy on real interest rates, real GDP, unemployment, and other real variables.

The article has argued that the beliefs that monetary policy is powerful in the short run and that money is superneutral in the long run may not be mutually consistent. The basic problem with most theories that reconcile these beliefs is that they rely directly or indirectly on the assumption that economic decisionmakers are victims of money illusion. If money illusion is the reason monetary policy has real effects, however, then its short-run real effects will be small and policymakers will not be able to exploit them systematically to achieve their goals. This point has been demonstrated in seminal work by Lucas (1972).

In the years since the 1970s, academic macroeconomics has slowly but surely embraced the neoclassical methodology pioneered by Lucas, which employs dynamic general equilibrium models and assumes that decisionmakers have rational expectations. The results of Lucas's work and that of a number of other neoclassical economists has served to further strengthen the monetarist position concerning long-run superneutrality of money.

In recent years, empirical studies of the impact of monetary policy have concentrated on identifying its short-run effects. This focus has been motivated, at least in part, by the conviction that money is long-run superneutral. Many researchers seem to believe that there is overwhelming empirical evidence in favor of long-run superneutrality. In reality, however, the proposition that monetary policy does not have long-run real effects is far from unequivocally established: indeed, an exploration of the empirical literature on long-run superneutrality could easily be the subject of a separate article. For the purposes of this article, it may suffice to cite a remark by Robert King and Mark Watson, two prominent macroeconomists whose empirical research has produced evidence both for and against long-run superneutrality. King and Watson (1992) report that for the United States during the postwar period the data do not appear to be consistent with the hypothesis that, over the long run, money is superneutral or that nominal interest rates move one-for-one with inflation.

The fact that the empirical evidence on the long-run superneutrality of monetary policy is not as overwhelming as some analysts believe suggests that there may be a need to look at theories that explore potential sources of long-run real effects for monetary policy. As this chapter has explained, Lucas's path-breaking work was an attempt to conduct a rigorous analysis of the logical consequences of the monetarist assumption that the real effects of monetary policy result from monetary surprises—a fact that has led Tobin, a leading Keynesian, to refer to Lucas's model as the "Monetarist Mark II" model. Lucas did not attempt to argue that every reasonable combination of assumptions would produce superneutrality, and it is consequently a mistake—albeit a very common mistake, even in the academic community—to equate neoclassical economics with the proposition that money is superneutral.

To repeat, Lucas's renowned 1972 paper employed innovative methodology to explore the implications of a very particular set of assumptions. The methodology is logically separate from the assumptions and can be used to analyze the consequences of very different assumptions. In fact, it is possible that monetary policy influences real economic activity for reasons completely different from the ones Lucas identified. In a recent interview in *New Yorker* magazine, Lucas acknowledges that the real effects of monetary policy may not result from unexpected policy changes. He comments, "Monetary shocks just aren't that important. . . . There's no question, that's a retreat in my views" (Cassidy 1996, 55).

Abandoning the assumption that policy surprises are the main reason monetary policy can have real effects leaves two options. One is to accept the view of the real business cycle theorists that Federal Reserve policy actions are essentially irrelevant. A second option is to attempt to identify alternative channels through which the monetary authority could affect real economic activity. This article has reviewed a small part of the recent academic literature that explores the second option. The results reported in this literature indicate that monetary policy may be a great deal more powerful than most academic economists believe. They also suggest that the possibility that monetary policy has substantial long-run real effects deserves more attention from economists and policymakers.

NOTES

1. Linton Weeks and John Berry, "The Shy Wizard of Money: Fed's Enigmatic Greenspan Moves Easily in His Own World." *Washington Post,* March 24, 1997, sec. A.
2. Money is said to exhibit long-run *neutrality* if permanent changes in the level of the supply of money have no long-run effects on real interest rates or the growth rate of real output. In this case, the levels to which prices and other nominal variables will increase are postulated to vary one for one with changes in the level of the money supply. Similarly, an economy is said to display long-run superneutrality if permanent changes in the rate of growth of the money supply have no long-run effects on either real interest rates or the rate of output growth, and the rates of inflation and other nominal variables are postulated to vary one for one with changes in the rate of growth of the money supply.
3. If monetary policy has real effects only because of money illusion then it is likely these effects will be very limited in scale. On the other hand, if monetary policy derives its real effects from other sources, then its short-run effects may be relatively large. Thus, limited short-run real effects may go hand in hand with long-run superneutrality while deviations from long-run superneutrality may produce powerful short-run effects.

4. In particular, as stated, this description of the quantity theory is really more of an accounting identity than a theory that qualitatively relates money to relevant macroeconomic variables. An accounting identity does not specify what is a given in the analysis and how different variables will change as a result of alternative policy changes. A theory or model, on the other hand, is specific about what is assumed to be exogenous to the model as well as what is determined within it and how different variables react to exogenous changes. A number of complimentary assumptions were really necessary for this equation to spell the list of properties Fisher attached to the quantity theory.
5. For this explanation to make sense some additional assumptions are required. See Espinosa and Russell (1997) for an explanation of these assumptions.
6. For a detailed nontechnical description of Lucas's contribution see Espinosa and Russell (1997).
7. Of course, if fiscal and monetary policy interactions led not only to long-run output level changes but to output growth changes, the Tobin effect could be of more significance.
8. In principle, of course, there exists the possibility that such a swap of liabilities results in no effects either real or nominal, either in the short or long term (a case made by Wallace 1984 and Sargent and Smith 1987 but not reviewed here), but under most circumstances it will.
9. The debt-GDP ratio cannot continue to grow forever. Otherwise, at some point the debt would get so large relative to households income that it would be impossible for them to save enough to hold it.
10. This term seems to have been first used by Miller and Sargent (1984).
11. Thus, if the real interest rate is 2 percent higher than the real growth rate then each dollar of debt costs the government two cents each year, adjusted for inflation.

REFERENCES

Ball, L., and Mankiw N.G. 1995. Relative-Price Changes as Aggregate Supply Shocks. *Quarterly Journal of Economics* (February):161–93.

Barro, R. J. 1995. Inflation and Economic Growth. Bank of England *Quarterly Bulletin* (May): 166–76.

Bencivenga, V. R., and Smith, B. D. 1991. Financial Intermediation and Endogenous Growth. *Review of Economic Studies* 58:195–209.

Bernanke, B. S., and Gertler, M. 1995. Inside the Black Box: The Credit Channel of Monetary Policy Transmission. NBER Working Paper No. 5146.

Bhattacharya, J., Guzman, M. G., Elisabeth Huybens, E., and Smith, B. D. 1997. Monetary, Fiscal, and Reserve Requirement Policy in a Simple Monetary Growth Model. *International Economic Review* 38:321–50.

Bruno, M., and Easterly, W. 1998. Inflation Crises and Long-Run Growth. *Journal of Monetary Economics* 41(1):3–26, February.

Bullard, J., and Keating, J. 1995. The Long-Run Relationship between Inflation and Output in Postwar Economies. *Journal of Monetary Economics* 36:477–96.

Bullard, J., and Russell, S. 1998a. How Costly Is Sustained Low Inflation for the U.S. Economy. Federal Reserve Bank of St. Louis and IUPUI Working Paper.

Bullard, J., and Russell, S. 1998b. "An Empirically Plausible Model of Low Real Interest Rates and Unbacked Government Debt." *Journal of Monetary Economics.* forthcoming.

Cassidy, J. 1996. The Decline of Economics. *The New Yorker,* December 2, 50–60.

Christiano, L., and Eichenbaum, M. S. 1991. Liquidity Effects, Monetary Policy, and the Business Cycle. Federal Reserve Bank of Minneapolis and Northwestern University Working Paper.

Christiano, L., and Eichenbaum, M. S. 1992. Liquidity Effects and the Monetary Transmission Mechanism. *American Economic Review* 82:346–53.

Danthine, J.-P., Donaldson, J. B. and Smith, L. 1987. On the Superneutrality of Money in a Stochastic Dynamic Macroeconomic Model. *Journal of Monetary Economics* 20:475–99.

DeGregorio, J. 1992. The Effects of Inflation on Economic Growth. *European Economic Review* 36:417–24.

Dow, J. P., Jr. 1995. The Demand and Liquidity Effects of Monetary Shocks. *Journal of Monetary Economics* 36 (December):91–115.

Espinosa-Vega, M., and Russell, S. 1997. History and Theory of the NAIRU: A Critical Review. Federal Reserve Bank of Atlanta *Economic Review* 82 (Second Quarter): 4–25.

Espinosa-Vega, M. and Russell, S. 1998a. Can Higher Inflation Reduce Real Interest Rates in the Long Run. *Canadian Journal of Economics* 31:92–103.

Espinosa-Vega, M. and Russell, S. 1998b. The Long-Run Real Effects of Monetary Policy: Keynesian Predictions from a Neoclassical Model. Federal Reserve Bank of Atlanta Working Paper 98-6, April.

Espinosa-Vega, M. and Yip, C. 1998. Fiscal and Monetary Interactions in an Endogenous Growth Model with Financial Intermediaries. *International Economic Review* 40(3)595–615, August.

Fisher, I. 1926. *The Purchasing Power of Money: Its Determination and Relation to Credit Interest and Crises.* New York: Macmillan Company.

Friedman, M. 1956. The Quantity Theory of Money—A Restatement. In *Studies in the Quantity Theory of Money,* edited by Milton Friedman. Chicago: University of Chicago Press.

Friedman, M. 1968. The Role of Monetary Policy. *American Economic Review* 68:1–17.

Fuerst, T. J. 1992. Liquidity, Loanable Funds, and Real Activity. *Journal of Monetary Economics* 29:3–24.

Greenspan, A. 1997. Monetary Policy Testimony and Report to the Congress Humphrey-Hawkins Testimony. Available on-line at ⟨http://www.bog.frb.fed.us/boarddocs/HH/9707Test.htm⟩ [August 12, 1998].

Haslag, J. 1998. Monetary Policy, Banking and Growth. *Economic Inquiry* 36:489–500.

Hoffman, D. L., and Raasche, R. H. 1991. Long-Run Income and Interest Elasticities of Money Demand in the United States. *Review of Economics and Statistics* 73:665–74.

Johnson, H. G. 1975. *On Economics and Society.* Chicago: University of Chicago Press.

Keynes, J. M. 1964. *The General Theory of Employment, Interest, and Money.* San Diego: Harcourt Brace and Company: Original edition, 1953.

King, R. and Watson, M. W. 1992. Testing Long Run Neutrality. National Bureau of Economic Research Working Paper No. 4156.

Kydland, F., and Prescott, E. 1982. Time to Build and Aggregate Fluctuations. *Econometrica* 50 (November): 1345–70.

Lucas, R. E. 1972. Expectations and the Neutrality of Money. *Journal of Economic Theory* 4:103–24.

Lucas, R. E. 1981. Tobin and Monetarism: A Review Article. *Journal of Economic Literature* 19:558–67.

Lucas, R. E. 1990. Liquidity and Interest Rates. *Journal of Economic Theory* 50 (April):237–64.

Metzler, L. A. 1951. Wealth, Saving, and the Rate of Interest. *Journal of Political Economy* 59 (April):93–116.

Meyer, L. H. 1997. Monetary Policy Objectives and Strategy. *Business Economics* 32 (January):17–20.

Miller, P. J., and Sargent, T. J. 1984. A Reply to Darby. Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall):21–26.

Nelson, C. R., and Plosser, C. I. 1982. Trends and Random Walks in Macroeconomic Time Series. *Journal of Monetary Economics* 10:139–62.

Phelps, E. 1967. Phillips Curves, Expectations of Inflation, and Optimal Unemployment over Time. *Economica* 34 (August):254–81.

Phillips, A. W. 1958. The Relationship between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. *Economica* 25 (November):283–99.

Sargent, T. J. 1996. Expectations and the Nonneutrality of Lucas. *Journal of Monetary Economics* 37:535–18.

Sargent, T. J., and Smith, B. 1987. Irrelevance of Open Market Operations in Some Economies with

Government Currency Being Dominated in Rate of Return. *American Economic Review* 77 (March): 78–92.

Sargent, T. J., and Wallace, 1976. Rational Expectations and the Theory of Economic Policy. *Journal of Monetary Economics* 2:169–83.

Sargent, T. J., and Wallace, N. 1981. Some Unpleasant Monetarist Arithmetic. Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall):1–17.

Schreft, S. L., and Smith, B. D. 1997. Money, Banking, and Capital Formation. *Journal of Economic Theory* 73:157–82.

Stiglitz, J. E. 1991. Alternative Approaches to Macroeconomics: Methodological Issues and the New Keynesian Economics. NBER Working Paper No. 3580.

Taylor, J. B. 1980. Aggregate Dynamics and Staggered Contracts. *Journal of Political Economy* 38, no. 1:1–23.

Taylor, J. B. 1994. The Inflation/Output Variability Trade-off Revisited. In *Goals, Guidelines, and Constraints Facing Monetary Policymakers.* Conference Series No. 38. Boston: Federal Reserve Bank of Boston.

Tobin, J. 1965. Money and Economic Growth. *Econometrica* 33:671–84.

Tobin, J. 1980. Stabilization Policy Ten Years After. *Brookings Papers on Economic Activity* no. 1:19–71.

Wallace, N. 1984. Some of the Choices for Monetary Policy. Federal Reserve Bank of Minneapolis *Quarterly Review* (Winter):15–24.

Wynne, M. A. 1995. Sticky Prices: What Is the Evidence? Federal Reserve Bank of Dallas *Economic Review* (First Quarter): 1–12.

This Page Intentionally Left Blank

68

Practical Issues in Monetary Policy Targeting

Stephen G. Cecchetti
Professor of Economics, Ohio State University, Columbus, Ohio

INTRODUCTION

What do monetary policymakers need to know and when do they need to know it? Textbook descriptions and academic discussions of policymaking usually ignore the practical problems faced by those who make the decisions and take the actions. While most economists would agree that monetary policy has real short-run effects and is most likely neutral in the long run, they could provide no more than informed speculation in helping decide at what level to set the target for a policy instrument and when to change it.

This chapter's purpose is to outline the type of information monetary policymakers need in practice, and to examine the data to see what we actually know.[1] Any policy rule must be formulated in several clearly defined steps. First, one must identify an operational instrument, best thought of as something policymakers can control precisely, like the federal funds rate or the monetary base. Next, there must be a target. Many central banks have stated that price stability is their goal, but an obvious alternative to targeting the aggregate price level is targeting nominal income.[2] In addition to choosing the target variable itself, formulating policy necessitates specifying a loss function: What is the relative importance of large and small, or positive and negative, deviations of aggregate prices from their target path? One might also assign a cost to large movements in the target variable. For example, it might be important for the Federal Reserve to have a reputation for changing the federal funds rate target smoothly, without large movements or sudden reversals, to avoid creating uncertainty in financial markets.

The next stage in devising a monetary rule is to link the operating instrument with the target. This requires specification and estimation of a macroeconomic model. One needs quantitative answers to questions of the form "If the federal funds rate is moved by one percentage point, what will be the path of the aggregate price level and real output over the following three years?" Not only do we require a point estimate of this response function, but it is also crucial that we know how precise our knowledge is in a statistical sense.

Reprinted from: *Federal Reserve Bank of Cleveland Economic Review,* v. 32, n. 1 (Quarter 1, 1996) 2–15.

Finally, policymakers need a timely estimate of their target variable's future path *in the absence of* any policy actions. In other words, they must know when external shocks hit the economy, how large they are, and what their impact on the time path of aggregate prices and real output will be.

The next section offers a detailed discussion of the modeling issue: How do we formulate and estimate the necessary simple, dynamic, empirical macroeconomic model? The section's first major part looks at econometric identification. What must we assume in order to disentangle the fluctuations in output and prices into their various components? How might we actually estimate the impact of monetary policy on macroeconomic quantities of interest?

The section's second major part discusses the issue of structural stability. Monetary policymakers change their emphasis fairly frequently, focusing on one indicator one year and another the next. How does this affect our attempt to estimate the relationship between the variables that policymakers control (like the federal funds rate) and the things we care about? Can we estimate a stable relationship between output, prices, and interest rates over any reasonable period?

The methodological discussion of modeling issues is followed by section II, in which I present a particular model and examine its properties. Several different types of results are included. First, I look at the impact of different sources of shocks on the variables of interest. Besides allowing answers to questions like "If the federal funds rate were to rise by 100 basis points, how much would output change over the next three years?" this approach makes it possible to examine the sources of fluctuations in output and prices. For example, has monetary policy been responsible for a significant share of output variation over the past decade?

Section III discusses how a policy rule can be formulated. The first step is to specify an objective function: What do policymakers actually want to stabilize? This discussion emphasizes the need for taking account of imprecision when forming a policy rule. We are uncertain how changes in the interest rate affect the size and timing of output and price movements. This means we cannot confidently predict policy actions' impact on target variables, so that policy actions differ from what they would be if we had perfect knowledge. From the theoretical discussion, I move on to examine several possible objective functions of policy and the interest rate paths implied by the combination of each rule and the estimated model. I focus throughout on the importance of employing rules that recognize the imprecision of our knowledge regarding the size of the linkages we need to estimate.

I reach three significant conclusions: First, since prices take time to respond to all types of economic shocks, the objective of price stability implies raising the federal funds rate immediately after a shock, instead of waiting for prices to rise. Second, and more important, comparing the results of price-level targeting with those of nominal-income targeting implies that the difficulties inherent in forecasting and controlling the former provide an argument for concentrating on the latter. Finally, it is possible to use policy rules to see how closely recent movements in the federal funds rate conform to those implied by either price-level or nominal-income targeting rules. The results show that the policy that is optimal in this limited sense involves faster, bigger movements than those shown by the actual federal funds rate path. This suggests that policymakers' actions have been based on something akin to nominal-income targeting, but with costs attached to interest rate movements.

MODELING ISSUES

The single biggest problem in formulating monetary policy rules is how to construct an empirical macroeconomic model that describes the critical aspects of the economy. It is important that the model be dynamic, summarizing the impacts of shocks to the economy—as well as those of intended policy actions—over time. The standard response to this challenge has been to construct various forms of vector autoregressions (VAR). A VAR can answer a question of the following type: "If the federal funds rate moves, when and by how much does the price level change?" Policymakers require quantitative answers to exactly these kinds of questions.

To construct any usable empirical model, a researcher must make a number of choices. I will describe four of these: 1) Which variables should be included in the model? 2) What is the appropriate measure of monetary policy? 3) How can the model be identified econometrically? and 4) Over what sample period should the model be estimated?

Variable Inclusion

When trying to discern the relationship between inflation, output, and monetary policy, should we include other variables in the model? Our answer is guided by the findings of Sims (1992), who estimates a model with prices, output, and an interest rate for several countries. His robust overall conclusion is that with this specification, increases in the interest rate (which should signal policy contractions) lead to prices that are higher than otherwise expected, not lower. This problem, which came to be known as the "price puzzle," can be eliminated by including commodity prices in the model. The reasoning is that the policymaker has additional knowledge about prices' future path that the three-variable model does not adequately summarize. Policy contractions, being based on this omitted information, signal that these other indicators are pointing toward higher prices.

More recent research, like that of Christiano, Eichenbaum, and Evans (1996a, 1996b), has shown that including commodity prices eliminates the puzzle. They suggest that higher commodity prices mean higher future overall prices, and that policymakers respond to this. In other words, an upward move in commodity prices precedes both a rise in the price level and a tightening of policy in the form of an increase in the federal funds rate. The omission of this information from the original Sims formulation led to a bias in which contractionary policy predicts higher aggregate prices. This is not a policy change, but simply a reaction to external events. The models of Christiano, Eichenbaum, and Evans do have the following property: Moving toward a more contractionary monetary policy drives prices down (relative to the trajectory they would follow without the policy change).

Choice of Policy Instrument

Beyond the question of which variables the model should include, it is necessary to specify a monetary policy instrument. Should one assume that policymakers are focusing on the federal funds rate itself (or behaving as if they were), or would it be more realistic to use nonborrowed reserves as the instrument? The literature takes up this issue in some detail.[3] Because events of the past 15 years suggest that the primary focus has been on the federal funds rate, I will assume that it contains the information necessary to gauge policy actions.[4]

Identification

A model builder's most complex decision is formulating a set of "identifying assumptions." This is also the subtlest issue and the one that has generated the most discussion in the literature. It is like the textbook question about estimating supply and demand curves: There, if data on the price and quantity of a good in a market both move, we cannot tell whether the root cause of the change was a shift in supply or a shift in demand. Here, things are a bit less transparent, because there are no clearly defined supply and demand curves in the standard microeconomic sense. Instead, it is necessary to distinguish whether prices, output, and interest rates moved as a result of policy shifts, or because of factors like changes in the price of oil (an aggregate supply shock) or in the demand for money (an aggregate demand shock).

To understand the problem and its solution more fully, we can begin by writing down a dynamic structural model in its moving-average form:

$$P_t = A_{11}(L)\varepsilon_{pt} + A_{12}(L)\varepsilon_{ct} + A_{13}(L)\varepsilon_{yt} + A_{14}(L)u_t \tag{1}$$

$$P_t^c = A_{21}(L)\varepsilon_{pt} + A_{22}(L)\varepsilon_{ct} + A_{23}(L)\varepsilon_{yt} + A_{24}(L)u_t \tag{2}$$

$$y_t = A_{31}(L)\varepsilon_{pt} + A_{32}(L)\varepsilon_{ct} + A_{33}(L)\varepsilon_{yt} + A_{34}(L)u_t \tag{3}$$

$$r_t = A_{41}(L)\varepsilon_{pt} + A_{42}(L)\varepsilon_{ct} + A_{43}(L)\varepsilon_{yt} + A_{44}(L)u_t, \tag{4}$$

where p_t, p_t^c, and y_t are the logs of the aggregate price level, commodity prices, and output, respectively, r_t is the policy indicator, the ε's are exogenous shocks, and u is the policy innovation. Equations (1)–(4) summarize the impact of all the shocks to the economy. The $A_{ij}(L)$'s are lag polynomials in the lag operator L. For example,

$$A_{11}(L)\varepsilon_{pt} = \sum_{i=0}^{\infty} a_{11i}L^i\varepsilon_{pt} = a_{110}\varepsilon_{pt} + a_{111}\varepsilon_{pt-1} + \ldots.$$

Because we do not observe the shocks, it is not possible to estimate the model (1)–(4) directly. Instead, we estimate the more familiar VAR form and place restrictions on the coefficients (the a_{ijk}'s) in order to recover estimates of the shocks.

Identification entails determining the errors in this four-equation system, that is, the actual sources of disturbances that lead to variation in prices, output, and interest rates. As the appendix to this paper describes, when there are four endogenous variables, six restrictions are required for complete identification.

All identification schemes involve assumptions about how these sources of variation are correlated. Researchers use two types of restrictions for this purpose. The first, based on the pioneering work of Sims (1980), is what I will call a "triangular identification," which assumes that a shock does not affect a variable contemporaneously, and so one or more of the a_{ij0}'s are zero. For example, it is commonly assumed that no variable other than policy itself responds to monetary shocks immediately, and so $a_{140} = a_{240} = a_{340} = 0$.

A more formal description of a triangular identification begins by writing the matrix $A(0)$ that is composed of all the coefficients of the $A_{ij}(0)$'s—that is, all the a_{ij0}'s. Triangular identification means assuming that six of these a_{ij0}'s are zero, and so

$$A(0) = \begin{bmatrix} a_{110} & 0 & 0 & 0 \\ a_{210} & a_{220} & 0 & 0 \\ a_{310} & a_{320} & a_{330} & 0 \\ a_{410} & a_{420} & a_{430} & a_{440} \end{bmatrix} \tag{5}$$

In other words, triangular identification means that the monetary policy shock u_t is identified by assuming that no variable other than the federal funds rate responds to it contemporaneously. The output shock, ε_{yt}, is identified by assuming that it is the portion of the error in the output equation that is orthogonal to the policy shock, while the commodity price shock, ε_{ct}, is the portion of the error in the commodity price equation that is orthogonal to these. The final part of the residual in the aggregate price equation that is orthogonal to all three of these is the aggregate price shock, ε_{pt}.

There are many other ways to constrain the four-variable VAR and achieve identification. One, based on the work of Galí (1992), combines two types of restrictions. The first are contemporaneous and resemble those used in the triangular method. The second, following Blanchard and Quah (1989), assume that some shocks have temporary, but not permanent, effects on some variables. For example, we might claim that monetary shocks have no long-run effects on real output, and so the impact of u_t on y_t dies out. Formally, this involves assuming that the a_{34k}'s sum to zero: $\sum_{k=0}^{\infty} a_{34k} = 0$.

Recalling that we need six restrictions, the Gall-style procedure begins with two contemporaneous restrictions based on the logic of data availability and the time people in the economy take to act. The first constraint is that monetary policy does not affect real output contemporaneously (within the month). In the notation used above, the assumption is that $a_{340} = 0$. This seems sensible, since production planning is unlikely to change suddenly after a policy innovation. The second constraint is that the aggregate price level does not enter the money supply rule. This also seems sensible, because the Bureau of Labor Statistics does not publicly release the Consumer Price Index (CPI) until the month following its price survey.

The Galí-style, long-run restrictions, based on Blanchard and Quah (1989), amount to assumptions that neither monetary policy nor aggregate price (other aggregate demand) shocks permanently affect real output or the real commodity price level.

Together, the two contemporaneous and four long-run restrictions allow us to estimate the impact of monetary policy shocks on prices and output.

Structural Stability

Variable inclusion and identification are related. The way in which we name various estimated shocks in a model obviously depends on the quantities being modeled in the first place. While connected to the other choices, the final, more general issue concerns the period over which the empirical model is estimated. The problem is that the reduced-form relationships in the data are unlikely to be stable over any reasonable sample.[5] The problem, known widely as the Lucas (1976) critique, is that policy rule changes alter the relationship among endogenous variables in the economy.

It is easy to see why this might happen. For the sake of discussion, assume that inflation is actually determined by the following structural model:

$$\pi_{t+1} = \alpha r_t + \beta_1 X_{1t} + \beta_2 X_{2t} + \omega_{t+1}, \tag{6}$$

where r_t is policy, ω_{t+1} is a random variable, and X_{1t} and X_{2t} are measures of general economic conditions, like things that influence aggregate supply and money demand.

Next, assume that we can write down a reaction function whereby policymakers automatically change their policy control variable when economic conditions change:

$$r_t = \gamma_1 X_{1t} + \gamma_2 X_{2t} + \nu_t. \tag{7}$$

The policymaker's role is to choose γ_1 and γ_2, the reaction of r_t to X_{1t} and X_{2t}. Since the γ's can be zero, a policy regime need not react to the X's. The term ν_t is a measure of the random component in the policy rule.

Now, consider the reduced-form regression:

$$\pi_{t+1} = \phi_1 X_{1t} + \phi_2 X_{2t} + \xi_t. \tag{8}$$

Since $\phi_i = \alpha\gamma_i + \beta_i$, changes in policy, which are changes in the γ's, will alter the correlation between the X's and π. In effect, the reduced-form inflation regression subsumes the monetary-policy reaction function (7), so that a change in the monetary authorities' policy rule—which may be a change in the relative weight placed on various indicators—will cause changes in (8).

As a practical matter, there are several ways to deal with the instability that may be caused by changes in monetary policy rules. First, one can use institutional information to restrict the data to a period when there were no large changes in policy procedure. Second, one can try to estimate the timing of structural breaks.[6] Alternatively, one can use time-varying parameter models, as Sims (1992) suggests. It is also possible to simply ignore the problem and use all of the available data.

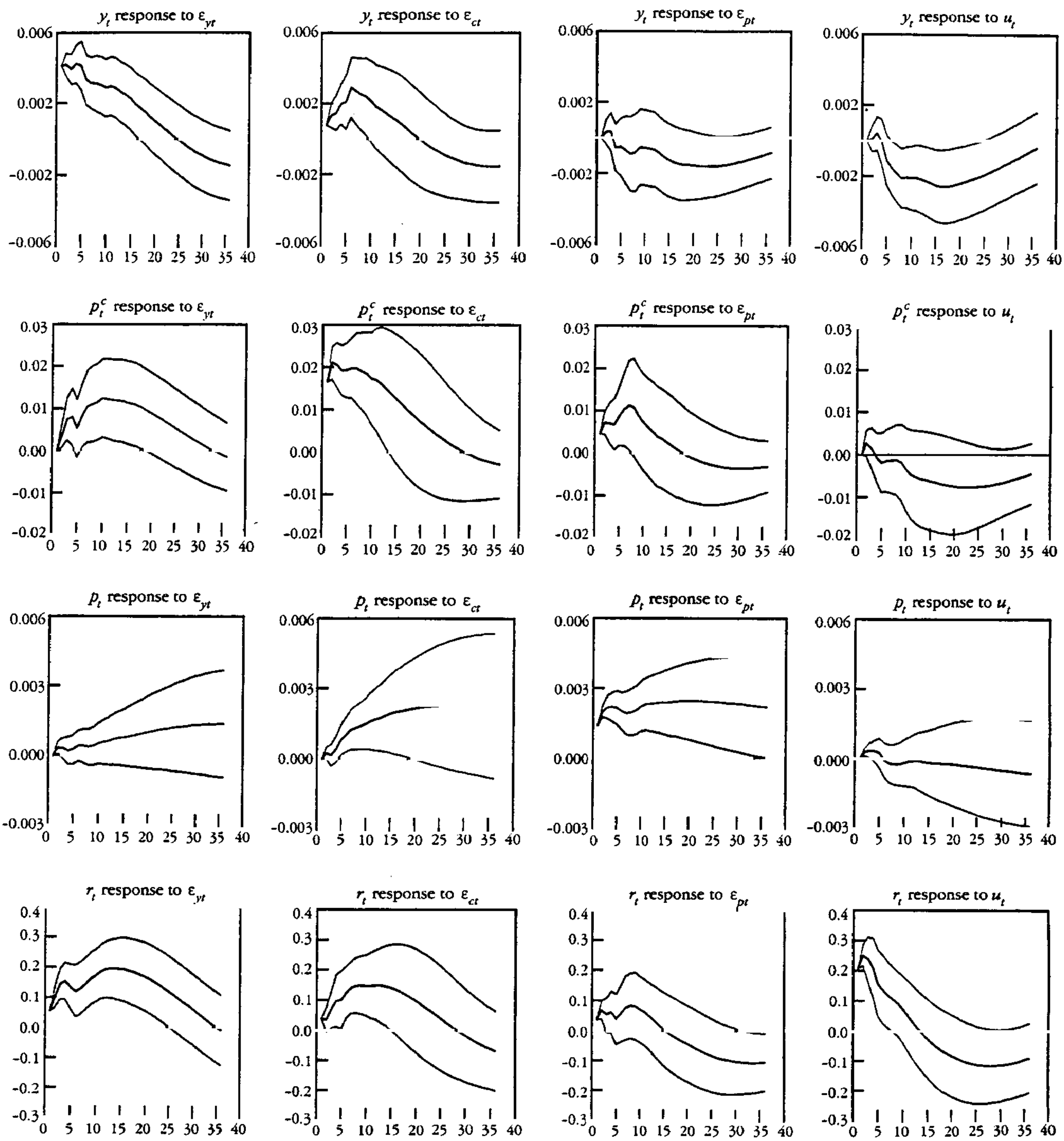
Following my earlier work, I use only the past decade's data, beginning in 1984. Excepting the truncated sample period, I will ignore the problems created by the Lucas critique in all of the calculations that follow. This is an unfortunate necessity if any progress is to be made.

RESULTS FROM ESTIMATING THE MODEL

Impulse Response Functions

Using monthly industrial production data for January 1984–November 1995, the CPI for urban wage earners (CPI-U), the *Journal of Commerce* index of industrial materials prices, and the federal funds rate, along with the triangular identification in equation (5), straightforward procedures yield estimates of the a_{ijk}'s, as well as a covariance matrix for these estimates. These are the time path of the impact of innovations on the model's endogenous variables. They tell us how any one of the four shocks will affect any of the four variables initially—and after several months.

It is easiest to present these results in a series of figures. Figure 1 shows estimates of 16 impulse response functions, plotted with two standard-error bands.[7] These are the response of output, aggregate prices, commodity prices, and the federal funds rate to a unit innovation to each of the four shocks.



a. Estimated response, with two standard-error bands.
NOTE: Horizontal axes are in months; vertical axes are in the change in the log per month.
SOURCE: Author's calculations.

Figure 1 Impulse response functions: triangular identification[a].

The impulse response functions are straightforward and easy to understand. Taking the policy innovation as an example, the last column of Figure 1 shows the result of an unanticipated 100-basis-point change in the federal funds rate for one month on y_t, p_t, p_t^c, and r_t over the next three years. For example, the fourth plot in the third row shows the impact of monetary policy shocks (u_t) on the aggregate price level (p_t). The estimates suggest that a one-time policy tightening—an increase in the federal funds rate—causes prices to rise slightly initially, then to fall below their original level after about six months. Over the next 30 months, the price level continues to fall. The standard-error bands on this figure imply that we are actually *very* unsure of the response. The data indicate a strong possibility that the policy tightening will result in a price-level increase.

Several additional features of Figure 1 are worth noting. First, in all cases, commodity prices (second row) respond more quickly and in the same direction as aggregate prices (third row). Second, for the three ε shocks, the output response seems to be more precisely estimated than the aggregate price response. This second conclusion is consistent with Cochrane's (1994) observation that real output is forecastable with high R^2 at horizons of several years, and with my finding (see Cecchetti [1995]) that inflation is difficult to forecast at horizons longer than a single quarter.

It is very tempting to seek a correspondence between the shocks in this four-variable VAR and those discussed in macroeconomics textbooks. In a simple model, the basic result is that aggregate supply shocks move prices and output in opposite directions, while aggregate demand shocks move them in the same direction. With this categorization, the impulse responses shown in figure 1 suggest that all of the shocks in this model come from the demand side. While this makes intuitive sense for the monetary policy shock, it renders the other classifications unsatisfactory.

One can either accept this at face value or ask whether it might result from the identification used to generate the estimates. Taking the second possibility seriously leads to examination of an alternative identification—the one proposed by Galí being a natural choice. Figure 2 plots the impulse response functions from such a model, estimated using exactly the same data. Because of the technical difficulty associated with their construction, I do not include standard-error bands. Here, the results differ markedly. It now appears that the output shock, ε_{yt}, behaves like an aggregate supply shock, while the three remaining shocks, representing the aggregate price-level shock, the raw material price shock, and the monetary policy shock, lead to reactions consistent with those expected from aggregate demand shocks.

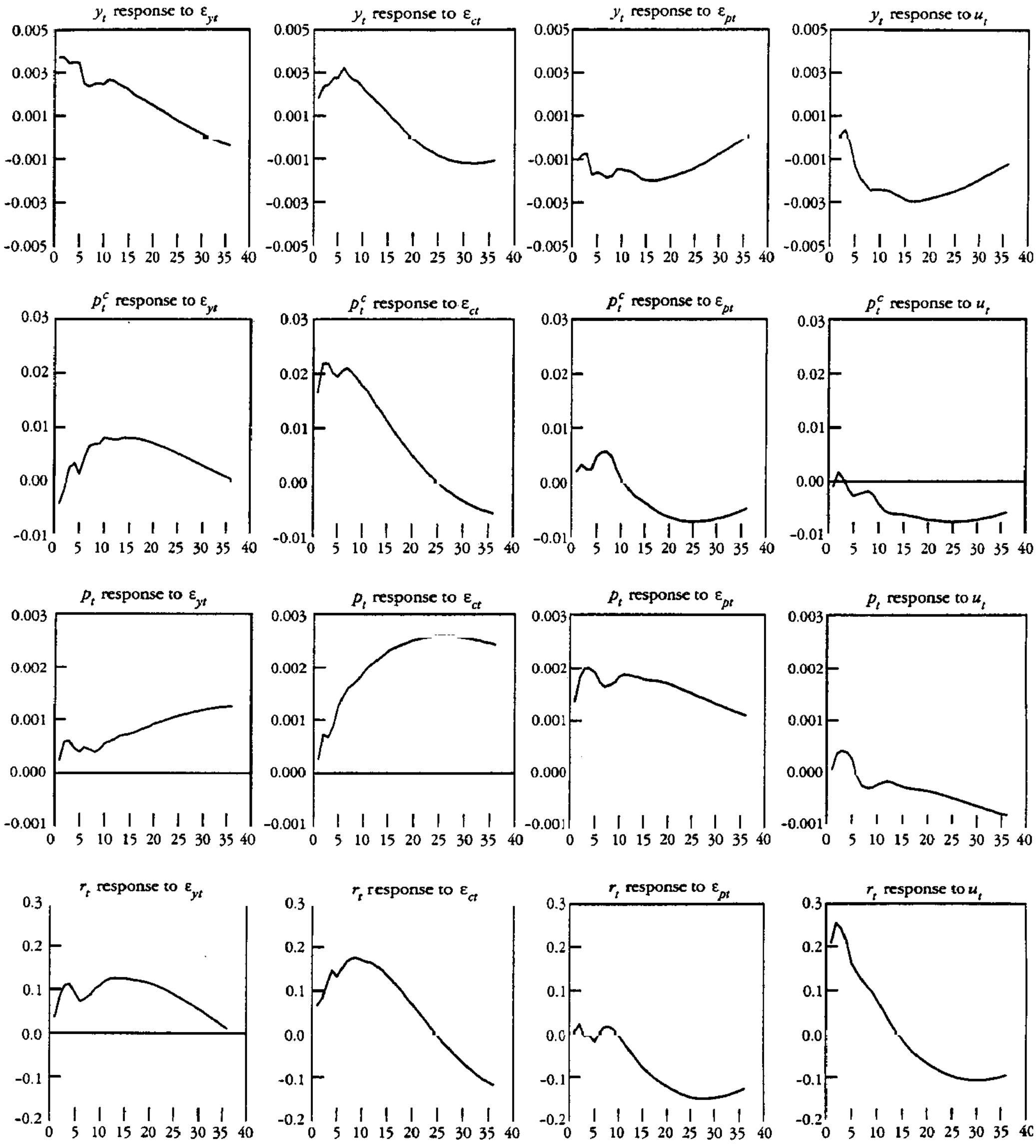
However, we can draw an important positive conclusion by comparing these two sets of identifying restrictions: The impulse response functions of the monetary policy shock are robust to changes in the identification procedure. Policy's impact on output and prices seems fairly robust to the exact methods used in estimation. Since they are easier to compute, I will now proceed using only the estimates obtained with the simpler triangular identification.

Historical Decompositions

While the ultimate goal is to use the estimated dynamic model to construct policy rules, the impulse response functions and structural innovations also allow us to compute the quantities known as "historical forecast decompositions." These allocate output and price movements into the portions accounted for by each of the structural shocks. It is easy to understand how these estimates are constructed from the structural model's equations (1)–(4):

Define the impact of the monetary policy shock on output as $H_{yu}(t)$. From equation (3), this is just

$$H_{yu}(t) = \sum_{i}^{\infty} a_{34i} u_{t-i}, \tag{9}$$



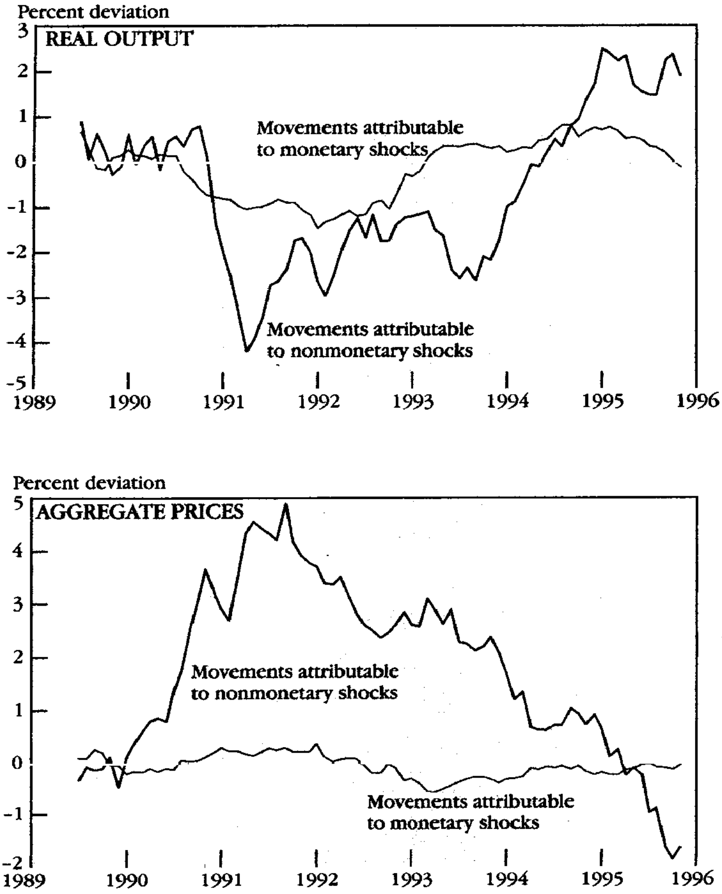
a. Estimated response.
NOTE: Horizontal axes are in months; vertical axes are in the change in the log per month.
SOURCE: Author's calculations.

Figure 2 Impulse response functions: Galí identification[a].

and analogously for the other shocks. Its estimated value, constructed from the parameter estimates, is

$$\hat{H}_{yu}(t) = \sum_{i}^{\infty} \hat{a}_{34i}\hat{u}_{t-i}, \tag{10}$$

where $\hat{u}_t$ is the estimated monetary policy innovation.



SOURCE: Author's calculations.

Figure 3 Forecast error attributable to various innovations.

Figure 3 plots the decomposition of the movements in real output and aggregate prices into the components attributable to monetary and nonmonetary shocks. In constructing these, I have truncated the sum in (9) at 60 months. Because of the difficulty in identifying innovations from nonmonetary sources, it seems prudent to simply sum them together. That is, I plot the fluctuations in y_t and p_t attributable to u_t, $\hat{H}_{yu}(t)$, and $\hat{H}_{pu}(t)$, and the portion not attributable to policy, $[y_t - \hat{H}_{yu}(t)]$ and $[p_t - \hat{H}_{pu}(t)]$.

The results show that, for the past seven years, important movements in both output and prices are largely accounted for by innovations other than those coming from monetary policy. The blue line representing $\hat{H}_{yu}(t)$ and $\hat{H}_{pu}(t)$ in the figure's two panels has much less variation than the green line representing the fluctuations in y_t and p_t that are attributable to nonmonetary policy shocks. This result is particularly striking for prices, where variation seems to be driven by innovations to output, raw materials prices, and the aggregate price level itself. Aggregate supply shocks and nonmonetary aggregate demand shocks account for most of the recent movements in key macroeconomic variables.

FORMULATING A POLICY RULE

Issues

The main use of the empirical model described in the first section of this chapter and estimated in second section is to provide quantitative answers to the questions required for implementing a policy rule. To see how this is done, first note that the model implies estimated values for the aggregate price level and real output:

$$\hat{p}_t = \hat{A}_{11}(L)\hat{\varepsilon}_{pt} + \hat{A}_{12}(L)\hat{\varepsilon}_{ct} + \hat{A}_{13}(L)\hat{\varepsilon}_{yt} + \hat{A}_{14}(L)u_t, \tag{11}$$

$$\hat{y}_t = \hat{A}_{31}(L)\hat{\varepsilon}_{pt} + \hat{A}_{32}(L)\hat{\varepsilon}_{ct} + \hat{A}_{33}(L)\ \hat{\varepsilon}_{yt} + \hat{A}_{34}(L)u_t. \tag{12}$$

A policy rule is a sequence of u_t's that is constructed to meet some objective. In other words, the policymaker is allowed to pick the path of the federal funds rate to meet a particular objective.[8]

The monetary policy literature includes many discussions of the efficacy of various objective functions. Mankiw (1994) includes several papers that deal with this topic explicitly. There are two primary candidates: price-level targets and nominal-income targets. One version of these involves setting the policy instrument—the u_t's in the model—to minimize the average expected mean square error (MSE) of either inflation or nominal-income growth over some future horizon. In the inflation case, the objective function can be written as

$$\min_{\{u_i\}} \frac{1}{b} \sum_{i=1}^{b} E(\hat{p}_i - p_o)^2, \tag{13}$$

where p_o is the log of the base-period price level and h is the policymaker's horizon. The expectation in (13) is over the sampling distribution of $\hat{p}$, which is related to the covariance matrix of the estimated coefficients in equation (11). Nominal-income targeting simply replaces the log price level in (13) with the sum of p_t and y_t.

One important distinction between the objective function (13) and more standard formulations is the treatment of parameter uncertainty. As the results in figure 1 clearly show, we are very unsure about the size and timing of price movements following innovations to the federal funds rate. When constructing a policy rule, it seems prudent to account for this lack of knowledge.

As Brainard (1967) originally pointed out, the presence of uncertainty has important implications. This is easily demonstrated in the present context. Consider a simplified version of the structural price and interest rate equations

$$p_t = \varepsilon_{pt} + \gamma u_t \tag{14}$$

$$r_t = u_t, \tag{15}$$

where γ is a parameter. Next, take the horizon in (13) to be one period ($h = 1$), and the initial log price level to be zero, $p_o = 0$. The policy control problem then reduces to

$$\min_{\{u_i\}} E[\hat{p}_i^2]. \tag{16}$$

Substituting in the expression for p_t, this is simply

$$\min_{\{u_i\}} E[\varepsilon_{pi} + \hat{\gamma} u_i]^2. \tag{17}$$

If we ignore that γ is estimated, then it is trivial to generate the policy rule. It is just

$$u_i^* = -\frac{1}{\hat{\gamma}}\varepsilon_{pi}. \tag{18}$$

Taking account of uncertainty in the estimate of γ, but continuing to assume that ε_{pt} is known, the minimization problem yields

$$u_i^* = -\frac{\hat{\gamma}}{[\hat{\gamma}^2 + \mathrm{Var}\,(\hat{\gamma})]}\,\varepsilon_{pi}. \tag{19}$$

For a given ε_{pt}, this leads to an unambiguously smaller response. In other words, imprecision creates caution, with policy reactions being muted in the face of uncertainty.

Reactions are further attenuated if policymakers attach a cost to the movement in instrument. Taking the same simple setup, imagine the modified objective function

$$\min_{\{u_i\}} E[\hat{p}_i^2 + \alpha \hat{r}_t^2]. \tag{20}$$

This produces the reaction function

$$u_i^* = -\frac{\hat{\gamma}}{[\hat{\gamma}^2 + \mathrm{Var}\,(\hat{\gamma}) + \alpha]}\varepsilon_{pi}, \tag{21}$$

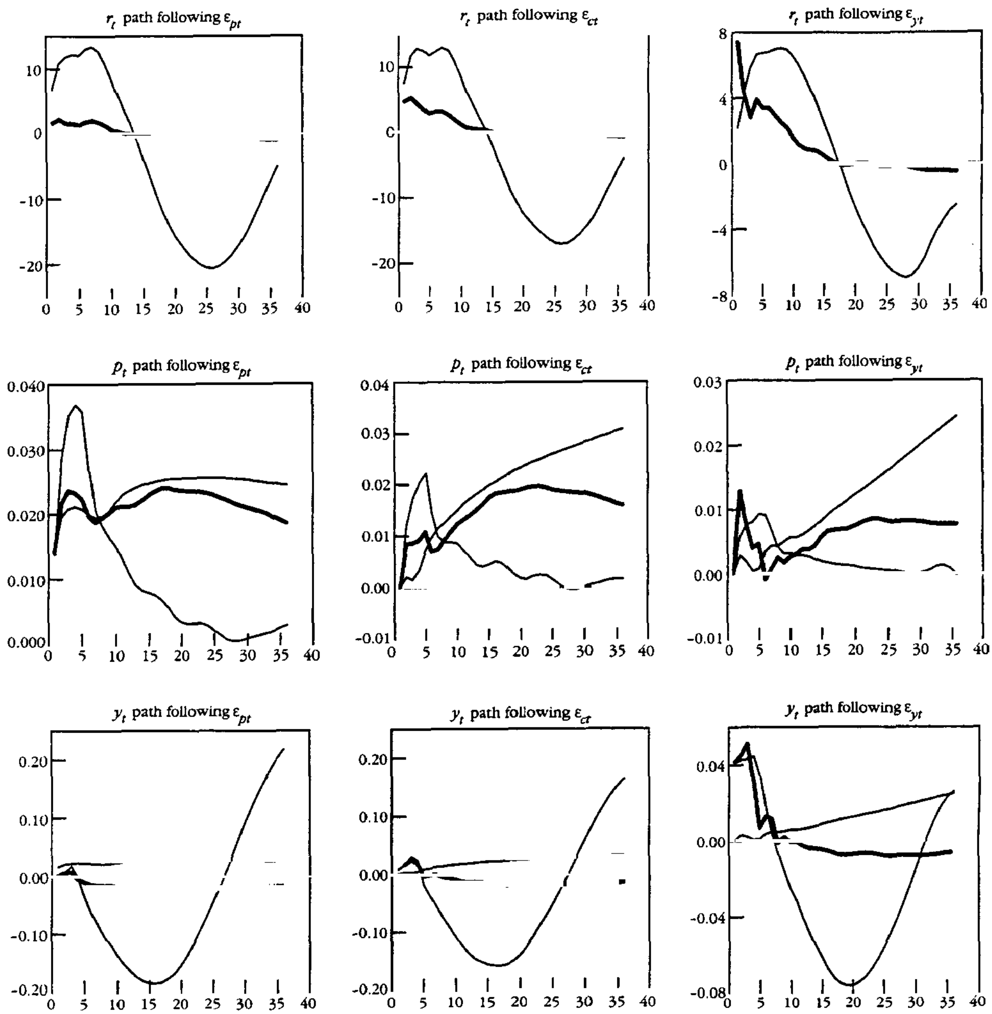
which will yield an even smoother path for the interest rate than does (19).

Results

I examine results based on several policy objectives. It is worth noting that the exercise described here appears to be a gross violation of the Lucas critique. That is to say, contrary to the implications of the discussion in section I, I assume that the reduced-form correlations among output, prices, and interest rates described by equations (11) and (12) are unaffected by the change in the policymaker's reaction function.

There are two ways to defend the procedure. The first is to take the view of Sims (1982)—that parameters in these models evolve slowly enough to make Lucas-critique considerations quantitatively unimportant. The second defense is to reinterpret the exercise as an attempt to recover the objective function that policymakers were implicitly using, by trying to match the actual federal funds rate path with that implied by an optimal rule.

I report results for three different policy rules. The first, which might be termed passive, holds the federal funds rate fixed in the face of the shock. (The model makes it clear that this is not really a passive policy, since it involves shocks to overcome the estimated reaction function.) The other two, which I will call active, minimize the average MSE of either the log of the price level or the log of nominal output over a 36-month horizon ($h = 36$).[9] For each rule, I examine three experiments—one for each structural shock. In each of the nine resulting cases, $\varepsilon_{jo} = 1$ and $\varepsilon_{lk} = 0$ for $l \neq j$ and $k \neq 0$. In other words, there is a unit innovation to one of the structural disturbances in the base period, and that is all. I then construct individual estimates for the optimal response of interest rates to each of the shocks.



—— Min MSE (π) policy —— Fixed interest rate policy ▬▬ Min MSE ($\pi + \dot{y}$)

NOTE: Horizontal axes are in months; vertical axes are in the change in the log per month.

SOURCE: Author's calculations.

Figure 4 Interest rate, output, and price paths following shocks, and the policy response.

Figure 4 reports the implied path of the federal funds rate, aggregate prices, and industrial production for each policy objective in response to each of the three structural shocks. The fixed federal funds rate policy results in consistently higher output and prices than does either of the other two policies. The activist policies both have the same profile, whatever the source of the shock. Output and prices both rise initially, and then fall, with output dropping more than prices.

Interestingly, both of the activist policies involve raising the funds rate immediately and then lowering it slowly. This follows directly from the fact that prices respond slowly to policy innovations (see the third row of Figure 1). The implication is that a policymaker who wishes to stabilize prices must respond to exogenous shocks quickly, in order to ensure that future price movements are minimized. That is the argument for the Federal Reserve's tightening up at the first sign of upward price pressure.

Comparing Targeting Objectives

These calculations have direct implications for the debate between advocates of price-level targeting and those who favor targeting nominal GDP. To see why, I have computed the implied root-mean-square error (RMSE) for inflation and nominal income for each policy. For the price-targeting case, these are the square root of the minimized objective function (13).

Table 1 shows the results. The computations suggest that nominal-income targeting has a certain robustness, since inclusion of real output in the objective function increases the RMSE for inflation only slightly. For the case of an output shock, the increase is from 0.24 to 0.61. However, when the output shock is the source of the instability, the move from price-level targeting to nominal-income targeting decreases the RMSE of nominal income substantially—from 4.12 to 0.69. In other words, the inability to estimate precisely either the impact of shocks on prices or prices' response to policy innovations argues strongly for including real variables in the objective function.

Comparing Actual and Implied Interest Rate Paths

Finally, one might ask how closely recent policy conforms to what would have been implied by either the price-level or nominal-income targeting rules plotted in Figure 5. A simulated interest-rate path can be calculated by taking the estimated structural innovations, the $\hat{\varepsilon}_{jt}$'s, and then computing the optimal policy responses implied by each rule before substituting the result into the equation for the federal funds rate, which is the equivalent of (11).[10]

Table 1 Comparison of Policy Responses

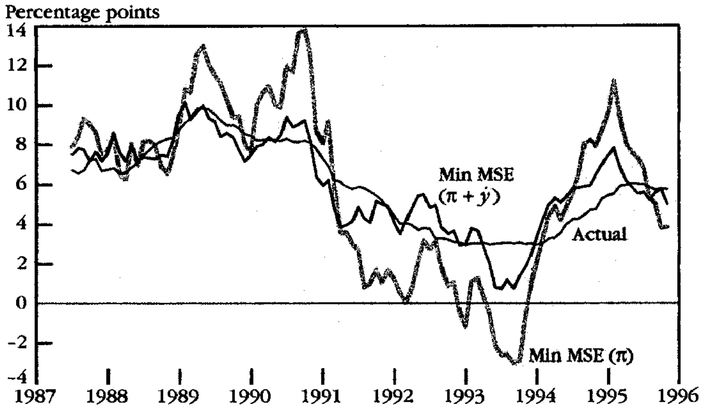
Average RMSE of inflation over a 36-month horizon

Policy rule	Source of shock: Aggregate price	Commodity price	Output
Fixed interest rate	2.35	1.98	1.14
Min MSE ($\pi + y$)	2.15	1.50	0.61
Min MSE (π)	0.99	0.51	0.24

Average RMSE of Nominal Income over a 36-Month Horizon

Policy rule	Source of shock: Aggregate price	Commodity price	Output
Fixed interest rate	1.86	4.89	6.19
Min MSE ($\pi + y$)	0.32	0.35	0.69
Min MSE (π)	0.99	10.85	4.12

Source: Author's calculations.



a. Monthly data, June 1987 to November 1995.
SOURCES: Author's calculations; and Board of Governors of the Federal Reserve System.

Figure 5 Comparison of optimal and actual federal funds rate paths[a].

Figure 5 compares the actual path of the federal funds rate with that implied by the estimated price-level and nominal-income targeting policies. When we examine the figure, several findings emerge. First, targeting the price level alone yields larger swings, as the funds rate reaches both higher and lower extremes. The actual funds rate is the least variable, looking like a smoothed version of the two simulated paths, but the general character of the plot suggests that the optimal policy response simply involves faster, bigger movements than those on the actual path.[11]

Figure 5, however, allows an even more interesting conclusion. From its results, it is possible to infer something about the procedures policymakers were actually following. Such a calculation does not violate the Lucas critique, since it is an attempt to recover the loss function implicit in the policy actions we actually observed.

The estimates imply that the actual funds-rate path was very similar to one that would have been implied by a nominal-income targeting procedure, only smoother. It is as if, over the past decade or so, the federal funds rate had been set to conform to a nominal-income targeting regime, but with policymakers attaching a cost to actually moving the funds rate. That is, the objective function that we can construct from the actual path of interest rates would minimize the sum of squared deviations in nominal income from a target path and squared movements in the federal funds rate, over a horizon of about three years.

SUMMARY

The information requirements for any policy rule are daunting. Not only do policymakers need timely information about current economic conditions, they also need forecasts of the future path of the variables they wish to control (aggregate prices and real output) and quantitative estimates of how their actions will affect these objectives.

This chapter's purpose is to suggest that much of our knowledge is very inexact, and that our inability to precisely forecast the results of policy changes should make us cautious. Even more important, the fact that we have a much better understanding of the impact of our policies on real output than on prices suggests that nominal-income targeting rules are more robust than price targeting rules. From a purely pragmatic viewpoint, someone who cares about nominal income is made substantially worse off by moving to a price-level target, which destabilizes real output considerably. Thus, practical issues make a strong argument for nominal-income targeting.

In addition, we have seen that the actual path of interest rates over the past decade is very similar to that implied by a nominal-income targeting rule, albeit one in which interest rate movements are viewed as costly. By comparing the actual interest-rate path with the path implied by the nominal-income targeting rule, we see that policymakers have smoothed interest rate movements more than the rule would have dictated, but not by much.

APPENDIX: IDENTIFICATION

To understand the more general issues of identification, it is useful to rewrite the four-equation model [(1)–(4)] in a more compact form:

$$x_t = A(L)e_t, \tag{A1}$$

where x_t and e_t are now vectors, and $A(L)$ is a matrix of lag polynomials. We can also write the model in its more familiar VAR reduced form as

$$R(L)x_t = \eta_t, \tag{A2}$$

where $R(0) = I$, the η_t's are i.i.d. (implying that they are orthogonal to the lagged x_t's), and $E(\eta\eta') = \Sigma$. It immediately follows that $A(L)e_t = R(L)^{-1}\eta_t$. This allows us to write $A(0)e_t = \eta_t$, and $A(L) = R(L)^{-1} A(0)$. As a result, given estimates of $A(0)$, $R(L)$, and η, we can recover estimates of both the structural innovations—the e_t's—and the structural parameters—the components of $A(L)$.

The issue of identification is the problem of estimating $A(0)$. To show how this is done, note that $A(0)E(ee')A(0)' = \Sigma$, where $E(ee')$ is diagonal by construction. Normalizing $E(ee') = I$, we obtain the result that $A(0)A(0)' = \Sigma$. In a system with n variables, Σ has $[n(n + 1)/2]$ unique elements, and so complete identification requires an additional $[n(n - 1)/2]$ restrictions. In a four-variable model, six more restrictions are needed. This is a necessary but not sufficient condition for identification. Sufficiency can be established by proving that the restrictions lead to construction of an $A(0)$ matrix that is invertible.

The long-run restrictions of the Galí-style identification can be understood by defining $A(1)$ as the matrix of long-run effects computed by summing the coefficients in $A(L)$. That is, the (i,j) element of $A(1)$ is

$$A_{ij}(1) = \sum_{k=0}^{\infty} a_{ijk}.$$

There are two long-run restrictions. The first is that the impact of ε_{pt} and u_t on y_t is transitory, and so $A_{31}(1) = A_{34}(1) = 0$. The second is that ε_{pt} and u_t have no permanent impact on the relative price of commodities, $(p_{ct} - p_t)$, that is, $A_{11}(1) - A_{21}(1) = A_{14}(1) - A_{24}(1) = 0$.

NOTES

1. This work is based on Cecchetti (1995).
2. I will not discuss the difference between price-level and inflation targeting. While this is a potentially important practical distinction, it is beyond the scope of this paper.
3. See, for example, discussions in Christiano, Eichenbaum, and Evans (1996a, 1996b) and Bernanke and Mihov (1995).
4. Most results are unaffected by the substitution of nonborrowed reserves, suggesting that the funds rate elasticity of reserve demand is relatively stable.
5. For a more detailed discussion, see section 4 of Cecchetti (1995).
6. This is the technique used in Cecchetti (1995).
7. The standard-error bands in the figure are constructed using the simple Taylor-series approximation:

$$F(\hat{\beta}) \approx F(\beta) + \frac{dF(\beta)}{d\beta} \mid_{\beta=\hat{\beta}} (\hat{\beta} - \beta),$$

where F is any differentiable function. The variance of F (β^) follows immediately as

$$E[F(\hat{\beta}) - F(\beta)]^2 \approx \left[\frac{dF(\beta)}{d\beta} \mid_{\beta=\hat{\beta}}\right]^2 Var(\hat{\beta}).$$

Here, we can think of the estimated impulse response functions, the $\hat{A}_{ij}$'s, as functions of the estimated reduced-form VAR coefficients, the elements of $\hat{R}(L)$. Given the estimated variance of these coefficient estimates, the variance of the $\hat{A}_{ij}$'s can be computed by numerical differentiation.
8. Feldstein and Stock (1994) examine an identical experiment, but without parameter uncertainty.
9. Because the model is estimated in logs, the minimum MSE of the nominal-income policy minimized the MSE of the sum of the log of industrial production and the log of the CPI.
10. Performing the calculations in this way ignores a number of elements. In particular, there is no guarantee that the policy rules generated from the artificial experiment of one unit shock in one ε_{jk} at a time will be robust to sequences of shocks in all the ε_{jk}'s simultaneously. One clear reason for this is that it ignores the covariance of estimated coefficients both within and across the elements of the $\hat{A}_{ij}(L)$'s.
11. As one would expect, these large policy innovations result in less stable real output, highlighting that the ultimate issue in policymaking is still the relative weight of prices and output in the objective function.

REFERENCES

Bernanke, B.S., and Mihov, I. Measuring Monetary Policy. Federal Reserve Bank of San Francisco, Working Paper No. 95-09, March 1995.

Blanchard, O.J., and Quah, D. The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review,* vol. 79, no. 4 (September 1989), pp. 655–73.

Brainard, W.C. Uncertainty and the Effectiveness of Policy. *American Economic Review,* vol. 57, no. 2 (May 1967), pp. 411–25.

Cecchetti, S.G. Inflation Indicators and Inflation Policy. In B.S. Bernanke and J.J. Rotemberg, eds., *NBER Macroeconomics Annual 1995.* Cambridge, Mass.: MIT Press, 1995, pp. 189–219.

Christiano, L.J., Eichenbaum, M., and Evans, C.L. The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds. *Review of Economics and Statistics,* vol. 78, no. 1 (February 1996a), pp. 16–34.

Christiano, L.J., Eichenbaum, M., and Evans, C.L. Identification and the Effects of Monetary Policy Shocks. In M.I. Blejer et al., eds., *Financial Factors in Economic Stabilization and Growth.* New York: Cambridge University Press, 1996b, pp. 36–74.

Cochrane, J.H. Shocks. *Carnegie–Rochester Conference Series on Public Policy,* vol. 41 (December 1994), pp. 295–364.

Feldstein, M., and Stock, J.H. The Use of a Monetary Aggregate to Target Nominal GDP. In N.G. Mankiw, ed., *Monetary Policy.* Chicago: University of Chicago Press, 1994, pp. 7–70.

Galí, J. How Well Does the IS-LM Model Fit Postwar U.S. Data? *Quarterly Journal of Economics,* vol. 107, no. 2 (May 1992), pp. 709–38.

Lucas, R.E., Jr. Econometric Policy Evaluation: A Critique. *Journal of Monetary Economics,* vol. 1, no. 2 (Supplementary Series 1976), pp. 19–46.

Mankiw, N.G., ed. *Monetary Policy.* Chicago: University of Chicago Press, 1994.

Sims, C.A. Macroeconomics and Reality. *Econometrica,* vol. 48, no. 1 (January 1980), pp. 1–48.

Sims, C.A. Policy Analysis with Econometric Models. *Brookings Papers on Economic Activity,* no. 1 (1982), pp. 107–52.

Sims, C.A. A Nine Variable Probabilistic Macroeconomic Forecasting Model," Cowles Foundation for Research in Economics at Yale University, Discussion Paper No. 1034, June 1992.

69

NAIRU
Is It Useful for Monetary Policy?

John P. Judd
Vice President and Associate Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

In recent years, a debate has re-emerged about whether the Federal Reserve should pay attention to the "NAIRU" in conducting monetary policy. NAIRU is an acronym for "non-accelerating-inflation rate of unemployment" (a closely related concept is the "natural rate of unemployment"). The NAIRU figures prominently in the Phillips curve, which is a relationship that incorporates a *temporary* trade-off between the unemployment rate and inflation. According to the Phillips curve, an unemployment rate that is below the level identified as the NAIRU (that is, a "tight" labor market) tends to be associated with an increase in inflation; conversely, an unemployment rate that is above the NAIRU tends to be associated with a decrease in inflation. It is well known that the trade-off between inflation and unemployment is only temporary and cannot be systematically exploited by monetary policies aimed at permanently lowering the unemployment rate. In the long run, attempts to do so end up generating higher inflation with no improvement in unemployment. However, the Phillips curve also implies that demand-induced changes in inflation tend to lag behind movements in the unemployment rate, which means that a comparison between the actual unemployment rate and the NAIRU may be helpful in forecasting future changes in inflation.

Tight labor (and product) markets were one reason for the Fed's "preemptive strike" against inflation in 1994 (see Judd and Trehan 1995). The federal funds rate was raised from 3% in early 1994 to 6% in early 1995 without actual increases in broad measures of inflation, like the CPI. This action was explained as a response to indications that inflation would rise in the future without policy action. Over the past year, however, the funds rate has not been raised despite a fall in the unemployment rate to 4¾–5%, below most estimates of the NAIRU. Some people have argued that policy action should be taken to prevent an upward creep in inflation, while others have asserted that there is no inflation threat on the horizon.

These recent experiences have stimulated the current debate about the NAIRU, with some economists arguing that it provides useful information for monetary policy and oth-

Reprinted from: *FRBSF Economic Letter,* No. 97-35 (Nov. 21, 1997).

ers arguing that it is dangerously misleading (*Journal of Economic Perspectives* 1997). This chapter discusses the key elements in this controversy.

FORECASTING

The lags in monetary policy present a problem for central banks, because a policy action taken today may not affect inflation for a year or two. Therefore, in attempting to control inflation, it is dangerous to look only at current rates of inflation. By the time inflation actually begins to rise, inflationary pressures may have been brewing for a year or two, and it may take a substantial tightening of policy (possibly leading to a recession) to head them off. The lag in policy explains why most central banks expend considerable effort in forecasting future economic developments. In fact, some central banks (for example, those in the United Kingdom, Canada, and New Zealand) use publicly announced forecasts as a key element in the formulation of their policies.

According to models of the economy that incorporate a Phillips curve, the unemployment rate plays a role in the transmission process from unanticipated changes in the aggregate demand for goods and services (called "demand shocks") to inflation. In these models, increases in demand raise real GDP relative to its potential level, which increases the demand for labor to produce the additional goods and services, and therefore lowers the unemployment rate relative to the NAIRU. Excess demand in goods and labor markets leads to higher inflation in goods prices and wages with a lag. Because of this, the unemployment rate can help in generating the inflation forecasts that are crucial in formulating monetary policy.

PROBLEMS WITH THE NAIRU

Critics of using the NAIRU concept to guide policy raise both empirical and theoretical arguments. On the empirical side, they point out that the estimated NAIRU for the U.S. has varied in the postwar period. In the 1960s, the NAIRU commonly was estimated at around 5%. By the mid-1970s, it had climbed to around 7%. And by the mid-1990s, it had fallen back to 5½% to 6% (Staiger, Stock, and Watson 1997). A number of factors can affect the NAIRU, including changes in labor force demographics, governmental unemployment programs, and regional economic disturbances.

A related empirical criticism is that the NAIRU cannot be estimated with much precision. Based upon comprehensive empirical analysis of Phillips curves, Staiger, Stock, and Watson conclude that their best fitting equation yields a 95% probability that the NAIRU falls within a range of 4.8 to 6.6%. Given this kind of uncertainty, the NAIRU can provide misleading signals for monetary policy at various times.

A theoretical objection to the use of the NAIRU for monetary policy is that the short-run trade-off between unemployment and inflation may be unstable over time. This trade-off is sensitive to the way in which expectations about inflation are formed, which in turn will depend upon the nature of the monetary policy regime itself. As noted above, for example, any trade-off would tend to disappear if a central bank attempted to exploit it systematically.

A further theoretic objection—one which has been discussed a lot recently—is that the NAIRU makes sense as an indicator of future inflation only when the economy is hit with demand shocks, like those described above for the Phillips curve model (Judd and Trehan 1990 and Chang 1997). However, the economy also may be affected by supply shocks, or unexpected changes in the aggregate supply of goods and services. An example of a supply shock would be a sudden increase in productivity. Initially, this kind of shock would raise the quantity of goods and services produced relative to the quantity demanded, and thus put *downward* pressure on prices. At the same time, the increase in real GDP would raise the demand for labor and reduce the unemployment rate. Thus, a falling unemployment rate would be associated with reduced pressure on prices. If a central bank were using the NAIRU to guide policy in this case, it might mistakenly see the lower unemployment rate as a reason to fear higher inflation in the future, and therefore might tighten policy.

Some observers argue that a supply shock is currently having an effect on the economy. Over the past couple of years, real GDP has increased rapidly, and the unemployment rate has fallen to a low rate of 4¾–5%, while inflation has come down a bit. Therefore, standard Phillips curves have over-forecasted inflation recently, although the errors generally have not been outside the historical range of errors. One explanation offered for recent developments is a surge in productivity due to the introduction of new computer-related technologies. While it is still too soon to know for sure what is driving recent developments, the possibility of a supply shock has to be taken seriously. This possibility illustrates the pitfalls in interpreting the implications of the unemployment rate for future inflation. At the same time, however, it is too soon to be sure that the current low level of the unemployment rate does not presage a rise in inflation in the future.

IS THERE A BETTER WAY?

The arguments presented above have been used to criticize what could be called a "trigger" strategy, in which the central bank would compare the unemployment rate to the latest estimate of the NAIRU and change the funds rate according to whether inflation was predicted to rise or fall in the future. This criticism of such a trigger strategy is well founded. However, it is doubtful that any central bank would base policy on such a simple response to any single variable.

A more relevant question is how forecasting models that incorporate the NAIRU concept perform relative to alternative models. Since all forecasting models are subject to error, the practical issue for central banks is which type of model provides the *best* forecasts. In other words, it is not enough to show that the NAIRU-based models are subject to error. It is also necessary to show that the uncertainties associated with them are bigger than those of alternative models.

The alternative models that have been used for this purpose include monetarist models that rely mainly on a measure of the money supply to forecast inflation, and vector autoregressions (VARs) that produce purely statistical forecasts without relying on any theory concerning what causes inflation. Both of these alternatives have drawbacks.

Since inflation is a monetary phenomenon, monetary models have an obvious theoretical advantage in forecasting inflation. However, the empirical problems with the monetary aggregates over the past 15 to 20 years are well known. In the early 1980s the Fed relied heavily on M1, a narrow aggregate; by the mid-1980s, however, M1's relationship with real GDP and inflation became too uncertain, and the Fed de-emphasized it in favor of the broader aggregates, M2 and M3. These aggregates retained some reliability until the 1990s when they began to experience serious problems. The main difficulty with all of these aggregates appears to have been the deregulation of the financial system in the 1970s and 1980s and the rapid financial innovation that has been going on in the U.S. and world economies for the past two decades or so. These difficulties help explain the results of studies by Stockton and Struckmeyer (1989) and Tallman (1995), which have found forecasting advantages with Phillips curve models compared with monetarist models, although both approaches involved considerable uncertainty.

With regard to VARs, it is well known that they do a good job of forecasting real GDP, but have more problems forecasting inflation (McNees 1986). The reliability of VARs appears to be particularly vulnerable to major changes in inflation regimes, such as the ones in the U.S. in the 1960s and late 1970s (Webb 1995).

CONCLUSIONS

Models that can forecast inflation are valuable to central bankers because monetary policy actions affect inflation with a lag. Models that incorporate a NAIRU concept have problems as forecasting devices, especially if the economy is hit with a supply shock. The current situation may be an example of such a case. Recent forecast errors, though not especially large by historical standards, nonetheless may provide a rationale for some deemphasis of the unemployment rate in policy deliberations. However, it is not clear that monetary models or VARs provide superior alternatives to NAIRU-based models. This consideration may help to explain the continued use of NAIRU-based models by many policymakers, despite well-known conceptual and empirical shortcomings.

As economists continue to work on these problems, advances in modeling may provide better alternatives. For example, Chang suggests that models along the lines of those developed by Bernanke (1986), which explicitly attempt to decompose relevant data into demand and supply shocks, might be useful. An important test of the usefulness of such models for monetary policy would be whether they offer advantages in forecasting inflation.

REFERENCES

Bernanke, B. S. 1986. Alternative Explanations of the Money-Income Correlation. *Carnegie-Rochester Conference Series on Public Policy* pp. 49–99.

Chang, R. 1997. Is Low Unemployment Inflationary? Federal Reserve Bank of Atlanta *Economic Review,* First Quarter, pp. 4–13.

Journal of Economic Perspectives. 1997. Symposium: The Natural Rate of Unemployment (Winter) pp. 3–108.

Judd, J. P., and Trehan, B. 1995. Has the Fed Gotten Tougher on Inflation? *FRBSF Weekly Letter* No. 95-13 (March 31).

Judd, J. P., and Trehan, B. 1990. What Does Unemployment Tell Us about Future Inflation? Federal Reserve Bank of San Francisco *Economic Review* (Summer) pp. 20–37.

McNees, S. K. 1986. The Accuracy of Two Forecasting Techniques: Some Evidence and an Interpretation. *New England Economic Review* (March) pp. 20–31.

Staiger, D., Stock, J. H. and Watson, M. W. 1997. The NAIRU, Unemployment and Monetary Policy. *Journal of Economic Perspectives* (Winter) pp. 33–49.

Stockton, D. J., and Struckmeyer, C. S. 1989. Tests of the Specification and Predictive Accuracy of Nonnested Models of Inflation. *Review of Economics and Statistics* pp. 275–283.

Tallman, E. W. 1995. Inflation and Inflation Forecasting: An Introduction. Federal Reserve Bank of Atlanta *Economic Review* (January/February) pp. 13–27.

Webb, R. H. 1995. Inflation Forecasts from VAR Models. *Journal of Forecasting* pp. 267–285.

70

An Alternative Strategy for Monetary Policy

Brian Motley
Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

John P. Judd
Vice President and Associate Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

In recent years considerable progress has been made in reducing inflation in the United States. During 1992, consumer inflation ran at about 3 percent, the lowest rate since the mid-1960s. Federal Reserve officials have made it clear that reducing inflation further, to near-zero rates, is the long-run goal of U.S. monetary policy. Maintaining stable prices is viewed as the main contribution that the central bank can make to a higher rate of economic growth over the long run.

However, the Fed's low-inflation goal does not appear to be fully credible to the public. The current unusually wide gap between long-term and short-term interest rates suggests that many participants in financial markets are concerned that inflation will pick up again later. This same concern is expressed in well-known surveys, which report expectations of higher inflation over the next ten years—around 5 percent for consumers, and 4 percent for financial decision-makers.

This lack of credibility is important because it means that the cost, in terms of lost output and employment, of further reductions in inflation may be larger than would otherwise be necessary. This chapter explores an approach to monetary policy, involving the targeting of nominal GDP, that could assist the Fed in achieving its low-inflation goal and, in the process, enhance the credibility of that goal.

THE MONETARY AGGREGATES

Why does the public continue to believe that inflation will revert to higher levels in the future? One possibility is that persistently large federal budget deficits may be fueling fears that ultimately pressure will build for the Fed to monetize them. Although the Fed was able to reduce inflation substantially in the 1980s in the face of large deficits, the public may not yet be convinced that the Fed's resolve will hold firm into the distant future.

Reprinted from: *FRBSF Weekly Letter*, No. 93-01 (Jan. 1, 1993).

A second possibility is that conflicting signals coming from the monetary aggregates in recent years have meant that the public has had a hard time monitoring the thrust of monetary policy to be sure that the Fed's commitment to low inflation remains intact. Since the mid-1970s, the Fed has established ranges for, and reported to Congress on, several monetary aggregates. The Fed followed the approach of gradually reducing the growth rate targets for the aggregates over time as a way to pursue gradual disinflation. Until 1983, a narrow aggregate, M1, was emphasized; since then, a broader aggregate, M2, has received most attention.

Unfortunately, the quality of the signals about the thrust of policy emanating from these aggregates has been garbled by financial innovation. The creation of new monetary and non-monetary instruments has altered the relationships between the various aggregates and spending in the economy, and this has made it difficult for the public to judge whether the Fed is maintaining its commitment to low inflation.

Developments in recent years illustrate this problem, since the various aggregates have sent conflicting signals. M1 has grown very rapidly, inducing some observers to worry about a surge in inflation in the future. At the same time, M2 has grown sluggishly, suggesting to others that policy has been too restrictive.

NOMINAL GDP

Given these problems with the aggregates, it seems worthwhile to investigate alternative approaches. One alternative would be for the Fed to set a target for *nominal* GDP, which is a measure of the aggregate demand for goods and services produced in the economy. This variable is appealing as an intermediate target because it has had a fairly consistent long-run relationship with price indices over the years. The only developments that could disrupt that relationship are unpredicted changes in the aggregate *supply* of goods and services, or *potential real GDP*. Although this variable is by no means perfectly predictable, it tends to evolve slowly. In particular, unexpected movements in aggregates supply generally are far smaller than those that have affected the relationship between the monetary aggregates and prices.

The major problem with adopting nominal GDP as an intermediate target is the difficulty in controlling it in the short to medium term. In part, this is because the relatively long lags from monetary policy actions to changes in nominal GDP make it difficult to offset the effects of unexpected shocks. As a consequence, the Fed could not be held accountable for controlling nominal GDP even over periods as long as a year. Hence it would be difficult for the public to monitor the Fed's actions by observing nominal GDP and comparing it to the target.

A solution to this problem can be found in a strategy of monetary policy originally proposed by Professor Bennett McCallum of Carnegie-Mellon University. First, the central bank would establish and announce a target path for nominal GDP that would be consistent with the economy operating at its long-run potential and with inflation equal to the targeted rate. Second, the central bank would define a rule that would specify what monetary policy actions it would take when nominal GDP misses its target. These monetary policy actions, in turn, would be defined by changes in a policy instrument, such as a short-term interest rate or a measure of bank reserves. It is important that the instrument be subject to the close control of the central bank in the short run, so that it can be held accountable for following the rule.

An example of such a rule would be that each quarter the central bank would change the federal funds rate by a specified proportion of the nominal GDP target "miss" in the prior quarter. If nominal GDP were above the announced target, whether because of faster-than-expected inflation or real GDP growth, the funds rate would be raised. Conversely, if nominal GDP were to fall below target, policy would be eased.

The rule would be set up such that if it were followed consistently over time, there would be a high probability of achieving the nominal GDP target (as well as the inflation goal) over the long term, even though there might be significant deviations in the short run. Thus for the public to be confident that the long-run inflation target would be achieved, the central bank would need only to achieve the instrument settings defined by the rule.

HISTORICAL ANALYSIS

We have examined how the U.S. economy might have performed over the past 30 years if such a rule had been in place (Judd and Motley 1992). We carried out counterfactual simulations under various sets of assumptions regarding the structure of the economy and how monetary policy was set. We first assumed that the instrument of monetary policy is the monetary base, which consists of the stocks of currency held outside the banking system and of bank reserves. Our empirical analysis suggested that the Fed could have maintained stable prices over the 30-year period if it had followed such a rule-based strategy. Even if the economy had been hit by random shocks that were as variable as those that actually occurred, there was a high probability that prices would have ended the period at about the same level as at the beginning.

However, except for a short period in the early 1980s, the Fed generally has preferred to conduct policy by manipulating a short-term interest rate. In part, this preference reflects concern that if it controlled the stock of bank reserves or the monetary base closely, interest rates would become excessively volatile. Moreover, a high proportion of U.S. currency is held abroad today, which casts doubt on a strategy that uses the base as an instrument to control aggregate demand in the U.S. economy.

Hence, we also analyzed the effects of a regime in which the policy instrument is a short-term interest rate, such as the federal funds rate. We found that a rule-based strategy using an interest rate instrument was likely to cause extreme fluctuations in the economy if the *level* of nominal GDP were targeted. The problem appears to be that the lags from changes in the interest rate to the level of nominal GDP are so long that policy was not able to prevent the economy from substantially overshooting its target in response to shocks. Such overshooting tends to set off cycles of ever increasing amplitude.

However, our results suggest that this volatility could be avoided if the interest rate were changed only in response to the *growth rate* of nominal GDP. Since the growth rate of nominal GDP could be returned to target more quickly than its level, the problems of instability are avoided. We found that a rule defined in terms of the growth rate of nominal GDP could, with a high degree of probability, keep inflation low without causing excessive fluctuations in output. In fact, the variance of real GDP would be about the same as under the policies actually followed over the last 30 years, while the volatility of interest rates would be reduced somewhat.

However, a rule that aimed at a target for the *growth rate,* rather than the *level* of nominal GDP, would automatically accommodate shocks to the level of nominal GDP, and thus would allow for the possibility that the price level might drift over time. In other words, it would not guarantee precise control of the average rate of inflation over long periods of time. Such drift would occur only if there were a prolonged series of positive or negative shocks. However, on the rare occasions when this occurs, it could be offset by one-time adjustments to the level of the targeted nominal GDP growth path. Thus this problem would not be insurmountable.

RULES AND DISCRETION

Given that it is impossible to foresee all future circumstances, central banks are understandably wary of committing themselves to rigid adherence to any rule, no matter how sensible that rule may seem to be. However, a rule such as the one under discussion may make a contribution to policy even if it were used only to modify the Fed's traditional discretionary approach. At present, the Fed evaluates alternative policy actions relative to a status quo policy of holding nominal interest rates unchanged. Naturally, the discussion tends to focus on whether the funds rate should be raised or lowered *from its recent level.* This may be misleading, because leaving the funds rate unchanged does not necessarily imply that the effect of policy on the economy also will remain unchanged. For example, if the economy is operating above its long-run capacity and nominal interest rates are held constant, a surge in inflation will have the effect of lowering real interest rates and so stimulating the economy even more.

If the Fed were to use a rule to specify its baseline policy, this problem could be mitigated. The rule would provide information as to the direction the interest rate should be moved to keep the economy on a stable growth path with low inflation. Although the Fed could, at its discretion, choose to modify the policy indicated by the rule, a debate that focused on whether policy should ease or tighten *relative to what the rule suggested* might be more useful than one that focused only on whether the interest rate should be changed from its recent level. In particular, putting the debate into such a framework could help policymakers make short-run discretionary decisions without losing sight of the long-run goal of controlling inflation.

So long as the instrument settings indicated by the rule were achieved on average over time, the rule would work to achieve the long-run inflation objective. Moreover, once the public began to see that the rule was being followed, the Fed's credibility would be enhanced. This would make it easier to maintain stable prices without causing undue disruption to real economic activity.

REFERENCE

Judd, J. P., and Motley, B. 1992. Controlling Inflation with an Interest Rate Instrument. Federal Reserve Bank of San Francisco Economic Review 3, pp. 3–22.

71
What Is the Optimal Rate of Inflation?

Timothy Cogley
Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

Central banks are now placing greater emphasis on maintaining low inflation, and this raises the question: How low should inflation be? Some say that the current level of inflation is acceptable, while others argue that inflation should be pushed toward zero. There are a number of ways to think about this problem, and one approach focuses on conditions for the optimal quantity of money. This chapter summarizes the lessons which this literature contains.

THE FRIEDMAN RULE

Milton Friedman (1969) provided a simple rule for determining the optimal rate of inflation. He started with the observation that money provides valuable services, as it makes it easier and more convenient for consumers to execute transactions. For example, while many establishments accept both money and credit cards as payment for goods and services, others accept only cash. A consumer could probably get by with credit cards alone, but this would entail spending more time seeking out establishments that accept them. Having some money in one's pocket saves the time and inconvenience of doing so.

While money is useful for carrying out transactions, it is also costly to hold. Monetary instruments generally earn less interest than securities such as Treasury bills. In fact, some forms of money, such as currency, earn no interest at all. The decision to hold more money means investing less in securities that pay more interest, and the cost of holding money depends on how much interest income is forgone. In order to decide how much money to hold, consumers must trade off the benefits of ease and convenience in carrying out transactions against the cost in terms of forgone interest earnings. In the end, people strike a balance between the two factors, holding more money when the cost is low and less when it is high. But as long as monetary instruments pay less interest than other securities, money will be costly to hold and consumers will have an incentive to economize on its use.

But economizing on money is somewhat wasteful from society's point of view. For while money is costly to hold, it is essentially costless for central banks to produce. Thus, by increasing the quantity of real balances (i.e., the nominal quantity of money divided by

---

Reprinted from: *FRBSF Economic Letter,* No. 97-27 (Sept. 19, 1997).

the price level), the central bank could make everyone better off at no cost to itself. That is, consumers would benefit from additional real balances because they would make it more convenient to carry out transactions, and it would cost the central bank nothing to provide this service. Robert Lucas (1987) writes that this is "one of the few legitimate 'free lunches' economics has discovered in 200 years of trying." Pursuing this idea to its logical conclusion, Friedman argued that the optimal policy involves eliminating incentives to economize on the use of money. To do so, the central bank should seek to eliminate the difference between interest rates on monetary instruments and on other securities, because then money would be costless to hold.

At the time Friedman wrote, money paid no interest, and the optimal policy called for setting nominal interest rates on bonds equal to zero. To a first approximation, the nominal interest rate equals the real interest rate plus expected inflation. To set the nominal interest rate to zero, it follows that the inflation rate must equal minus the real interest rate. If the latter were around 2 to 3%, Friedman's arguments suggest that the central bank should seek to deflate at a rate of 2 to 3%. This would involve reducing the nominal quantity of money, but this would fall at a slower rate than the price level, and the quantity of real balances would increase.

An alternative way to eliminate the interest differential between money and bonds would be to pay interest on money. While monetary instruments such as checkable deposits now earn interest, narrow measures such as currency do not, and it is difficult to imagine a low-cost way to begin paying interest on cash. But since currency is still essential for some purchases, it would still be optimal for the central bank to make it costless to hold. And since no interest is paid on currency, the Friedman rule carries through.

SEIGNORAGE AND OTHER TAXES

Edmund Phelps (1973) criticized the Friedman rule on the grounds that it ignores considerations related to taxation. Phelps pointed out that inflation is a source of tax revenue for the government and that if inflation were reduced other taxes would have to be increased in order to replace the lost revenue. He also argued that some inflation would be desirable if distortions associated with inflation taxes were less costly than distortions associated with other taxes to which the government might resort.

Phelps's argument raises a number of important issues. First, in what sense is inflation a source of tax revenue? The government can finance its purchases in three ways: by levying direct taxes, by borrowing from the public, and by printing money. When the government pays for goods and services with newly printed money, it increases the money supply, which raises the price level, and the increase in the price level diminishes the real value of pre-existing money held by the public. Thus, inflation works like a tax on people who hold money: the newly printed money enables the government to buy goods and services at the expense of everyone who held pre-existing money balances, which fall in real value as a consequence of the increase in the price level. Revenue raised in this manner is called "seignorage."

If a central bank were to follow the Friedman rule, it would reduce inflation and force the government to rely on other taxes to pay its bills. Phelps argued that those taxes introduce distortions of their own, which may outweigh the benefits of reduced inflation. What is the nature of these distortions?

To a first approximation, taxes on income, consumption, capital gains, and so on are levied on a proportional basis. Thus, for example, workers who earn higher levels of income pay higher income taxes. This distorts private economic decisions because it creates incentives to alter behavior in order to avoid the tax. For example, a tax on labor income raises the pre-tax wage that firms have to pay but lowers the after-tax wage that workers receive. An increase in the pre-tax wage reduces the quantity of labor that firms want to hire, and a fall in the after-tax wage reduces the quantity of labor that workers are willing to supply. The tax makes firms worse off because they pay higher wages but employ fewer workers and produce less output. The tax also makes workers worse off because they earn lower wages and work fewer hours. The government collects revenue from the tax and uses it to provide public goods and services, but the losses of those who pay the tax generally exceed the revenue collected. The difference between the losses of those who pay the tax and the revenue raised is known as the "deadweight loss" or "excess burden" of the tax, and one principle of public finance is that taxes should be administered in a way that minimizes these losses.

The Friedman rule would certainly not be optimal if seignorage revenue were replaced by other tax increases that were even more distortionary. On the other hand, the fact that governments must choose among distortionary taxes does not necessarily invalidate the Friedman rule. The optimal mix depends on how distortionary the various taxes are. Phelps argued that at low rates of inflation, distortions associated with the inflation tax might be minor and that substituting other taxes for the inflation tax might result in greater dead-weight losses. But contrary to Phelp's argument, there are reasons to believe that the inflation tax is highly distortionary.

One is related to the fact that money is an intermediate good rather than a final good. Intermediate goods are those goods which are used in the production of other goods and services, and taxes on these commodities are regarded as undesirable because they introduce two sets of distortions. On the one hand, they reduce production efficiency and increase the cost of producing final goods. And on the other, since this increase in cost is reflected in final goods prices, they distort final goods markets as well. Alternatively, the government could raise the same amount of revenue by taxing final goods directly, and while this would still distort final goods markets, it would not distort production efficiency.

V. V. Chari, Lawrence Christiano, and Patrick Kehoe (1996) apply this principle to the inflation tax. They point out that money is intrinsically useless and should not be classified as a final good. Instead, since money is useful for facilitating transactions, it should be regarded as an intermediate good. This line of reasoning suggests that the inflation tax is really an indirect tax on other goods and that taxing those goods directly might be more efficient. In fact, Chari, et al., construct a number of theoretical examples in which this is so. In their examples, the Friedman rule remains optimal even though seignorage revenue must be replaced by distortionary taxes on labor income or consumption. That is, although the replacement taxes are distortionary, their calculations suggest that the inflation tax is even more distortionary.

Furthermore, even when the Friedman rule isn't strictly optimal, there are two reasons why it is likely to be nearly so. First, Robert Lucas (1994) estimates that even low rates of inflation involve substantial welfare costs. Second, seignorage accounts for only a small fraction of tax revenue raised in the United States, usually less than 3%, and the adjustments in other taxes needed to replace lost seignorage revenue are likely to be minor. Taken together, these two facts suggest that the optimal inflation tax is not likely to be far from the Friedman rule. For example, Casey Mulligan and Xavier Sala-i-Martin (1997) estimate that taking replacement taxes into account raises the optimal inflation rate by perhaps as much as 1%. R. Anton Braun (1994) estimates that tax considerations raise the optimal inflation rate by between 1 and 6%. Taking the average of Braun's estimates and subtracting a real interest rate of, say, 2.5% yields an optimal inflation rate of 1%. Thus, one can rationalize a goal of (approximate) price stability by appealing to optimal tax arguments.

OTHER CONSIDERATIONS

Other factors might further increase the optimal rate of inflation. For example, the underground economy uses cash intensively and is difficult to reach with other tax instruments. Some additional inflation might be desirable as a tax on activity in this sector. Similarly, much of the stock of U.S. currency is held abroad, and it might be desirable to collect seignorage from foreigners. Quantitative information on the amount of money held abroad or used in the underground economy is sparse, however, and it is difficult to say how much these factors might raise the optimal inflation rate.

REFERENCES

Braun, R. A. 1994. How Large Is the Optimal Inflation Tax? *Journal of Monetary Economics* 34, pp. 201–214.

Chari, V. V., Christiano, L. J. and Kehoe, P. J. 1996. Optimality of the Friedman Rule in Economies with Distorting Taxes. *Journal of Monetary Economics* 37, pp. 203–223.

Friedman, M. 1969. *The Optimal Quantity of Money and Other Essays.* Chicago: Aldine.

Lucas, R. E., Jr. 1987. *Models of Business Cycles.* Oxford: Basil Blackwell.

Lucas, R. E. 1994. On the Welfare Cost of Inflation. CEPR Working Paper No. 394.

Mulligan, C. B., and Sala-i-Martin, X. X. 1997. The Optimum Quantity of Money: Theory and Evidence. NBER Working Paper No. 5954.

Phelps, E. S. 1973. Inflation in the Theory of Public Finance. *Scandinavian Journal of Economics* 75, pp. 67–82.

72

Monetary Policy and the Well-Being of the Poor

Christina D. Romer and David H. Romer
Professor of Economics, University of California, Berkeley, California

Poverty is arguably the most pressing economic problem of our time. And because rising inequality, for a given level of income, leads to greater poverty, the distribution of income is also a central concern. At the same time, monetary policy is one of the modern age's most potent tools for managing the economy. Given the importance of poverty and the influence of monetary policy, it is natural to ask if monetary policy can be used as a tool to help the poor.

It is this possibility that we pursue in this chapter. We examine the influence of monetary policy on poverty and inequality both over the business cycle in the United States and over the longer run in a large sample of countries. Our analysis suggests that there are indeed important links between monetary policy and the well-being of the poor in both the short run and the long run, but that the short-run and long-run relationships go in opposite directions. Expansionary monetary policy aimed at rapid output growth is associated with improved conditions for the poor in the short run, but prudent monetary policy aimed at low inflation and steady output growth is associated with enhanced well-being of the poor in the long run.

The existing literature on monetary policy and the poor focuses almost exclusively on the short run. Monetary policy can affect output, unemployment, and inflation in the short run. As a result, if poverty and inequality respond to these variables, monetary policy can affect the well-being of the poor. Furthermore, because unanticipated inflation can redistribute wealth from creditors to debtors, monetary policy can also affect distribution through this channel.

In the first section of the chapter, we provide some up-to-date estimates of the cyclical behavior of poverty and inequality. We confirm the common finding that poverty falls when unemployment falls. In contrast to earlier authors, however, we find no evidence of important effects of cyclical movements in unemployment on the distribution of income. We find some evidence that unanticipated inflation narrows the income distribution, though we can detect no noticeable impact on poverty. Finally, using the Federal Reserve's

Reprinted from *Federal Reserve Bank of Kansas City Economic Review,* v. 84, n. 1 (First Quarter 1999) 21–49.

Survey of Consumer Finances, we find that the potential redistributive effects of unanticipated inflation on the poor through capital gains and losses are very small.

Because of the short-run cyclicality of poverty, some authors have concluded that compassionate monetary policy is loose or expansionary policy. We, however, argue that this view misses the crucial fact that the cyclical effects of monetary policy on unemployment are inherently temporary. Monetary policy can generate a temporary boom, and hence a temporary reduction in poverty. But, as unemployment returns to the natural rate, poverty rises again. Furthermore, the expansionary policy generates inflation. If a monetary contraction is used to reduce inflation, the adverse effects on poverty offset even the temporary reduction in poverty during the earlier boom.

In the long run, monetary policy most directly affects average inflation and the variability of aggregate demand. Therefore, the important question from the perspective of monetary policy makers concerned with the condition of the poor is whether there is a link between these variables and poverty and inequality. We investigate such long-run relationships in the second section of the paper.

We use data for a large sample of countries from the 1970s and 1980s to see if there is a systematic relationship between poverty and the variables directly affected by monetary policy in the long run. We find that there are indeed important negative relationships between the income of the poor and both average inflation and macroeconomic instability. These relationships are quantitatively large and robust to permutations in samples and control variables.

Looking at the components behind the reduced-form correlations provides insight into the source of these relationships. Our own estimates and those in the literature suggest that high inflation and macroeconomic instability are correlated both with less rapid growth of average income and with lower equality. We also find that it is primarily the long-run link between monetary policy and the behavior of average income that is driving the negative correlations of both inflation and variability with poverty.

Researchers and policymakers should obviously interpret correlations such as the ones we report with caution. They could, for example, result from some third factor, such as education or government effectiveness, that affects both poverty and monetary policy. Nevertheless, they are certainly consistent with the notion that controlling inflation and output variability through sound monetary policy is likely to result in higher income for those at the bottom of the distribution in the long run. For this reason, we conclude that compassionate monetary policy is, most likely, simply sound monetary policy. Monetary policy that aims to restrain inflation and minimize output fluctuations is the most likely to permanently improve conditions for the poor.

THE EFFECTS OF MONETARY POLICY ON THE POOR IN THE SHORT RUN

The Channels through which Monetary Policy Affects the Poor

Expansionary monetary policy raises both output and inflation in the short run. These short-run effects of monetary policy can influence the well-being of the poor through three channels.

First, and most important, the rise in average income in a cyclical expansion directly reduces poverty. For a given distribution of income around its mean, an increase in the mean reduces the number of people below a fixed cutoff. That is, a rise in all incomes together increases the incomes of the poor, and raises some of their incomes above the poverty level. Since expansionary monetary policy raises average income in the short run, this is a powerful mechanism through which monetary policy can immediately benefit the poor.

Second, there may be cyclical changes in the distribution of income. The declines in unemployment and increases in labor force participation and in real wages in an expansion are likely to be concentrated disproportionately among low-skilled workers. Thus the income distribution may narrow. In this case, there are short-run benefits of expansionary policy to the poor beyond its effect on average income. On the other hand, transfers are less cyclical than labor income, and the poor receive a larger fraction of their income from transfers than do the remainder of the population. If this effect predominates, the income distribution could widen in a boom. In this case, the benefits of expansionary policy to the poor are smaller than what one would expect given the impact on mean income.

Third, the inflation created by expansionary monetary policy has distributional effects. Inflation can harm the poor by reducing the real value of wages and transfers. For example, the fact that real welfare benefits fell in the 1970s may have been partly due to inflation. The pension income of the poor, on the other hand, is insulated from inflation: well over 90 percent of the pension income of the elderly poor comes from Social Security, which is indexed (U.S. House of Representatives, 1996, Table A-10). Finally, unanticipated inflation benefits nominal debtors at the expense of nominal creditors. If the poor are net nominal debtors, inflation can help them through this channel.

With these general considerations in mind, we turn to the empirical evidence to examine the impact of cyclical fluctuations and inflation on poverty. We also examine these variables' impact on the distribution of income. Our approach follows such authors as Blinder and Esaki (1978), Blank and Blinder (1986), Cutler and Katz (1991), Blank (1993), and Blank and Card (1993). We differ from these authors in focusing on the absolute rather than the relative well-being of the poor, in emphasizing the distinction between unanticipated and anticipated inflation, and in considering more recent data.

Because the income measures that we examine do not include capital gains and losses, these data may miss some of the short-run effects of monetary policy on the poor. Therefore, after examining the impact of unemployment and inflation on poverty and income distribution, we examine the financial balance sheets of the poor to see if unanticipated inflation is likely to have any substantial effect on them through this channel.

Poverty and the Macroeconomy

We examine the relationship of poverty with unemployment and inflation in the postwar United States. Because data on poverty and income distribution are only available annually, we use annual data throughout. Our basic sample period is 1969–94; this is the longest period for which all of the series we use are available. Our dependent variable is the poverty rate—that is, the fraction of the population living in households with incomes below the poverty level. We use the unemployment rate for men aged 20 and over as our cyclical indicator; for simplicity, we refer to this as "unemployment" in what follows. Our measure of inflation in year t is the change in the logarithm of the GNP deflator from the fourth quarter of year t-1 to the fourth quarter of year t. To separate inflation into its anticipated and unanticipated components, we use the inflation forecasts from the Survey of Professional Forecasters (formerly the ASA/NBER survey). Specifically, our measure of expected inflation in year t is the median forecast in November of year t-1 of inflation over the next four quarters.

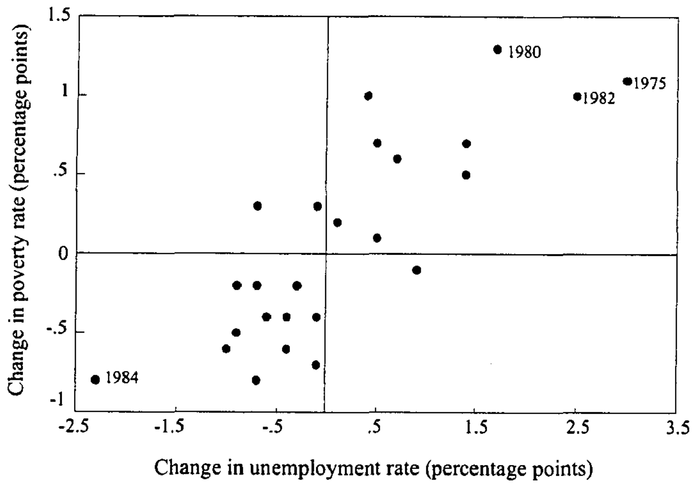


Figure 1 Poverty and unemployment.

Figures 1 through 3 show the basic relationships. Figure 1 is a scatter plot of the change in the poverty rate against the change in unemployment. There is a strong positive relationship. That is, increases in unemployment are associated with increases in poverty. Figures 2 and 3 are scatter plots of the change in poverty against the unanticipated change and the anticipated change in inflation, respectively. Figure 2 shows no clear relationship between changes in poverty and unanticipated inflation. Figure 3, on the other hand, shows a moderate tendency for poverty to fall when there are anticipated increases in inflation.

The corresponding regressions are reported in the first three columns of Table 1. The regression of the change in the poverty rate on the change in unemployment yields a t-statistic of almost seven. The point estimate implies that a rise in unemployment of one percentage point is associated with a rise in the poverty rate of 0.4 percentage points. The regression of the change in poverty on the unanticipated change in inflation produces a

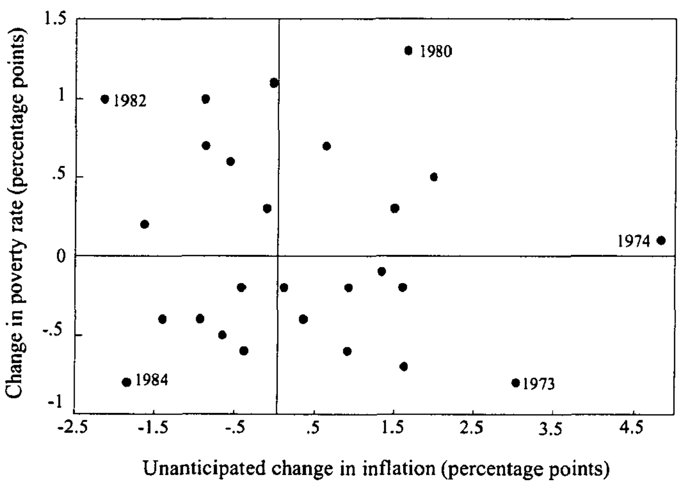


Figure 2 Poverty and unanticipated inflation.

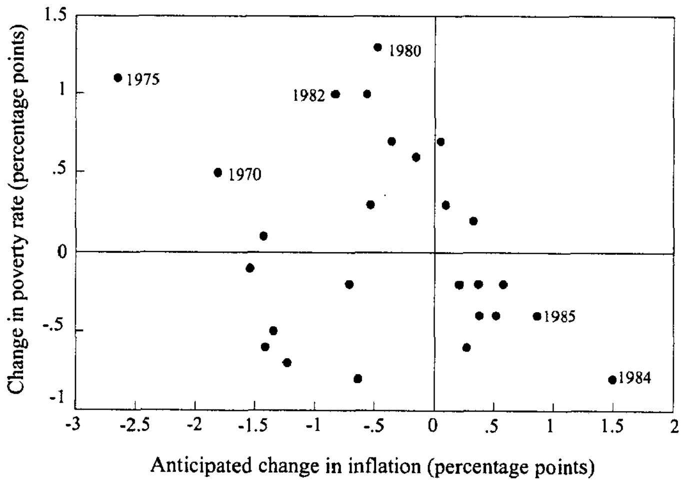


Figure 3 Poverty and anticipated inflation.

coefficient that is small and insignificant. Finally, the relationship between the change in poverty and the anticipated change in inflation is close to significant. The point estimate implies that an anticipated increase in inflation of one percentage point is associated with a decline in poverty of 0.2 percentage points.

Column 4 considers all three variables together. In addition, because poverty fell on average less rapidly in the 1980s and 1990s than in earlier decades, this specification includes a trend. As before, there is a quantitatively large and overwhelmingly statistically significant relationship between unemployment and poverty. The point estimate on the change in unemployment is similar to that in the univariate regression. The estimated coefficient on the unanticipated change in inflation continues to be small and statistically insignificant. The one important change is that the coefficient on the anticipated change in inflation is now close to zero and not at all significant. That is, the multivariate regression suggests a strong relationship between unemployment and poverty, and essentially no relationship between inflation and poverty.

The reason the univariate and multivariate specifications yield different results for anticipated inflation is that anticipated increases in inflation are correlated with falls in unemployment. When the change in unemployment is omitted from the regression, the antic-

Table 1 Poverty and the Macroeconomy

	(1)	(2)	(3)	(4)
Constant	.01 (.15)	.08 (.60)	–.02 (.16)	–.79 (1.39)
Change in unemployment	.44 (6.91)			.49 (5.71)
Unanticipated change in inflation		–.04 (.44)		.03 (.52)
Anticipated change in inflation			–.21 (1.64)	.05 (.36)
Trend				.02 (1.54)
R^2	.67	.01	.10	.75
S.e.e.	.37	.64	.61	.35

The dependent variable is the change in the poverty rate. The sample period is 1969–94. Absolute values of *t*-statistics are in parentheses.

ipated change in inflation serves as a noisy proxy for this variable. The result is a modest negative coefficient. But when the change in unemployment is included, the negative coefficient on the anticipated change in inflation disappears. That is, there is no evidence of a direct impact of anticipated inflation on poverty.

Poverty has fallen relatively little since 1985 despite the large fall in unemployment. Blank (1993) therefore suggests that cyclical expansions may have a smaller impact on poverty today than in the past. To explore this possibility, were-estimate the regression in column 4 allowing the constant term and the coefficient on the change in unemployment to take on different values beginning in 1983 (the date suggested by Blank). This exercise provides no support for Blank's suggestion. The point estimates of unemployment's impact on poverty for the two periods are essentially identical (0.479 versus 0.475), and the t-statistic for the null hypothesis that the effect has not changed is virtually zero (0.04). That is, the reason that poverty has not fallen greatly in the past 15 years is not that cyclical expansions are much less effective in reducing poverty than before, but that other forces—most obviously the long-term trend of rising inequality—have roughly offset the effects of the large fall in unemployment.[1]

Income Distribution and the Macroeconomy

Cyclical fluctuations clearly affect poverty through their impact on average income. But they may also affect poverty by changing the distribution of income around its mean. To investigate this possibility, we consider the relationship between income distribution and macroeconomic performance.

We consider three measures of income distribution: the Gini coefficient for family incomes, the fraction of income going to the poorest fifth of families, and the fraction of income going to the poorest fifth of households. The last two measures differ only in the population they consider: the family-based measure is based on groups of two or more individuals living together related by blood or marriage, while the household-based measure is based on all individuals.

For simplicity, we focus on the specification like that in our multivariate regression in Table 1. That is, we regress the change in the relevant measure of income distribution on a constant, the change in unemployment, the unanticipated and anticipated changes in inflation, and a trend.

Table 2 reports the results. The point estimates suggest that unemployment has little impact on income distribution. The estimated impact of unemployment on the Gini coefficient is close to zero and highly insignificant. For the share of income going to the poorest fifth of families, a one-percentage-point rise in the unemployment rate is associated with a fall in the poor's income share of 0.05 percentage points. This estimate is marginally significant, but quantitatively small. For example, this group's income share fell by 1.4 percentage points from its peak in 1969 to 1994. And when we consider the income share of the poorest fifth of households rather than the poorest fifth of families, the estimates imply that an increase in unemployment is associated with a slight rise in the poor's income share.

While unemployment appears to have no noticeable effect on distribution, the results do suggest that inflation may narrow the income distribution slightly. The estimates imply that unanticipated inflation is associated with a higher income share of the poor (by either measure) and with a lower Gini coefficient. However, only the correlation between inflation and the income share of the poorest fifth of households is statistically significant. And even in that case, the omission of a single year (1974) reduces the t-statistic to 1.5. More

Table 2 Income Distribution and the Macroeconomy

	(1) Change in Gini coefficient	(2) Change in lowest quintile's share (families)	(3) Change in lowest quintile's share (households)
Constant	–.40 (.43)	–.09 (.58)	.00 (.04)
Change in unemployment	–.02 (.15)	–.05 (1.99)	.02 (.95)
Unanticipated change in inflation	–.10 (1.12)	.02 (1.08)	.03 (2.77)
Anticipated change in inflation	–.15 (.62)	–.01 (.37)	.03 (.81)
Trend	.02 (.79)	.00 (.21)	–.00 (.26)
R^2	.13	.29	.32
S.e.e.	.57	.09	.08

The sample period is 1969–94. Absolute of *t*-statistics are in parentheses.

important, the estimated coefficients are small. For example, the point estimate implies that one percentage point of unanticipated inflation is associated with a fall in the Gini coefficient (measured on a scale of 0 to 100) of just 0.10. For comparison, the rise in the Gini coefficient from 1969 to 1994 was 7.70. Finally, the point estimates for changes in anticipated inflation are similar to those for unanticipated inflation. The coefficients are estimated less precisely, however.[2]

Previous Studies

Various other authors have examined the impact of macroeconomic performance on poverty and income distribution using U.S. time-series data. Essentially everyone who has examined the issue has found, as we do, that economic expansions reduce poverty (see, for example, Anderson, 1964; Perl and Solnick, 1971; and Blank, 1993). And the results of Blinder and Esaki (1978), Blank and Blinder (1986), and Cutler and Katz (1991) are consistent with our finding that inflation leaves the income distribution essentially unchanged or causes it to narrow slightly.

Previous work has found a stronger impact of unemployment on income distribution than our results suggest (Metcalf, 1969; Blinder and Esaki, 1978; Blank and Blinder, 1986; Cutler and Katz, 1991; and Castañeda, Díaz-Giménez, and Ríos-Rull, 1998). For example, Blank and Blinder find that a one-percentage-point rise in unemployment is associated with a fall in the income share of the poorest fifth of families of 0.19 percentage points. For comparison, our point estimate is a fall of 0.05 percentage points (and a rise of 0.02 percentage points for the poorest fifth of households). The key to the difference is the sample periods: increases in unemployment are associated with widening of the income distribution in the 1950s and 1960s, but with essentially no change in the 1970s and 1980s. Previous work examines earlier sample periods. Blank and Blinder, for example, consider 1948–83, and Castañeda, Díaz-Giménez, and Ríos-Rull consider 1948–86. As a result, these papers find a relationship between unemployment and the distribution of income. But that relationship is largely absent when more recent data are analyzed.

There are two other important types of evidence concerning economic aggregates and the welfare of the poor in addition to the U.S. time-series evidence. The first is the U.S. regional evidence examined by Blank and Card (1993). Blank and Card investigate the relationship between regional poverty rates and measures of regional economic activity.

They focus on time-series cross-section regressions that include both year and region dummies; thus they do not use either the aggregate time-series variation or the overall cross-region variation in their estimation. Their findings provide further support for the proposition that increases in overall economic activity reduce poverty. For example, they estimate that a one-percentage-point fall in a region's unemployment rate is associated with a reduction in the poverty rate of 0.28 percentage points. And they find, as we do, no discernable change over time in the impact of economic activity on poverty.

Blank and Card also find little impact of overall activity on income distribution. For their baseline specification, they find that a change in a region's unemployment rate has virtually no impact on the poor's income share. When they include control variables or use the growth of median income rather than unemployment as their cyclical indicator, they find that economic expansions cause a slight rise in the poor's share. These weak effects arise from a combination of two offsetting forces: the poor's labor earnings are much more responsive than other groups' to overall activity, but labor earnings are a considerably smaller fraction of their income.

The second kind of additional evidence is that from other countries. Guitián (1998) reports time-series estimates of inflation's impact on the poor's income share for ten countries. The estimated effect is positive in four cases and negative in six, and in most cases it is not clearly significantly different from zero. Thus again there is no evidence of an important systematic short-run effect of inflation on income distribution.

Inflation and the Balance Sheets of the Poor

One of the most commonly cited effects of inflation is that it causes redistributions from creditors to debtors. Unanticipated inflation reduces the real value of nominal assets and liabilities. It therefore causes real capital losses for nominal creditors, and real capital gains for nominal debtors. If the poor are net nominal debtors, these effects on net benefit them.

The income measures we consider above do not include any capital gains and losses. To investigate inflation's impact on the poor's balance sheets, we therefore examine the balance sheet data from the Federal Reserve's 1995 Survey of Consumer Finances. (These data are available on-line from the Board of Governors of the Federal Reserve System.) Specifically, we examine the financial assets and liabilities of the poor to see if they are likely to be affected substantially by unanticipated inflation.

We focus on the quintile reporting the lowest total income. Some households in this group, however, cannot reasonably be considered poor. For example, some have very high wealth but low or negative income for the survey year because of large losses from their businesses. Since some of these households have extremely high assets and liabilities, classifying them as poor would distort the averages severely. We therefore exclude households with net worth over $100,000. These are households whose net worth puts them in the top 36 percent of the population as a whole. This criterion eliminates about 12 percent of the low-income households from our sample.

Table 3 summarizes the financial balance sheets of this group. We divide financial assets into three categories: transactions accounts, whole life insurance, and other financial assets.[3] We divide financial liabilities into four categories: real estate debt, credit card balances, installment debt, and other liabilities.

The data confirm the conventional view that the poor are nominal debtors. The average poor household has $3,385 of financial assets and $5,201 of debts, and thus has nega-

Table 3 The Financial Balance Sheet of the Poor

	Mean	Fraction with positive amount	Mean excluding top 10 percent
Assets			
Transactions accounts	$1,237	58%	$342
Whole life insurance	729	13	13
Other financial assets	1,418	21	89
All financial assets	3,385	66	1,070
Liabilities			
Real estate debt	2,660	11	9
Credit card balances	440	25	67
Installment debt	1,590	29	317
Other debt	511	8	0
All debt	5,201	50	1,372
Financial net worth	–1,816		

tive financial net worth. And most of the poor's debts are medium and long term: the two most important categories of debt are real estate and installment debt.

But the more important message of Table 3 is that the potential redistributive effects of unanticipated inflation on the poor through capital gains and losses are small. This is true in two senses. First, the mean levels of financial assets and liabilities among the poor are too small to be greatly affected by inflation. A back-of-the envelope calculation demonstrates this. Shiller (1997) reports that the standard deviation of the 10-year inflation rate for the postwar United States is 32 percentage points. Suppose then that inflation over a 10-year period is 32 percentage points higher than anticipated. In addition, suppose that the real value of the poor's financial assets is fully insulated from this inflation, while the real value of their debts falls by half the amount of the unexpected inflation; that is, suppose that the real value of their debts falls by 16 percent. These assumptions surely understate the impact of inflation on the poor's assets, and almost surely overstate the impact on their debts. With these assumptions, the inflation causes a real capital gain to the average poor household of about $800 over the 10-year period, or about $80 per year. For comparison, average annual income in this group is $6,882. Thus, even this generous calculation of the redistributive benefits of inflation to the poor does not yield a large estimate.

Second, the vast majority of the poor have very few financial assets and liabilities at all. For example, 56 percent have less than $500 of all financial assets, and 76 percent have less than $500 of financial assets other than transactions accounts. Similarly, 61 percent have liabilities of less than $500, and 89 percent have no real estate debt (which is the only category that includes any substantial long-term debt). More generally, the average levels of assets and liabilities cited above are driven by a small number of households. Average debts excluding the 10 percent of the poor with the highest debts are just $1,372, and average financial assets excluding the 10 percent with the highest financial assets are just $1,070. Thus for the vast majority of the poor, the potential redistributive effects of inflation are much smaller than the already low figure computed above. We conclude that the traditional redistributive effects of unanticipated inflation are of little importance for the poor.

Implications

Although the cyclical behavior of poverty and income distribution is interesting, it is in fact of little relevance to monetary policy. The reason is simple and well known: monetary policy cannot cause a permanent boom.

To see the difficulty facing monetary policy makers who are concerned about the poor, suppose that output and unemployment are at their normal or natural levels, and that policymakers undertake expansionary policy. The result is a period of below-normal unemployment and above-normal output. This cyclical expansion raises the incomes of the poor and lowers the poverty rate.

To gauge the possible size of this effect on poverty, consider an expansionary monetary policy that reduces the unemployment rate from the natural rate to two points below and keeps it low for two years. Based on the estimates in Table 1 (column 4), such a reduction would lower the poverty rate by almost exactly one percentage point the first year and keep it at that level the second year. Since a reduction in the poverty rate of one percentage point is substantial, such a policy would clearly benefit the poor in the short run.

But the boom cannot last. Monetary policy can push unemployment below normal and output above normal only temporarily. The low unemployment and high output cause inflation to rise. For example, using the usual rule of thumb that unemployment one percentage point below the natural rate for a year raises the inflation rate by one-half of a percentage point, the two-year, two-percentage-point reduction in unemployment described above would lead to inflation that is two percentage points higher than before. Output and unemployment, however, inevitably return to their normal levels. When this happens, poverty returns to its initial level. Even if policymakers are willing to tolerate the higher inflation, all the expansionary policy has achieved is a temporary period of below-normal poverty at the cost of permanently higher inflation.

A more likely outcome is that policymakers will choose not to accept the higher inflation. In this case, they will adopt contractionary policies to bring inflation back to its initial level. The result is a period of below-normal output and above-normal unemployment and poverty. In this case, policy has had no impact on the average level of poverty; it has only rearranged its timing.

In addition, some recent evidence suggests that the output-inflation tradeoff is asymmetric: above-normal output causes inflation to rise more rapidly than the same amount of below-normal output causes it to fall (Clark, Laxton, and Rose, 1996, and Debelle and Laxton, 1997). In this case, the contraction needed to decrease inflation is larger than the expansion that increased it, and so the boom-bust cycle raises average poverty.

We have described the dilemma facing compassionate policymakers in terms of the decision of whether to undertake expansionary policy in an economy operating at normal capacity. But the problem is general. Suppose, for example, that concern about the poor causes monetary policymakers to err on the side of preventing recessions. Such a policy results in output being above normal more often than it is below normal. Since above-normal output raises inflation and below-normal output lowers it, the result is that inflation is on average rising. But then policymakers are in the same position as before. At some point they must switch to a policy of keeping output on average equal to normal. Thus a policy of erring against contraction can produce at most a temporary period of below-normal poverty. And in the more likely case where policymakers eventually decide to reverse the rise in inflation, the policy does not succeed in lowering average poverty at all.

In summary, the cyclical aspects of poverty are not central to the question of how concern about poverty and income distribution should affect monetary policy. Monetary policy cannot permanently reduce poverty and inequality by creating booms or preventing recessions.·

THE EFFECTS OF MONETARY POLICY ON THE POOR IN THE LONG RUN

The Channels through which Monetary Policy Affects the Poor

What monetary policy can control in the long run is average inflation and the variability of aggregate demand. These can affect the well-being of the poor both by influencing long-run growth and by influencing the distribution of income.

High inflation creates uncertainty, generates expectations of future macroeconomic instability and distortionary policies, disrupts financial markets, and creates high effective tax rates on capital. It thereby discourages investment of all types: physical capital accumulation, human capital accumulation, innovation and research and development, and foreign direct investment and technology transfer. As a result, it can retard growth. Because macroeconomic instability is also likely to discourage investment, it can have similar effects. Furthermore, to the extent that high inflation and high variability generate uncertainty about the return to productive activities and increase the scope for activities that are privately but not socially beneficial, they may lower work effort and lead to rent seeking. This can also erode a country's average standard of living.

High inflation and macroeconomic volatility can also affect the poor through the distribution of income around its average. There are at least five channels through which monetary policy can affect long-run income distribution. First, the redistributions caused by swings in unanticipated inflation directly affect inequality. Second, the reductions in physical capital investment caused by uncertainty and financial-market disruptions raise the average return on capital and depress wages; thus they widen the income distribution. Third, offsetting this, inflation may shift the burden of taxation away from labor and toward capital. Fourth, the uncertainty and reduced effectiveness of financial markets caused by inflation and macroeconomic instability reduce not just physical capital investment, but human capital investment. This thwarts an important mechanism by which inequality can be mitigated. And finally, inflation and macroeconomic volatility may harm some sectors of the economy disproportionately. For example, they may be particularly harmful to simple manufacturing or export-oriented industries. Depending on the relative position of the workers in these industries, this can either increase or decrease inequality.

To investigate how inflation and macroeconomic instability affect the poor, we examine the cross-country relationship between these variables and the poor's standards of living. Because the effects of inflation and volatility are likely to be gradual and cumulative, little can be learned from looking at variation over time within a country. Across countries, in contrast, there is a great deal of variation in the long-term performance of monetary policy. Thus the cross-country evidence has the greatest potential to be instructive.

We begin by examining the relationship between the long-run performance of monetary policy and the overall well-being of the poor. We then turn to monetary policy's relationship with the two determinants of that overall well-being, the average income of the population as a whole and the distribution of income.

As is well known, cross-country regressions must be interpreted with caution. There are inevitably a host of important omitted factors, and the search for useful instruments for macroeconomic variables has had little success. Thus such regressions can show only correlations, not causation. Nonetheless, we think it is useful to ask how the poorest segment of society fares in countries where monetary policy has kept inflation low and demand stable relative to countries where policy has produced high inflation and unstable demand.

Data

The key variable in our analysis is the average income of the poorest quintile of a country's population. We derive this measure by multiplying the average real income in each country times the share of income going to the poorest fifth of the population, times five.

The data on the income share of the poorest fifth of the population come from the comprehensive database on inequality assembled by Deininger and Squire (1996). This database is the result of a careful and exhaustive search of country-level inequality data. We restrict our attention to data that meet Deininger and Squire's criteria for high quality: the data must be based on household surveys and have comprehensive coverage of the population and of income sources. Deininger and Squire are also the source of the data on the Gini coefficient that we analyze later.[4]

Our empirical work focuses on data for 1988. We choose this date on the basis of data availability: using more recent data requires large reductions in the sample. Inequality and poverty evolve sufficiently slowly, however, that it is unlikely that the specific year we consider is crucial to the results. Thus our share data are for 1988 whenever possible. If no data are available for a country for that year, we use as close a year as possible, but in any event not before 1983 or after 1993.

The data on average real income per person are from Summers and Heston's well-known data set. These data are described by Summers and Heston (1991). Updated versions are available on-line from the National Bureau of Economic Research; we use Mark 5.6 of the data. All of the real income data are for 1988.

As described above, we focus on two indicators of the long-run performance of monetary policy: average inflation and the variability of aggregate demand. We measure inflation as the average change in the logarithm of the GDP deflator over the period 1970–90, and demand variability as the standard deviation of the change in the logarithm of nominal GDP over the same period. These data are from the World Bank's World Data CD-ROM (1995 edition).

We consider two basic samples. The first consists of all countries for which we can obtain data. This sample consists of 66 countries when we analyze the average income of the poor. The second sample consists of industrialized countries. Specifically, we consider the countries that were in the Organization for Economic Cooperation and Development (OECD) as of 1973. This is a simple way of excluding the less industrialized countries that have joined the OECD in the past few years (the Czech Republic, Hungary, Korea, Mexico, and Poland). This sample has 19 countries when we analyze the average income of the poor.

Monetary Policy and the Well-Being of the Poor

Figure 4 is a scatter plot of the logarithm of the average income of the poorest fifth of the population against average inflation. Figure 5 replaces average inflation with demand variability. Both plots suggest a negative relationship: the average income of the poor tends to be lower in countries where monetary policy has produced higher average inflation and

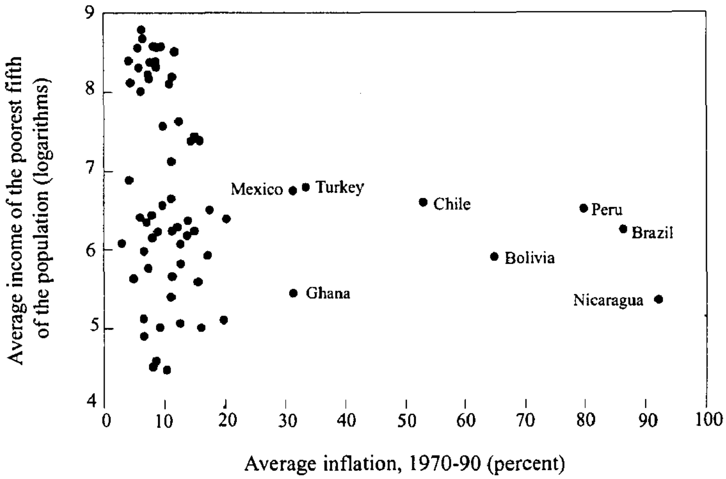


Figure 4 The income of the poor and average inflation.

greater macroeconomic volatility. Both charts also show that there are a handful of outliers that are likely to be important to any estimated relationship. Figures 6 and 7 therefore repeat Figures 4 and 5 without the outliers. Specifically, we omit countries with average inflation above 25 percent from Figure 6, and countries with a standard deviation of nominal GDP growth above 30 percent from Figure 7. Again, both plots suggest negative relationships.

Table 4 reports regression results. Column 1 is a regression of the logarithm of the average income of the poor on a constant and average inflation; thus it is the regression corresponding to Figure 4. The point estimate implies that a one-percentage-point rise in average inflation is associated with a reduction in the poor's average income of about one and one-half percent. Thus, a country with inflation that is one standard deviation (18.5 percentage points) above the mean is predicted to have average income of the poorest quintile

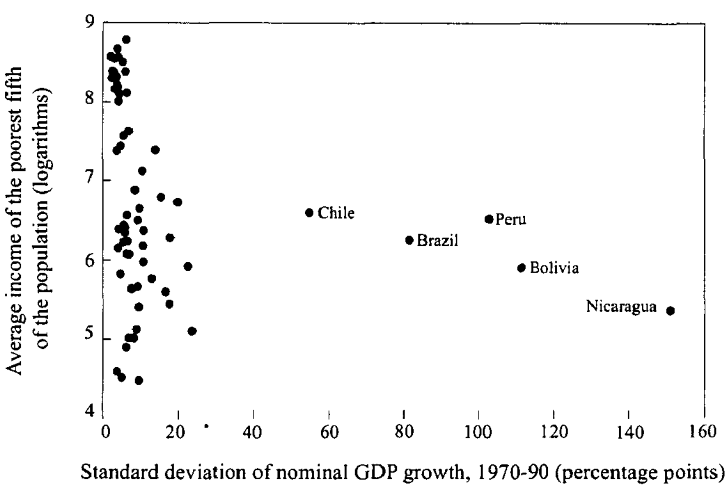


Figure 5 The income of the poor and aggregate demand variability.

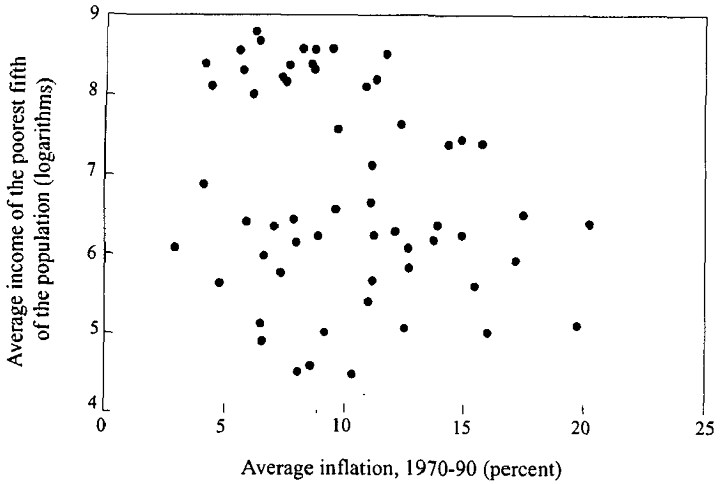


Figure 6 The income of the poor and average inflation excluding outliers.

that is 25.6 percent below the mean. That is, it suggests a quantitatively important relationship. The relationship is not estimated very precisely, however. For example, the null hypothesis of no association is only moderately rejected.

Column 2 considers demand variability rather than average inflation; thus it corresponds to Figure 5. Again the point estimate implies a large relationship. A one-percentage-point rise in the standard deviation of nominal GDP growth is associated with a one-percent fall in the poor's average income. This implies that a country with demand variability one standard deviation (26.7 percentage points) above the mean is predicted to have average income of the poorest quintile 28.6 percent below the mean. But again the estimate is imprecise.

Columns 3 and 4 exclude the outliers; thus they correspond to Figures 6 and 7. The point estimates rise sharply. They now imply that a one-percentage-point rise in average in-

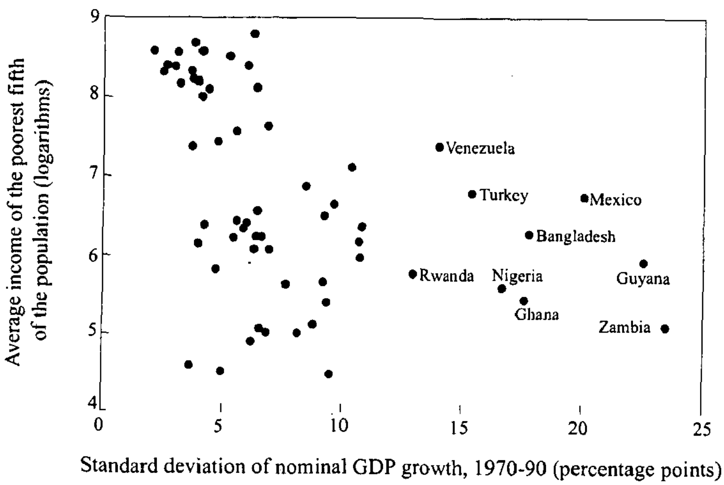


Figure 7 The income of the poor and aggregate demand variability excluding outliers.

Table 4 Monetary Policy and the Income of the Poor

	(1)	(2)	(3)	(4)	(5)
Constant	6.93 (34.68)	6.87 (39.97)	7.64 (16.99)	7.62 (27.59)	6.83 (29.73)
Average inflation	−1.38 (1.68)		−8.58 (2.05)		.57 (.24)
Standard deviation of nominal GDP growth		−1.07 (1.89)		−11.18 (3.70)	−1.44 (.87)
Outliers excluded?	No	No	Yes	Yes	No
Sample size	66	66	58	61	66
R^2	.04	.05	.07	.19	.05
S.e.e.	1.23	1.22	1.26	1.16	1.23

The dependent variable is the logarithm of the average income of the poorest fifth of the population. Absolute values of *t*-statistics are in parentheses.

flation is associated with a fall in the poor's income of 9 percent, and that a one-percentage-point rise in the standard deviation of nominal GDP growth is associated with a fall of 11 percent. That is, the results suggest that the relationship between the long-run performance of monetary policy and the poor's well-being is greater at low levels of average inflation and demand variability. As a result, even though the standard deviations of both inflation and variability are much smaller in the reduced samples, a country with inflation one standard deviation (4.0 percentage points) above the mean is predicted to have average income of the poorest quintile 34.4 percent below the mean, and a country with demand variability one standard deviation (5.0 percentage points) above the mean is predicted to have average income of the poor 55.4 percent below the mean. Excluding the outliers greatly increases the estimated coefficients' standard errors, however. As a result, the coefficient on average inflation is still only marginally significant. But despite the rise in the standard errors, the coefficient on demand variability is now highly significant.

Average inflation and the standard deviation of nominal GDP growth are highly correlated. In the full sample of 66 countries for which we have average income for the poorest quintile, for example, their correlation is 0.94. As a result, the data are not able to distinguish the relationship of the poor's average income with average inflation from its relationship with demand variability. Column 5 shows the results of including both variables in the regression. The standard errors of both coefficients are large, and neither is close to statistically significant.

As described above, these simple cross-country regressions leave out many other factors that influence the incomes of the poor, and these omitted factors may be correlated with the long-run performance of monetary policy. One way to address this problem is to add dummy variables for different regions to the regressions. There may be important differences across parts of the world in such factors as the quality of institutions and cultural attitudes toward thrift and entrepreneurship. By including regional dummies, we can eliminate the possibility that such differences are the source of our results. At the same time, including the dummies has the disadvantage that we no longer use the large cross-region variation in the long-run performance of monetary policy to estimate the coefficients of interest.

Table 5 reports the results of re-estimating the regressions in Table 4 with a dummy variable for each continent.[5] The addition of the continent dummies does not change the basic character of the results. For the full-sample regressions (columns 1 and 2), the main ef-

Table 5 Monetary Policy and the Income of the Poor with Continent Dummies

	(1)	(2)	(3)	(4)	(5)
Average inflation	−1.47 (2.23)		−5.71 (1.95)		−1.65 (1.03)
Standard deviation of Nominal GDP growth		−.85 (1.96)		−3.80 (1.64)	.13 (.13)
Outliers excluded?	No	No	Yes	Yes	No
Sample size	66	66	58	61	66
R2	.67	.66	.68	.66	.67
S.e.e.	.75	.76	.77	.78	.76

The dependent variable is the logarithm of the average income of the poorest fifth of the population. All equations include continent dummies. Absolute values of t-statistics are in parentheses.

fect of including the dummies is to reduce the standard errors of the coefficients on the monetary policy variables slightly. For the regressions excluding the outliers (columns 3 and 4), including the dummies reduces the point estimates considerably. They are, nevertheless, still quite large: for both average inflation and variability, a country that has a value for the independent variable one standard deviation above the mean is still predicted to have average income of the lowest quintile roughly 20 percent below the mean. The coefficient on demand variability, however, is no longer clearly significant. And when we include both average inflation and demand variability (column 5), we again find that neither coefficient can be estimated with any precision.

Table 6 reports the results for the traditional OECD. Again outliers are an important concern: Turkey is by far the poorest country in this sample, and has by far the highest inflation and the most volatile demand. Thus we report the results both with and without Turkey.

The regressions show that among industrialized countries, there is a powerful relationship between average inflation and the well-being of the poor. For the full sample, the point estimate is that a one-percentage-point rise in average inflation is associated with a fall in the poor's average income of 7 percent. Thus, a country with inflation one standard deviation (6.3 percentage points) above the mean for industrialized countries is predicted to have average income for the poorest quintile 42.6 percent below the mean. The null hypothesis of no relationship is overwhelmingly rejected. When Turkey is excluded, the point estimate is even larger. It is not as precisely estimated as before, but is still highly significant.

Table 6 Monetary Policy and the Income of the Poor in Industrialized Countries

	(1)	(2)	(3)	(4)	(5)
Constant	8.87 (64.51)	8.71 (48.90)	9.12 (39.23)	8.49 (22.94)	8.84 (65.24)
Average inflation	−6.74 (5.79)		−9.69 (3.85)		−9.87 (3.90)
Standard deviation of nominal GDP growth		−11.79 (3.46)		−5.98 (.64)	7.76 (1.38)
Outliers excluded?	No	No	Yes	Yes	No
Sample size	19	19	18	18	19
R^2	.66	.41	.48	.02	.70
S.e.e.	.31	.41	.31	.42	.30

The dependent variable is the logarithm of the average income of the poorest fifth of the population. Absolute values of t-statistics are n parentheses.

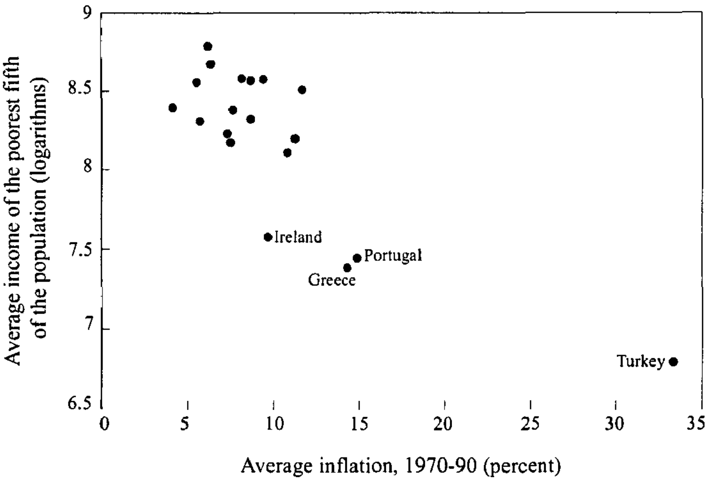


Figure 8 The income of the poor and average inflation in industrialized countries.

The relationship between demand variability and the income of the poor in industrialized countries, on the other hand, is not clear. When Turkey is included, there is a large and highly significant negative association. When Turkey is omitted, however, the relationship is estimated so imprecisely that a two-standard-deviation confidence interval includes both very large negative and very large positive coefficient values.

For the industrialized countries, in contrast to the full sample, attempting to distinguish the relationship of the poor's incomes with average inflation from their relationship with demand variability produces a clear result: it is average inflation that is associated with the poor's incomes. As column 5 shows, the point estimate on average inflation is large and highly significant, while the coefficient on demand variability is not estimated with any precision. And excluding Turkey has no important effect on the estimates or their standard errors.

Figures 8 and 9 show the source of the estimates. Figure 8 shows that there is a strong

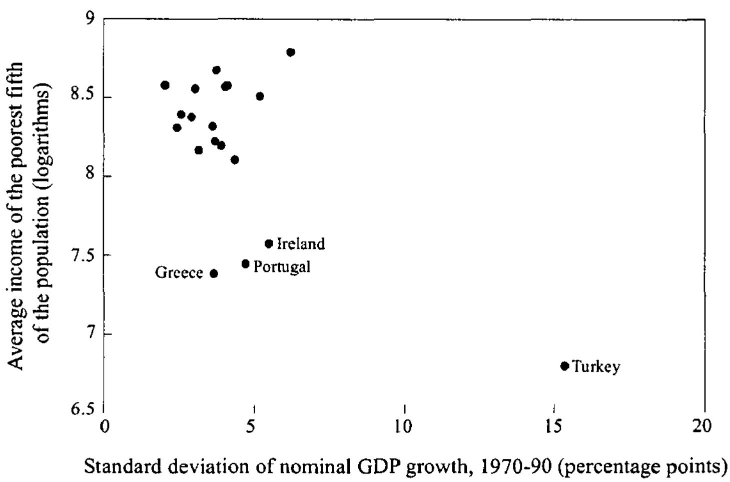


Figure 9 The income of the poor and aggregate demand variability in industrialized countries.

negative relationship between the poor's average income and average inflation in the industrialized countries either with or without Turkey. Figure 9, on the other hand, shows that there is no clear relationship between the poor's incomes and demand variability in these countries beyond the fact that Turkey has highly volatile demand and particularly low incomes among its poor.

We conclude that the data point to an important relationship between the long-run performance of monetary policy and the well-being of the poor. On average, the poor are much better off in countries where monetary policy has kept inflation low and aggregate demand growth stable.

There are two important caveats to this conclusion, both of which are common to this type of cross-country exercise. First, the estimates are imprecise. For example, although the point estimates imply a large relationship, the data do not provide compelling evidence against the view that there is no systematic relationship between the long-run performance of monetary policy and the poor's incomes. Second, the regressions do not establish causation. There may be omitted variables that are correlated with the performance of monetary policy that are, in fact, the key determinants of the poor's incomes.

For the conduct of monetary policy, the issue of causation is in fact less important than it appears. High inflation cannot be eliminated in isolation. If there is high inflation because a lack of fiscal discipline or of an effective tax system is leading the government to rely on money finance, for example, reducing inflation requires eliminating the underlying fiscal problem. More generally, inflation reduction is often part of a comprehensive package of policies involving fiscal discipline, macroeconomic stabilization, and microeconomic liberalization. If the package raises the standards of living of the poor in the long run, the question of whether it was the reduction in inflation or the other policy changes that was key is of secondary importance.

Monetary Policy and Average Income

As a matter of definition, the average income of the poor is determined by the average income of the full population and how the poor's incomes compare with that average. Thus, to investigate the relationship between the poor's incomes and monetary policy further, we examine the relationships of average income and of income distribution with monetary policy.

Of these two determinants of the poor's average income, the average income of the full population is by far more important. As described above, the average income of the poorest fifth of the population equals the product of overall average income and the poorest fifth's income share, times 5:

$$\bar{Y}_{POOR} = \bar{Y} \times Q \times 5, \tag{1}$$

where $\bar{Y}_{POOR}$ is the poor's average income, $\bar{Y}$ is overall average income, and Q is the lowest quintile's income share. Thus:

$$ln\bar{Y}_{POOR} = lnY + lnQ + ln5. \tag{2}$$

The variance of the logarithm of the poor's average income therefore equals the sum of the variance of the logarithm of overall average income, the variance of the logarithm of the lowest quintile's share, and a covariance term:

$$Var(ln\bar{Y}_{POOR}) = Var(ln\bar{Y}) + Var(lnQ) + 2Cov(ln\bar{Y}, lnQ). \tag{3}$$

Computing the three terms on the right-hand side of equation (3) shows that the large majority of the variation across countries in the poor's average income arises from variation in overall average income. For our full sample of 66 countries, for example, over two-thirds of the variance in $ln\bar{Y}_{POOR}$ is due to variance in $ln\bar{Y}$. Less than one-eighth comes from variance in lnQ. The remaining one-fifth comes from the fact that the overall average and the lowest quintile's share are moderately correlated.

Considering the possibilities for changing the poor's average income within a country rather than examining the variation across countries reinforces the view that average income is the prime determinant of the poor's well-being. The cross-country record provides many examples of countries where misguided policies have severely lowered average incomes and of countries where sound policies appear to have significantly raised average incomes. As Li, Squire, and Zou (1998) show, however, large changes in income distribution within a country are rare. The variation in inequality within countries over time is vastly smaller than the variation across countries, and statistically or quantitatively significant trends in inequality within a country are uncommon.

Thus, for monetary policy to have an important impact on the well-being of the poor in the long run, it must have an important effect on the long-run behavior of average income. This relationship between inflation and average income has been the subject of considerable research (see, for example, Fischer, 1993; Cukierman and others, 1993; Barro, 1996; and Bruno and Easterly, 1998). An examination of these findings provides an important check on our previous findings concerning inflation and poverty. This is especially true because studies of the general inflation-income link can typically analyze much larger samples: many countries that do not keep statistics on poverty or distribution do have income and inflation data.

The basic facts about the relation between inflation and the long-run behavior of average income are similar to those concerning inflation and the incomes of the poor. Lower inflation is on average associated with higher growth, but the data do not allow the relationship to be pinned down with much confidence. Consider, for example, a simple cross-country regression of average annual growth in income per person over the period 1970–88 on a constant and average inflation over 1970–90 for the 104 countries for which we can obtain data on both variables. This regression produces a coefficient on average inflation of −0.022, with a *t*-statistic of 2.2. Thus a country with average inflation one standard deviation (19.0 percentage points) above average is predicted to have an annual growth rate 0.41 percentage points below average. The findings are robust to the inclusion of continent dummies. Excluding countries with high average inflation raises the point estimate considerably; as in our other regressions, however, it also raises the standard error sharply.

Barro (1996) conducts a more detailed examination of the relationship between the long-run performance of monetary policy and long-run growth. He creates a panel data set of 251 observations by constructing separate observations for the periods 1965–75, 1975–85, and 1985–90 for as many countries as possible. He regresses growth in a country in a given period on its average inflation in the period and a large number of control variables, including measures of physical and human capital accumulation. The inclusion of the controls means that the estimates may understate the effects of inflation. If inflation reduces growth by lowering investment, for example, the estimated coefficient on inflation will not capture this.

Barro's point estimates are very similar to those in our simple cross-section. In his baseline specification, for example, the coefficient on inflation is −0.024. Because of the larger sample and the control variables, however, the standard error is much smaller than in the cross-section. In his basic specification, the t-statistic for the null hypothesis of no relationship is almost five.

Barro reports three other results of interest. First, in his sample it is average inflation and not variability (which he measures as the standard deviation of inflation) that is related to growth. Second, excluding the high-inflation observations has little impact on the estimates. In that sense, the results do not depend on these observations. But excluding these observations again raises the standard errors greatly. As a result, the null hypothesis of no relationship can no longer be rejected.

Barro's final result concerns causation. He proposes using dummy variables for countries' prior colonial status as instruments for inflation. Former French and British colonies inherited anti-inflationary norms and institutions, including the fixed-exchange-rate regimes of France's African colonies and the currency boards of many of Britain's colonies. The former Spanish and Portuguese colonies had no such legacies, and their inflation rates have on average been much higher. Thus, prior colonial status is correlated with inflation. Unfortunately, it may also be correlated with factors other than inflation that influence growth: the different colonizers may have affected future growth in ways other than through their impact on inflation. But Barro argues that it is nonetheless interesting to see how using the measures of prior colonial status as instruments affects the estimated relationship between inflation and growth. The answer is that it increases the relationship slightly. Moreover, these variables are not just proxying for Latin American countries: adding a dummy for these countries to the regression has no great effect.

Cukierman and others (1993) also propose instruments for inflation. Specifically, they construct two measures of nonindependence of the central bank: the fraction of political transitions that are accompanied or quickly followed by replacement of the central bank governor, and the frequency of replacement of the central bank governor in times of political stability. Both measures, like Barro's, are correlated with inflation. But there are again reasons that they may be correlated with other determinants of growth. For example, they may be higher in countries where political changes are more disruptive or the rule of law is weaker. Nonetheless, the results are instructive: as with Barro's study, moving from ordinary least squares to instrumental variables increases the magnitude of the estimated relationship between inflation and growth.

Thus, the data suggest that higher inflation is associated with lower growth in overall incomes. More important, two attempts to tackle the issue of causation find no evidence that this correlation is the result of omitted variables that are correlated with inflation. That is, they find no evidence that the correlation does not reflect an effect of inflation on long-run growth. Since the growth of overall income is the key determinant of the poor's well-being in the long run, these results corroborate our earlier finding that inflation appears to be detrimental to the average income of the poor.

Monetary Policy and Income Distribution

The second determinant of the poor's average income is the distribution of income. As our final step, we therefore examine the relationship between the long-run performance of monetary policy and income distribution. We use the Gini coefficient as our measure of income distribution. Because Gini coefficients are available for slightly more countries than are data on average income of the poor, our primary sample in this analysis includes 76 (rather than 66) countries.

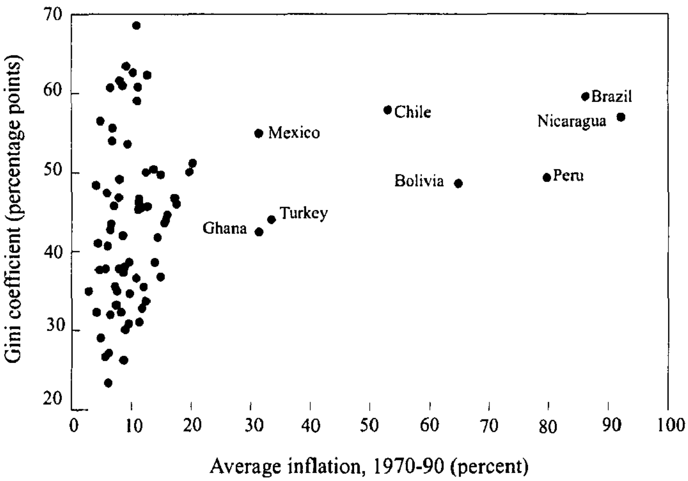


Figure 10 Inequality and average inflation.

Figures 10–13 show the basic relationships. The first two are scatter plots of the Gini coefficient against average inflation and the standard deviation of nominal GDP growth for all countries for which we have data. The next two exclude the outliers. All four charts suggest positive relationships. That is, the distribution of income is less equal in countries with higher average inflation and greater macroeconomic volatility.

Table 7 reports the basic regressions. Column 1 shows that a one-percentage-point rise in average inflation is associated with a rise in the Gini coefficient of 0.2 points, and that the null hypothesis of no relationship is rejected. This relationship is substantial. For example, a country with average inflation one standard deviation (17 percentage points) above average is predicted to have a Gini coefficient 3.3 points above average. For comparison, the standard deviation of the Gini across our sample of countries is 10.6 points.

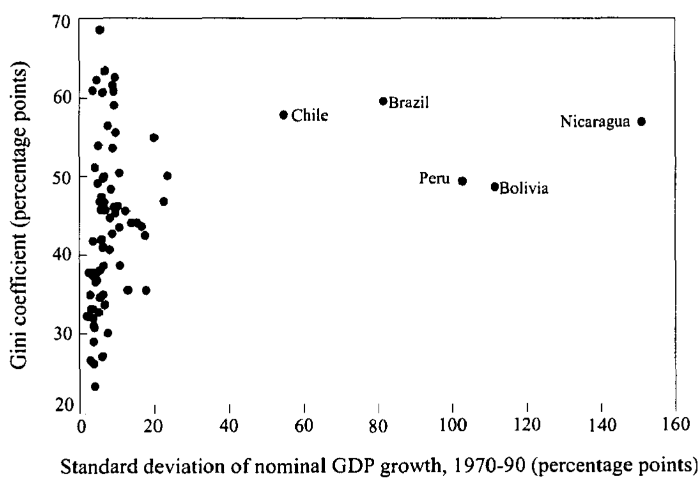


Figure 11 Inequality and aggregate demand variability.

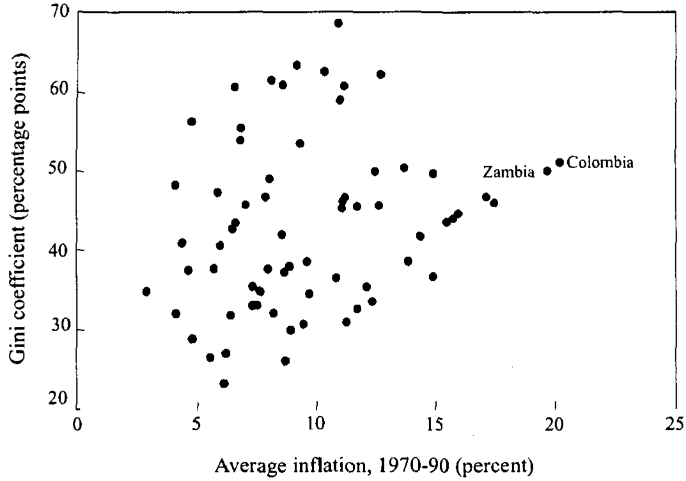


Figure 12 Inequality and average inflation excluding outliers.

Column 2 shows a similar result for volatility. A one-standard-deviation difference between countries in demand variability (25 percentage points) is associated with a 2.9-point difference in Gini coefficients, and the null hypothesis of no relationship is rejected.

Omitting the outliers greatly increases both the point estimates and their standard errors. But with the outliers excluded, the variation in the performance of monetary policy across countries is much smaller: the standard deviation across countries of average inflation is now 3.9 percentage points, and that of the standard deviation of nominal GDP growth is now 4.7 percentage points. As a result, the predicted differences in inequality associated with differences of one standard deviation in average inflation and volatility are roughly the same as before. The estimates imply that a country with average inflation one standard deviation above average has a Gini coefficient 2.5 points above average, and that a country with demand variability one standard deviation above average has a Gini 2.9

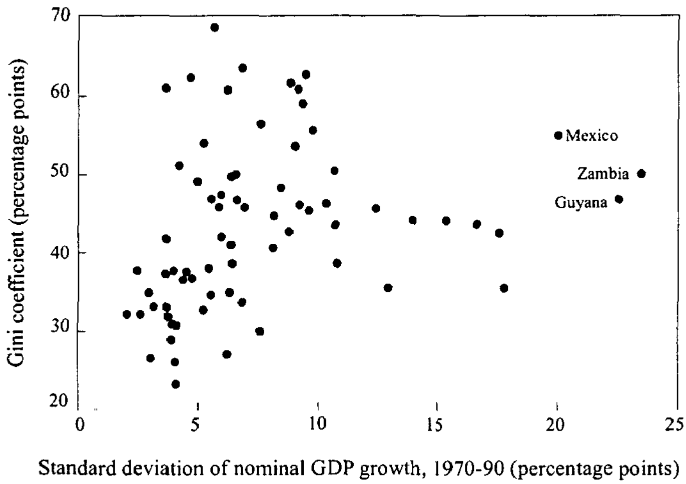


Figure 13 Inequality and aggregate demand variability excluding outliers.

Table 7 Monetary Policy and Inequality

	(1)	(2)	(3)	(4)	(5)
Constant	.41 (27.07)	.43 (31.70)	.37 (10.85)	.39 (16.64)	.41 (23.57)
Average inflation	.19 (2.79)		.64 (1.94)		.24 (1.22)
Standard deviation of nominal GDP growth		.12 (2.49)		.61 (2.35)	−.04 (.27)
Outliers excluded?	No	No	Yes	Yes	No
Sample size	76	76	68	71	76
R^2	.09	.08	.05	.07	.10
S.e.e.	.10	.10	.10	.10	.10

The dependent variable is the Gini coefficient. Absolute values of *t*-statistics are in parentheses.

points above average. Both estimates are only marginally significant, however. Finally, column 5 shows that it is again not possible to distinguish separate relationships with average inflation and with demand variability.

The results for inequality, in contrast to those for the poor's average income, are sensitive to the inclusion of continent dummies. This is shown in Table 8. The only statistically significant result is that for the full sample, either excluding or including variability, higher average inflation is associated with greater inequality. All the other estimates are sufficiently imprecise that it is not possible to reject either the hypothesis of no relationship or the hypothesis of a quantitatively important one.

Finally, Table 9 reports the results for the traditional OECD. There is a quantitatively large and statistically significant positive association between inequality and average inflation. This is true regardless of whether Turkey is included in the sample, and regardless of whether the regression also includes variability. For variability, in contrast, only the simple regression for the full sample shows a significant relationship. In the other cases, the estimates are too imprecise to be informative.

We conclude that there is some evidence of an important positive relationship of inequality with average inflation and demand variability. This finding is consistent with the results of Al-Marhubi (1997). Al-Marhubi finds a positive correlation between inequality and average inflation similar to the one we report in column 1 of Table 7. He also finds that this result is robust to controlling for political stability, central bank independence, and openness.

Table 8 Monetary Policy and Inequality with Continent Dummies

	(1)	(2)	(3)	(3)	(5)
Average inflation	.12 (2.00)		.24 (.93)		.34 (2.29)
Standard deviation of nominal GDP growth		.05 (1.15)		−.10 (.48)	−.15 (1.59)
Outliers excluded?	No	No	Yes	Yes	No
Sample size	76	76	68	71	76
R^2	.57	.55	.57	.54	.59
S.e.e.	.07	.07	.07	.07	.07

The dependent variable is the Gini coefficient. All equations include continent dummies. Absolute values of *t*-statistics are in parentheses.

Table 9 Monetary Policy and Inequality in Industrialized Countries

	(1)	(2)	(3)	(4)	(5)
Constant	.29 (19.07)	.30 (16.65)	.29 (11.11)	.35 (9.46)	.30 (18.60)
Average inflation	.46 (3.41)		.55 (1.89)		.72 (2.66)
Standard deviation of nominal GDP growth		.75 (2.12)		−.45 (.48)	−.69 (1.10)
Outliers excluded?	No	No	Yes	Yes	No
Sample size	21	21	20	20	21
R^2	.38	.19	.17	.01	.42
S.e.e.	.04	.04	.04	.04	.04

The dependent variable is the Gini coefficient. Absolute values of *t*-statistics are in parentheses.

Once again, the finding of a correlation does not establish causation. Sachs (1989) argues that inequality arising from sources other than monetary policy leads to distributional conflicts, which in turn lead to fiscal stalemates, macroeconomic instability, and reliance on money finance. Thus our correlations may reflect causation from inequality to monetary policy rather than the reverse. Indeed, Al-Marhubi's regressions have inflation on the left-hand side and inequality on the right.

Even with this important caveat in mind, we believe this analysis of the correlation between inequality and monetary policy provides further corroboration of our key finding. Our analysis shows that low inflation and macroeconomic stability are associated with higher income for the poor. While this correlation is due primarily to the correlation between prudent monetary policy and growth, the link is augmented by the correlation between prudent policy and greater equality.

CONCLUSIONS

Deriving implications about the impact of alternative policies from admittedly imperfect regressions is always risky. Nevertheless, we believe two conclusions about the interaction between monetary policy and the well-being of the poor are warranted.

First, our analysis suggests that the usual emphasis on the short-run effects of monetary policy on poverty is fundamentally misguided. It is certainly true that expansionary policy can generate a boom and reduce poverty temporarily. But the effect is unquestionably just that—temporary. Monetary policy cannot generate a permanent boom. When output returns to the natural rate, poverty will return to its initial level. Moreover, the cost of such a boom is that inflation is permanently higher. If the higher inflation creates a consensus for tight policy to reduce inflation, the resultant rise in unemployment leads to a rise in poverty that offsets even the temporary reduction generated by the boom.

Second, the cross-country relationship between monetary policy and poverty suggests that monetary policy that aims at low inflation and stable aggregate demand is the most likely to result in genuinely improved conditions for the poor in the long run. It is, of course, completely possible that the relationship between prudent monetary policy and higher incomes for the poorest quintile that we find is not causal. Nevertheless, we strongly suspect that the typical package of reforms that brings about low inflation and macroeconomic stability will also generate improved conditions for the poor and more rapid growth for all.

ENDNOTES

1. We have investigated the robustness of our findings along a large number of dimensions: omitting the trend; including lagged as well as contemporaneous changes in unemployment and inflation; considering longer sample periods (which requires us to not distinguish between anticipated and unanticipated inflation); allowing the effects of inflation as well as unemployment to change in 1983; and estimating the regressions in levels rather than changes (and including the lagged dependent variable on the right-hand side). In all cases, the qualitative picture is the same: there is a strong relationship between unemployment and poverty, and no clear relationship between inflation and poverty. In two of the variants (omitting the trend and including lags), there is a modest tendency for increases in inflation to be associated with increases in poverty. But the coefficients on inflation are never significantly different from zero.
2. We have investigated the robustness of the results for the distribution of income along the same dimensions that we examined the results for poverty. These results support our findings that unemployment has no systematic impact on the distribution of income, and that inflation may narrow it slightly.
3. The two most important components of our other financial assets category are certificates of deposit and the survey's residual category (which includes loans, future proceeds, royalties, futures, nonpublic stock, deferred compensation, oil/gas/mineral investment, cash not elsewhere classified).
4. As Deininger and Squire describe, the inequality measures for some countries are based on spending rather than income. We adjust these observations as suggested by Deininger and Squire to make them comparable to the income-based measures. Specifically, we add 6.6 points to the spending-based Gini coefficients, and we subtract 1.2 percentage points from the spending-based figures for the share of the poorest fifth of the population.
5. We use Summers and Heston's definitions of the continents.

REFERENCES

Al-Marhubi, F. 1997. A Note on the Link between Income Inequality and Inflation. *Economics Letters* 55 (September), pp. 317–19.

Anderson, W. H. L. 1964. Trickling Down: The Relationship between Economic Growth and the Extent of Poverty among American Families. *Quarterly Journal of Economics* 78 (November), pp. 511–24.

Barro, R. J. 1996. Inflation and Growth. Federal Reserve Bank of St. Louis *Review* 78 (May/June), pp. 153–69.

Blank, R. M. 1993. Why Were Poverty Rates So High in the 1980s? In Dimitri B. Papadimitriou and Edward N. Wolff, ed., *Poverty and Prosperity in the USA in the Late Twentieth Century.* New York: St. Martin's Press, pp. 21–55.

Blank, R. M., and Blinder, A. S. 1986. Macroeconomics, Income Distribution, and Poverty. In Sheldon H. Danziger and Daniel H. Weinberg, eds., *Fighting Poverty: What Works and What Doesn't.* Cambridge: Harvard University Press, pp. 180–208.

Blank, R. M., and Card, D. 1993. Poverty, Income Distribution, and Growth: Are They Still Connected? *Brookings Papers on Economic Activity* (2), pp. 285–339.

Blinder, A. S., and Esaki, H. Y. 1978. Macroeconomic Activity and Income Distribution in the Postwar United States. *Review of Economics and Statistics* 60 (November), pp. 604–9.

Bruno, M., and Easterly, W. 1998. Inflation Crises and Long-Run Growth. *Journal of Monetary Economics* 41 (February), pp. 3–26.

Castañeda, A., Díaz-Giménez, J. and Ríos-Rull, 1998. Exploring the Income Distribution Business Cycle Dynamics. *Journal of Monetary Economics* 42 (August), pp. 93–130.

Clark, P., Laxton, D. and Rose, D. 1996. Asymmetry in the U.S. Output-Inflation Nexus. *IMF Staff Papers* 43 (March), pp. 216–51.

Cukierman, A., Kalaitzidakis, P., Summers, L. H. and Webb, S. B. 1993. Central Bank Independence, Growth, Investment, and Real Rates. *Carnegie-Rochester Conference Series on Public Policy* 39 (December), pp. 95–140.

Cutler, D. M., and Katz, L. F. 1991. Macroeconomic Performance and the Disadvantaged. *Brookings Papers on Economic Activity* (2), pp. 1–74.

Debelle, G., and Laxton, D. 1997. Is the Phillips Curve Really a Curve? Some Evidence for Canada, the United Kingdom, and the United States. *IMF Staff Papers* 44 (June), pp. 249–82.

Deininger, K., and Squire, L. 1996. A New Data Set Measuring Income Inequality. *World Bank Economic Review* 10 (September), pp. 565–91.

Fischer, S. 1993. The Role of Macroeconomic Factors in Growth. *Journal of Monetary Economics* 32 (December), pp. 485–512.

Guitián, M. 1998. Monetary Policy: Equity Issues in IMF Policy Advice. In Vito Tanzi and Ke-young Chu, eds., *Income Distribution and High-Quality Growth.* Cambridge: MIT Press, pp. 333–50.

Li, H., Squire, L. and Zou, H. 1998. Explaining International and Intertemporal Variations in Income Inequality. *Economic Journal* 108 (January), pp. 26–43.

Metcalf, C. E. 1969. The Size Distribution of Personal Income during the Business Cycle. *American Economic Review* 59 (September), pp. 657–68.

Perl, L. J., and Solnick, L. M. 1971. A Note on 'Trickling Down'. *Quarterly Journal of Economics* 85 (February), pp. 171–8.

Sachs, J. D. 1989. Social Conflict and Populist Policies in Latin America. National Bureau of Economic Research Working Paper No. 2897 (March).

Shiller, R. J. 1997. Public Resistance to Indexation: A Puzzle. *Brookings Papers on Economic Activity* (1), pp. 159–228.

Summers, R., and Heston, A. 1991. The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988. *Quarterly Journal of Economics* 106 (May), pp. 327–68.

U.S. House of Representatives, Committee on Ways and Means. 1996. *1996 Green Book: Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means.* Washington, D.C.: Government Printing Office.

73

What Are the Lags in Monetary Policy?

Glenn Rudebusch

Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

The economist's adage about monetary policy actions is that they affect the economy with "long and variable" lags. This chapter considers some empirical evidence on how long and how variable those lags may have been in the past. To the extent that past effects will be similar to those today, such estimates may shed light on the magnitude and timing of the effects of the tightening of monetary policy last year.

This chapter reviews empirical estimates from four models that are based on particular theoretical structures of the economy; in addition, I examine estimates from a purely statistical model. All of these estimates use movements in short-term interest rates to identify the stance of monetary policy. Such an identification is now widely accepted (for example, see any of the references cited below), in part, because recent financial innovation has reduced the reliability of traditional measures of the money supply to indicate policy actions.

Overall, all the models appear to provide fairly consistent evidence that a monetary tightening or loosening has the greatest effect on the growth of output during the first eight quarters, and that this effect is fairly evenly distributed between the first and second years. Nevertheless, it should be stressed that there is substantial uncertainty associated with the models' estimates of the dynamic response of output to any particular monetary policy episode.

STRUCTURAL MODEL ESTIMATES

An estimate of the effect of a monetary policy action on output can be easily obtained from a structural macroeconometric model. Simply measure the difference between two dynamic simulations of the model—one with and one without the policy action. This section examines the responses of four different structural models: (1) the MPS model, which is maintained by the staff of the Federal Reserve Board (see Mauskopf 1990); (2) the DRI model, which is a commercially available product (see Probyn and Cole 1994); (3) the

Reprinted from: *FRBSF Weekly Letter,* No. 95-05 (Feb. 3, 1995).

FAIR model, which is maintained at Yale University by Ray Fair (see Fair 1994); and (4) an FRBSF model, which was developed at this Bank.

These four models vary considerably in size, ranging from about 30 stochastic equations in the FRBSF and FAIR models to almost 1,000 equations in the DRI model. However, all of the models specify the structure of the economy in fairly similar traditional Keynesian terms. In each model, the short-term interest rates most closely associated with monetary policy actions are linked to long-term interest rates via a backward-looking term structure equation. Thus, changes in short rates are assumed to change expectations about future short rates only gradually; hence, short rate changes become embedded with a lag in long rate. These changes in long and short nominal interest rates generally imply changes in real rates as well, because wage and price expectations are assumed to adjust only sluggishly. These interest rate movements also affect household financial wealth in the models because of an arbitrage among bond, equity, and other asset prices. In addition, for two of the models—MPS and DRI—the foreign exchange rate also responds to interest rate movements.

In the models, these financial market repercussions of monetary policy actions affect all categories of real economic activity. Typically, inventory investment is linked most closely to short rates, business fixed investment and residential construction are linked to long rates, household spending on durable goods depends on financial wealth as well as interest rates, and net exports are tied to the exchange rate. However, the response of production and real spending to the changes in the financial environment is modeled with a lag that reflects, for example, the delay from new spending plans to new orders and contracts and finally to new construction and production. The length of this lag varies with the category of spending but is typically at least several quarters in duration.

The first four lines of the table display the lagged monetary policy responses of output in the structural models. The fifth line gives the average response. Each line shows the estimated effect on the four-quarter growth rates of output during each of the first three years after a 1 percentage point increase in short-term interest rates.

Overall, the responses of the models to a monetary shock are fairly similar during the first eight quarters. Averaging across the models, the 1 percentage point increase in the short rate slows output growth by about four-tenths of a percentage point in each of the succeeding two years. Thus, two years after a tightening, the level of real GDP is about three-quarters of a percent lower than it would have been otherwise. The average effect of a monetary shock on the growth of output for these models is greatest during the first eight quarters. However, the models do differ significantly in describing the effect of a monetary shock past two years, with the MPS model showing a further significant reduction in the growth of output and the other models showing little further response.

NONSTRUCTURAL MODEL ESTIMATES

In the last decade, most empirical research on the dynamic response of output to monetary policy has eschewed structural models. Instead, researchers have investigated the effects of policy using vector autoregressions (VARs), which are small, purely statistical models. These VARs typically contain no more than half a dozen variables and are constructed simply by regressing each variable, in turn, on all of the other variables. The VARs contain essentially no theoretical economic structure, and, in particular, they take no stand on the nature of a monetary transmission mechanism. However, VARs are particularly adept at

Table 1 The Effect on GDP Growth of a 1 Percentage Point Increase in Short-Term Interest Rates (In Percentage Points)

	1st year	2nd Year	3rd year
Structural Models			
MPS	–.20	–.70	–1.10
DRI	–.47	–.53	–.13
FAIR	–.24	–.25	+.03
FRBSF	–.55	–.19	+.04
Average	–.37	–.42	–.29
Nonstructural Model			
VAR	–.64	–.26	+.08

summarizing the dynamic correlations found in the data; hence, they can provide useful information on the response of output to movements in interest rates. (For an introduction and references to the literature, see Balke and Emery 1994.)

The bottom row in the table gives the response of output to a tightening of monetary policy as estimated from a VAR. (The VAR used is fairly typical of those in the literature and contains four variables: real GDP, the GDP deflator, a commodity price index, and the federal funds rate.) A 1 percentage point positive shock to short-term interest rates shaves about six-tenths of a percentage point off the growth of GDP in the first year after the shock and another quarter of a percentage point in the second year. The VAR's overall response is broadly in line with the estimates from the structural models; however, in the VAR, the bulk of the impact on output growth occurs a bit earlier than for the structural models.

UNCERTAINTY ABOUT THE ESTIMATES

The imprecision associated with the above numerical estimates should not be underestimated. Uncertainty about the effect of policy arises because the specification of the model is not known to be correct. At the very least, there is uncertainty about the appropriate variables to be included in the model as well as the appropriate values of the parameters in the equations of the model.

For example, to appreciate the degree of uncertainty surrounding the estimates in the table, we can begin by considering the uncertainty in the models' parameter estimates. Taking into account such model uncertainty for the VAR, a 90 percent confidence interval for the effect of the policy shock on output growth during the first four quarters is about half a percentage point in size and ranges from –0.9 to –0.4 of a percentage point. In the second and third years after the policy action, the ranges are even larger—about six-tenths of a percentage point in size. Similar confidence intervals incorporating parameter uncertainty are harder to obtain for the structural models' estimates. However, based on a related investigation in Fair (1994), it appears likely that such estimates also are plagued by about the same amount of imprecision from parameter uncertainty.

Assessing the effects of other types of model misspecification is more difficult. For the structural models, some of this uncertainty can be gauged from the range of estimates presented in the table. For example, in the second year after the increase in interest rates, the four models give a range of estimates of the effect on output growth of about one-half of a percentage point in size. Given that these models share a similar intellectual framework, this range should be viewed as just a starting point regarding the effect of structural uncertainty.

For the VAR, one area of structural uncertainty involves the number of lags of variables to be included in the equations. The VAR in the analysis above used six lags. Letting the number of lags vary from four to eight, the 90 percent confidence interval for the second-year effect on growth is boosted in size to about eight-tenths of a percentage point (including parameter uncertainty), while the first- and third-year confidence intervals are little affected. However, some modest experimentation suggests that all of the confidence intervals would be enlarged by varying the variables included or the estimation sample period.

THE LAGS IN RECENT POLICY

The level of short-term interest rates rose by about 2½ percentage points during 1994 as a result of a monetary tightening. This increase was fairly evenly distributed over the year. Can the above results aid in assessing the timing of the effects of these policy actions? Taken at face value and ignoring their imprecision, the model estimates suggest that as a rough average approximation, over half of the total effect of the increase in rates will be felt in 1995, with the residual impact fairly evenly distributed in 1994 and 1996. However, several important caveats are in order.

For the structural models, the applicability of the results depends crucially on the validity of the models' specifications. However, these models may not be accurate guides to judging the timing of the effects of the most recent policy tightening. Perhaps the most significant departure of recent history from the models is in bond yields. The term structure equations in the structural models, which imply that long-term rates react with a lag to changes in the short rate, are an important element in the models' lagged output responses. In contrast, during the past year, the actual increases in short rates have been matched contemporaneously and, at times, even have been anticipated by equivalent increases in long rates. This deviation from the models suggests caution in relying on the structural model estimates.

For the VAR, there is a more subtle caveat. The VAR output effects shown in the table are in response to a positive *innovation* in interest rates, which is defined as an (exogenous) change in rates that is not in response to changes in any other variable. An endogenous change in interest rates, that is, one which is typical of past changes and is predictable, might well be followed by an output path that was much different from the one given in the table. Thus, any assessment of the effect of recent policy on output with a VAR requires a decomposition of recent interest rate increases into exogenous and endogenous changes. Such decompositions are model-dependent and are likely to be contentious.

REFERENCES

Balke, N. S., and Emery, K. M. 1994. Understanding the Price Puzzle. *Economic Review,* Federal Reserve Bank of Dallas (fourth quarter) pp. 15–26.

Fair, R. 1994. *Testing Macroeconometric Models.* Cambridge, Mass.: Harvard University Press.

Mauskopf, E. 1990. The Transmission Channels of Monetary Policy: How Have They Changed? *Federal Reserve Bulletin* (December) pp. 985–1008.

Probyn, C., and Cole, D. 1994. The New Macroeconomic Model. *DRI/McGraw–Hill U.S. Review* (July) pp. 39–49.

74

Monetary Policy and the Long Boom

John B. Taylor
Professor of Economics, Stanford University, Stanford, California

I regret that I never had the opportunity to work or study with Homer Jones. But I know people who worked and studied with him, and I have enjoyed talking with them and reading about their recollections of Homer Jones. What is most striking to me, of all that has been said and written about Homer Jones, is his incessant striving to learn more about economics and his use of rigorous economic research to improve the practical operation of economic policy. As a college student at Rutgers, Milton Friedman studied under Homer Jones. Friedman credits Jones as an essential influence on his own decision to study economics, and I want to begin this lecture with a quote from Friedman (1976, p. 436) describing certain features of Homer Jones character:

> The hallmark of his contribution is throughout those same traits that exerted so great an influence on me in my teens: complete intellectual honesty; insistence on rigor of analysis; concern with facts; a drive for practical relevance; and, finally, a perpetual questioning and reexamination of conventional wisdom.

I am going to return to these character traits of Homer Jones later in this lecture for they are part of the story I want to tell.

DEFINING "THE LONG BOOM"

I must begin by explaining what I mean by the term "The Long Boom" in the title of this lecture. This month (April 1998) the United States economy celebrates seven years of economic expansion. By definition an economic expansion is the period between recessions; that is, a period of continued growth without a recession. The last recession in the United States ended in April 1991, so as of this April we have had seven years of expansion and we are still going. This current expansion is a record breaker: to be exact it is the second longest peacetime expansion in American history.

But what is more unusual is that this current expansion was preceded by the first longest peacetime expansion in American history. That expansion began in November

Reprinted from: *Federal Reserve Bank of St. Louis Review* (November/December 1998) 3–11. This chapter is a reprint of The Homer Jones Lecture delivered at the Federal Reserve Bank of St. Louis on April 16, 1998. Kent Koch provided research assistance.

1982 and continued through August 1990. It lasted seven years and eight months. Although the 1980s expansion was the first longest peacetime expansion in American history, the current expansion may very well continue long enough to become the first longest peacetime expansion. Either way, we are now experiencing back-to-back the first and second longest peacetime expansions in American history.

There is something even more extraordinary. The recession that occurred between these two record-breaking expansions was—at least for the economy as a whole—short-lived and relatively mild. Hence, during the last 15 years not only did we have the two longest peacetime expansions in American history, but the sole recession we had during these 15 years was remarkably short and mild. This *15-year period of unprecedented stability* and virtually uninterrupted growth is what I refer to as the The Long Boom.

The Long Boom has another characteristic which I briefly mention now and discuss in more detail later, for it, too, is a part of the story I want to tell. The inflation rate has been very low and very stable during The Long Boom, much lower and much more stable than during the years immediately prior to The Long Boom.

Consider this long boom period in comparison with other periods in American history. First, go back to the 15-year period before The Long Boom—the late 1960s and 1970s. In the same span of time we had four recessions, not one. The economy was much more unstable with many ups and downs. We also had the longest inflation in American history—a very unstable and high inflation—a remarkably different experience compared to the last 15 years. The economy was not performing well.

For another historical comparison, go back exactly 100 years, to the 1890s. Compare the 1890s with the 1990s. If the 1990s were like the 1890s, we would have had a recession in 1990 or 1991 that was bigger, quite a bit bigger, than the one we actually had. And that would not have been the end of it. We would have had another recession in 1993, a big one, right when President Clinton was taking office. In 1996 we would have had yet another recession. We would have just been coming out of that recession now. So you can see how dramatically different the economy has been during The Long Boom. Times have changed.

Figure 1 helps visualize this. The upper part of Figure 1 shows real gross domestic product (GDP) in the United States from 1955 to the present. You can see that the economy is growing. You can also see the ups and downs: the 1981–82 recession and recovery; and the 1990–91 recession and recovery. The trend line in Figure 1 shows where the economy is going in the longer term. The lower part of Figure 1 nicely illustrates the large change in economic stability. It's like a microscope that focuses on the fluctuations in real GDP around trend GDP. It shows the GDP gap, which is defined as the percentage difference of real GDP from trend GDP. On the right is where we have been recently. On the left is where we were before. You can see clearly in Figure 1 that the ups and downs—recession, expansion, recession, expansion—are much milder and much less volatile in the latter period than in the earlier period. During The Long Boom, there is obviously greater stability. This greater stability is one reason why the stock market has boomed during the same period, 1982 to the present.

EXPLAINING THE LONG BOOM

What are the underlying reasons for this remarkable and unprecedented period of economic performance in America? Many explanations have been offered. Some have to do with inevitable changes in the structure of the economy that have had the fortunate by-product of

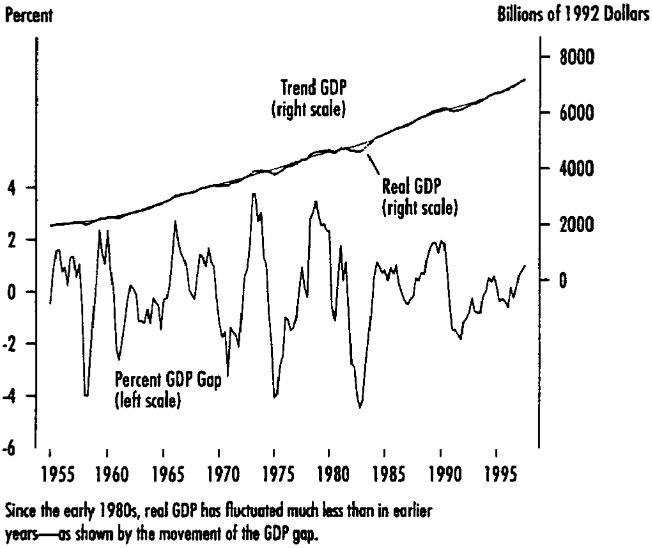


Figure 1 A more stable economy.

a more stable economy. Other explanations are related to economic policy—deliberate decisions of economic policymakers to change policy. So which is it? Good fortune or good economic policy? Let us consider each in turn.

Good Fortune?

Many have noted that the U.S. economy is much more service-oriented now than it was in the past. Services—educational services, legal services, financial services—are generally not as cyclical as manufactured goods, such as automobiles or airplanes. Recessions hit manufacturing sectors quite hard. But in the service industries, business-cycle fluctuations have been typically small. So maybe The Long Boom, with its greater stability, is due to the economy becoming more service-oriented. The problem with this explanation is that the move to a service-oriented economy has been a very gradual change occurring over many decades. It could not explain the sudden shift toward greater economic stability in the early 1980s shown in Figure 1. Hence, a more service-oriented economy is unlikely to be an explanation.

Others have noted better control of inventories. It is true that inventory sales ratios are lower now because inventories are being managed better. The just in time approach to inventory management is now much more common. During most ups and downs in the economy, inventories fluctuate widely. As the economy starts to dip, firms want to cut their inventories; they reduce their orders, and production falls even more rapidly. Thus, better control of inventories may be an explanation for this greater stability that defines The Long Boom. But this explanation also has problems. If you take out the fluctuations of inventories and look at what is left over—final sales—you see virtually the same amount of improvement in economic stability. That is, if I replaced real GDP with final sales of goods and services in the economy in Figure 1, it would look essentially the same.

A very common explanation of The Long Boom is that the U.S. economy has been lucky with respect to the shocks hitting the economy. Recall that in the 1970s we had several large oil shocks. In the 1980s and 1990s we seem to have had fewer oil shocks. But for two reasons I must reject this explanation, too. First, the poor economic performance, along with the higher inflation, that we experienced in the 1970s really got started before the oil shocks began. Second, the U.S. economy had serious shocks in the 1980s and 1990s, including the savings and loan crisis, the oil shock when Iraq invaded Kuwait, and the East Asian crisis.

Now let us move on to economic policy as a possible explanation of The Long Boom. There are, of course, two aspects of economic policy one should focus on: fiscal policy and monetary policy. Let us consider fiscal policy first.

Fiscal Policy

Has fiscal policy seen a major change that could have led to The Long Boom? What about government budget deficits? Budget deficits were huge throughout the 1980s and much of the 1990s. Budget deficits were smaller in the late 1960s and 1970s. Therefore a smaller budget deficit does not seem plausible as an explanation for The Long Boom.

What about the ability of government fiscal policy to respond to recessions by lowering taxes or increasing spending? Has that response gotten larger, quicker, or more efficient? No. In fact, if anything, the ability of the federal government to make discretionary fiscal policy changes to mitigate or offset recessions has diminished. President Bush proposed a small economic stimulus package to be put in place at the end of the 1990–91 recession, but Congress rejected it. In 1993 President Clinton also proposed an economic stimulus package and again the Congress rejected it. Hence, I have to rule out fiscal policy—either smaller budget deficits or better counter-cyclical policy—as a possible explanation of The Long Boom.

Monetary Policy

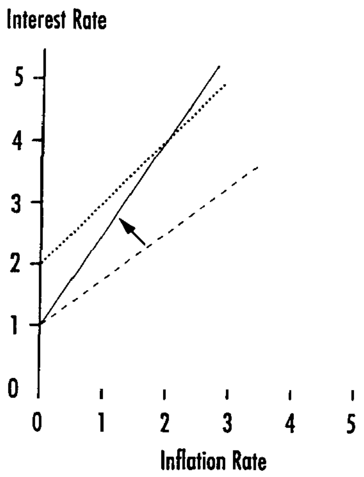
Now consider monetary policy. As is probably already obvious, I will argue that monetary policy is the key factor behind The Long Boom. To do so, I must first discuss briefly the *monetary transmission* process in the United States. Monetary policy is, of course, the responsibility of the policymakers who serve on the Federal Open Market Committee (FOMC). The ultimate tool of monetary policymakers is the money supply. An increase in money supply growth will lead eventually to an increase in inflation. However, the FOMC actually carries out its money supply decisions by focusing on the short-term interest rate—the federal funds rate—that commercial banks charge when they loan funds to each other. The federal funds rate is the instrument of policy that the FOMC members vote on. The federal funds rate, of course, has a big impact on all other interest rates, especially other short-term interest rates, but also on longer-term interest rates. Interest rates have a big effect on the economy. Higher interest rates tend to slow down the economy; lower interest rates tend to stimulate the economy. To be sure, these changes in interest rates are closely related to the money supply, and if the FOMC tried to keep interest rates too low for too long, the money supply would have to increase and this would cause inflation to rise.

When the members of the FOMC make decisions about the interest rate—whether back in the 1970s or today—they look at a number of factors, including the inflation rate and real GDP. Is it possible to see a change in this decision-making process that could explain The Long Boom? Has anything important changed about monetary policy? To explore these questions, consider a numerical example that illustrates how the FOMC might respond to a change in the economy, such as an increase in the inflation rate. First, imagine that the FOMC gets reports that the inflation rate is 1 percentage point higher. Let us suppose that in this circumstance the FOMC decides to raise the federal funds rate by .75 percentage points. The inflation rate is up by 1 percentage point, and the interest rate is up by three-quarters of a percentage point. The FOMC has raised the interest rate in response to inflation. But what happened to the *difference* between the federal funds rate and the inflation rate, a measure of the *real* interest rate? According to this measure, the real interest rate has gone down by a quarter of a percentage point. In other words, the federal funds rate did not go up by enough to raise the real interest rate. It is the real interest rate that affects spending. So by allowing the real interest rate to fall, the FOMC would be doing exactly the opposite of what it should do when the inflation rate rises. The FOMC members would be voting to stimulate the economy just when they should be trying to cool off an inflationary surge. So this policy, with a response of .75 percentage points, is not a good policy. It adds fuel to the inflation fire.

Now consider an alternative monetary policy response. Again, start with the same scenario: The inflation rate rises by 1 percentage point. But now suppose the FOMC, instead of raising the federal funds rate by .75 percentage points, raises it by 1.5 percentage points. In other words, as the inflation rate goes up by 1 percentage point, the interest rate goes up by 1.5 percentage points. The *real* interest rate now goes up by half a percentage point, and that is the right thing for the FOMC to do. The FOMC has removed fuel from the inflation fire because the higher real interest rate is going to reduce demand for automobiles, houses, and other goods. Even though in both cases the FOMC raised the interest rate when inflation rose, in one case the policy was right, and in the other case the policy was wrong.

In fact, these two cases are not hypothetical examples. They turn out to be actual descriptions of monetary policy before and after the start of The Long Boom, respectively. The first case characterizes monetary policy during the late 1960s and 1970s period, while the second case characterizes monetary policy during The Long Boom of the 1980s and 1990s.

By responding in this more aggressive way during The Long Boom, the Federal Reserve (the Fed) has been able to keep the inflation rate lower and much more stable than in the earlier periods. In my view, that change in policy has been the key to keeping the real economy stable. Every recession in the post–World War II economic history of the United States has been preceded by a run-up of inflation. So by keeping the inflation rate low and stable through this policy (taking the fuel off the fire when inflation heats up) the Fed has succeeded in stabilizing the economy, and making recessions less frequent, smaller, and shorter. That has made all the difference, which Figure 2 illustrates. The graph has the interest rate—in particular, the federal funds rate, the variable that the FOMC is making decisions about—on the vertical axis. On the horizontal axis is the inflation rate. The dashed line is the bad policy, where the coefficient is .75. For this policy, when the inflation rate rises by 1 percentage point, the FOMC raises the interest rate by only .75 percentage points. For the solid line, when the inflation rate rises by 1 percentage point, the FOMC raises the interest rate by 1.5 percentage points. For reference, I also show a dotted line in Figure 2 for which the response is exactly one. The real interest rate is constant along that line. By being more responsive than that dotted line, the Fed lets the real interest rate rise when inflation rises. By being less responsive (following the dashed line), the Fed lets the real interest rate fall when inflation rises. These two lines characterize the decisions of the Fed. Although this may sound like an overly simple description of how the Fed makes its deci-



The response of the federal funds rate to changes in inflation is more aggressive (solid line) in The Long Boom than in the prior 15 years (dashed line). The dotted line has a slope of 1 and represents a constant real interest rate.

Figure 2 A change in monetary policy.

sions, it is actually very accurate. Empirical estimates of the Fed's reaction function (regressions of the interest rate on inflation and other variables) show that the dashed line corresponds to the late 1960s and 1970s period (the great inflation and all those business cycles) and the solid line refers to the more peaceful economic times associated with The Long Boom.

What are the implications of this assessment of policy for the future? Put as simply as possible, the Fed should continue to respond to inflation according to a 1.5 response coefficient. If it continues doing that, it will be able to keep the economy stable, making future long booms more common, not avoiding recessions completely, of course, but making recessions smaller and less frequent. Focusing on keeping the inflation rate low and stable and responding aggressively with interest rates is the most important thing the Fed can do to keep the economy stable.

WHY THE CHANGE IN MONETARY POLICY?

Now, what has caused this shift in monetary policy? Did economic research play a role? To answer these questions we must look at some of the history of the Federal Reserve. Go back to the 1950s. The 1951 Accord between the Federal Reserve and the Treasury released the Fed from the job of assisting the Treasury borrowing by keeping interest rates low, as it had done during World War II. But after the Accord, the Fed actually had to decide what to do with the interest rate. One widely discussed suggestion was to lean against the wind, raising the interest rate when the economy grew more rapidly or inflation started to pick up. Leaning against the wind seemed to have the directions of interest rate adjustments right, but it had nothing to say about the size of the adjustments. What is the wind? How do you measure it? What do you lean with and by how much? There were many important, but unanswered, questions.

Meigs (1976, p. 440) described the situation in the 1950s as follows, ". . . The Manager [of the open market account] generally tried to keep free or net borrowed reserves . . . at a level, he thought would satisfy FOMC members' desires for a little more, or a little less, or about the same 'degree of restraint.' " But the degree of restraint was not quantitatively defined and the impact of changes in the degree of restraint was uncertain. As Meigs (1976) put it, "We were as uncertain about how monetary policy worked as were our colleagues at the other Reserve Banks and the Board." A similar accounting comes from Board member Sherman Maisel (1973, p. 77), who admitted "that after being on the Board for eight months and attending twelve open market meetings, I began to realize how far I was from understanding the theory the Fed used to make monetary policy . . . Nowhere did I find an account of how monetary policy was made or how it operated."

This was the situation when Homer Jones arrived at the research department of the Federal Reserve Bank of St. Louis in 1958. He and others were uncomfortable with this vagueness and uncertainty about the operations of monetary policy, and he tried to make the FOMC decisions more specific. The vagueness was why, according to Meigs (1976), Homer Jones "undertook the extraordinary program of monetary research to which all of us are indebted today."

The research program undertaken by Homer Jones helped change this situation in several ways. First, the research improved the money supply statistics. This allowed policymakers to see how money supply growth targeting would work and to measure the movement in interest rates that would accompany money growth targets. The St. Louis Fed Model of the U.S. economy provided an analytical structure through which different monetary policy procedures could be explored. The research was also essential in helping policymakers distinguish between real interest rates and nominal interest rates. The neglect of this distinction is at the heart of the poor monetary policy performance in the late 1960s and 1970s when the interest rate reaction was too small. One can easily imagine how useful Homer Jones' character traits (quoted from Milton Friedman at the start of this lecture) were for stimulating and motivating others to carry out this important research agenda.

In my view, the research directed by Homer Jones was an essential part of a gradual process through which the Fed learned more about the conduct of monetary policy. Of course, others participated in this process—the staff at the Federal Reserve Board and other District banks as well as academic economists. In my view, the result of this gradual learning process was a recognition by the 1980s that changes in interest rates had to be larger and quicker if they were to keep inflation and the overall economy stable. Focusing on the monetary aggregates—especially during the disinflation period of the late 1970s and early 1980s when interest rates had to rise by a very large amount—and emphasizing the distinction between the real and nominal interest rate were part of the means towards this end. Of course, there were other factors that led to the change in monetary policy. The Fed learned from the great inflation experience of the 1970s and discovered through increasing evidence that there was no long-run Phillips curve tradeoff.

A PROBLEM WITH THE LONG BOOM

So far in this lecture I have emphasized the many good features of The Long Boom and the role of economic policy in helping to bring them about. Now let me move on to some not-so-good features and the role that economic policy might have in alleviating them. These policies are what we should focus on in the future.

The main problem with The Long Boom is that productivity growth is much lower than during past periods of U.S. history. By definition, *productivity* is the amount that workers produce on average during a given time at work. Labor productivity *growth* means that for the same number of hours of work, a worker can produce more. Productivity growth is the means through which people improve their living standards, because more production per worker means that workers can earn more. Productivity growth is why the standard of living is now so much higher than it was in the days before the industrial revolution.

Productivity growth in the 1950s and 1960s was about 2 percent per year. During the period of The Long Boom, productivity has been about 1 percent per year. Steady 2 percent per year productivity growth means that the average worker in America can produce 2 percent more next year compared to this year for the same amount of time on the job, and 2 percent more the next year and so on. The 2 percent accumulates and compounds. Unfortunately, the data indicate that productivity growth is only half as much as it was in the 1950s and 1960s.

Now, to be sure, there are some signs that productivity growth has increased recently. In fact, a recent buzzword in the financial press is the "New Economy." The New Economy is characterized by higher productivity growth. Others argue that there is a problem in measuring productivity, especially in a computer age, and that productivity is actually higher than we think. Still others have argued that an even bigger productivity spurt is about to happen: We have all this great technology—computers, biotech, and telecommunications—ready to be used in the workplace to make people more productive. But, in my view, it is too soon to conclude that we are now in a period of persistently higher productivity growth. If productivity has not, or does not, pick up, then improvements in living standards are going to be much less than in most of U.S. history.

If we could get productivity growth up by 1 percent per year, from 1 percent to 2 percent per year, it would make a huge difference for the future. As I noted, that would take us to where productivity growth was in the 1950s and 1960s. Such a growth rate would remove many problems facing us in the future: Social security would no longer be the problem it is today, the income distribution would improve, and poverty rates would decline. That is what happened during the 1950s and early 1960s when productivity growth was very strong. The income distribution narrowed, because the productivity gains spread across the whole population.

A GOAL

I propose that we adopt a goal. The goal is much easier to write down than to carry out. I write in bold-faced characters: +1%. The goal is to raise productivity growth by 1 percentage point per year, so that productivity growth averages 2 percent per year for the 21st century. In other words, we would see 2 percent per year productivity growth rather than 1 percent per year productivity growth for the next 100 years.

Now, I will be the first to admit that this is a goal that is very difficult to achieve. But I think goals are useful for bringing attention to a problem, for focusing policymakers efforts, and simply for getting things done. Even if we got halfway to that +1% goal, it would make a tremendous difference. The U.S. economy reached that 2 percent goal in the past, so it may indeed be achievable.

What can we do to help achieve the +1% goal? By way of offering some examples, I would like to mention four economic policy measures that—taken together—would achieve the goal. But the important point is that setting such a goal would improve the debate about policy alternatives that have the best chance of achieving the goal.

Let us consider budget policy first. The debate about the federal budget has changed remarkably in the last two years, from debate about how to end the deficit to a debate about how to use the surplus. Simply running a budget surplus would help achieve the productivity growth goal. Why? Because by running a budget surplus, the federal government can add saving to the economy rather than subtract saving from the economy. More saving means more investment, raising capital, and increasing productivity. However, it is nearly impossible for politicians to run a budget surplus for any length of time. Every time we have seen projections of surpluses in the past, we also have seen proposals for spending the surpluses. And, of course, spending the surplus means we do not have a surplus.

Thus, a better approach is to establish a fiscal policy that will increase economic growth that does not require running a surplus. One idea is a recent proposal to partly privatize social security. It would use the projected surplus to allow people to put funds equal to a fraction of their payroll tax into a private savings account. Those funds would then be part of national saving and would increase investment, and thereby increase productivity growth. Other people might have other ideas; with the +1% goal in mind, those alternatives can be discussed and debated.

Second, consider tax reform. Several proposals for tax reform would increase productivity growth. A flat tax, by exempting investment from taxation, would stimulate investment. Replacing the income tax with a retail sales tax would increase the tax on consumption relative to saving, encouraging people to do more saving. Permitting people to save in more tax exempt accounts—extensions of the educational savings accounts and IRAs—would increase saving. Again there are alternatives, but with +1% goal, we would have to adopt one of them.

Next, consider education reform. I think this is the most important thing we can do for productivity growth, though the payoff will not start right away. Education reform would address the problem of workers not having enough skills or enough training to make use of new technology in the workplace. Education reform is controversial. I am in favor of greater choice of schools through vouchers, but again the point here is to discuss, debate, and adopt some reform that will help achieve the +1% goal.

Finally, consider regulatory policy. The United States has had deregulation of entry and price in a number of industries including trucking, airlines, and telecommunications, and that has been good for the economy. But there is another kind of regulation called social regulation—which includes environmental regulation, work and safety regulation—that has not been reduced. These types of regulations affect everyone, not only business firms. At my university, for example, one can see many examples where regulations, not just from the federal government but also from the state and local governments, interfere with production—in our case the production of research and teaching. To reduce this interference and stimulate productivity growth, we should have a regulatory policy that applies an effective cost/benefit criterion. If there is going to be a regulation imposed on a university, on a private business, on any other organization, it should pass the criterion that any other good policy would pass—that the benefits outweigh the costs.

CONCLUSION

In summary, in this chapter I have pointed out both the good and the bad of this period that I have called The Long Boom. As I see it, monetary policy deserves much of the credit for what is good—the unprecedented degree of economic stability. Of course, there will be recessions in the future, but if the Fed can maintain the kind of monetary policy rule we have had during The Long Boom and if the same type of monetary policy can be used by other central banks in the world, then more long booms will occur in the future.

Even with such stability, however, if the United States is not successful in developing and implementing economic policies to raise productivity growth, then the 21st century will not be one of much progress and the current problems of social security and income distribution will get worse. If the United States is in a period of a New Economy, as some argue, then it only just arrived and we should adopt policies to maintain it. If the United States is not yet in a period of a New Economy, which is more likely in my view, then we should adopt policies to establish it. That is the purpose behind the +1% goal for productivity growth that I suggested.

I am sure rigorous economic research of the kind Homer Jones advocated, directed, promoted, and carried out will be essential to developing and adopting policies to raise productivity growth and achieve such a goal. We need a *new* Homer Jones to help us find policies for economic *growth* just as we were lucky to have had the *original* Homer Jones to help us find policies for economic *stability*.

REFERENCES

Friedman, M. Homer Jones, A Personal Reminiscence. *Journal of Monetary Economics* (November 1976), pp. 433–36.

Maisel, S. J. *Managing the Dollar,* W. W. Norton & Company (1973).

Meigs, A. J. Campaigning for Monetary Reform: The Federal Reserve Bank of St. Louis in 1959 and 1960. *Journal of Monetary Economics* (November 1976), pp. 439–53.

75

Interest Rates and Monetary Policy

Glenn Rudebusch
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

In the postwar period, the ultimate objectives of the Federal Reserve—namely full employment and stable prices—have remained unchanged; however, the Fed has modified its operational and intermediate objectives for monetary policy several times in response to changes in the economic environment. For example, in 1970, the Federal Reserve formally adopted monetary targets in an attempt to use an intermediate nominal objective or anchor to resist slowly rising inflation. Furthermore, from 1979 through the early 1980s, a narrow monetary reserve aggregate was ostensibly used as the operational instrument of policy. This period, however, was the high-water mark for money.

Over the past 15 years, the Federal Reserve and many other central banks have increasingly relied on interest rates, to the almost complete exclusion of monetary or reserve aggregates, both as sources of information for determining policy and as operating instruments for conducting policy. For example, when announcing its policy action on March 25, 1997, the Federal Open Market Committee stated that it had "decided today to tighten money market conditions slightly, expecting the federal funds rate to rise $\frac{1}{4}$ percentage point to around $5\frac{1}{2}$ percent." This explicit characterization by the FOMC of a monetary policy action in terms of a change in the overnight federal funds rate is just one signal of the current preeminence of interest rates in the conduct of monetary policy. This latest shift in the conduct of policy from money to interest rates has been spurred by two developments: first, the breakdown of traditional relationships between money and economic activity largely brought on by innovations in payment and transactions technologies; second, the increasing sophistication of financial markets and central banks regarding information about the future as embedded in financial instruments (including, for example, the emergence of derivatives and inflation-indexed debt).

One key aspect of interest rates that has become particularly important for the operation of monetary policy is the term structure relationship of short- and long-term rates. This Chapter reviews some of the issues involved in answering two crucial questions for central banks: (1) How should information in the term structure be interpreted and used for conducting monetary policy? and (2) How will central bank actions, especially those expressed as changes in a short-term interest rate, affect the term structure of interest rates and, in turn, the rest of the economy?

Reprinted from: *FRBSF Economic Letter,* No. 97-18 (June 13, 1997).

INTERPRETING THE TERM STRUCTURE

One way in which interest rates appear to be playing a larger role in monetary policy is as informational indicators. For example, current expectations about future inflation may help determine how the economy will perform in later years. Therefore, central banks are interested in obtaining information about current expectations from forward-looking financial markets in order to help predict future paths for inflation and output.

In obtaining such information from financial markets, central banks have relied on the "Expectations Theory" of the term structure. This theory states that longer-term interest rates are set according to market expectations of future shorter-term rates; specifically, rates will be set so that a representative investor is indifferent between holding a long-term bond or a sequence of short-term bonds covering the same length of time. For example, as a first approximation, the current two-month interest rate should equal the average of the current one-month rate and the market's expectation of the one-month rate that will prevail one month from now—the so-called one-month forward rate.

The short end of the term structure, say maturities of less than six months, is one area of particular interest for central banks. At this horizon, according to the expectations theory, interest rates primarily reflect market expectations about very near-term monetary policy settings of the overnight rate (as described in Rudebusch 1995a, b). Central banks are interested in forward rates at this short horizon in part to understand market expectations of the immediate path of the policy rate. Given such expectations, central banks can evaluate whether their near-term policy intentions are being appropriately communicated to markets. In the U.S., the market for federal funds futures, which has traded only since 1988, provides particularly clear readings on forward policy rates over the next few months (see Rudebusch 1996).

Besides obtaining near-term *interest rate expectations,* central banks also are interested in the term structure at the five- to ten-year horizon in order to get an indication of the market's *inflation expectations.* According to common wisdom, the nominal yield on a bond equals, to a first approximation, the real yield plus the average expected inflation rate (the so-called Fisher equation). Assuming that changes in real interest rates are known (or can be ignored), then changes in nominal rates can be translated into changes in inflation expectations. Central banks are keenly aware of the importance of such inflation expectations both as inputs to forecasts of future inflation and economic activity and as measures of the credibility of the central bank's current stance of monetary policy in achieving the long-run goal of price stability. Goodfriend (1993), for example, argues that inflation expectations obtained from the term structure have had a major influence on the conduct of monetary policy by the Federal Reserve.

Still, it should be stressed that interpreting the term structure is not without some ambiguity, in part, because the application of the expectations theory to obtain interest rate expectations from the term structure is not always straightforward. For example, an investor considering the choice between a long-term bond and a sequence of short-term bonds may demand a premium in the latter case for facing the interest rate uncertainty involved in the period-by-period rollover of debt. Thus, in general, the two-month rate equals the average of the current and future one-month rates plus a (possibly negative) term premium. An unobservable term premium that varies over time certainly hinders the process of interpreting the term structure.

Although the evidence is not unambiguous (see, for example, Rudebusch 1995b and Campbell 1995), it appears that a time-varying term premium is not too severe a problem for obtaining interest rate expectations at short horizons—especially with high-frequency (say, daily) data—which are often the focus of particular interest to central banks. However, a time-varying term premium is more likely to be an important consideration at the long maturities used to obtain inflation expectations. Furthermore, movements in real interest rates at long horizons may be unclear, so that the translation of nominal forward rates to inflation expectations may be especially uncertain. There is, however, one recent development that may help alleviate this second problem. The U.S. Treasury has started to issue inflation-indexed debt, which should help pin down movements in the real interest rate. Indeed, the Bank of England has used indexed debt, which has been issued in Great Britain for over a decade, to obtain estimates of real rates and inflation expectations. As described by Deacon and Derry (1994), the Bank of England has found that the difference between the nominal and real term structure provides a useful measure of inflation expectations.

AFFECTING THE TERM STRUCTURE

Besides interpreting the term structure of interest rates, central banks also may be interested in altering it through shifts in monetary policy. In the common textbook description of the transmission of monetary policy, as encapsulated for example in the so-called IS–LM model, the supply of money plays an important role. The equilibrium of money supply by the central bank and money demand by the public (the LM curve) provides an interest rate, which in turn helps to determine the demand for output (via the IS curve). Currently, however, many central banks appear uninterested in the quantity of money and instead focus directly on interest rates. For example, the Federal Reserve Board's new large-scale macroeconomic model of the U.S. economy that is designed to aid in understanding the effects of monetary policy contains roughly 300 equations but includes not a single money supply variable (see Brayton, et al. 1997).

Many central banks have simply taken a short-term interest rate as their direct operating instrument. (For example, the popular Taylor (1993) Rule description of Federal Reserve behavior assumes that the stance of monetary policy is well represented by the federal funds rate.) In this case, the monetary transmission mechanism operates from the short-term rate to real spending on goods and services (that is, simply via the IS curve). Of course, none of the important sectors of real spending—housing, investment, or consumption—depends directly on the overnight federal funds rate. Instead, spending depends on longer-term interest rates. In this way, gauging how changes in the short rate induced by the central bank affect the entire term structure of longer-term rates will be a crucial link in understanding the monetary transmission mechanism.

Cook and Hahn (1989) provide some of the earliest information on the effects of central bank actions on the term structure. They searched for the days on which the *Wall Street Journal* reported that the Federal Reserve had changed the federal funds rate. Then, for those days, they correlated the actual changes in longer-term rates with the funds rate changes. They found a substantial correlation that diminished, but never disappeared, as the maturity of the longer-term security was increased. For example, even the yield on a 10-year bond would typically rise 10 to 15 basis points on the day that the funds rate was increased by a percentage point. In a sense then, the federal funds rate, as the instrument of Fed policy, is the tip of the term structure tail that wags the dog of the economy.

Of course, the movements in longer rates following a policy action are not always the same. According to the expectations theory, these movements reflect both the immediate change in the funds rate as well as market expectations about future policy actions, which may vary with the exact circumstances. For example, as described in Campbell (1995), the 10-year rate jumped by almost twice as much as the increase in the funds rate at the time of the Fed tightening in February 1994 instead of the typical muted response described by Cook and Hahn. Such variability in financial market responses is an important source of the uncertainty associated with the real effects of monetary policy actions.

CONCLUSION

In the U.S. and other countries, interest rates are a key feature of the conduct of monetary policy; therefore, central banks are concerned about both how to interpret information from the term structure of interest rates and how their actions affect the term structure. Research suggests that, while short-term forward rates can give fairly accurate readings of interest rate expectations in the short run, longer-term rates give less clear readings of inflation expectations. As for monetary policy's effects on the term structure, although research shows that longer-term rates do tend to react when the fed funds rate moves, the size of this response can vary substantially.

REFERENCES

Brayton, F., Mauskopf, E. Reifschneider, D. Tinsley, P. and Williams, J. 1997. The Role of Expectations in the FRB/US Macroeconomic Model. *Federal Reserve Bulletin* 83 (April) pp. 227–245.

Campbell, J. Y. 1995. Some Lessons from the Yield Curve. *Journal of Economic Perspectives* 9, no. 3, pp. 129–152.

Cook, T., and Hahn, T. 1989. The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s. *Journal of Monetary Economics* 24, pp. 331–351.

Deacon, M., and Derry, A. 1994. Estimating market interest rate and inflation expectations from the prices of UK government bonds. *Bank of England Quarterly Bulletin* (August) pp. 232–240.

Goodfriend, M. 1993. Interest Rate Policy and the Inflation Scare Problem: 1979–1992. Federal Reserve Bank of Richmond *Economic Review* 79, no. 1, pp. 1–24.

Rudebusch, G. D. 1995a. Federal Reserve Policy and the Predictability of Interest Rates. *FRBSF Weekly Letter* (June 23).

Rudebusch, G. D. 1995b. Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure. *Journal of Monetary Economics* (April) pp. 245–274.

Rudebusch, G. D. 1996. Do Measures of Monetary Policy in a VAR Make Sense? Federal Reserve Bank of San Francisco Working Paper 96–05.

Taylor, J. 1993. Discretion vs. Policy Rules in Practice. *Carnegie–Rochester Conference Series on Public Policy* (February), pp. 195–214.

76

Monetary Policy and Long-Term Interest Rates

Yash P. Mehra
Federal Reserve Bank of Richmond, Richmond, Virginia

The standard view of the transmission mechanism of monetary policy assigns a key role to long-term interest rates. According to this view, a monetary policy tightening pushes up both short and long interest rates, leading to less spending by interest-sensitive sectors of the economy and therefore to lower real growth. Conversely, a monetary easing results in lower interest rates that stimulate real growth. An open question in discussions of this view is whether monetary policy has significant empirical effects on long-term interest rates.[1]

In this chapter, I provide new evidence on the quantitative effect of monetary policy on the long-term interest rate. The federal funds rate is used as a measure of monetary policy.[2] The work extends the previous research in two main directions. First, following Goodfriend's (1993) description of funds rate policy actions, it distinguishes empirically between the long- and short-run sources of interaction between the funds rate and the long rate; this distinction is absent in previous studies. Second, the analysis in Goodfriend also suggests that the near-term effects of funds rate policy actions on the long rate may be quite variable. Hence the work examines the temporal stability of such effects, an issue also virtually ignored in previous research.

The empirical work focuses on the behavior of the nominal yield on ten-year U.S. Treasury bonds during the period 1957Q1 to 1995Q2. The results here indicate that the bond rate moves positively with the funds rate in the long run. However, this comovement arises because the bond rate automatically moves with trend inflation (the Fisher relation) and the Federal Reserve (Fed) keeps the level of the funds rate in line with the going trend rate of inflation (the long-run Fed reaction function). Apart from the correlation that occurs through the inflation channel, I find that empirically there is no other source of long-run interaction between the bond rate and the funds rate. This result arises because the bond rate's other component—the expected long real rate—is mean stationary and therefore unrelated to the level of the funds rate in the long run. These results have the policy implication that monetary policy can permanently lower the bond rate only by lowering the trend rate of inflation.

Reprinted from: *Federal Reserve Bank of Richmond Economic Quarterly,* v. 82, No. 3 (Summer 1996) 27–49.

The short-run stance of monetary policy is measured by the spread between the funds rate and the trend rate of inflation (the funds rate spread). The results indicate that movements in the funds rate spread have a statistically significant effect on the bond rate and that the magnitude of its near-term effect on the bond rate has increased significantly since 1979.[3] In the pre-1979 period, the bond rate rose anywhere from 14 to 29 basis points whenever the funds rate spread widened by one percentage point. In the post-1979 period, however, the estimate of its near-term response ranges from 26 to 50 basis points.

The short-run results thus suggest that, ceteris paribus, a monetary policy tightening measured by an increase in the funds rate spread does result in higher bond rates in the short run, in line with the traditional view of the transmission mechanism. However, this increase in the short-run sensitivity of the bond rate to policy actions may itself be due to the way the Fed has conducted its monetary policy since 1979. The Fed's post-1979 efforts to bring down the trend rate of inflation, coupled with lack of credibility, may have amplified the near-term effects of funds rate changes on the bond rate.

The plan of this chapter is as follows. The first section presents the methodology that underlies the empirical work, the second section contains empirical results, and concluding observations are in the third section.

THE MODEL AND THE METHOD

This section describes the methodology that underlies the empirical work in this article.

The Fisher Relation, the Bond Rate, and the Federal Funds Rate

In order to motivate the empirical work, I first describe how monetary policy may affect the bond rate in the short run and the long run.[4] As indicated before, the federal funds rate is used as a measure of the stance of monetary policy. Thus a monetary policy action is defined as a change in the funds rate, and a monetary policy strategy is defined as the reaction function that would lead to policy actions. The Fisher relation for interest rates provides a convenient framework within which effects of policy actions can be described. The Fisher relation is

$$BR_t = rr_t^e + \dot{p}_t^e, \tag{1}$$

where BR_t is the bond rate, rr_t^e is the expected long real rate, and $\dot{p}_t^e$ is the expected long-term inflation rate. Equation (1) relates the bond rate to expectations of inflation and the real rate.

The Fisher relation indicates that policy actions could conceivably affect the bond rate by altering expectations of inflation, the real rate, or both. Since policy actions may not always move the real rate and inflation components in the same direction, the near-term responses of the bond rate to such actions cannot be determined a priori. Much may depend upon the nature of the strategy being pursued by the Fed. Goodfriend (1993) discusses three different strategies that may lie at the source of interaction between the bond rate and the funds rate. Consider, first, pure cyclical strategies in which the Fed routinely raises (lowers) the funds rate in response to cyclical expansions (downturns) without attempting to affect the current trend rate of inflation expected by the public. Under that strategy, a funds rate increase will tend to raise the bond rate by raising current and expected future short real rates (i.e., by raising the rr_t^e component in [1]). This cyclical comovement is short run in nature.

The second strategy discussed by Goodfriend considers the response of the Fed to an exogenous change in the trend rate of inflation. If the trend rate of inflation increases, the bond rate automatically moves with inflation (equation [1]). The Fed may choose to keep the short real rate steady and will therefore move the funds rate in line with the rising or falling inflation rate. In this case the bond rate comoves with the funds rate because the Fed is responding to changing inflation in a neutral fashion. I refer to this source of comovement as long run in nature.

Finally, consider an aggressive strategy that could be taken either to promote real growth or to reduce the going trend rate of inflation. Under that strategy, the net impact of policy actions on the bond rate is complex because they can move the real rate (rr_t^e) and the inflation expectations ($\dot{p}_t^e$) in opposite directions. The real rate effect moves the bond rate in the same direction as the funds rate, while the inflation effect moves the bond rate in the opposite direction. Thus the net effect of policy actions on the long rate cannot be determined a priori.

To illustrate, consider an aggressive increase in the funds rate intended to reduce the trend rate of inflation. Such a tightening can shift both components of the bond rate. If the Fed's disinflation policy is credible, then short rates rise and expected inflation falls. The fall in expected inflation may thus offset somewhat the immediate response of the bond rate to the funds rate. If the decline in expected inflation persists, then the Fed may soon bring down the funds rate consistent with the lower trend rate of inflation. However, if the public does not yet have full confidence in the Fed's disinflation, then the Fed may have to persist with a sufficiently high funds rate until real growth slows and inflation actually declines. In this case, the immediate and near-term effects of the funds rate on the bond rate may be large relative to the previous case. These policy actions generate correlations between the bond rate and the funds rate which are both short and long run in nature.

Empirical Specifications of the Bond Rate Regressions: Short- and Long-Run Effects

The discussion in the previous section suggests the following observations. First, the bond rate may be correlated with current and past values of the funds rate, but the strength and duration of that correlation is a matter for empirical analysis.[5] Second, the correlation between the funds rate and the bond rate induced by pure cyclical and/or aggressive policy actions is likely to be short run, appearing over business cycle periods. In contrast, the correlation induced by the Fed's reaction to shifts in the trend rate of inflation may be long run. A rise in the trend rate of inflation that permanently raises the bond rate will also result in a higher funds rate if the Fed is trying not to induce any cyclical or aggressive element into its policy. Third, other economic factors such as inflation, the deficit, or the state of the economy may also influence the bond rate. Such correlations are apart from the one induced by monetary policy actions.

Given the above-noted considerations, the empirical work here examines the relationship between the bond rate and the funds rate in a multivariable framework. The other economic determinants included in the analysis are the actual inflation rate and the output gap that measures the cyclical state of the economy.[6] The work identifies the short- and long-run sources of correlation between the funds rate and the bond rate using cointegration and error-correction methodology. In particular, I proceed under the assumption, whose validity I do examine, that levels of the empirical measures of the long rate, the inflation rate, the funds rate, and other economic determinants each have unit roots (stochastic trends) and that there may exist cointegrating relationships among these variables. I interpret cointegrating regressions as measuring the long-run equilibrium correlations and error-correction regressions as measuring short-run correlations among variables.

To illustrate, assume that we are examining the interaction between the bond rate and the funds rate in a system that also includes inflation. Assume further that tests for cointegration indicate the presence of the following two cointegrating regressions in the system:

$$BR_t = a_0 + a_1\dot{p}_t + U_{1t}, \quad a_1 = 1, \text{and} \tag{2}$$

$$NFR_t = b_0 + b_1\dot{p}_t + U_{2t}, \quad b_1 = 1, \tag{3}$$

where BR_t is the bond rate, $\dot{p}_t$ is actual inflation, NFR is the nominal funds rate, and U_{1t} and U_{2t} are two stationary disturbances. Equation (2) indicates that the bond rate moves one-for-one with inflation and that the long real rate is mean stationary. Equation (2) indicates that the bond rate moves one-for-one with inflation and that the long real rate is mean stationary. Equation (3) indicates that the funds rate moves one-for-one with inflation and that the real funds rate is mean stationary. These two cointegrating regressions are consistent with the presence of long-run equilibrium correlations between variables. If in the cointegrating regression, say, equation (2), $\dot{p}_t$ is weakly exogenous, then the long-run correlation can be given a causal interpretation, implying that the bond rate is determined by the (trend) rate of inflation.[7] The hypothesis that inflation is weakly exogenous in (2) can be tested by examining whether in regressions (4) and (5)

$$\Delta BR_t = a_2 + \delta_1(BR - \dot{p})_{t-1} + \sum_{s=1}^{n1} a_{3s}\Delta BR_{t-s} + \sum_{s=1}^{n2} a_{4s}\Delta\dot{p}_{t-s} \tag{4}$$

$$\Delta\dot{p}_t = b_2 + \delta_2(BR - \dot{p})_{t-1} + \sum_{s=1}^{n1} b_{3s}\Delta BR_{t-s} + \sum_{s=1}^{n2} b_{4s}\Delta\dot{p}_{t-s}, \tag{5}$$

where $\delta_1 \neq 0$ but $\delta_2 = 0$.[8] That result indicates that it is the bond rate, not inflation, that adjusts in response to deviations in the long-run relationship.

The cointegrating regressions discussed above identify the long-run comovements among variables. In order to estimate the short-run responses of the bond rate to the funds rate, the empirical work uses the following error-correction model of the bond rate:

$$\begin{aligned} \Delta BR_t = d_0 + \lambda_1 U_{1t-1} + \lambda_2 U_{2t-1} + \sum_{s=1}^{n} d_{1s}\Delta BR_{t-s} \\ + \sum_{s=0}^{n} d2s\Delta NFR_{t-s} + \sum_{s=0}^{n} d_3\Delta\dot{P}_{t-s} + \epsilon_t, \end{aligned} \tag{6}$$

where all variables are as defined before and where Δ is the first difference operator. U_{1t-1} and U_{2t-1} are one-period lagged values of the residuals from the cointegrating regressions. If we substitute for U_{1t} and U_{2t} from (2) and (3) into (6), we can rewrite (6) as in (7):

$$\begin{aligned} \Delta BR_t = \tilde{d} + \lambda_1(BR_{t-1} - a_1\dot{p}_{t-1}) + \lambda_2(NFR_{t-1} - b_1\dot{p}_{t-l}) \\ + \sum_{s=1}^{n} d_{1s}\Delta BR_{t-s} + \sum_{s=0}^{n} d_{2s}\Delta NFR_{t-s} + \sum_{s=0}^{n} d_{3s}\Delta\dot{p}_{t-s} + \epsilon_t, \end{aligned} \tag{7}$$

where $\tilde{d} = d_0 - \lambda_1 a_0 - \lambda_2 b_0$. The short-run regression (7) includes levels as well as differences of variables. The empirical effects of changes in the inflation rate and the funds rate on the bond rate may occur through two distinct channels. First, those changes may affect the bond rate directly by altering future expectations of the inflation rate and the real rate of interest. The parameters d_{is}, $i = 2, 3$, $s = 0, n$, measure near-term responses of the bond rate to changes in its economic determinants. But, as noted before, signs and magnitudes of those parameters are a matter for empirical analysis because they depend upon factors such as the strategy of policy actions, the credibility of the Fed, and the nature of persistence in data. Lagged values of changes in the bond rate are included in order to capture better its own short-run dynamics.

The second focuses on disequilibrium in the long-run relations which may be caused by changes in the inflation rate and the funds rate. For example, aggressive funds rate changes taken to affect real growth or inflation may result in the level of the funds rate that is out of line with its value determined by the long-run equilibrium relation ($NFR_t \lesseqgtr b_0 + \dot{p}_t$ in [3]). Such short-run disequilibrium can also occur if the Fed adjusts the funds rate with lags in response to rising or falling inflation. Similarly, even though the bond rate moves automatically with inflation, short-run influences from other economic factors may result in the level of the bond rate that is out of line with its long-run equilibrium value ($R_t \lesseqgtr a_0 + \dot{p}_t$ in [2]). Such transitory perturbations in long-run equilibrium relations may have consequences for short-run changes in the bond rate. The parameters λ_1 and λ_2 in (7) thus measure the responses of the bond rate to such disequilibrium. The expected sign for λ_1 is negative, because the presence of the error-correction mechanism implies that the bond rate should decline (increase) if it is above (below) its long-run equilibrium value. In contrast, the sign of λ_2 is expected to be positive. But note all these disequilibrium effects are short-run (cyclical) in nature because in the long run (defined here in the equilibrium sense) they disappear and the bond rate is at its long-run equilibrium value determined by (2), i.e., $a_0 + \dot{p}_t$.

Data and Estimation Procedures

The empirical work in this article focuses on the behavior of the long rate during the sample period from 1957Q1 to 1995Q2. The long rate is measured by the nominal yield on ten-year U.S. Treasury bonds (*BR*). In most previous studies a distributed lag on the actual inflation rate is used as proxy for the long-run anticipated inflation, and actual inflation is generally measured by the behavior of the consumer price index. I also use actual inflation as proxy for anticipated inflation. I, however, measure inflation as the average of change in the consumer price index, excluding food and energy, over the past three years ($\dot{p}$).[9] The output gap (gaph) is the natural log of real GDP minus the log of potential GDP, which is generated using the Hodrick-Prescott filter. The interest rates are monthly averages for the last month of the quarter.

The stationarity properties of the data are examined using tests for unit root and mean stationarity. The unit root test used is the augmented Dickey-Fuller test and the test for mean stationarity is the one in Kwiatkowski et al. (1992). The test used for cointegration is the one proposed in Johansen and Juselius (1990).[10]

EMPIRICAL FINDINGS

In this section I describe cointegration test results for a system that includes the bond rate (*BR*), the inflation rate ($\dot{p}$), and the nominal funds rate (*NFR*). I also discuss short-run results from error-correction regressions for the full sample period as well as for several subperiods. The section concludes with an explanation of different pre- and post-1979 sample results.

Cointegration Test Results

Test results for unit roots and mean stationarity are summarized in the appendix. They indicate that the bond rate (*BR*), the inflation rate ($\dot{p}$), and the nominal funds rate (*NFR*) each have a unit root and thus contain stochastic trends. The output gap variable by construction is stationary.

Test results for cointegration are also summarized in the appendix. I first focus on the bivariable systems (*BR*, $\dot{p}$), (*NFR*, $\dot{p}$), and (*BR*, *NFR*). Test results are consistent with the presence of cointegration between variables in each system, indicating that the bond rate is cointegrated with the inflation rate and the funds rate. The funds rate is also cointegrated with the inflation rate. Thus the bond rate comoves with each of these nonstationary variables, including the funds rate.

The presence of cointegration between two variables simply means that there exists a long-run stochastic correlation between them. In order to help determine whether such correlation can be given a causal interpretation, Table 1 presents test results for weak exogeneity of

Table 1 Cointegrating Regressions and Test Results for Weak Exogeneity

Equation number	Panel A cointegrating regressions	Panel B error-correction coefficients (t-value) in equations for		
		ΔBR	$\Delta\dot{p}$	ΔNFR
1	$BR_t = 3.3 + 0.93\,\dot{p}_t = U_{1t}$ (10.0)	−0.18 (3.4)	−0.01 (0.6)	
2	$BR_t = 1.2 + 0.91\,NFR_t + U_{2t}$ (50.7)	−0.17 (2.5)		0.21 (1.5)
3	$NFR_t = 1.8 + 1.1\,\dot{p}t + U_{3t}$ (6.2)		0.00 (0.4)	−0.27 (3.2)
4	$(BR - \dot{p})_t = 2.2 + 0.09\,NFR_t$ (1.3) $\chi_1^2 = 1.7$			

Notes: Cointegrating regressions given in panel A above are estimated by the dynamic OLS procedure given in Stock and Watson (1993), using leads and lags of first differences of the relevant right-hand side explanatory variables. Eight leads and lags are included. Parentheses that appear below coefficients in cointegrating regressions contain t-values corrected for the presence of moving average serial correlation. The order of serial correlation was determined by examining the autocorrelation function of the residuals. χ_1^2 is the Chi-square statistic that tests the hypothesis that the coefficient that appears on NFR in equation 4 is zero.

Panel B above contains error-correction coefficients from regressions of the form

$$\Delta X_{1t} = \delta_1 U_{t-1} + \sum_{s=1}^{4} a_s \Delta X_{1t-s} + \sum_{s=1}^{4} b_s \Delta X_{2t-s}$$

$$\Delta X_{2t} = \delta_2 U_{t-1} + \sum_{s=1}^{4} c_s \Delta X_{1t-s} + \sum_{s=1}^{4} d_s \Delta X_{2t-s},$$

where U_{t-1} is the lagged value of the residual from the cointegrating regression that is of the form

$$X_{1t} = d_0 + d_1 X_{2t} + U_t,$$

and where X_{1t} and X_{2t} are the pertinent nonstationary variables. The relevant cointegrating regressions are given in panel A above. Parentheses that appear below error-correction coefficients contain t-values.

the long-run parameters. In the system (BR, $\dot{p}$) inflation is weakly exogenous but the bond rate is not, indicating that it is the bond rate that adjusts in response to deviations in the long-run relationship. Thus the long-run equilibrium relationship between the bond rate and the inflation rate can be interpreted as a Fisher relation in which the bond rate is determined by the (trend) rate of inflation. In the system (NFR, $\dot{p}$) inflation is again weakly exogenous but the funds rate is not. Here again the long-run relation can be interpreted as one in which the inflation rate drives the interest rate: in this case the short-term rate. Hence, I interpret the long-run equilibrium relationship between the funds rate and the inflation rate as a kind of reaction function.

The test results for weak exogeneity discussed above for systems (BR, $\dot{p}$) and (NFR, $\dot{p}$) also imply that inflation causes the comovement between the bond rate and funds rate. The bond rate comoves with the funds rate because the bond rate moves automatically with inflation and the Fed keeps the funds rate in line with the trend rate of inflation in the long run.

The analysis above, based on bivariable systems, thus suggests that in the Fisher relation the funds rate should not be correlated with the bond rate once we control for the correlation that is due to inflation. I test this implication by examining whether the ex post real rate (BR–$\dot{p}$) is correlated with the funds rate. I do so by expanding the Fisher relation to include the funds rate while maintaining the Fisher restriction that the bond rate adjusts one-for-one with inflation. In that regression the funds rate is not significant (see Table 1, equation 4). The Chi-square statistic for the null hypothesis that the ex post real rate is not correlated with the funds rate is small, consistent with no correlation.

The result that the bond rate is cointegrated with the actual inflation rate and thus the ex post real rate (R_t–$\dot{p}_t$) is stationary also implies that the expected long real rate is stationary. This can be seen if we express the Fisher relation (1) as

$$BR_t - \dot{p}_t = rr_t^e + (\dot{p}_t^e - \dot{p}_t) = rr_t^e + U_t,$$

where all variables are defined as before and where U_t is the disturbance term. This disturbance arises because the long-term expected inflation rate may differ from the three-year inflation rate. As is clear, the stationarity of (BR_t–$\dot{p}_t$) implies the stationarity of the expected long real rate.

Error-Correction Regressions

Since the ex ante long real rate is mean stationary, not constant, cyclical and aggressive funds rate changes discussed before may still affect the bond rate by altering expectations of its real rate and inflation components. I now explore those short-run effects by estimating the error-correction equation.

The empirical results discussed in the previous section are consistent with the following two cointegrating relationships:

$$BR_t = a_0 + a_1\dot{p}_t + U_{1t} \quad \text{and} \tag{8}$$

$$NFR_t = b_0 + b_1\dot{p}_t + U_{2t}. \tag{9}$$

Equation (8) is the Fisher relation, and equation (9) the Fed reaction function. The latter captures that component of the funds rate that comoves with trend inflation. The residual U_{2t} is then the component that captures the stance of cyclical and aggressive funds rate policy actions. Consider then the following error-correction equation:

$$\Delta BR_t = d_0 + \lambda_1 U_{1t-1} + \lambda_2 U_{2t-1} + \sum_{s=1}^{n} d_{1s}\Delta BR_{t-s} + \sum_{s=1}^{n} d_{2s}\Delta\dot{p}_{t-s} + \sum_{s=1}^{n} d_{3s}NFR_{t-s} + \sum_{s=1}^{n} d_{4s}gaph_{t-s} + \epsilon_t. \tag{10}$$

The parameter λ_2 in (10) measures the response of the bond rate to the lagged value of the funds rate spread $(NFR - b_0 - b_1\dot{p})$. Since the short-run equation is in first differences, the near-term response of the bond rate to the funds rate spread can be calculated as $-\lambda_2/\lambda_1$.[11] Also, equation (10) includes only lagged values of economic determinants and hence can be estimated by ordinary least squares. In order to examine subsample variability, I estimate equations (8) through (10) over several sample periods, all of which begin in 1961Q2 but end in different years from 1972 through 1995.

Table 2 presents some key coefficients (λ_1, λ_2, $-\lambda_2/\lambda_1$, a_1, b_1) from these regressions. If we focus on full sample results, then it can be seen that all these key coefficients appear with expected signs and are statistically significant. In cointegrating regressions the bond rate adjusts one-for-one with inflation and so does the funds rate. The hypotheses that $a_1 = 1$ and $b_1 = 1$ cannot be rejected and thus are consistent with our priors about the interpretation of (8) as the Fisher relation and of (9) as the Fed reaction function. In the error-correction regression λ_2 is positive and its estimated value indicates a one percentage point rise in the funds rate spread raises the bond rate by 13 basis points in the following

Table 2 Short-Run Error-Correction Regressions Using Residuals

	Panel A: error-correction coefficients				Panel B: coefficients from cointegrating regressions	
Estimation ends in year	$\lambda_2(\tilde{t})$	$\lambda_1(\tilde{t})$	$-\lambda_2/\lambda_1$	Mean lag in quarters	$a_1(\tilde{t})$	$b_1(\tilde{t})$
1995Q2	0.13 (2.7)	-0.31 (4.8)	0.42	3.2	0.93 (10.0)	1.10 (6.0)
1994	0.12 (2.4)	-0.29 (4.5)	0.41	3.4	0.93 (10.1)	1.10 (6.2)
1992	0.14 (2.6)	-0.30 (4.5)	0.46	3.3.	0.92 (9.7)	1.10 (6.5)
1990	0.15 (2.6)	-0.31 (4.3)	0.47	3.3	0.92 (9.8)	1.10 (6.7)
1988	0.14 (2.4)	-0.30 (4.1)	0.48	3.3	0.94 (10.9)	1.00 (6.6)
1986	0.16 (2.7)	-0.34 (2.7)	0.48	2.9	0.94 (10.8)	1.10 (6.9)
1984	0.13 (1.9)	-0.30 (2.6)	0.46	3.3	0.93 (11.1)	1.10 (6.7)
1982	0.19 (2.5)	-0.54 (3.6)	0.36	1.8	0.76 (17.1)^c	0.84 (7.6)
1980	0.18 (2.1)	-0.71 (2.2)	0.26	1.4	0.68 (30.8)^c	0.60 (5.6)^c
1979Q3	0.17 (3.1)	-0.67 (3.3)	0.25	1.5	0.69 (31.1)^c	0.77 (8.9)^c
1978	0.16 (2.8)	-0.64 (3.1)	0.26	1.5	0.71 (28.4)^c	0.78 (9.9)^c
1976	0.17 (3.2)	-0.65 (3.0)	0.26	1.5	0.71 (21.4)^c	1.00 (15.9)
1974	0.23 (2.9)	-0.98 (3.6)	0.23	1.0	0.78 (17.9)^c	0.97 (14.6)
1972	0.11 (1.2)	-0.75 (2.6)	0.14	1.3	0.73 (6.9)^c	0.68 (2.5)

^{c} indicates the relevant coefficient (a_1 or b_1) is significantly different from unity.
Notes: The coefficients reported above are from the following regressions:

$$BR_t = a_0 + a_1\dot{p}_t + U_{1t}, \quad \text{(a)}$$

$$NFR_t = b_0 + b_1\dot{p}_t + U_{2t}, \text{ and} \quad \text{(b)}$$

$$\Delta BR_t = d_0 + \lambda_1 U_{1t-1} + \lambda_2 U_{2t-1} + \sum_{s=1}^{n} d_{1s}\Delta BR_{t-s} + \sum_{s=1}^{2} d_{2s}\Delta\dot{p}_{t-s} + \sum_{s=1}^{2} d_{3s}\Delta NFR_{t-s} + \sum_{s=1}^{2} d_{4s}\, gaph_{t-s} + \varepsilon_t, \quad \text{(c)}$$

where gaph is the output gap and other variables are defined as in Table A1. Equations (a) and (b) are estimated by the dynamic OLS procedure given in Stock and Watson (1993) and equation (c) by ordinary least squares. The estimation period begins in 1961Q2 and ends as shown in the first column. $\tilde{t}$ is the t-statistic. The mean lag is calculated as $-1/\lambda_1$.

period. The net increase totals 42 basis points. The mean lag $(-1/\lambda_1)$ in the short-run effect of the funds rate spread on the bond rate is approximately 3.2 quarters, indicating that these near-term responses dissipate quite rapidly.[12]

If we focus on subsample results, it can be seen that all key coefficients still appear with expected signs and are statistically different from zero. However, there are some major differences between pre- and post-1979 regression estimates. In pre-1979 cointegrating regressions the hypotheses that $a_1 = 1$ and $b_1 = 1$ are generally rejected. In contrast, that is not the case in most post-1979 regressions. The rejection, however, is more common in the Fisher relation than it is in the Fed reaction function.

In pre-1979 error-correction regressions the bond rate does respond to the funds rate spread—the one-period response (λ_2) ranges from 11 to 23 basis points. But the one-period response is generally quite close to the near-term net response $(-\lambda_2/\lambda_1)$, indicating that the effect of policy actions on the bond rate did not persist too long. In post-1979 error-correction regressions, however, the near-term response of the bond rate to the funds rate spread is larger than the one-period lagged response. In particular, the immediate response of the bond rate to the funds rate ranges from 13 to 16 basis points and the near-term response from 36 to 48 basis points. Together these estimates imply that the near-term response of the bond rate to the funds rate spread has increased since 1979.[13] I argue below that these different results may be due in part to the way the Fed has conducted its monetary policy since 1979. In particular, I focus on the Fed's disinflation policy.

Before 1979 the Fed did not aggressively attempt to bring down the trend rate of inflation. In the long-run Fed reaction function estimated over 1961Q2 to 1979Q3, the parameter b_1 is less than unity, indicating that the Fed did not adjust the funds rate one-for-one with inflation (see Table 2). Moreover, the short-run reaction functions estimated in Mehra (1996) also indicate that in the pre-1979 period the Fed did not respond to accelerations in actual inflation. Hence, during this early period a monetary policy tightening measured by a widening in the funds rate spread may have affected the bond rate primarily by altering its expected real rate component. Because the funds rate increase alters only near-term expectations of future short real rates, its impact on the bond rate is likely to be modest, as confirmed by low estimates of λ_2 in Table 2. When the bond rate rises above the current inflation rate, both the bond rate and actual inflation rises, speeding up the adjustment, as confirmed by high estimates of λ_1 in Table 2. As a result, the immediate effect of policy on the bond rate dissipates quickly and hence the near-term effect is close to the immediate impact.

In the post-1979 period, however, the Fed made a serious attempt to bring down the trend rate of inflation and to contain inflationary expectations. The descriptive analysis of monetary policy in Goodfriend (1993) and the short-run reaction function estimated in Mehra (1996) are consistent with this observation. Hence short-run increases in the funds rate spread may also have affected the bond rate by altering the long-term expected inflation. If the Fed's disinflation policy had been credible, then increases in the funds rate spread that raise the bond rate's real component may also lower expectations of the long-term expected inflation rate, thereby offsetting somewhat the immediate or the very-near-term response of the bond rate to the funds rate spread. The evidence reported in Table 2, however, indicates that the estimated coefficient $(-\lambda_2/\lambda_1)$ that measures the near-term response shows no tendency to fall in the post-1979 period. On the other hand, if the public does not have full confidence in the Fed's disinflation policy and if the Fed has to persist with sufficiently high short real rates to reduce the trend rate of inflation or contain inflationary expectations, then the estimated effect of a policy action on the bond rate would last longer. In this case, the near-term effect of the funds rate spread on the bond rate would be larger and the mean lag in the effect of such policy on the bond rate would also be higher. Both these implications are consistent with results in Table 2, where both the estimated short-run effect $(-\lambda_2/\lambda_1)$ and the mean lag $(-1/\lambda_1)$ are higher in the post-1979 period than they were in the period before.

Additional Empirical Results

In this section I discuss and report some additional test results which confirm the robustness of conclusions reached in the previous section. I consider several changes in the specification of the short-run equation (10). First, I reproduce the empirical work in Table 2 under the alternative specification that the short-run stance of monetary policy is measured by the real funds rate. I then consider additional changes in specification to address concerns raised by the potential endogeneity of monetary policy actions, alternative measures of inflation, and the potential stationarity of data. For these latter experiments I focus on results that pertain to the short-run effect of the funds rate on the bond rate over two sample periods only, 1961Q1 to 1995Q2 and 1961Q2 to 1979Q3. Hence I report two key coefficients, λ_2 and $(-\lambda_2/\lambda_1)$.

The results in Table 2 discussed in the previous section use the residual from the long-run Fed reaction function as a measure of the short-run stance of monetary policy. I now examine results if the short-run stance of policy is measured instead by the real funds rate $(NFR - \dot{p})$. Furthermore, I now estimate the Fisher relation (8) jointly with the short-run error-correction equation. This procedure allows for richer short-run dynamics in estimating the long-run effect of inflation on the bond rate. The short-run equation that incorporates these two new changes can be derived by replacing the residuals U_{1t-1} and U_{2t-1} in (10) by lagged levels of the variables and then by setting $b_1 = 1$. The resulting short-run equation is

$$\Delta BR_t = \tilde{d} + \lambda_1 BR_{t-1} + \lambda_3 \dot{P}_{t-1} + \lambda_2 (NFR - \dot{p})_{t-1} + \sum_{s=1}^{n} d_{1s} \Delta BR_{t-s} + \sum_{s=1}^{n} d_{2s} \Delta \dot{p}_{t-s} + \sum_{s=1}^{n} d_{3s} \Delta (NFR - \dot{p})_{t-s} + \sum_{s=1}^{n} d_{4s}\, gaph_{t-s} + \epsilon_t, \tag{11}$$

where $\lambda_3 = -\lambda_1 a_1$. In (11), if $a_1 = 1$, then the coefficients on BR_{t-1} and $\dot{p}_{t-1}$ sum to zero $(\lambda_1 + \lambda_3 = 0$ in (11)). I impose this restriction only if it is not rejected. Table 3 reports some key coefficients $(\lambda_1, \lambda_2, -\lambda_2/\lambda_1, \lambda_3)$. These estimated coefficients confirm the qualitative nature of results in Table 2. First, the real funds rate is a significant predictor of the bond rate and this result is fairly robust over several subsamples. Second, the Chow test indicates that two key parameters, λ_1 and λ_2, are unstable only between pre- and post-1979 periods (the date of the break is 1980Q2). This result is consistent with an increase in the near-term effect of policy on the bond rate in the post-1979 period.

The short-run error-correction equation (11) was alternatively re-estimated using the consumer price index and the GDP deflator to measure inflation, that is, the average inflation rate over the past three years. As in a few previous studies, I also used the Livingston survey data on one-year-ahead inflationary expectations. The results continue to indicate that the funds rate spread generally does help predict the bond rate and that the near-term response of the bond rate to the funds rate spread has increased since 1979 (see rows 1 through 3 in Table 4).[14]

The funds rate spread here measures the short-run stance of monetary policy because it is that component of the funds rate that does not comove with inflation. This spread, however,

Table 3 Short-Run Error-Correction Regressions Using the Level of the Funds Rate Spread

Estimation ends in year	$\lambda_2(\tilde{t})$	$\lambda_1(\tilde{t})$	$-\lambda_2/\lambda_1$	Mean lag in quarters	$\lambda_3(\tilde{t})$	$-\lambda_3/\lambda_1$	F1
1995Q2	0.12 (2.6)	-0.29 (4.7)	0.40	3.5	0.29 (4.7)	1.00	0.89
1994	0.10 (2.2)	-0.26 (4.3)	0.35	3.8	0.26 (4.3)	1.00	0.89
1992	0.11 (2.4)	-0.28 (4.4)	0.39	3.6	0.28 (4.4)	1.00	0.64
1990	0.11 (2.2)	-0.28 (4.1)	0.39	3.6	0.28 (4.1)	1.00	0.68
1988	0.11 (2.2)	-0.27 (3.9)	0.41	3.7	0.27 (3.9)	1.00	0.63
1986	0.13 (2.5)	-0.30 (4.2)	0.41	3.2	0.30 (4.2)	1.00	0.55
1984	0.10 (1.7)	-0.22 (2.3)	0.41	4.5	0.22 (2.3)	1.00	0.11
1982	0.21 (3.4)	-0.53 (3.5)	0.40	1.9	0.48 (3.5)	0.90	3.70*
1980	0.12 (1.9)	-0.24 (2.0)	0.50	4.2	0.24 (2.0)	1.00	1.00
1979Q3	0.13 (2.5)	-0.41 (2.4)	0.31	2.4	0.32 (2.5)	0.80	3.60*
1978	0.12 (2.4)	-0.41 (2.3)	0.29	2.4	0.32 (2.3)	0.78	3.40*
1976	0.14 (2.6)	-0.53 (2.6)	0.26	1.9	0.39 (2.5)	0.74	6.60**
1974	0.22 (3.1)	-0.92 (3.8)	0.23	1.1	0.67 (3.7)	0.73	14.20**
1972	0.21 (2.3)	-0.97 (3.4)	0.22	1.0	0.69 (3.3)	0.71	10.10**

* Significant at the 10 percent level.
** Significant at the 5 percent level.
Notes: The coefficients reported are from regressions of the form

$$\Delta BR_t = d_0 + \lambda_1 BR_{t-1} + \lambda_3 \dot{P}_{t-1} + \lambda_2 (NFR - \dot{p})_{t-1} + \sum_{s=1}^{2} d_{1s} \Delta BR_{t-s}$$
$$+\sum_{s=1}^{2} d_{2s} \Delta \dot{p}_{t-s} + \sum_{s=1}^{2} d_{3s} \Delta (NFR - \dot{p})_{t-s} + \sum_{s=1}^{2} d_{4s}\, gaph_{t-s},$$

where gaph is the output gap and other variables are defined in Table A1. All regressions are estimated by ordinary least squares. The estimation period begins in 1961Q2 and ends in the year shown. The mean lag is calculated as $-1/\lambda_1$. $\tilde{t}$ is the t-statistic. F1 tests the null hypotheses that λ_1 and λ_3 sum to zero, indicating that the Fisher restriction is consistent with data.

is still endogenous because, as noted before, the Fed routinely raises the funds rate during cyclical expansions and lowers it during cyclical downturns. The potential problem created by such endogeneity is that if the bond rate is directly influenced by variables that reflect the cyclical state of the economy and if those variables are omitted from short-run regressions, then the funds rate spread may be picking up the influence of those variables on the bond rate rather than the influence of monetary policy on the real component of the bond rate.

The short-run regressions reported in Tables 2 and 3 already include many of those variables such as the output gap that measures the cyclical state of the economy and changes in inflation, the bond rate, and the funds rate spread itself. While it is difficult to know all the information that the Fed may be using in setting its short-run funds rate policy, I re-estimated (11) alternatively including additional variables such as nonfarm payroll employment, sensitive materials prices, the deficit, and real growth. Those additional variables, when included in (11), are generally not significant and therefore do not change the qualitative nature of results in Table 3 (see rows 4a through 4d in Table 4).

It is sometimes argued that unit root tests used here have low power in distinguishing whether the variables are stationary or integrated. Hence the cointegration and error-correction methodology used here to distinguish between long- and short-run sources of comovement between the bond rate and the funds rate is suspect. However, the evidence presented above that the bond rate and the funds rate adjust one-for-one with inflation in

Table 4 Sensitivity to Changes in Specification

Changes in specification	Panel A: 1961Q2–1995Q2 $\lambda_2(\tilde{r})$	$(-\lambda_2/\lambda_1)$	f_2	$\tilde{f}_2$	Panel B: 1961Q2–1979Q3 $\lambda_2(\tilde{r})$	$(-\lambda_2/\lambda_1)$	f_2	$\tilde{f}_2$
1. CPI	0.08 (1.7)	0.42			0.09 (1.6)	0.16		
2. GDP deflator	0.11 (2.2)	0.49			0.10 (2.0)	0.19		
3. Livingston survey	0.08 (1.7)	0.35			0.07 (1.6)	0.23		
4. CPIEXFE Additional Variables								
a. Δ ln *PEM*	0.14 (2.9)	0.48			0.15 (2.8)	0.35		
b. d_t	0.10 (2.3)	0.37			0.13 (2.0)	0.28		
c. Δ ln *SMP*	0.10 (2.4)	0.38			0.13 (2.5)	0.29		
d. Δ ln ry_t	0.12 (2.5)	0.48			0.10 (1.5)	0.25		
5. Stationary: Level Regressions			0.07 (1.9)	0.42			0.09 (2.1)	0.29

Notes: The coefficients reported are from regressions of the form given in Table 3. The Fisher restriction is imposed in regressions estimated over 1961Q2–1995Q2 but not in those estimated over 1961Q2 to 1979Q3. CPI is the consumer price index; CPIEXFE is the consumer price index excluding food and energy; PEM is the nonfarm payroll employment; Δ ln *ry* is real GDP growth; *d* is federal government deficits scaled by nominal GDP, and SMP is the sensitive materials prices. The coefficients reported in the row labeled 5 are from the regression (14) of the text. f_2 is the coefficient that measures the contemporary response of the bond rate to the funds rate spread and $\tilde{f}2$ the near-term.

the long run is confirmed even if I treat the bond rate, the funds rate, and inflation as stationary variables. The stationary versions of the long-run regressions (8) and (9) can be expressed as in (12) and (13):

$$BR_t = a_0 + \sum_{s=0}^{n} a_{1s}\dot{p}_{t-s} + \sum_{s=1}^{n} a_{2s} BR_{t-s} + U_{1t} \qquad (12)$$

$$NFR_t = b_0 + \sum_{s=0}^{n} b_{1s}\dot{p}_{t-s} + \sum_{s=1}^{n} b_{2s} NFR_{t-s} + U_{2t}. \qquad (13)$$

The net response of the bond rate to inflation is $\left(\sum_{s=0}^{n} a_{1s}/1 - \sum_{s=1}^{n} a_{2s}\right)$ and to the funds rate is $\left(\sum_{s=0}^{n} b_{1s}/1 - \sum_{s=1}^{n} b_{2s}\right)$. One cannot reject the hypotheses that these net responses each are unity. As for short-run correlations, consider the following stationary version of the short-run equation:

$$BR_t = f_0 + f_1\dot{p}_t + f_2\,(NFR - \dot{p})_t + f_3\,gaph_t + \sum_{s=1}^{n} f_{4s}\,BR_{t-s} + \epsilon_t, \qquad (14)$$

where all variables are as defined before. Equation (14) already incorporates the long-run restriction that short-run funds rate policy actions affect the bond rate by altering the spread between the funds rate and the inflation rate. The other restriction can be imposed by the requirement that coefficients f_1 and $\sum_{s=1}^{n} f_{4s}$ in (14) sum to unity. The parameter f_2 in (14) measures the contemporaneous response of the bond rate to the funds rate spread and its net response can be calculated as $(f_2/1 - \sum_{s=1}^{n} f_{4s})$.

I estimate equation (14) by instrumental variables that are just lagged values of the right-hand side explanatory variables. In such regressions the funds rate spread still helps predict the bond rate (see row 5 in Table 4).

A Comparison with Some Previous Studies

In this section I discuss some previous studies that use entirely different methodologies but reach conclusions regarding the short-run impact of policy on long-term rates which are qualitatively similar to those reported here.

The first set consists of studies by Cook and Hahn (1989), Radecki and Reinhart (1994), and Roley and Sellon (1995). All three of these studies examine the response of long-term interest rates to changes in a measure of the federal funds rate target. They differ, however, with respect to the sample period studied and the length of the interval over which the interest rate response is measured. In Cook and Hahn the sample period studied is September 1974 to September 1979 and the interest rate response is measured on the day of the target change. In the other two studies the sample periods examined fall within the post-1979 period: 1989 to 1993 in Radecki and Reinhart (1994) and October 1987 to July 1995 in Roley and Sellon (1995). The measurement interval in Radecki and Reinhart spans the first ten days following the policy change, whereas in Roley and Sellon the time interval spans the period from the day after the previous policy action to the day after the current policy action. The economic rationale for the use of a wider measurement interval as in Roley and Sellon is that many times monetary policy actions are already anticipated by the markets, so that long-term interest rates move ahead of the announced change in the funds rate target. The relative magnitudes of interest rate responses before, during, and after the policy action, however, depend upon whether policy actions are anticipated or unanticipated and upon the degree of persistence in anticipated policy actions.

The measurement interval is the narrowest in Cook and Hahn and Radecki and Reinhart; the results there indicate that a one percentage point increase in the funds rate target induces 12 to 13 basis points movement in the ten-year bond rate (increases in longer maturity bond rates are somewhat smaller). The size of the interest rate response during and after the change in policy action as measured by Roley and Sellon is also modest; the 30-year Treasury bond yield rises by 10 basis points following one percentage point increase in the effective federal funds rate target. However, when the measurement interval includes days before the change in policy action, the measured interest rate response rises to 38 basis points. Thus a significant part of the response occurs before policy action is announced, indicating the presence of anticipated effect. What needs to be noted is that the magnitude of the total interest rate response measured by Roley and Sellon is quite close to the near-term response that I have estimated using an entirely different estimation methodology. Recall that for the complete sample period 1961Q2 to 1995Q2 the estimated near-term response of the ten-year bond rate to the funds rate spread is 42 basis points (see Table 2). The estimated near-term response is 36 basis points if instead I use the 30-year bond yield in my empirical work.

The other recent study showing that in the short run the long real rate comoves with the short nominal rate is by Fuhrer and Moore (1995). According to the expectational theory of the term structure of interest rates, the ex ante long-term real rate can be viewed as a weighted moving average of future short real rates. Fuhrer and Moore define the short real rate as the nominal yield on three-month Treasury bills minus the actual quarterly inflation rate. They then use a vector autoregression to construct long-horizon forecasts of the time paths of the three-month Treasury bill rate and the inflation rate. Given those forecasts they compute the 40-quarter-duration long-term real rate, since the average duration (maturity) of Moody's BAA corporate bond rate is 40 quarters. What they find is that over the period 1965 to 1992 the sample path of the ex ante long real rate tracks closely that of the short-term nominal rate (Fuhrer and Moore 1995, Figure 1, p. 224). The ex ante long real rate is still relatively stable, however, and only about one-fourth of the increase in the short nominal rate is reflected in the long real rate.

CONCLUDING OBSERVATIONS

This chapter has investigated empirically the immediate, near-term, and long-run effects of monetary policy on the bond rate. The federal funds rate is used as a measure of monetary policy, and the long run is viewed as the period during which trend relationships emerge. The results indicate that the long-run effect of monetary policy on the bond rate occurs primarily through the inflation channel.

In the short run, however, monetary policy also affects the bond rate by altering its expected real rate component. The short-run stance of monetary policy is measured by the spread between the funds rate and the ongoing trend rate of inflation. The results indicate that the near-term effect of the funds rate spread on the bond rate has increased considerably since 1979. In the pre-1979 period, the bond rate rose anywhere from 14 to 29 basis points whenever the funds rate spread widened by one percentage point. In the post-1979 period, however, the estimate of its near-term response ranges from 26 to 50 basis points.

This increase in the short-run sensitivity of the bond rate to monetary policy actions is consistent with the way the Fed has conducted its monetary policy since 1979. Since then the Fed has made a serious attempt to bring down the trend rate of inflation and contain inflationary expectations. If the public does not have full confidence in the Fed's disinflation policy, and if the Fed has to persist with sufficiently high short real rates to reduce the trend rate of inflation or contain inflationary expectations, then the estimated effect of a policy action on the bond rate would last longer. As a result, the near-term effect of policy on the bond rate would be stronger than would be the case if the disinflation policy were fully credible.

APPENDIX A

The stationarity properties of data are investigated using both unit roots and mean stationarity tests. The test for unit roots used is the augmented Dickey-Fuller test and the one for mean stationarity is the one in Kwiatkowski et al. (1992). Both these tests are described in Mehra (1994).

Table A1 Tests for Unit Roots and Mean Stationarity

Series	Panel A Tests for Unit Roots			Panel B Tests for Mean Stationarity
X	$\hat{\rho}$	$t_{\hat{\rho}}$	k	$\hat{n}u$
BR	0.96	-1.7	5	0.83*
$\dot{p}$	0.99	-2.0	1	0.56*
NFR	0.89	-2.9[a]	5	0.46[a]
BR–$\dot{p}$	0.87	-3.2*	3	0.04
NFR–$\dot{p}$	0.70	-5.0*	5	0.03
ΔBR	-0.10	-5.6*	4	0.19
$\Delta\dot{p}$	0.60	-4.3*	7	0.18
ΔNFR	-0.30	-4.9*	7	0.07

[a] The test statistic is close to the relevant 5 percent critical value.
* Significant at the 5 percent level.
Notes: BR is the ten-year Treasury bond rate; $\dot{p}$ is the average inflation rate over the past three years; and NFR is the nominal funds rate. Δ is the first difference operator. Inflation is measured by the behavior of the consumer price index excluding food and energy. This price series begins in 1957; therefore the effective sample period studied is 1960Q1 to 1995Q2. The values for ρ and t-statistics ($t_{\hat{\rho}}$) for $\rho = 1$ in panel A above are from the augmented Dickey-Fuller regressions of the form

$$X_t = a_0 + \rho X_{t-1} + \sum_{s=1}^{k} as\, \Delta X_{t-s}, \qquad \text{(a)}$$

where X is the pertinent series. The number of lagged first differences (k) included in these regressions are chosen using the procedure given in Hall (1990). The procedure starts with some upper bound on k, say k max chosen a priori (eight quarters here). Estimate (a) above with k set at k max. If the last included lag is significant, select $k = k$ max. If not, reduce the order of the autoregression by one until the coefficient on the last included lag is significant. The test statistic $\hat{n}_u$ in panel B above is the statistic that tests the null hypothesis that the pertinent series is mean stationary. The 5 percent critical value for $\hat{n}_u$ given in Kwiatkowski et al. (1992) is 0.463 and for $t_{\hat{\rho}}$ given in Fuller (1976) is −2.89.

Table A1 presents test results for determining whether the variables BR, $\dot{p}$, NFR, $BR - \dot{p}$ and $NFR - \dot{p}$ have a unit root or are mean stationary. As can be seen, the t-statistic ($t_{\hat{\rho}}$) that tests the null hypothesis that a particular variable has a unit root is small for BR, $\dot{p}$, and NFR, but large for $BR - \dot{p}$ and $NFR - \dot{p}$. On the other hand, the test statistics ($\hat{n}_u$) that tests the null hypothesis that a particular variable is mean stationary is large for BR, $\dot{p}$, and NFR, but small for $BR - \dot{p}$ and $NFR - \dot{p}$. These results indicate that BR, $\dot{p}$, and NFR have a unit root and are thus nonstationary in levels. In contrast $BR - \dot{p}$ and $NFR - \dot{p}$ do not have a unit root and are thus stationary in levels. Table A1 also presents unit roots and mean stationary tests using first differences of BR, $\dot{p}$, and NFR. As can be seen, the test results indicate that first differences of these variables are stationary.

The test for cointegration used is the one proposed in Johansen and Juselius (1990). The test procedure is described in Mehra (1994). Two test statistics—the trace test and the maximum eigenvalue test—are used to evaluate the number of cointegrating relationships. Table A2 presents these two test statistics for determining whether in bivariable systems like (BR, $\dot{p}$), (BR, NFR) and (NFR, $\dot{p}$) there exist a cointegrating vector. Those test results are consistent with the presence of cointegration between variables in each system.

Table A2 Cointegration Test Results

System	Trace test	Maximum eigenvalue test	k
$(BR, \dot{p})$	16.2*	11.3	8
(BR, NFR)	36.4*	32.5*	6
$(NFR, \dot{p})$	22.3*	17.1*	8

*Significant at the 5 percent level.
Notes: Trace the maximum eigenvalue tests are tests of the null hypothesis that there is no cointegrating relation in the system. The test used for cointegration is the one proposed on Johansen and Juselius (1990). The lag length in the relevant VAR system is k and is chosen using the likelihood ratio test given in Sims (1980). In particular, the VAR model initially was estimated k set equal to a maximum number of eight quarters. This unrestricted model was then tested against a restricted model, where k is reduced by one, using the likelihood ratio test. The lag length finally selected is the one that results in the rejection of the restricted model.

ACKNOWLEDGMENTS

The author thanks Peter Ireland, Roy Webb, and Marvin Goodfriend for their many helpful comments. The views expressed are those of the author and do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

NOTES

1. In most previous studies using an interest-rate-based measure of monetary policy, a short-term money market rate (the three-month or one-year Treasury bill rate) is used. Most of those studies surveyed recently in Akhtar (1995) find significant and large effects of short rates on long rates. In those studies the long-run response of nominal long rates ranges from about 22 to 66 basis points for every one percentage point change in nominal short rates. However, there is considerable skepticism about the reliability and interpretation of those effects. One main reason for such skepticism is even though monetary policy has its strongest effect on a short-term money market rate, the latter is also influenced by nonmonetary forces. Hence changes in short rates do not necessarily reflect changes in the stance of monetary policy.

   There are a few other empirical studies that use the federal funds rate as a measure of monetary policy. But most of those studies examine the effect of policy on the long rate in a bivariable framework. In such studies the estimated impact of policy on the long rate is quantitatively modest and temporally unstable (see Akhtar 1995, Table 3). An exception is the recent work in Mehra (1994) which uses a multivariable framework and finds a significant effect of the real funds rate on the long rate. However, Mehra (1994) does not investigate the robustness of those results to alternative specifications or to different sample periods.
2. Recent research has shown that the federal funds rate is a good indicator of the stance of monetary policy (Bernanke and Blinder 1992; Bernanke and Mihov 1995).
3. In this chapter near-term effects refer to responses of the bond rate to recent past values of the funds rate spread. The immediate effect is the response to the one-period lagged value of the spread and the near-term effect is the cumulative response to all such past values. What I call the near-term effect is sometimes referred to as the long-run effect in previous studies. As indicated later, I use the long-run effect to measure the effect that arises from the existence of equilibrium or trending relationships among nonstationary variables.

4. The discussion in this section draws heavily from Goodfriend (1993).
5. This lag arises not because financial markets adjust slowly but rather because funds rate strategy puts considerable persistence in the funds rate. Such a lag can also arise if the bond rate depends upon anticipated policy moves which in turn are influenced partly by current and past values of the policy variable.
6. This framework differs somewhat from the ones used in Goodfriend (1993) and Mehra (1994). Goodfriend describes interactions between the bond rate and the funds rate, taking into account the behavior of actual inflation and real growth, whereas in Mehra (1994) the deficit also is included. That work, however, indicates that the deficit variable is not a significant determinant of the bond rate once we control for the influences of inflation and real growth. Hence the deficit variable is excluded from the work here. I use the output gap rather than real growth as a measure of the state of the economy because the bond rate appears more strongly correlated with the former than with the latter. The qualitative nature of results, however, is the same whether the output gap or real growth is used as a measure of the state of the economy. Moreover, I do examine the sensitivity of results to some changes in specification in the subsection entitled "Additional Empirical Results."
7. The concept of weak exogeneity is introduced by Engle et al. (1983). The hypothesis that inflation is weakly exogenous with respect to the parameters of the cointegrating vector simply means that inferences on such parameters can be efficiently carried out without specifying the marginal distribution of inflation. More intuitively, inflation in equation (2) could be considered predetermined in analyzing the response of the bond rate to inflation.
8. This test is proposed in Johansen (1992).
9. I get similar results if instead the consumer price index or the GDP deflator is used to measure actual inflation (see the subsection entitled "Additional Empirical Results").

   In a couple of recent studies (Hoelscher 1986; Mehra 1994) the Livingston survey data on one-year-ahead inflationary expectations are used to measure long-run anticipated inflation. The results in Mehra (1994), however, indicate that the near-term impact of the funds rate on the bond rate remains significant if one-year-ahead expected inflation (Livingston) data are substituted for actual inflation in the empirical work (see Mehra 1994, Table 4). That result continues to hold in this article also (see the subsection entitled "Additional Empirical Results").
10. These tests are described in Mehra (1994).
11. This can be shown as follows. Assume, for example, the level of the bond rate is related to inflation and the funds rate spread as in

$$BR_t = a_0 + a_1\dot{p}_t + a_2(NFR - b_1\dot{p})_t + V_t, \tag{a}$$

   where the stationary component is $a_0 + a_2(NFR - b_1\dot{p})_t$ and the nonstationary component is $a_1\dot{p}_t$. The error-correction regression is

$$\Delta BR_t = \lambda_1 V_{t-1} + \sum_{s=1}^{n} \text{(other lagged differences of variables)} + \epsilon_t, \tag{b}$$

   where λ_1 is negative. Substituting for V_{t-1} from (a) into (b) yields (c):

$$\Delta BR_t = \lambda_1 BR_{t-1} - \lambda_1 a_1 \dot{p}_{t-1} - \lambda_1 a_2 (NFR - b_1\dot{p})_{t-1} + \textit{other terms}. \tag{c}$$

   Equation (c) can be estimated and a_2 can be recovered as $\lambda_1 a_2/\lambda_1$, which is the minus of the coefficient on $(NFR - b_1\dot{p})_{t-1}$ divided by the coefficient on BR_{t-1}. The coefficient a_2 then measures the near-term response of the bond rate to the funds rate spread. I do not label a_2 as measuring the long-run effect because the spread is stationary. The long run is defined as the period over which trend relationships emerge. In the long run the funds rate spread $(NFR - \dot{p})$ is constant.
12. In estimated short-run regressions the coefficients that appear on lagged differences of the bond rate are very small. If we ignore those coefficients, then the short-run equation (c) given in foot-

note 11 can be expressed as

$$BR_t = \frac{-\lambda 1 a 1}{1 - (1 + \lambda 1)L} \dot{p}_{t-1} + \frac{\lambda_2}{1 - (1 + \lambda_1)L} (NFR - \dot{p})_{t-1} + \textit{other terms},$$

where L is the lag operator and where λ_2 is $-\lambda_1 a_2$. The coefficients (w_i) that appear on lagged levels of $NFR - \dot{p}$ are then of the form λ_2, $\lambda_2(1 + \lambda_1)$, $\lambda_2(1 + \lambda_1)^2$, etc. The mean lag then can be calculated as follows:

$$\textit{Mean Lag} = \sum_{i=1}^{\infty} w_i\, i / \sum_{i=1}^{\infty} w_i$$

$$= \frac{1}{(1 - 1 - \lambda_1)^2} \times \frac{1}{\lambda_2 \frac{1}{1 - 1\lambda_1}} = -\frac{\lambda_1}{\lambda_1^2} = -\frac{1}{\lambda_1}.$$

13. This is confirmed by results of the formal Chow test that is discussed in the next subsection.
14. When inflation is measured by the behavior of the consumer price index or the Livingston survey, I get some mixed results. The statistical significance of the coefficient that appears on the funds rate spread is not robust over different sample periods.

REFERENCES

Akhtar, M. A. Monetary Policy and Long-Term Interest Rates: A Survey of Empirical Literature. *Contemporary Economic Policy,* vol. XIII (July 1995), pp. 110–30.

Bernanke, B. S., and Blinder, A. S. The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review,* vol. 82 (September 1992), pp. 901–21.

Bernanke, B. S., and Mihov, I. Measuring Monetary Policy. NBER Working Paper No. 5145. June 1995.

Cook, T. and Hahn, T. The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s. *Journal of Monetary Economics,* vol. 24 (November 1989), pp. 331–51.

Engle, R. F., Hendry, D. F. and Richard, J-F. Exogeneity. *Econometrica,* vol. 51 (March 1983), pp. 277–304.

Fuhrer, J. C., and Moore, G. R. Monetary Policy Trade-offs and the Correlation between Nominal Interest Rates and Real Output. *American Economic Review,* vol. 85 (March 1995), pp. 219–39.

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

Goodfriend, M. Interest Rate Policy and the Inflation Scare Problem: 1979–1992. Federal Reserve Bank of Richmond *Economic Quarterly,* vol. 79 (Winter 1993), pp. 1–24.

Hall, A. Testing for a Unit Root in Time Series with Pretest Data Based Model Selection. Manuscript. North Carolina State University, 1990.

Hoelscher, G. New Evidence on Deficits and Interest Rates. *Journal of Money, Credit, and Banking,* vol. XVIII (February 1986), pp. 1–17.

Johansen, S. Cointegration in Partial Systems and the Efficiency of Single Equation Analysis. *Journal of Econometrics,* vol. 52 (June 1992), pp. 389–402.

Johansen, S. and Juselius, K. Maximum Likelihood Estimation and Inference on Cointegration—With Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics,* vol. 52 (May 1990), pp. 169–210.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics,* vol. 54 (October–December 1992), pp. 159–78.

Mehra, Y. P. A Federal Funds Rate Equation. Federal Reserve Bank of Richmond *Economic Inquiry* 35(3):621–30, July.

Mehra, Y. P. An Error-Correction Model of the Long-Term Bond Rate. Federal Reserve Bank of Richmond *Economic Quarterly,* vol. 80 (Fall 1994), pp. 49–68.

Radecki, L. J., and Reinhart, V. The Financial Linkages in the Transmission of Monetary Policy in the United States. In *Bank for International Settlements, National Differences in Interest Rates Transmission.* Basle, Switzerland, 1994.

Roley, V. V., and Sellon, G. H., Jr. Monetary Policy Actions and Long-Term Interest Rates. Federal Reserve Bank of Kansas City *Economic Review,* vol. 80 (Fourth Quarter 1995), pp. 73–89.

Sims, C. A. Macroeconomics and Reality. *Econometrica,* vol. 48 (January 1980), pp. 1–48.

Stock, J. H., and Watson, M. W. A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica,* vol. 61 (July 1993), pp. 783–820.

This Page Intentionally Left Blank

77

Monetary Policy and Long-Term Real Interest Rates

Timothy Cogley
Senior Economist, Federal Reserve Bank of San Francisco, San Francisco, California

For many years, the Federal Reserve has relied on M2 as an indicator of current economic conditions. Over the last few years, however, the historical relation between M2 velocity and interest rates has broken down, as velocity has risen sharply while interest rates have fallen. This development has made it difficult to interpret the information in M2. For example, over the last few years, M2 growth often has fallen below its target range. Based on historical relations, the slow growth in M2 would be a sign of serious weakness in the economy. But in this case, although the economy has not been strong, the weakness in M2 mainly reflected the fact that M2 velocity turned out to be higher than expected. In other words, if historical relationships had held, the recent slow growth in M2 would have implied a much weaker economy than actually was observed.

Since it has become more difficult to interpret movements in M2, a variety of other indicators of economic and financial conditions have been explored. For example, Chairman Greenspan has suggested that real interest rates might be a useful indicator of economic conditions. The Federal Reserve System cannot peg real interest rates, because an attempt to do so would run the risk of generating a cumulative inflationary or deflationary process. But movements in real interest rates might provide timely and useful information about economic and financial conditions and thus might provide a useful guidepost for monetary policy makers.

To make this operational, we need to address two sets of issues. The first set concerns how to define a benchmark with which to compare movements in real interest rates. That is, how do we decide whether real rates are unusually high or low? The second set concerns the measurement of real interest rates. Trehan (1993) discusses these issues in terms of short-term real interest rates. This chapter complements his analysis by discussing these issues in terms of long-term interest rates.

DEFINING A BENCHMARK REAL INTEREST RATE

The first problem is to define a benchmark real interest rate and to determine how and why it varies over time. Chairman Greenspan defines the benchmark real interest rate as the level that, if sustained, would keep the economy at its productive potential over time. This

Reprinted from: *FRBSF Weekly Letter,* No. 93-42 (Dec. 3, 1993).

is conceptually similar to Milton Friedman's (1968) natural rate of unemployment, so I will refer to it as the "natural" real interest rate.

The natural real interest rate will vary over time, because the economy is subject to non-monetary shocks that affect its actual and potential output. Furthermore, since it is difficult to track the variation in the economy's productive potential, it may also be difficult to track the natural real rate, especially over short horizons.

To some extent, this difficulty can be mitigated by tracking the long-term, rather than the short-term, natural rate. Long-term real interest rates can be decomposed into two parts. One is the expected real return earned by rolling over short-term bills, and the other is the expected excess return earned by holding long-term bonds. The first component is motivated by the Expectations Hypothesis of the term structure of interest rates. According to this theory, arbitrage between alternative long- and short-term financial investments will ensure that the long-term rate will equal the expected rollover return on short-term bills. However, empirical studies show that there are predictable excess returns on long-term bonds, and this is often interpreted as a manifestation of time-varying risk, since arbitrage would also eliminate risk-adjusted excess returns. Thus the second component can be interpreted as a risk premium.

Both of these components are likely to be less variable over longer holding periods than they are over short holding periods. The first component, the rollover return, is equal to a weighted average of expected future short rates, and the averaging process smooths out much of the variation in short rates. Thus long-horizon rollover returns ought to be less variable than short-horizon roll-over returns. Similarly, the long-term risk premium is a weighted average of expected short-term excess holding returns. Again, because of the averaging process, long-term premia ought to be smoother than short-term excess returns. Therefore the long-term natural rate is likely to be less variable than the short-term natural rate.

One problem that remains to be solved is defining the natural risk premium. This is important because much of the variation in long-term real interest rates appears to be due to variation in risk-premia (for example, see Cogley 1993). The problem is that current models of the term structure do not generate empirically plausible risk premia. Specifically, if we assume that financial markets eliminate arbitrage opportunities, then risk-adjusted excess returns should be *unpredictable.* However, when confronted with data, theoretical models of the term structure imply that "risk-adjusted" excess returns are *predictable,* and this is generally interpreted as a sign that the models do not correctly adjust for risk. Without an adequate model of risk, it would be difficult to know what is meant by the natural level of the risk premium. The problem of modeling risk premia remains an active area of research.

MEASURING LONG-TERM REAL INTEREST RATES

The market real rate of interest is equal to the nominal interest rate minus the expected inflation rate (this is sometimes call the *ex ante* rate). Since market expectations of inflation are not directly observable, the real interest rate also is unobservable. Thus, real interest rates must be estimated. There are several ways to do this, corresponding to different estimates of expected inflation.

One approach is to substitute actual, realized inflation rates for expected inflation. This measure is known as the *ex post* real interest rate. Since ex post real interest rates are based on the actual inflation rate over the holding period, they cannot be computed until the holding period has ended. Thus ex post real interest rates on long-term bonds are available only after long lags. For example, the most recently available ex post 10 year real interest rate is the one for November 1983. Since ex post real rates are not available on a timely basis, they are not likely to be useful for current policy analysis.

Furthermore, at the long end of the maturity spectrum, ex post real rates are not likely to match ex ante real rates closely. The ex ante real interest rate is equal to the ex post real rate plus the error in forecasting inflation over the holding period. Thus the ex post real rate can be regarded as the sum of the true ex ante value plus a measurement error. Since long-term inflation forecast errors are highly correlated over time, there can be persistent differences between ex ante and ex post real interest rates. As a consequence, the measurement error in ex post real interest rates often obscures the true ex ante value, and this makes it difficult to use long-term ex post rates for historical analysis.

A second approach is to estimate expected inflation using survey data on inflation forecasts. Various surveys of long-term inflation expectations are available on a sporadic basis going back to 1979. For example, Figure 1 reports the 10-year real interest rate based on the Hoey survey of 10-year inflation expectations. Survey data have three limitations. First, since respondents may have little at stake when filling in the survey, there is some concern that survey data may not provide an accurate measure of inflation expectations. Second, surveys of long-term inflation forecasts go back only to the late 1970s. Since this period covers only a few business cycles, there may be too little data to learn much about the cyclical properties of long-term real rates. Third, the early surveys have missing observations, which greatly complicates statistical analysis. The Federal Reserve Bank of Philadelphia has begun to collect a survey of long-run inflation forecasts on a regular basis. While this seems like a worthwhile long-run investment, it may take a number of business cycles before there are enough observations to use the data to analyze the cyclical properties of long-term real interest rates.

A third way to measure long-term real interest rates is to estimate them using an econometric model. For example, the model discussed in Cogley (1993) can be used for this purpose. This model estimates long-term real interest rates by forecasting short-term real interest rates and excess holding returns over long time horizons and then discounting them back

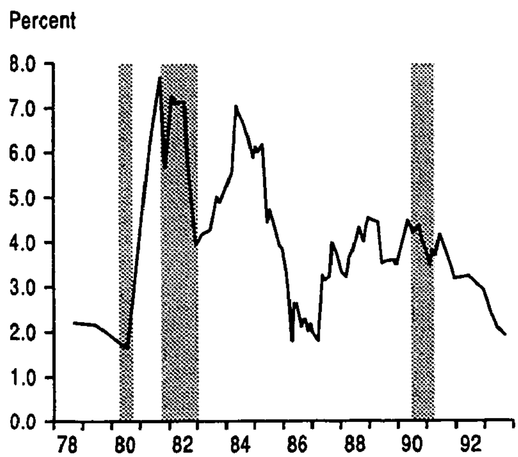


Figure 1 Real 10-year treasury bond yield derived from Hoey survey of inflation.

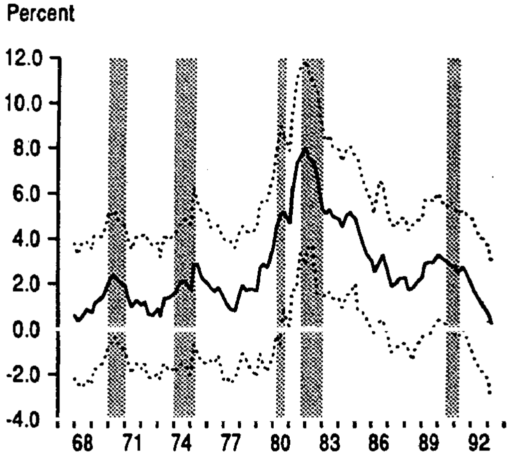


Figure 2 Econometric estimates of the long-term real rate.

to the present. The forecasting model includes lagged values of the 3-month Treasury bill rate, the inflation rate, the unemployment rate, and the ex post excess holding return on 10-year Treasury bonds. The model was estimated over the period 1968.Q1 to 1993.Q2, and ex ante forecasts were generated by iterating through every quarter in the sample.

The resulting long-term real rate is shown by the solid line in Figure 2. The shaded areas mark the dates of recessions, as determined by the National Bureau of Economic Research. The long-term real rate appears to be countercyclical, although it does not systematically lead or lag the business cycle. For example, in the 1974–1975 recession, the real rate peaked near the trough of the cycle, while in the 1981–1982 recession it peaked shortly after the downturn.

For our purposes, it is important to try to quantify the uncertainty about this measure of the long-term real rate. This measure is based on estimates of a forecasting model, and errors in estimating the model will translate into errors in measuring the long-term real rate. A Monte Carlo simulation was conducted in order to quantify the degree of uncertainty, and the results are shown by the dotted lines in Figure 2. These mark the margin of error associated with the estimated real rate. Specifically, at any given date, there is a 5 percent chance that the real interest rate could be as high as the upper curve as well as a 5 percent chance that it could be as low as the lower curve. In other words, the figure tells us that there is a 9 in 10 chance that the real rate lies somewhere between the dotted lines. The average distance between the upper and lower margins of error is roughly 6 percent; thus it is difficult to pin down long-term real interest rates with a great deal of precision.

Furthermore, this interval understates the true uncertainty about the long-term real rate because it only accounts for uncertainty in the estimates of the parameters of the forecasting model, not for uncertainty about the model's specification. Presumably, if this were taken into account, the margin of error would be even larger.

CONCLUSION

Conceptually, it is difficult to define a natural long-term real interest rate because we do not yet have satisfactory models of risk. Empirically, it is difficult to estimate long-term real interest rates because there is a great deal of uncertainty about long-horizon forecasts. Thus, while long-term real interest rates may prove to be a useful indicator of economic and financial conditions, we need to confront a number of difficult issues in order to make this operational.

REFERENCES

Cogley, T. 1993. Interpreting the Term Structure of Interest Rates. *Federal Reserve Bank of San Francisco Weekly Letter* (April 16).

Friedman, M. 1968. The Role of Monetary Policy. *American Economic Review* (March) pp. 1–17.

Trehan, B. 1993. Real Interest Rates. *Federal Reserve Bank of San Francisco Weekly Letter* (November 5).

This Page Intentionally Left Blank

78
Real Interest Rates

Bharat Trehan
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

In his latest Humphrey-Hawkins testimony before Congress, Federal Reserve Chairman Greenspan suggested that real interest rates could come to play a larger role in the formulation of monetary policy. This chapter discusses what that role might be and describes what we know about the behavior of short-term real rates in the U.S. since the 1960s. Our review suggests that while it is unlikely that real interest rates will provide information about how to conduct policy in the short run, they are likely to be useful in helping to avoid policy settings that are untenable in the long run.

REAL RATES AND POLICY

The real rate of interest can be defined as the nominal rate minus the expected rate of inflation. Thus, if the nominal rate on 3-month T-bills is 4 percent, and the expected rate of inflation is 3 percent, then the real rate on 3-month T-bills would be 1 percent. Real rates are important because they provide a measure of the value of resources today versus tomorrow (abstracting from inflation), so they play a central role in people's decisions about saving, consuming, and investing. Therefore, real rates figure prominently in most kinds of models of the macroeconomy, as well as in discussions of how monetary policy affects the economy.

In thinking about the role of real interest rates in monetary policy it is useful to have a concept of the equilibrium real rate. For our purposes the equilibrium real rate can be defined as the rate that would equate economy-wide demand and supply in the long run, once short-run disturbances have worked themselves out. Factors that shift demand or supply could cause this rate to vary over time. For instance, the equilibrium real rate would rise if individuals decided to consume more today and reduce the amount they had been saving for the future.

The Fed cannot keep the real interest rate away from its equilibrium level in the long run without generating either an acceleration or a deceleration in inflation. Consider, for example, what would happen if the Fed tried to keep the real rate artificially low. An increase in the rate of money growth initially would push down both real and nominal interest rates,

---

Reprinted from: *FRBSF Weekly Letter,* No. 93-38 (Nov. 5, 1993).

given that the inflation rate would not adjust at once to the higher money growth. However, faster money growth would raise inflation expectations eventually, so the inflation rate would adjust as would interest rates. To keep real interest rates artificially low, the Fed would need to raise the rate of money growth further; of course, then the rate of inflation would rise further and the whole cycle would begin again. It turns out that this description of what would happen if the Fed tried to keep real interest rates too low is not entirely hypothetical; the Fed's attempts to stimulate the economy during the 1970s (through what amounted to a policy of extremely low real interest rates) led to steadily rising inflation that was finally checked at great cost during the 1980s. Thus, historical experience suggests that the Fed would not find the real interest rate a useful target. This does not mean that the real rate cannot be used in the policymaking process; the Fed could still use it as an indicator—that is, the Fed could use the information contained in the real interest rate to help determine the appropriate stance of policy.

However, using the real rate of interest as a monetary policy indicator is not a straightforward task either. For one thing, it is difficult to determine what the equilibrium rate is at any point in time, because real interest rates are affected by many factors. To get a sense of how many different factors may be involved, consider how many different explanations have been offered for the unusually high real rates of the 1980s. Among the candidates: tighter monetary policy, easier fiscal policy, an increase in the rate of return to capital, a slow adjustment of expectations to declining inflation, as well as the savings and loan crisis.

SHORT-TERM REAL RATES SINCE 1960

In view of these problems it may seem that we could get a better sense of the equilibrium level of real rates simply by looking at the historical data we have. But things are not so straightforward there either. One problem is that we cannot observe the real rate that is likely to be most relevant to decisionmaking. An individual's decision about whether to borrow money, for example, is based on the real rate of interest she expects to pay at the time she borrows the money; that is, it is the nominal rate less the rate of inflation expected over the life of the loan. Economists call this the *ex ante* real rate. Since we cannot observe this rate directly, we must resort to making an estimate about the public's expectation of inflation over the relevant horizon and then subtracting this estimate from the observed nominal rate. Errors in estimating expected inflation translate into errors in estimating the real rate. Further, since inflation in the near term (say over the next three months) is easier to predict than inflation in the long run (say over ten years), estimates of long-term real rates are likely to be more problematic than estimates of short-term real rates.

Keeping these caveats in mind, we now turn to the data itself. Figure 1 plots the nominal and estimated real interest rates on 3-month Treasury bills since 1960. The estimate of the real rate was constructed by subtracting expected inflation from the nominal rate, with expected inflation modeled as a function of lagged inflation alone. The shaded areas represent recessions.

Over this period, the nominal rate has been rather volatile, ranging between 3 and 16 percent. Although the real rate is less volatile than the nominal rate, it does tend to move around quite a bit. Further, shocks to the 3-month real rate appear to persist for a long time, and the rate shows little tendency to return to a central value or mean; that is, it appears to

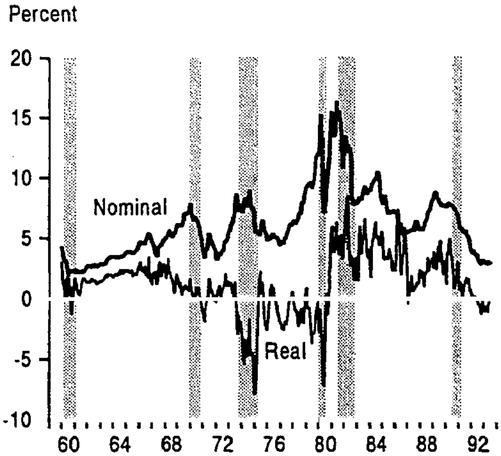


Figure 1 3-month T-bill rate.

be "nonstationary." Empirical studies on the issue are divided, with some concluding that the real rate is nonstationary while others disagree.

If real rates *are* nonstationary, then the concept of the equilibrium real rate would have little operational significance for monetary policy. For example, a fall in the measured real rate could be the result of a permanent decline in the equilibrium real rate, in which case it would have no implication for the stance of monetary policy; or it could be the result of easy monetary policy, in which case it would suggest that monetary policymakers would need to reverse direction at some point in the future.

Even if real rates are *not* strictly nonstationary, the figure indicates that it still would be difficult to devise an operational measure of the equilibrium real rate. This is because the short-term rate shows a tendency to hover around different values for extended periods of time. One can make out at least three different phases in the figure. The first phase coincides roughly with the 1960s, when real rates were low but positive, and averaged 1.7 percent. The second covers the 1970s, when real rates were generally negative, and averaged–0.9 percent. Short real rates were noticeably higher in the 1980s, when they averaged 3.1 percent. Real rates have averaged 0.9 percent in the 1990s, although it is too soon to tell whether this marks a new phase. Some researchers have suggested that the change in the behavior of real rates (especially from the 1970s to the 1980s) is associated with a change in the monetary policy regime; however, others have pointed out that similar regime changes in other countries have not had the same impact. (See Bonser-Neal 1990 for a discussion.)

Short-term real interest rates show little evidence of systematic variation over the cycle, once these longer-run patterns are allowed for. Thus, real rates were low but positive during the two recessions over the 1960s (averaging less than 1 percent in the two recessions), negative during the 1973–1975 recession (averaging–4.2 percent), and noticeably high during the 1981–1982 recession (averaging 4.4 percent). This behavior is not surprising, given that different factors are likely responsible for different recessions. Thus, it is generally agreed that the 1981–1982 recession was caused by tighter monetary policy, which would tend to push up real rates in the short run. By contrast, adverse supply shocks are believed to have played a larger role in the 1973–1975 recession; these shocks as well as the generally easier policy followed by the Fed during this period would tend to reduce real rates in the short run. In general, then, it seems difficult to predict how short real rates would behave over the cycle without further information about what was going on.

Could nominal rates be used to infer the underlying behavior of real rates? A look at Figure 1 shows that the answer is no, since the two do not move in the same way relative to each other. Once again, this is not surprising, since different factors will have different effects on the two rates. Rising inflation in the 1970s, for example, was associated with higher nominal rates but lower real rates. This negative relationship between real rates and inflation has been noted by various researchers for other time periods and other countries as well, but has not been satisfactorily explained. (See, for instance, Mishkin 1988.) By contrast, real and nominal rates moved closely together over the 1979–1982 period, when the Fed moved nominal rates to combat inflation. Note also that real and nominal rates have moved closely together since the mid-1980s, implying that inflation expectations have not changed much over the period.

CONCLUSIONS AND POLICY IMPLICATIONS

Real interest rates are affected by a large number of factors, and it is difficult to know where the "equilibrium rate" would be at any point in time. The problem is made worse by the fact that the *ex ante* real rate cannot be observed directly. Having to estimate these rates naturally introduces error into this process.

The fact that we are uncertain about the equilibrium rate does not mean that real rates have no role to play in setting monetary policy. While it is difficult to determine the correct level of real interest rates at any point in time, it is easier to tell whether a given level of rates is outside some reasonable range. For example, if there is one clear lesson for monetary policy from the 1970s, it is that short-term real rates should not be forced below zero for long periods. Thus, while it is unlikely that real interest rates can be used to provide a day-to-day guide for monetary policy, they can provide warnings about extreme policy settings.

REFERENCES

Bonser-Neal, C. 1990. Monetary Regime Changes and the Behavior of Real Interest Rates: A Multi-country Study. *Journal of Monetary Economics* pp. 329–359.

Mishkin, R. F. 1988. Understanding Real Interest Rates. NBER Working Paper No. 1204.

79

Taylor's Rule and the Fed 1970–1997

John P. Judd
Vice President and Associate Director of Research, Federal Reserve Bank of San Francisco, San Francisco, California

Glenn D. Rudebusch
Research Officer, Federal Reserve Bank of San Francisco, San Francisco, California

This chapter estimates a simple model of the Federal Reserve's "reaction function"—that is, the relationship between economic developments and the Fed's response to them. We focus on how this estimated reaction function has changed over time. Such changes are not surprising given compositional changes in the Federal Open Market Committee, and we consider three subsamples delineated by the terms of recent Fed Chairmen. We find that the estimated reaction functions for each period vary in ways that seem broadly consistent with the success or failure during the period at controlling inflation. These results suggest that a Taylor-rule framework is a useful way to summarize key elements of monetary policy.

Macroeconomists have long been interested in modeling the Federal Reserve's "reaction function"—that is, modeling how the Fed alters monetary policy in response to economic developments. The Fed's reaction function plays an important role in a wide variety of macroeconomic analyses. It can provide a basis for forecasting changes in the Fed's policy instrument—namely, short-term interest rates. Also, within the context of a macro model, the reaction function is an important element in evaluating Fed policy and the effects of other policy actions (e.g., fiscal policy) or economic shocks (e.g., the 1970s oil embargo). Finally, when rational expectations are assumed in macro models, knowing the correct reaction function is an important element in estimating the entire model. For example, with forward-looking expectations, estimates of a parameter such as the one linking real spending to the policy instrument will likely depend crucially on expected monetary policy and the nature of the monetary policy regime.

Numerous reaction functions have been estimated. Khoury (1990), for example, surveys 42 such empirical Fed reaction functions from various studies. Moreover, the large numbers of monetary policy vector autoregressions (VARs) (e.g., Bernanke and Blinder 1992) that have been estimated recently also include an equation that can be interpreted as a Fed reaction function (see Rudebusch 1998).

Reprinted from: *FRBSF Economic Review*, n. 3 (1998) 3–16.

However, despite this work, researchers have not been particularly successful in Lproviding a definitive representation of Fed behavior. Khoury finds little consistency in the significance of various regressors in the reaction functions she surveys. She states (p. 28): "One who [examines] just one of these reaction functions may feel convinced that one has learned how the Fed responds to economic conditions, but that seeming knowledge disappears as one reads a large number of these studies." Overall, it appears that there have not been any great successes in modeling Fed behavior with a single, stable reaction function. As an illustration, McNees (1992) compares his latest estimate of a Fed reaction function to his previous estimate (1986) and states (p. 11): "The number of modifications to the original specification required to make it track the past six years serve[s] as a clear illustration that policy reaction functions can be fragile."

There are a number of plausible explanations for such instability. For example, a central bank's reactions may be too complex to be adequately captured by a simple linear regression. Another factor may be changes in the composition of the Federal Open Market Committee (FOMC). Such compositional changes may bring to the fore policymakers with different preferences and different conceptions of the appropriate operation and likely transmission of monetary policy. While many people and events influence policy, arguably one of the more important and *identifiable* compositional changes is in the Fed Chairmanship. Changes associated with different Chairmen may be exogenous, but there also may be an endogenous element that represents an adaption to "lessons" learned from prior experiences. Indeed, Chairmen may be chosen who are seen as likely to avoid the mistakes of the past.[1] For example, part of the backing for Paul Volcker as Chairman in the high-inflation environment of 1979 may have come from the expectation that he would be tough on inflation.[2]

Accordingly, in this chapter we estimate reaction functions for three separate empirical subsamples delineated by the identity of the Fed Chairman: specifically, we consider the terms of Arthur Burns (1970.Q1–1978.Q1), Paul Volcker (1979.Q3–1987.Q2), and Alan Greenspan (1987.Q3–present). We omit any discussion of policy under Chairman Miller (1978.Q2–1979.Q2) because of his short tenure. This delineation gives us three subsamples of moderate and approximately equal length that are selected in an *a priori* fashion.

The organizing principle for our investigation is the Taylor rule, which we use as a rough gauge for characterizing and evaluating the broad differences in the relative weights given to monetary policy goal variables between periods. The rule specifies that the real federal funds rate reacts to two key goal variables—deviations of contemporaneous inflation from an inflation target and deviations of real output from its long-run potential level. These variables would appear to be consistent with the Fed's legislated mandate.[3] Moreover, Taylor (1993) argued that this rule represents "good" policy, in the sense that it relates a plausible Fed instrument to reasonable goal variables, and it stabilizes both inflation and output reasonably well in a variety of macroeconomic models. More recent model simulation studies (e.g., Rudebusch and Svensson 1998 and Levin, Wieland, and Williams 1997) have reinforced the latter conclusion.

Moreover, these recent studies suggest that although Taylor-type rules are very simple, they may be capable of capturing the essential elements of more realistic regimes in which the central bank "looks at everything." Simple Taylor-type reaction functions were found to perform almost as well as optimal, forecast-based reaction functions that incorporate all the information available in the models examined. In addition, the simple specification was found to perform almost as well as reaction functions that explicitly include a variety of additional variables. These results appear to be fairly robust across a variety of macroeconomic models. Thus, the general form of the Taylor rule may be a good device for capturing the key elements of policy in a variety of policy regimes.

The rule as originally specified by Taylor serves as a useful starting point for our investigation below. After briefly examining this rule, we focus on econometrically estimating a dynamic version of the rule for the three periods defined above. We find that, overall, the estimated dynamic Taylor-type reaction functions do provide a way to capture important elements of the policy regimes in place during these periods. The key elements of the estimated reaction functions for each period also vary in ways that seem broadly consistent with the success or failure during the periods at controlling inflation. This conclusion is reinforced at the end of the paper by explicit evaluations of the reaction functions in the three periods in the context of a small macro model.

We do not regard the results of this investigation as providing a complete representation of Fed behavior, in part because we have controlled for only one source of sample instability. However, we hope our results serve as a springboard for a discussion of some of the salient features—and changes—in Federal Reserve behavior over time. Also, as noted above, it is important to develop a better understanding of how the Fed's reaction function has changed over time for macroeconometric research.

TAYLOR'S ORIGINAL RULE

Taylor (1993) suggests a very specific and simple rule for monetary policy. His rule sets the level of the nominal federal funds rate equal to the rate of inflation (in effect, making it an equation for the ex post *real* funds rate) plus an "equilibrium" real funds rate (a "natural" rate that is seen as consistent with full employment) plus an equally weighted average of two gaps: (1) the four-quarter moving average of actual inflation in the GDP deflator less a target rate, and (2) the percent deviation of real GDP from an estimate of its potential level:

$$i_t = \pi_t + r^* + 0.5(\pi_t - \pi^*) + 0.5\,(y_t) \qquad (1)$$

where

i = federal funds rate,
r^* = equilibrium real federal funds rate,
π = average inflation rate over the contemporaneous and prior three quarters (GDP deflator),
π^* = target inflation rate
y = output gap (100.(real GDP–potential GDP) ÷ potential GDP).[4]

Taylor did not econometrically estimate this equation. He assumed that the weights the Fed gave to deviations of inflation and output were both equal to 0.5; thus, for example, if inflation were 1 percentage point above its target, the Fed would set the real funds rate 50 basis points above its equilibrium value. Furthermore, Taylor assumed that the equilibrium real interest rate and the inflation target were both equal to 2 percent. We shall examine these assumptions below; however, it is instructive to consider the interest rate recommendations from the original Taylor rule.[5]

Figure 1 illustrates the original Taylor rule during 1970–1998. The top panel shows the recommendations of the rule on a quarterly basis. The bottom two panels show the variables that enter the rule—the GDP gap and inflation. As explained earlier, higher levels of

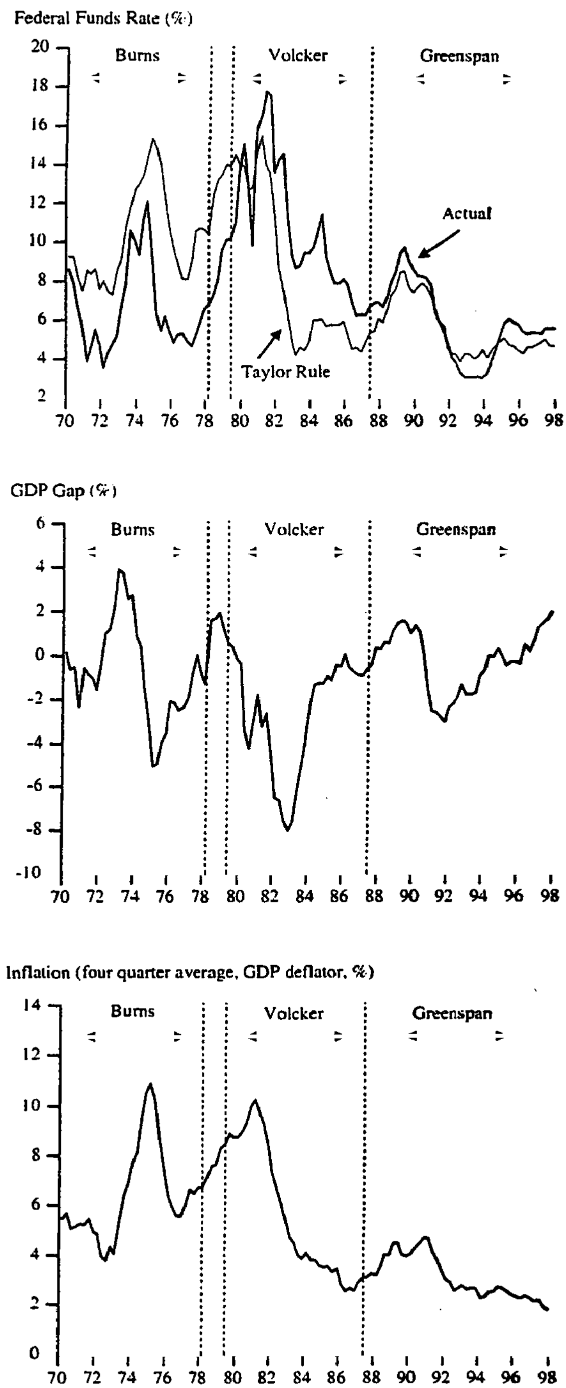


Figure 1 The Taylor rule and its components.

both variables lead to a higher level of the recommended funds rate. In 1979, for example, the rule recommended a high funds rate mainly because inflation was quite high, and to a lesser extent, because real GDP exceeded its potential level by a small amount.

As shown in Figure 1, the original Taylor rule fits reasonably well to the actual funds rate during the Greenspan period. It captures the major swings in the funds rate over the period, but with less amplitude. The R^2 for the period is 87 percent for quarterly levels of the nominal funds rate, and 52 percent for quarterly changes. The arguments in the rule—inflation and the GDP gap—roughly correspond with goals legislated for U.S. monetary policy—stable prices and full employment. In this spirit, Governor Meyer (1998) stresses that stabilizing real GDP around its trend in the short run and controlling inflation in the longer term are important concerns of the Fed. Although U.S. policymakers look at many economic and financial indicators, the two gaps specified in the rule may be highly stylized measures of important short- and long-run concerns. Also, the GDP gap can be interpreted not only as a measure of business cycle conditions, but also as an indicator of future inflation in the context of a Phillips curve model. Measures of the productive capacity of the U.S. economy, whether measured by potential GDP, industrial capacity, or the "natural" rate of unemployment, appear to figure prominently in Fed forecasts of future inflation (Greenspan 1995).[6] Overall, by focusing on policy responses to the Fed's basic goal variables, the Taylor rule implicitly captures policy responses to the many economic factors that affect the evolution of those goal variables.[7]

Judd and Trehan (1995) argue that the Taylor rule also provides some perspective on policies during the Burns and Volcker periods. With regard to the Burns period, although the movement of the actual funds rate was highly correlated with the rule's prescriptions, the funds rate itself was consistently lower than the rule's recommended rate (Figure 1). This result is consistent with the overall increase in inflation during this period, and it confirms that the rule, with its explicit 2 percent inflation target, might have held inflation to a much lower level than policy actually did. During the Volcker period, when the Fed significantly reduced inflation, the funds rate was consistently higher than what the rule recommended, suggesting that the Fed was more aggressive in reducing inflation than the rule would have been.

ISSUES IN ESTIMATING TAYLOR'S RULE

The original Taylor rule appears to provide a rough description of policy during the Greenspan period, as well as a useful benchmark for discussing the policy regimes in place during the Burns and Volcker periods. While the original rule provides a reasonable starting point, this section examines alternatives to Taylor's simple specification by econometrically estimating the reaction function weights, rather than simply choosing parameters as Taylor did. Estimating Taylor-type equations may provide a better description of Fed policy. We consider several issues in estimating a reaction function based on the Taylor rule, including the specification of dynamics, the equilibrium real rate, the inflation target, and the output gap.[8]

Estimating the Taylor Rule with Dynamics

It is fairly straightforward to estimate a Taylor rule as in equation (1). Simply replace the rule-based recommended nominal funds rate with the historical series, add a residual error term to capture deviations from the rule, and estimate the weights as coefficients. One complication to this procedure is that central banks often appear to adjust interest rates in a gradual fashion—taking small, distinct steps toward a desired setting (see, e.g., Rudebusch 1995). We allow for such interest rate smoothing by estimating the Taylor rule in the context of an error correction model. This approach allows for the possibility that the funds rate adjusts *gradually* to achieve the rate recommended by the rule.[9]

In our specification, we replace equation (1) with

$$i_t^* = \pi_1 + r^* + \lambda_1 (\pi_t - \pi^*) + \lambda_2 y_t + \lambda_3 y_{t-1} \tag{2}$$

where i^*_t is explicitly denoted as the recommended rate that will be achieved through gradual adjustment. Also, equation (2) includes an additional lagged gap term along with the contemporaneous gap. This is a general specification that allows for the possibility that the Fed responds to a variety of variables proposed as reasonable monetary policy targets, including inflation alone ($\lambda_2 = \lambda_3 = 0$, as in Meltzer 1987), nominal GDP growth ($\lambda_1 = \lambda_2 = -\lambda_3$, as in McCallum 1988), inflation and real GDP growth with different weights ($\lambda_1 \neq \lambda_2 = -\lambda_3$), as well as inflation and the GDP gap in level form (as in Taylor 1993).

The dynamics of adjustment of the actual level of the funds rate to i^*_t are given by

$$\Delta i_t = \gamma(i^*_t - i_{t-1}) + \rho\Delta i_{t-1}. \tag{3}$$

That is, the change in the funds rate at time t partially corrects the "error" between last period's setting and the current recommended level (the first term), as well as maintaining some of the "momentum" from last period's funds rate change (the second term).[10]

By substituting equation (2) into (3), we obtain the equation to be estimated:

$$\Delta i_t = \gamma\alpha - \gamma i_{t-1} + \gamma(1 + \lambda_1)\pi_t + \gamma\lambda_2 y_t + \gamma\lambda_3 y_{t-1} + \rho\Delta i_{t-1}, \tag{4}$$

where $\alpha = r^* - \lambda_1\pi^*$. This equation provides estimates of the weights on inflation and output in the rule and on the speed of adjustment to the rule.

Determining r* and π*

As is clear from equation (4), our estimation cannot pin down *both* the equilibrium real funds rate (r^*) and the inflation target (π^*) simultaneously. These two terms are combined in the constant term (α) and cannot be identified separately. The economics of this lack of identification are clear in the original Taylor rule of equation (1): The contemporaneous arithmetic effect on the recommended policy rate is the same for a 1 percentage point increase in r^* and for a 2 percentage point decrease in π^*. If both of these magnitudes are unknown, then neither can be individually identified from the estimate of the single parameter α.

Of course, if we assume a particular value for the equilibrium funds rate, then, through the estimates of α and λ_1, we can obtain an estimate of the inflation target. Conversely, an assumption about the inflation target can yield an estimate of the equilibrium rate. Table 1 sheds some light on plausible estimates of these quantities. One simple benchmark for the equilibrium real funds rate is the average real rate that prevailed historically over periods with a common start and end inflation rate.[11] As shown in the first column of Table 1, over a long sample from the early 1960s to the present, inflation edged up, on net, only slightly, while the real funds rate averaged 2.39 percent, which appears to be in the range of reasonable estimates.[12] During the Greenspan period (column 2), the real rate averaged 2.82, which is a bit higher. However, this higher level is consistent with the fact that

Table 1 Interest Rates and Inflation

	Long sample (61.Q1–97.Q4)	Greenspan (87.Q3–97.Q4)	Volcker (79.Q3–87.Q2)	Burns (70.Q1–78.Q1)
Average real interest rate (%)	2.39	2.82	5.35	0.02
Percentage point change in inflation	0.38	−1.32	−5.81	1.23
Average inflation (%)	4.38	3.03	5.35	6.47
End-of-sample inflation (%)	1.77	1.77	3.07	6.69

NOTE: The change in inflation (in percentage points) is calculated as the difference in four-quarter inflation from the first quarter to the last quarter of the sample. End-of-sample inflation is average inflation over the final four quarters of the sample. Inflation is measured as the four-quarter change in the GDP deflator, and the interest rate is the federal funds rate.

inflation fell more than 1 percentage point during the Greenspan sample while it rose slightly during the long sample. Certainly, Taylor's (1993) suggestion that 2 percent was a reasonable guess for the value of the equilibrium rate during the Greenspan period seems plausible. It is more difficult to pin down the equilibrium real rate in the Volcker period. During this period, the real rate averaged over 5 percent, but there was also a large decline in inflation, so this average rate is likely much higher than the equilibrium real rate.[13] Conversely, real rates averaged about zero during the Burns period, but during this time inflation and inflationary pressures were rising, so the equilibrium rate was most likely higher than the average.

It is less clear how to obtain implicit inflation targets from the historical data. Table 1 provides the average levels of inflation over the various samples. However, given the persistence of inflation, the assumption that the target inflation rate of policymakers has been achieved on average in various samples seems less plausible than the assumption regarding the real funds rate (i.e., that cyclical fluctuations have averaged out over time). Policymakers, notably in the early part of the Volcker period, could have "inherited" persistent inflation rates much different from their own target rate, which could then skew their sample averages. More interesting perhaps is the end-of-sample inflation rate, which gives a reading on what policymakers were able to achieve by the end of their tenure. Note that this rate for Greenspan is close to the 2 percent target assumed in Taylor (1993).

As this discussion should make clear, there is much uncertainty in choosing values for r^* and π^*.[14] Therefore, we will show results below under a variety of assumptions about these magnitudes.

Estimating Potential Real GDP

One final issue to consider is the specification of the real output gap, which is defined as the percentage difference between real GDP and potential GDP. Potential output is unobserved and must be estimated. An atheoretical method to do this is to fit a trend to the data—again, on the assumption that over time cyclical fluctuations average out. For example, Taylor (1993) simply used a linear trend of log real GDP over a short sample period (1984–1992) as a proxy for potential output. One also could use a segmented linear trend (following Perron 1989) or a quadratic trend (as in Clarida, Gertler, Gali 1997a, 1997b) or other nonstructural methods (see Cogley 1997).

We believe a structural approach to estimating the output gap is more appropriate conceptually than an atheoretic approach, since the presence of output in the policy rule not only may reflect an interest in stabilizing real fluctuations but also may provide policymakers some information on future inflation. The structural approach is also the one typically used by policymakers at the Fed and elsewhere. In this paper, we use a structural definition of potential GDP that was developed at the Congressional Budget Office (1995). It is denoted Y^*, and the associated gap is shown in Figure 1.[15] This measure of potential output is not a simple fitted GDP trend, but is estimated in terms of a relationship with future inflation similar to the way a time-varying NAIRU is estimated within the context of a Phillips curve. We examine the robustness of our results to alternative measures of the gap in the Appendix.

ESTIMATES OF REACTION FUNCTIONS

Our main hypothesis is that taking account of changes in Fed Chairmen helps to account for changes in the Fed's reaction function. Accordingly, we conduct Chow tests on equation (4) for two breaks during the 1970.Q1–1997.Q4 period corresponding to the terms of Chairmen Burns, Volcker, and Greenspan (the Miller term, 1978.Q2–1979.Q2, was excluded). While a finding of significant breaks in the data would not be strong evidence in favor of our hypothesis, it would at least be a reasonable initial step that should be taken before proceeding to estimate separate reaction functions for those periods.

The test gives the null hypothesis of no structural change a p-value of 0.00 (i.e., it rejects stability at the 0.00 percent level of significance). In addition, we looked at the Burns/Volcker period and tested for a break between their terms, and similarly at the Volcker/Greenspan period and tested for a break between their terms. These tests rejected stability at significance levels of 0.00 and 0.07, respectively.

In the remainder of this section, we present three exhibits that detail the estimates of separate reaction functions for each of the three Chairmen. We estimate the basic equation (equation (4)) using OLS and then re-estimate the equation after eliminating insignificant terms.[16]

The Greenspan Period, 1987.Q3–1997.Q4

The lagged gap is insignificant in A (see Figure 2), the basic regression, so it is eliminated in B. Regression B explains 71 percent of the quarterly variation in the change in the funds rate (with an adjusted $R^2 = 0.67$), and has a standard error of 27 basis points. Not surprisingly, this regression has a closer fit with the data than Taylor's original specification.[17]

Several interesting issues arise from regression B. First, the estimates suggest gradual, rather than instantaneous, adjustment of the funds rate to the rule. The funds rate typically adjusts enough to eliminate 28 percent of the difference between the lagged actual and rule-recommended funds rate each quarter. Second, the estimated weight on the GDP gap of 0.99 is higher than Taylor assumed (0.50). In this regard, some researchers have found that a larger weight on the output gap than Taylor assumed produces a lower output variance for a given inflation variance in model simulations (e.g., Rudebusch and Svensson 1998 and Williams 1997).

	α	γ	λ_1	λ_2	λ_3	ρ	$\bar{R}^2$	SEE	Q
A	1.21	0.27	0.57	1.10	−0.12	0.43	0.67	0.27	20.32
	(1.79)	(4.87)	(2.72)	(2.83)	(−0.31)	(3.94)			(0.03)
B	1.31	0.28	0.54	0.99		0.42	0.67	0.27	20.78
	(2.26)	(5.27)	(2.96)	(7.46)		(4.00)			(0.02)

GREENSPAN REGRESSION INFORMATION

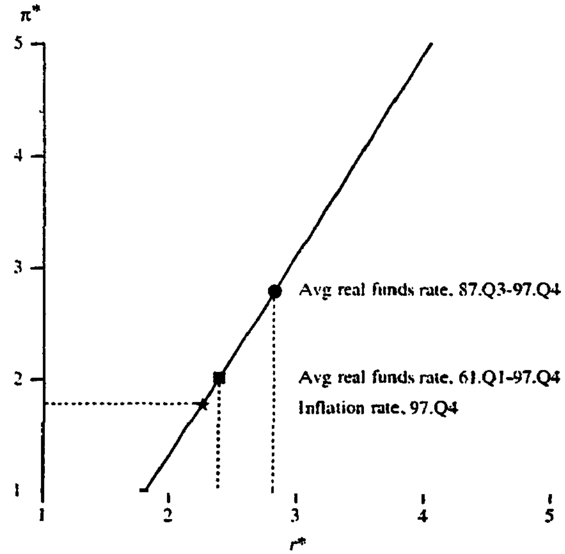


Figure 2 Regression results for Greenspan period, 1987.Q3–1997.Q4.

Finally, the data provide a fairly narrow range of estimates of the equilibrium real funds rate and the inflation target. The various estimates of the equilibrium funds rate and the inflation target that are consistent with the estimated constant term can be seen in Figure 2. The average long-sample and Greenspan-sample real funds rates and the end-of-sample inflation rate are consistent with a fairly tight range of tradeoffs on the line. The estimates of both the inflation target and the real equilibrium funds rate all lie in a range from 1.8 to 2.8 percent—not far from Taylor's assumption of 2 percent.

The Volcker Period, 1979.Q3–1987.Q2

As with the prior regression, estimation of the basic equation (A) finds evidence of partial adjustment of the funds rate to the rule (see Figure 3). However, the dynamic pattern is somewhat different in that the lagged dependent variable is not significant; thus, in regression B, we drop this term. Regression A also suggests that the Volcker period involved a response to the change in, rather than the level of, the GDP gap. Indeed, this restriction cannot be rejected at any conventional significance level. In regression B, the change in the GDP gap enters. The coefficient on the inflation gap in B is very close to

	α	γ	λ_1	λ_2	λ_3	ρ	$\bar{R}^2$	SEE	Q
A	2.04	0.36	0.69	2.40	−2.04	−0.08	0.47	1.33	10.75
	(0.87)	(2.25)	(1.32)	(1.35)	(−1.42)	(−0.44)			(0.22)
B	2.42	0.44	0.46	1.53*	−1.53*	—	0.48	1.31	9.43
	(1.56)	(3.64)	(1.79)	(1.92)	(−1.92)				(0.31)

* restriction: $\lambda_2 + \lambda_3 = 0$

Volcker Regression Information

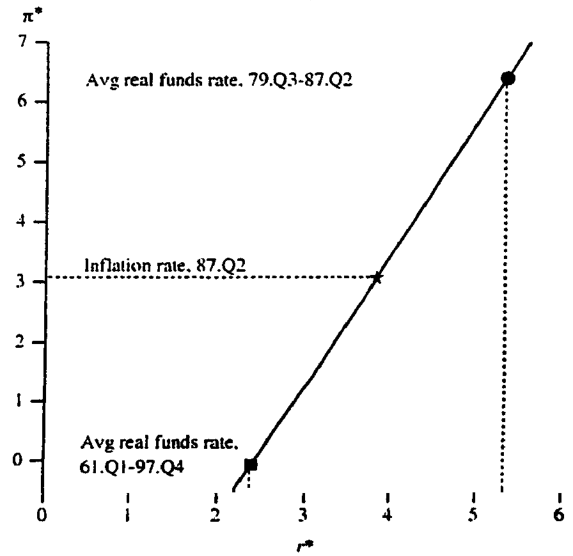


Figure 3 Regression results for Volcker period, 1979.Q3–1987.Q2.

the 0.5 assumed by Taylor, although the estimated coefficient is only very marginally statistically significant. Overall, our results suggest that policy was concerned with the rate of inflation relative to a target and with the growth rate of real GDP relative to the growth rate of potential GDP.

However, the equation is estimated with much less precision for the Volcker period than for the Greenspan period. The coefficients on λ_1, λ_2, and λ_3 are individually significant only at the 6 to 8 percent level (although they are jointly significant at the 1 percent level), and the standard error is 1.31 percentage points compared to 0.27 percentage point in the Greenspan period. In part, this could be because the policy problem in 1979 was so clear: double-digit inflation then was so far above any reasonable inflation target that policy did not need to be as concerned with the rather refined judgments about funds rate settings provided by a Taylor-style reaction function. Instead, policy could make gross judgments about keeping the real funds rate at a "high" level until inflation began to come down. Alternatively, the imprecision could reflect noisy movements in the funds rate under a nonborrowed reserves operating procedure.

As shown in Figure 3, we obtain a wider range of estimates for the implicit inflation target during the Volcker period than during the Greenspan period. This occurs because the average real funds rate during Volcker's tenure (5.35 percent) differs substantially from the average over the entire sample (2.39 percent). The corresponding estimates of the inflation target range from 6.4 percent to–0.1 percent. These estimates bracket the end-of-Volcker-sample inflation rate of 3.07 percent, which corresponds to an r^* of 3.8 percent. The initial tightening of monetary policy could be justified by any of these inflation targets, since inflation was almost 9 percent at the beginning of the Volcker period. Thereafter, it is not possible to tell if the high real funds rates (relative to the Greenspan period) reflects a very low inflation target or a belief that the equilibrium real interest rate was unusually high, possibly because of a perceived need to offset the effects of highly expansionary fiscal policy.

The Burns Period, 1970.Q1–1978.Q1

A key feature of Figure 4 is the insignificance of the coefficient on the inflation gap in the general regression A. Note that this does not mean that inflation considerations are entirely absent from the regression for the Burns period. As is clear in equation (2), even when $\lambda_1 = 0$, the nominal funds rate is affected by movements in inflation; however, these movements are simply the one-for-one changes that are necessary to hold the level of the real funds rate unchanged in the face of changes in inflation. Thus, the regressions suggest that the *real* funds rate was *not* adjusted on the basis of changes in inflation.

The lack of a response of the real funds rate to deviations between inflation and an inflation target will be a critical failing for a monetary policy rule. Without the "anchor" of an inflation target to moor the economy, nominal quantities, like inflation and aggregate demand, will be allowed to drift. Indeed, the lack of an implicit inflation target appears to be consistent with the increase in inflation during the Burns period (Figure 1). Of course, other factors may have played a role as well. In particular, there were two large oil shocks in the Burns period. These events no doubt contributed to the inflationary problems of the period, although a consistent policy response to these inflation shocks most likely would have reduced their effects. We address this issue in more detail in the next section.

When the insignificant inflation and contemporaneous gap terms are dropped, we obtain regression B, which shows partial adjustment of the funds rate to a rule that includes only the lagged GDP gap.[18] Since the inflation gap is not in the regression, the constant term (α) provides an estimate of the equilibrium real funds rate (r^*) implicit in Fed policy during the period. (This can be seen in equation (4) by setting $\lambda_1 = 0$, which makes $\alpha =$

	α	γ	λ_1	λ_2	λ_3	ρ	$\bar{R}^2$	*SEE*	*Q*
A	1.68	0.56	−0.15	0.16	0.72	0.25	0.53	0.84	15.92
	(1.43)	(4.34)	(−0.80)	(0.45)	(2.59)	(1.67)			(0.04)
	$\alpha = r^*$								
B	0.71	0.58	—	—	0.89	0.26	0.52	0.84	14.81
	(2.68)	(4.78)			(5.85)	(1.76)			(0.06)

Figure 4 Regression results for Burns period, 1970.Q1–1978.Q1.

r^*.) One interpretation of this estimate (i.e., that $r^* = 0.7$ percent) is that policy was predicated on the belief that the equilibrium real funds rate was well below postwar experience in the U.S.

A perhaps more plausible interpretation is that the level of the output gap prevailing at the time was consistently mis-estimated during the Burns period. If, for example, the *average* level of the output gap were estimated to be around $1\frac{1}{2}$ percentage points lower than our current estimate for that period, then the estimates in regression B would be consistent with an *average* equilibrium funds rate of around 2 percent. The existence of such large mistakes in the contemporaneous estimates of the output gap have been given an important role during the period by many analysts (e.g., Blinder 1979, p. 35). Such a consistent string of mistakes would not be too surprising. During the Burns period, productivity and potential output both exhibited a surprising (and still largely unexplained) slowdown in growth, and demographic factors, especially the entrance of the baby boom generation into the labor force, conspired to create an increase in the natural rate of unemployment that also was unexpected. Indeed, during the Burns period, there was a widespread view that an unemployment rate of 4 to 5 percent was a suitable benchmark rate for policy. In contrast, recent (time-varying) estimates of the natural rate that prevailed during the Burns period are in the 6 percent range (e.g., Gordon 1997). Such a difference could account for the consistently easy policy during the Burns period. (With an Okun's Law coefficient of 2, the unemployment gap error translates into an underestimation of the output gap on the order of 2 to 4 percent, which would put the funds rate too low.)

MODEL-BASED EVALUATION OF ALTERNATIVE POLICY RULES

It has become common to evaluate the effectiveness of policy rules or reaction functions like the ones estimated above in terms of the volatility of inflation and output that might result if the rule were used by policymakers. Estimates of this volatility can be obtained from simulations of macro models that include the rule to be evaluated (or by similar analytical methods). See, for example, Rudebusch and Svensson (1998) and Levin, Wieland, and Williams (1997) for recent examples. While exercises of this type can provide useful information for evaluating alternative rules, they are not likely to provide conclusive answers. The results depend upon the particular model employed in the analysis, and because there is no single consensus model in macroeconomics, results from any one model will be subject to debate. (Also, in many cases, the relative rankings of alternative rules are not clear because a tradeoff exists between a rule that has a lower inflation variance and another rule that has a lower real GDP variance.)

As an initial step in evaluating the reaction functions estimated in this paper, we have used a simple model from Rudebusch and Svensson (1998). It includes an aggregate supply equation (or "Phillips curve") that relates inflation to the output gap:

$$\tilde{\pi}_{t+1} = 0.68\tilde{\pi}_t - 0.09\tilde{\pi}_{t-1} + 0.29\tilde{\pi}_{t-2} + 0.12\tilde{\pi}_{t-3} + 0.15y_t + \varepsilon_{t+1}, \tag{5}$$

(where $\tilde{\pi}_t$ is the quarterly, not four-quarter, inflation rate) and an aggregate demand equation (or "IS curve") that relates output to a short-term interest rate:

$$y_{t+1} = 1.17y_t - 0.27y_{t-1} - 0.09(\bar{i}_t - \pi_t) + \eta_{t+1} \tag{6}$$

(where $\bar{i}_t$, is the average funds rate over the current and prior three quarters). This simple model produces transparent results, captures the spirit of many practical policy-oriented

Table 2 Model-Based Volatility Results

Monetary policy reaction function	Standard Deviation	
	π_t	y_t
Taylor rule	3.86	2.23
Greenspan period	3.87	2.18
Volcker period	4.80	2.73
Burns period	Does not converge	

macroeconomic models, and fits the data quite well.[19] In addition to these equations, the estimated reaction functions for the three periods were included one at a time (as well as the original Taylor rule), and the unconditional standard deviations of inflation and the output gap were calculated.

The results are presented in Table 2. In the Rudebusch-Svensson model, the estimated reaction function for the Greenspan period has an advantage over the function for the Volcker period: the former function produces a lower standard deviation for the real output gap and about the same standard deviation for (four-quarter) inflation. However, we would not want to emphasize this comparison too much because the differences are not large and may be reversed in a different model. The function for the Greenspan period produces about the same volatility of both inflation and real GDP as the original Taylor rule.

The results for the Burns period seem more telling, since the model did not converge when that reaction function was included. This dynamic instability reflects the fact that inflation is not anchored in the Burns period. This result is likely to show up in a variety of models when the reaction function for the Burns period is used. Indeed, Clarida, Gali, and Gertler (1997a) use a calibrated forward-looking model to show that their estimated pre-1979 reaction function is unstable.

The contrast between the three estimated reaction functions is demonstrated in Figure 5 with counterfactual simulations of the Burns period. These are simulations of equations (5) and (6) along with, in turn, the Burns, Volcker, and Greenspan reaction functions. The actual historical shocks to equations (5) and (6) in the Burns period are used, so in all three cases inflation is pushed up by unfavorable shocks. Still, the difference between the Burns reaction function and the other two is striking, for only with the Burns reaction function does inflation remain at a high level.

CONCLUSIONS

The estimates in this paper indicate that a Taylor-type reaction function seems to capture some important elements of monetary policy during Alan Greenspan's tenure to date as Federal Reserve Chairman. This regression implies that movements in the funds rate-over that period have been broadly consistent with a policy regime aimed at low inflation in the long run and a stable level of output around trend in the short run. However, the results differ somewhat from Taylor's original specification of the rule in two main ways. The funds rate appears to have reacted about twice as strongly to the GDP gap as Taylor assumed, and it appears to have moved gradually, rather than instantaneously, into rough accord with the estimated Taylor rule.

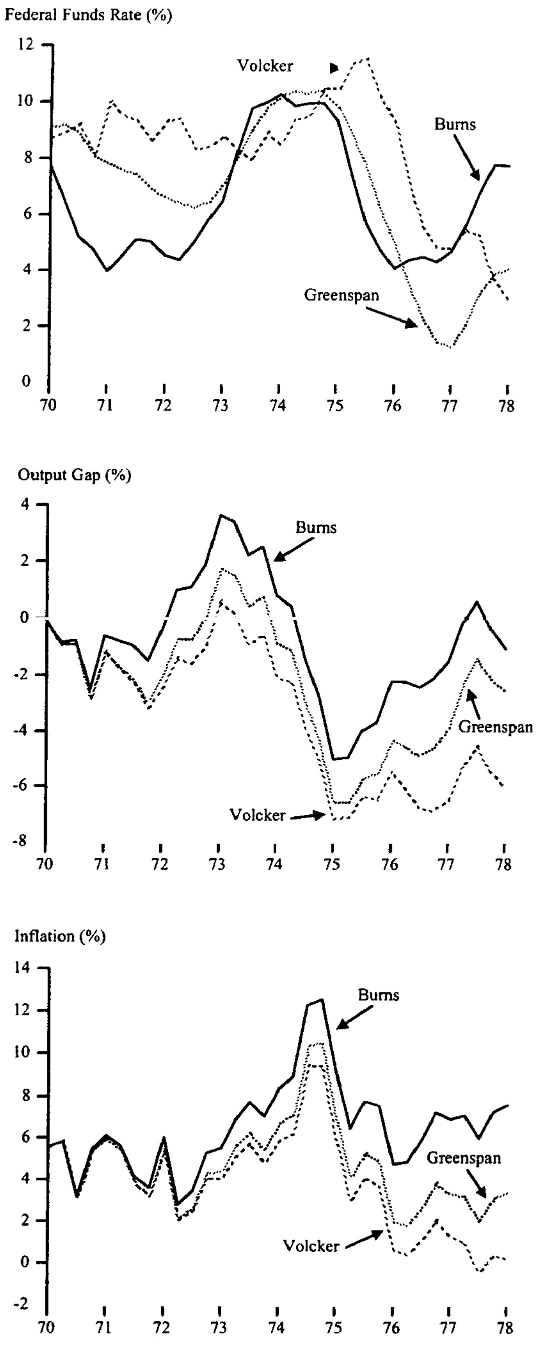


Figure 5 Counterfactual simulations: 1970.Q1–1978.Q1.

The estimates for the Volcker period are less precise than those for the Greenspan period. Nonetheless, they suggest that the Fed adjusted the funds rate gradually in response to concerns with achieving an inflation target well below the rate inherited by the FOMC in the late 1970s. This result is consistent with the substantial progress achieved in reducing inflation during the period. Policy also appears to have given weight to cyclical considerations, but this concern came in the form of reactions to the growth rate rather than to the level of real GDP.

In the Burns period, we find a weak policy response to inflation. Instead, policy seems to have been geared mainly toward gradual responses to the state of the business cycle. Moreover, some evidence suggests that policy either was oriented around an unusually low estimate of the equilibrium real funds rate or around an estimate of potential output that appears to have been too high in retrospect. These results seem consistent with the key feature of Burns's tenure as Chairman of the Fed—rising inflation—and they appear to show up as dynamic instability in our model simulations.

Overall, the dynamic Taylor-type reaction functions estimated during the Burns, Volcker, and Greenspan periods, appear to have differed in important ways from one another. As noted above, this investigation has not provided a complete representation of changes in Fed behavior, in part because we have controlled for only one source of sample instability. This may account for the sensitivity of some of the results to alternative specifications as shown in the Appendix. However, we hope our results represent a step in the direction of uncovering the key elements—and changes—in Federal Reserve behavior over time.

The finding that the monetary policy regime may have changed in significant ways over time has implications for at least two strands of literature in macroeconomics. First, the finding raises questions about attempts to estimate monetary policy shocks using identified VARs estimated over long sample periods. If the implicit reaction functions in these VARs do not properly capture the changes in the way policy was formulated, then the estimated shocks will not properly measure the "surprises" in policy. Thus, our results reinforce the conclusions of Rudebusch (1998) that such VARs may be misspecified. Second, in macroeconomic models with rational expectations, parameters throughout the models depend upon the monetary policy regime in place. If the policy regime has changed frequently in the postwar period, it may be difficult to obtain good estimates of these rational expectations models, in part because we may not have long enough sample periods under a consistent policy regime.

APPENDIX: ALTERNATIVE SPECIFICATIONS

We examine the robustness of the results presented in the text by looking at regressions using alternative measures of inflation and the GDP gap. The results are presented in Table A1. With regard to inflation, the estimated regressions show little sensitivity to these alternative measures.

With regard to the GDP gap, we estimate reaction functions using three estimates of potential GDP, namely, Y^*, which is described in the text, a segmented linear trend with one break in 1973.Q1, and a quadratic trend. Figure A1 shows the alternative estimates of potential output and the corresponding GDP gaps. The GDP gap measured in terms of Y^* has cross-correlations of 0.99 and 0.80 with the quadratic and linear trend gaps, respectively. A recent example of a divergence among these series occurred in the

Table A1 Reaction Functions—Alternative Specifications Inflation

	α	γ	λ_1	λ_2	λ_3	ρ	r^*	$\bar{R}_2$	*SEE*	*Q*
Greenspan										
GDP deflator	1.31	0.28	0.54	0.99	—	0.42	—	0.67	0.27	20.78
	(2.26)	(5.27)	(2.96)	(7.46)	—	(4.00)	—	—	—	(0.02)
PCE price index	2.39	0.23	0.07	1.02	—	0.44	—	0.62	0.29	12.12
	(3.56)	(4.30)	(0.37)	(5.66)	—	(3.89)	—	—	—	(0.28)
Core CPI	1.00	0.25	0.37	1.15	—	0.49	—	0.64	0.28	16.24
	(1.28)	(4.56)	(1.79)	(7.20)	—	(4.47)	—	—	—	(0.09)
Volcker										
GDP deflator	2.42	0.44	0.46	1.53	−1.53	—	—	0.48	1.31	9.43
	(1.56)	(3.64)	(1.79)	(1.92)	(−1.92)	—	—	—	—	(0.31)
PCE price index	1.46	0.29	0.54	2.58	−2.58	—	—	0.40	1.41	9.06
	(0.52)	(2.59)	(1.21)	(1.65)	(−1.65)	—	—	—	—	(0.34)
Core CPI	1.32	0.35	0.35	2.57	−2.57	—	—	0.43	1.37	8.70
	(0.57)	(3.03)	(1.11)	(2.07)	(−2.07)	—	—	—	—	(0.37)
Burns	$\alpha = r^*$									
GDP deflator	0.71	0.58	—	—	0.89	0.26	0.71	0.52	0.84	14.81
	(2.68)	(4.78)	—	—	(5.85)	(1.76)	(2.68)	—	—	(0.06)
PCE price index	0.95	0.51	—	—	0.86	0.17	0.95	0.48	0.87	16.97
	(2.98)	(4.33)	—	—	(4.67)	(1.14)	(2.98)	—	—	(0.03)
Core CPI	1.30	0.56	—	—	1.14	0.34	1.30	0.40	0.94	10.72
	(4.17)	(3.46)	—	—	(6.12)	(1.99)	(4.17)	—	—	(0.22)
Potential GDP										
Greenspan										
Y^*	1.31	0.28	0.54	0.99	—	0.42	—	0.67	0.27	20.78
	(2.26)	(5.27)	(2.96)	(7.46)	—	(4.00)	—	—	—	(0.02)
Segmented	5.35	0.8	−0.99	0.90	—	0.54	—	0.58	0.31	10.71
linear trend	(3.52)	(3.56)	(−1.93)	(3.99)	—	(4.88)	—	—	—	(0.38)
Quadratic trend	1.09	0.28	0.37	0.82	—	0.52	—	0.64	0.28	19.10
	(1.80)	(4.77)	(1.90)	(7.05)	—	(4.94)	—	—	—	(0.04)
Volcker										
Y^*	2.42	0.44	0.46	1.53	−1.53	—	—	0.48	1.31	9.43
	(1.56)	(3.64)	(1.79)	(1.92)	(−1.92)	—	—	—	—	(0.31)
Segmented	2.29	0.43	0.48	1.57	−1.57	—	—	0.48	1.31	9.28
linear trend	(1.39)	(3.46)	(1.75)	(1.84)	(−1.84)	—	—	—	—	(0.32)
Quadratic trend	2.23	0.43	0.50	1.60	−1.60	—	—	0.48	1.31	9.23
	(1.33)	(3.46)	(1.78)	(1.84)	(−1.84)	—	—	—	—	(0.32)
Burns	$\alpha = r^*$									
Y^*	0.71	0.58	—	—	0.89	0.26	0.71	0.52	0.84	14.81
	(2.68)	(4.78)	—	—	(5.85)	(1.76)	(2.68)	—	—	(0.06)
Segmented	1.54	0.52	—	—	1.07	0.25	1.54	0.46	0.89	13.88
linear trend	(3.76)	(4.15)	—	—	(4.83)	(1.59)	(3.76)	—	—	(0.09)
Quadratic trend	−0.15	0.59	—	—	0.88	0.26	−0.15	0.52	0.84	14.57
	(−0.58)	(4.78)	—	—	(5.90)	(1.78)	(−0.58)	—	—	(0.07)

1990s, when the segmented linear trend showed output consistently below potential, while the other two measures showed a rising gap that became positive toward the end of the sample. Differences like these can have noticeable effects on Fed policy concerns. The regression results for the reaction function using the linear trend differ from those using the other gap measures in Table A1. In fact, in the Greenspan period, the introduction of the linear trend actually changes the sign of the response to the inflation gap. The alternative measures of the gap have little effect on the results for the Burns or the Volcker periods.

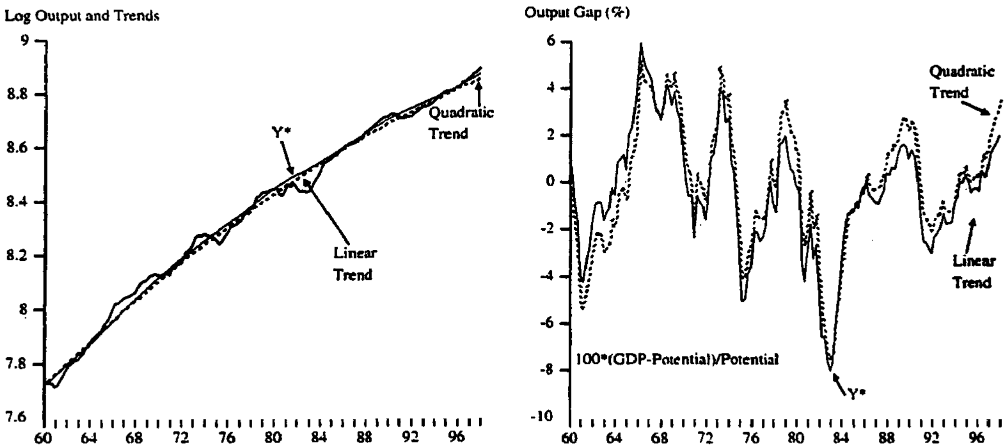


Figure A1 Alternative estimates.

ACKNOWLEDGMENTS

We thank Kenneth Kasa, Bharat Trehan, and Carl Walsh for useful comments and Alison Wallace for excellent research assistance; they are not responsible for any errors or omissions.

NOTES:

1. However, attempts to avoid the mistakes of the past sometimes may lead to new mistakes. De Long (1997, p. 250) argues that "... at the deepest level, the truest cause of the inflation in the 1970s was the shadow cast by the great depression . . . ."
2. De Long (1997, p. 274) argues, "A mandate to fight inflation by inducing a significant recession was in place by 1979, as a result of a combination of fears about the cost of inflation, worry about what the 'transformation of every business venture into a speculation on monetary policy' was doing to the underlying prosperity of the American economy, and fear that the structure of expectations was about to become unanchored and that permanent double-digit inflation was about to become a possibility."
3. The 1977 amendment to the Federal Reserve Act requires the Fed to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." The Humphrey-Hawkins Act of 1978 affirms the responsibility of the federal government in general to promote "full employment and production. . . . and reasonable price stability," among other things.
4. Taylor (1993) used a log linear trend of real GDP over 1984.Q1 to 1992.Q3 as a measure of potential GDP. As discussed below, we have used a more flexible structural estimate.
5. For a complementary analysis, see Taylor (1997).
6. Given the lags in the monetary transmission mechanism, an explicitly forward-looking version of the Taylor rule—with inflation and output *forecasts* as arguments—also might be appropriate. Clarida, Gali, and Gertler (1997a, 1997b) estimate a rule using inflation forecasts and ob-

tain results similar to our own, and Rudebusch and Svensson (1998) examine the theoretical properties of such a rule.

7. The Taylor rule has gained the attention of some Fed policymakers (Blinder 1996, *Business Week* 1996, Meyer 1998, and Yellen 1996), who have used it as a helpful, broad characterization of U.S. monetary policy. In addition, it has gained some acceptance outside the Fed as a way to think about how the Fed might react to economic and inflationary developments (Prudential Economics 1996, and Salomon Brothers 1995a, 1995b). Of course, there are always questions about the reliability of any current implications of the rule because of uncertainty about the level of potential GDP. Some analysts argue that increased productivity, due to computer and other technological developments, means that potential output is being mis-measured. See Trehan (1997) for a discussion of the debate about productivity.
8. We use current data throughout this paper. It would be preferable to use the original data that policymakers actually were looking at when decisions about the funds rate were being made. Unfortunately, we do not have access to these data for our full 1970–1997 sample period. See Orphanides (1997) for an analysis of the effects of original versus final data in estimating a Taylor rule for the 1987–1992 period.
9. Mehra (1994) employs a similar dynamic specification.
10. We think that this "error correction" framework is a useful one for the consideration of dynamics. However, although the funds rate, the output gap, and the inflation rate are highly persistent, we make no claims that they are nonstationary (consistent with Rudebusch 1993).
11. This is analogous to using the average unemployment rate over periods with no net change in inflation to estimate a constant "natural" rate of unemployment (or NAIRU).
12. The real rates in Table 1 are calculated on an ex post basis as in equation (1), but similar results were obtained using ex ante rates constructed with the one-year-ahead inflation forecasts from the Philadelphia Fed's inflation expectations survey.
13. Still, it is possible that the equilibrium rate was elevated during the Volcker period given the large federal budget deficits. For a model-based definition of a time-varying equilibrium rate, see Bomfim (1997).
14. There is also, of course, the issue of time variation in π^* and r^* (as noted in footnote 13). Even during a given Chairman's term, there may well be changes in the target inflation rate. Indeed, this is the essence of the opportunistic approach to monetary policy described by Bomfim and Rudebusch (1998).
15. This series is conceptually similar and highly correlated with the Q^* series in Braun 1990; Hallman, Porter, and Small 1991; and Orphanides (1997).
16. Given the lags in the transmission process of monetary policy, there is little danger of reverse causation from i_t to π_t and y_t.
17. The Q-statistic suggests the possibility of autocorrelation in the regression. Much of this may be due to our use of time-aggregated data. When we respecified the regression using end-of-quarter funds rate data, the Q-statistic did not show signs of autocorrelation, the lagged change in the funds rate became statistically insignificant, and the other coefficients were close to the results in the original specification. This result adds to our confidence in the specification of the right-hand-side variables in regression B, which we retained in the interest of obtaining an equation that can be used in a quarterly macroeconomic model with quarterly average measurement of the funds rate.
18. This regression shows signs of autocorrelation (the Q-statistic has a p-value of 0.6 percent). As with the Greenspan regression, when the funds rate is defined in terms of the level of the last week in the quarter, rather than as a quarterly average, the coefficients in the equation change very little, but the Q-statistic becomes insignificant.
19. The equations were estimated from 1961.Q1 to 1996.Q4. See Rudebusch and Svensson (1998) for details. The estimates in (5) and (6) differ very slightly from those in that paper because of the longer sample and data revisions.

REFERENCES

Bernanke, B. S., and Blinder, A. S. 1992. The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review* 82, pp. 901–921.

Blinder, A. S. 1979. *Economic Policy and the Great Stagflation.* San Francisco: Academic Press.

Blinder, A. S. 1996. Remarks at the Senior Executives Conference of the Mortgage Bankers Association, New York, N.Y. (January 10).

Bomfim, A. 1997. The Equilibrium Fed Funds Rate and the Indicator Properties of Term-Structure Spreads. *Economic Inquiry* 35(4) pp. 830–846.

Bomfim, A., and Rudebusch, G. 1998. Opportunistic and Deliberate Disinflation Under Imperfect Credibility. Federal Reserve Board, FEDS Working Paper 98-01.

Braun, S. N. 1990. Estimation of Current Quarter Gross National Product by Pooling Preliminary Labor-Market Data. *Journal of Business and Economic Statistics* 8(3) pp. 293–304.

Brayton, F., and Tinsley, P. A. 1996. A Guide to FRB/US: A Macroeconomic Model of the United States. Federal Reserve Board, FEDS Working Paper 96-42.

Business Week. 1996. How Low Should Inflation Be? (October 9) pp. 68–69.

Clarida, R., Gali, J. and Gertler, M. 1997a. Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. Unpublished paper, Columbia University (May).

Clarida, R., Gali, J., and Gertler, M. 1997b. Monetary Policy Rules in Practice: Some International Evidence. NBER Working Paper Series, No. 6254 (November).

Cogley, T. 1997. Evaluating Non-Structural Measures of the Business Cycle. Federal Reserve Bank of San Francisco *Economic Review* 97-03, pp. 3–21.

Congressional Budget Office. 1995. CBO's Method for Estimating Potential Output. *CBO Memorandum* (October).

De Long, J. B. 1997. America's Peacetime Inflation: The 1970s. In *Reducing Inflation,* eds. Christina D. Romer and David Romer. NBER Studies in Business Cycles, Volume 30.

Gordon, R. 1997. The Time-Varying NAIRU and its Implications for Economic Policy. *Journal of Economic Perspectives* 11, pp. 11–32.

Greenspan, A. 1995. Testimony before the Committee on Banking, Housing, and Urban Affairs, United States Senate (February 22).

Hallman, J. J., Porter, R. D. and Small, D. H. 1991. Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run? *American Economic Review* 81, pp. 841–858.

Judd, J. P., and Trehan, B. 1995. Has the Fed Gotten Tougher on Inflation? *FRBSF Weekly Letter* 95-13 (March 31).

Khoury, S. S. 1990. The Federal Reserve Reaction Function: A Specification Search. In *The Political Economy of American Monetary Policy,* ed. Thomas Mayer. Cambridge, England: Cambridge University Press, pp. 27–41.

Levin, A., Wieland, V. and Williams, J. C. 1997. Are Simple Monetary Policy Rules Robust to Model Uncertainty? Unpublished paper, Federal Reserve Board.

McCallum, B. 1988. Targets, Indicators, and Instruments of Monetary Policy. In *Monetary Policy in an Era of Change.* Washington D.C.: American Enterprise Institute (November 16–17).

McNees, S. K. 1986. Modeling the Fed: A Forward-Looking Monetary Policy Reaction Function. *New England Economic Review* (November/December) pp. 3–8.

McNees, S. K. 1992. A Forward-Looking Monetary Policy Reaction Function: Continuity and Change. *New England Economic Review* (November/December), pp. 3–13.

Mehra, Y. P. 1994. A Federal Reserve Reaction Function. Unpublished paper, Federal Reserve Bank of Richmond.

Meltzer, A. 1987. Limits on Short-Run Stabilization Policy. *Economic Inquiry* 25(1) pp. 1–14.

Meyer, L. H. 1998. The Strategy of Monetary Policy. The Alan R. Holmes Lecture, Middlebury College, Middlebury, Vermont (March 16).

Orphanides, A. 1997. Monetary Policy Rules Based on Real-Time Data." Unpublished paper, Board of Governors of the Federal Reserve System.

Perron, P. 1989. The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica* 57, pp. 1361–1401.

Prudential Economics. 1996. Feature: Federal Reserve Policy and the Taylor Rule. *Economic and Investment Analysis,* Volume 12, Number 11 (November) pp. 9–12.

Rudebusch, G. 1993. The Uncertain Unit Root in Real GNP. *The American Economic Review* (83) pp. 264–272.

Rudebusch, G. 1995. Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure. *Journal of Monetary Economics* 24, pp. 245–274.

Rudebusch, G. 1998. Do Measures of Monetary Policy Shocks in a VAR Make Sense? *International Economic Review* 39(4) pp. 907–941.

Rudebusch, G. and Svensson, L. E. O. 1998. Policy Rules for Inflation Targeting. Federal Reserve Bank of San Francisco Working Paper 98-03 (January).

Salomon Brothers, Inc. 1995a. Policy Rules Shed Light on Fed Stance. *Perspectives for Financial Markets* (June 26).

Salomon Brothers, Inc. 1995b. Disinflation's Surprising Progress. *Perspectives for Financial Markets* (December).

Taylor, J. B. 1993. Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, pp. 195–214.

Taylor, J. B. 1997. An Historical Analysis of Monetary Policy Rules. Unpublished paper, Stanford University (December).

Trehan, B. 1997. A New Paradigm? *FRBSF Economic Letter* 97-29 (October 10).

Williams, J. C. 1997. Simple Rules for Monetary Policy. Unpublished paper, Board of Governors of the Federal Reserve System (August 18).

Yellen, J. L. 1996. Remarks at the National Association of Business Economists, Washington D.C. (March 13).

Index