The GSEs: Where Do We Stand?

William Poole

Milton Friedman and U.S. Monetary History: 1961-2006

Edward Nelson

The Lower and Upper Bounds of the Federal Open Market Committee’s Long-Run Inflation Objective

Daniel L. Thornton

Granger Causality and Equilibrium Business Cycle Theory

Yi Wen
The GSEs: Where Do We Stand?
William Poole

Milton Friedman and U.S. Monetary History: 1961-2006
Edward Nelson

The Lower and Upper Bounds of the Federal Open Market Committee's Long-Run Inflation Objective
Daniel L. Thornton

Granger Causality and Equilibrium Business Cycle Theory
Yi Wen
The GSEs: Where Do We Stand?

William Poole

This article was originally presented as a speech to the Chartered Financial Analysts of St. Louis, St. Louis, Missouri, January 17, 2007.


One of the Federal Reserve’s most important responsibilities is maintenance of financial stability. The job obviously, and sometimes dramatically, encompasses crisis response. However, the very existence of a crisis, when one occurs, often demonstrates a failure of some sort, on the part of the firms involved, the government, or the Federal Reserve. It would not be difficult to cite examples of such failures.

Not long after coming to the St. Louis Fed in 1998, I became interested in government-sponsored enterprises, or GSEs. My interest arose when I began digging into aggregate data on the financial markets and discovered how large these firms are. The bulk of all GSE assets are in the housing GSEs—Fannie Mae, Freddie Mac, and the 12 federal home loan banks (FHLBs). Using information as of September 30, 2006—the latest available as of this writing—these 14 firms have total assets of $2.67 trillion; given their thin capital positions, their total liabilities are only a little smaller. Just two firms—Fannie Mae and Freddie Mac—account for $1.65 trillion of the assets, or 62 percent of all housing GSE assets. Moreover, Fannie Mae and Freddie Mac have guaranteed mortgage-backed securities outstanding of $2.82 trillion. Thus, the housing GSE liabilities on their balance sheets and guaranteed obligations off their balance sheets are about $4.47 trillion, which may be compared with U.S. government debt in the hands of the public of $4.83 trillion.

In what follows, I’ll confine most of my comments to Fannie Mae and Freddie Mac, where the largest issues arise. My purpose is to make the case once again that failure to reform these firms leaves in place a potential source of financial crisis. Although there is pending legislation in Congress, a major restructuring of these firms and genuine reform appear to be as distant as ever.

My initial curiosity about the GSEs was stoked simply by the size of these firms. As I investigated further, I became concerned about their thin capital positions and the realization that if any of them got into financial trouble the markets and the federal government would look to the Federal Reserve to deal with the problem. As I worked through the issues, I began to speak on the subject; my first such speech was in October 2001 (Poole, 2001). I last spoke on a GSE topic two years ago, before the St. Louis Society of Financial Analysts. My title then was “GSE Risks” (Poole, 2005).

Given that the risks did not seem likely to disappear any time soon, about six months ago I settled on a GSE topic once again.

Today I want to look back over the past few years to summarize a few of the changes that have occurred at the GSEs and in the regulatory envi-
rnen t they face. It is no exaggeration to say these have been event-filled years for the GSEs, primarily because of disclosures of accounting irregularities at Fannie Mae and Freddie Mac. Although these firms stopped growing when the irregularities were disclosed, I will emphasize that once they get their houses in good order they will likely resume rapid growth because of the special advantages they enjoy in the marketplace from their ties to the federal government. I remain hopeful that Congress will eventually pass meaningful GSE reform legislation. Private sector financial firms ought to have an intense interest in reform legislation. Still, given that there seems to be so little appreciation of the importance of the GSE issue, where do they—and we—go from here?

Before proceeding, I want to emphasize that the views I express here are mine and do not necessarily reflect official positions of the Federal Reserve System. I thank my colleagues at the Federal Reserve Bank of St. Louis for their comments—especially Bill Emmons, senior economist, who provided special assistance.

THE HOUSING GSEs SINCE JUNE 9, 2003

Although the housing GSEs are less obscure than they used to be, they are not much discussed in recent months. A year ago I would have noted that it was not unusual to find stories about the GSEs on the front pages of major financial newspapers. They were the subject of substantial debate in Congress and among financial policy experts. They had escaped from obscurity, primarily because of publicity in recent years over their accounting irregularities. But today they seem to be returning to obscurity.

For Fannie Mae and Freddie Mac, the two stockholder-owned housing GSEs, history can be divided into two distinct eras—before June 2003 and after. June 9, 2003, was the day the board of directors of Freddie Mac announced discovery of significant accounting irregularities. The stock prices of both Freddie Mac and Fannie Mae plunged, as investors immediately realized that something might have gone terribly wrong with both GSEs. Subsequent investigations by private experts and public authorities confirmed the fears of many investors and financial supervisors. These giant, fast-growing firms had poor accounting systems and financial controls.

Because it is important for my analysis later, keep in mind these facts: First, the effect of disclosure of accounting irregularities at Freddie Mac on June 9, 2003, led to a decline of 16 percent in Freddie’s stock price and 5 percent in Fannie’s stock price that day. However, as I’ll document later, the effect of these disclosures on the mortgage market was negligible. Similarly, when Fannie’s accounting irregularities were disclosed on September 22, 2004, its stock fell by 6.5 percent that day and by a total of 13.5 percent over a three-day period; the mortgage rate was again unaffected.

Fortunately for financial stability, the accounting irregularities at Freddie Mac had been designed, as we later learned, to understate earnings by a total of about $9 billion over a period of years. Thus, there was no question of Freddie Mac defaulting on any of its obligations and immediately unleashing unpredictable effects on its counterparties or the financial system. In 2004, we learned that Fannie Mae’s accounting was revealed to be faulty. In December 2006, Fannie restated its earnings for 2002, 2003, and the first half of 2004, revealing that it had overstated its earnings by a total of about $6 billion.

Fannie and Freddie are supervised by the Office of Federal Housing Enterprise Oversight, or OFHEO. In both cases, OFHEO’s early response to disclosure of the accounting irregularities was to declare the enterprises “significantly undercapitalized” because their extremely high leverage makes uncertainty of any kind about the true capital backing of their portfolios a risk to their own safety and soundness, as well as the stability of the financial system. Beginning in the first quarter of 2004, OFHEO required Freddie Mac to hold capital at least 30 percent above the statutory minimum level; OFHEO imposed the identical requirement on Fannie Mae in the third quarter of 2005. In addition, OFHEO required the firms to correct their accounting; undertake a thorough
review of corporate governance, incentives, and compensation; appoint an independent chief risk officer; and refrain from increasing their retained portfolios.

The stunning accounting irregularities at Freddie Mac and Fannie Mae served as wake-up calls both to the GSEs themselves and to the supervisory and legislative communities. Freddie Mac fired virtually all of its top-level management immediately in June 2003 and then, a few months later, fired the new CEO it had hired to replace the original disgraced CEO.\(^1\) Barely a year-and-a-half later, Fannie Mae ejected its own top managers, who had repeatedly declared that, unlike Freddie’s, its own books were clean. The boards of both companies agreed to a series of governance reforms designed to bring the GSEs into line with other large financial firms. Hundreds of millions of shareholder dollars were committed to rebuilding accounting and control systems at both firms. Both firms agreed to restate earnings for the past few years; so massive was this undertaking that neither firm is current on its financial reporting. Freddie did release its annual report for 2005 but, according to its press release of January 5, 2007, may revise its results materially for the first nine months and the third quarter of 2006. Nor is Fannie filing current reports. In December 2006, Fannie filed its Form 10-K for 2004 with the Securities and Exchange Commission (SEC). Currently, investors in common stock or debt obligations issued by both companies rely on partial and incomplete information subject to material revision.

The GSE accounting scandals constituted a rude awakening for OFHEO and Congress. OFHEO was caught napping at Freddie Mac but, to its credit, then identified Fannie Mae’s shortcomings on its own. Once alerted to the problems, OFHEO’s tenacious investigations into wrongdoing at both Freddie Mac and Fannie Mae spurred investigations by the SEC and the Department of Justice. Congressional hearings were held, and GSE reform legislation was passed in oversight committees of both houses of Congress in 2004 and 2005, although no final legislation has been enacted as of this time. I’ll have more to say about reform legislation later, because I think this is an important missing piece of the overall puzzle.

Meanwhile, the FHLBs—the “other housing GSEs”—were enduring accounting and control crises of their own. Two of the twelve FHLBs signed written regulatory agreements in 2004 with their supervisor, the Federal Housing Finance Board (FHFB), to rectify portfolio risk-management deficiencies. Then, in 2005, 10 of the 12 FHLBs failed to meet their agreed deadline to register their stock with the SEC. Like Fannie Mae and Freddie Mac, all of the FHLBs restated their earnings for recent years; all have now returned to timely filing of accounting statements.

So where do we stand? I would characterize the current situation as a period of uneasy waiting. The GSEs have grown much more slowly, and they have been more reticent in public in recent quarters than they had been during the pre-2003 decade. It appears that they want to pursue a low-key strategy while memories of their accounting and control failures gradually fade. Their aim, apparently, is to return to the environment before heightened scrutiny arose in 2003.

### WHAT HAS BEEN ACCOMPLISHED: ANALYSIS OF GSE RISKS

Although I think much more needs to be done, it would be a mistake to believe that nothing useful was done after severe accounting problems surfaced in June 2003. In general terms, the most important achievement is a much broader and better-informed discussion of the risks to financial stability posed by the GSEs.\(^2\) We were fortunate that the GSE accounting and governance scandals did not threaten the immediate solvency of the enterprises and that the problems surfaced when the economy and financial markets were strong.

I will point to six major contributions to the public investigation into, and debate about, the risks posed by the GSEs. There have been other

---

\(^1\) Freddie Mac’s board of directors had misjudged at first how deeply ingrained the internal-control and governance problems were and had hired the former CFO to become the new CEO.

\(^2\) For a more detailed discussion of this topic, see Poole (2005).
contributors, to be sure, but this list provides what I think is a good overview of the issues and what we have learned so far:

- a 2003 study by Dwight Jaffee of interest rate risks run by the GSEs;
- a 2003 study by OFHEO of the potential systemic risks posed by the GSEs;
- a series of testimonies and speeches by Federal Reserve Board Chairman Alan Greenspan;
- a series of research papers prepared by Federal Reserve System staff members;
- the results of a Federal Reserve ad hoc study group investigating counterparty exposures and risks in the over-the-counter interest rate derivatives markets;
- and an economic-capital analysis of Fannie Mae and Freddie Mac prepared by Kenneth Posner, an equity analyst at Morgan Stanley.

These bullet points provide the flavor of some of the recent work on the GSEs. The appendix to this speech provides a brief summary of each of these items and citations.

Considering these results as a whole, we have learned a great deal in recent years about the way the GSEs operate, the risks they are taking and how they attempt to manage them, and what effects the GSEs have on financial markets during normal times as well as during periods of market turbulence. Armed with this knowledge, lawmakers and policymakers are in a much better position to make needed improvements in the statutory and regulatory environment in which the GSEs operate.

THE CASE FOR FUNDAMENTAL REFORM

I continue to believe that the nation would be well-served by turning the GSEs into genuinely private firms, without government backing, implied or explicit. If they bolster their capital, they can function perfectly well as purely private firms.

A key issue for many is whether privatizing Fannie and Freddie would raise mortgage rates paid by borrowers. We now have some solid evidence on how the mortgage market would function if the housing GSEs became fully private firms. A careful econometric investigation by three economists at the Board of Governors last year (Lehnert, Passmore, and Sherlund, 2006, abstract) reached this conclusion: “We find that GSE portfolio purchases have no significant effects on either primary or secondary mortgage rate spreads.” Put another way, the 30-year mortgage rate fluctuates in tandem with the rate on 10-year Treasury bonds and the spread over the Treasury rate is not affected by portfolio purchases by Fannie and Freddie.

Another approach to acquiring evidence on the effects on the mortgage rate of mortgage purchases by Fannie and Freddie is to examine what happened when their portfolios stopped growing in the wake of disclosures of accounting irregularities. Those disclosures led OFHEO to impose 30 percent temporary surcharges on the firms’ required minimum capital levels. Freddie Mac’s capital surcharge was imposed in January 2004, whereas Fannie Mae’s capital surcharge became effective in September 2004.3 To meet the higher capital ratio, the two firms had to do some combination of raising new capital and reducing their portfolios.

The retained portfolios of mortgages and mortgage-backed securities (MBS) held by Fannie and Freddie grew strongly in the years preceding the OFHEO orders. For example, if we look at year-end figures for 2002 and 2003, we see that over the course of 2003 the two firms’ retained portfolios grew by a net of 12.3 percent and, at the end of 2003, they held 22 percent of outstanding mortgages on 1- to 4-family properties. Net growth of their retained portfolios then stopped; over the course of both 2004 and 2005, their total portfolios of mortgages and MBS fell slightly. In 2006, their retained portfolios continued to decline and by the end of the third quarter their portfolios were below year-end 2005. Meanwhile, the total market continued to expand. The com-

3 See www.ofheo.gov/media/pdf/capclass93004.pdf.
bined market share of Fannie and Freddie fell from 22 percent at the end of 2003 to 14 percent at the end of the third quarter of 2006.

What happened to the mortgage spread when the GSEs stopped accumulating ever-larger portfolios? Nothing. Because fixed-rate mortgages are subject to prepayment risk, whereas the 10-year Treasury bond is not, there is a degree of variability of the mortgage spread. But if the cessation of the GSEs’ portfolio growth had made a difference, it surely would have shown up in the data. The annual average of the spread in 2003, before the OFHEO orders that restricted Fannie and Freddie’s portfolio growth, was 180 basis points; the spread was 157 basis points in both 2004 and 2005.

Nor did we observe any sort of shock to the market when the accounting irregularities at Freddie were disclosed in June 2003. The spread was 196 basis points in May 2003, 198 basis points in June, and 196 basis points in July. Consider also January 2004, when OFHEO imposed a capital surcharge on Freddie. That month, the mortgage spread was 159 basis points. The month before, the spread was 161 basis points; the month after, 156 basis points. The OFHEO order applying to Fannie came in September 2004. That month the spread was 163 basis points; the month before, 159; the month after, 162.

Toward the beginning of my remarks, I noted that disclosure of the accounting irregularities did affect the stock prices of the two firms. Now we see that there was no effect on the mortgage market. The issue, clearly, is the profitability of the firms and not effects on the mortgage market. The effects of problems at Fannie and Freddie on the mortgage market have been minimal because the market contains many competent and well-capitalized competitors that can readily pick up the slack when other players stumble.

Financial firms throughout the economy ought to have an intense interest in reforming the GSEs. One reason is simply that banks and other financial firms, and many nonfinancial firms, hold large amounts of GSE obligations and GSE-guaranteed MBS. I believe that many risk managers simply accept that GSEs are effectively backstopped by the Federal Reserve and the federal government without ever thinking through how such implicit guarantees would actually work in a crisis. The view seems to be that someone, somehow, would do what is necessary in a crisis. Good risk management requires that the “someone” be identified and the “somehow” be specified. I have emphasized before that if you are thinking about the Federal Reserve as the “someone,” you should understand that the Fed can provide liquidity support but not capital.4 As for the “somehow,” I urge you to be sure you understand the extent of the president’s powers to provide emergency aid, the likely speed of congressional action, and the possibility that political disputes would slow resolution of the situation.

There is a long-run issue that goes beyond that of today’s systemic risk. The fact is that it is very profitable for a firm to be able to borrow at close to the Treasury rate, lend at the market rate, and hold little capital. That is why the promise of constraints on the portfolio growth at Fannie and Freddie had a significant effect on their stock prices. Any firm with such a privileged position will want to extend its scope of operations. Over the past 15 years, Fannie Mae and Freddie Mac have grown much more rapidly than has the stock of mortgages outstanding and, as a consequence, now hold or guarantee a large fraction of U.S. home mortgages. At the end of 1990, they held in their portfolios 5 percent of the mortgages for 1- to 4-family properties; the share peaked at 22 percent at the end of 2003; and, at the end of the third quarter of 2006, the share was 14 percent. Given the powerful incentive Fannie and Freddie have to grow, the systemic risk they pose to the economy will also grow.

Once their current accounting problems are fully resolved, Fannie and Freddie will want to resume their growth. It is simply very profitable to be able to borrow at close to the Treasury rate and invest in mortgages while holding minimal capital. Banks maintain capital ratios double or more the ratios that Fannie and Freddie maintain. Banks pay deposit insurance premiums to the

4 For a discussion of Federal Reserve emergency powers, see Poole (2004).
Federal Deposit Insurance Corporation, whereas Fannie and Freddie pay no insurance premiums. Assuming that the implied guarantee would, in a crisis, lead to a federal bailout, U.S. taxpayers bear the risk while the shareholders and managers of Fannie and Freddie enjoy the profits. This situation encourages these firms to grow vigorously.

These two firms, however, cannot meet their growth targets in the long run if they confine their operations to conforming home mortgages. Their interest in increasing the conforming mortgage limit is clear. Moreover, in my opinion, it is inevitable that they will look for ways to extend their operations into new areas. They have that clear incentive because of the implicit federal guarantee they enjoy. For them to extend their operations into market segments already well served by existing private firms will not enhance the efficiency of mortgage markets or reduce costs to mortgage borrowers.

There are two possible ways to constrain the operations of the GSEs to areas with a clear public purpose. One is to end the implied federal guarantee so that Fannie Mae and Freddie Mac compete on an equal basis with other fully private firms. The other is to place restrictions on the size of their owned portfolios if they retain their privileged position. Their owned portfolios should be limited to mortgages held temporarily in the process of securitization.

Absent complete privatization, or on the way to it, Congress should strengthen the powers of OFHEO or a successor regulator. OFHEO has weaker powers than provided by law to the federal bank regulators—the Office of the Comptroller of the Currency, the Federal Reserve, and the Federal Deposit Insurance Corporation. The GSE supervisory framework remains fragmented and weak, as the GAO has pointed out on numerous occasions. Thus, structural change of the GSEs and their supervision should be at the top of the reform agenda. There is a glaring need for legislation to clarify the bankruptcy process should a GSE fail. At present, there is no process and no one knows what would happen if a GSE becomes unable to meet its obligations.

Freddie Mac and Fannie Mae both got into trouble with accounting irregularities in part because of the complexities under generally accepted accounting principles for derivatives positions and rules determining which assets should be reported at market value and which should be reported at amortized historical cost. Sound risk management practices require that GSE management base decisions on market values, or estimates as close to market values as financial theory and practice permit. The reason is simple: Fannie Mae and Freddie Mac pursue policies that inherently expose the firms to an extreme asset/ liability duration mismatch. They hold long-term mortgages and MBS financed by short-term liabilities. Given this strategy, they must engage in extensive operations in derivatives markets to create synthetically a duration match on the two sides of the balance sheet. These operations expose the firm to a huge amount of risk unless the positions are measured at market value.

Almost all the assets and liabilities of the GSEs are either traded actively in excellent markets or have values that can be accurately measured by prices in such markets. For this reason, the financial condition of the GSEs ought to be measured through fair-value accounting and such accounts ought to be the principal yardstick of condition and performance.

**CONCLUSIONS**

Since the GSE accounting scandals emerged in mid-2003, one thing has remained rock-solid: The GSEs have continued to borrow at yields only slightly higher than those of the U.S. government and noticeably lower than those available to any other AAA-rated private company or entity. In other words, despite the vast recent accumulation of knowledge about the significant risks run by the GSEs, as well as their inability (or unwillingness) to manage these risks, investors in GSE debt securities appear unmoved. Upon reflection, the lack of market discipline evident during this crisis period is striking—like a dog that did not bark. This fact indicates to me that there still is a significant problem with the GSEs that needs to be fixed.

---

5 Most recently, the GAO criticized GSE oversight in Walker (2005).
The obvious answer to why the dog did not bark is that the so-called “implicit guarantee”—that is, the belief by investors that the U.S. government would not allow the GSEs to default on their debt obligations—has not been removed. Indeed, the talk of increased GSE regulation and the failure of structural-reform legislation to become law may actually have reinforced the belief of many that, overall, the government is perfectly happy with the situation as it is. The GSEs remain politically powerful, if less strident than they were a few years ago.

Three essential reforms are needed to eliminate the GSEs’ threat to financial stability. First is a limit on their portfolio growth, second is an increase in their minimal required capital, and third is satisfactory bankruptcy legislation so that, should the worst happen, federal authorities can deal with the problem in an orderly way.

Freddie Mac apparently does not expect any significant increases in constraints on its operations. Funds that could have been used to build capital to better protect taxpayers have instead been used to increase common stock dividends. Freddie set a quarterly dividend of $0.22 in the fourth quarter of 2002 and has increased the dividend every year since. As of the fourth quarter of 2006, the dividend stands at $0.50 per quarter, more than twice its level four years earlier. Fannie Mae cut its dividend in half in early 2005 to build capital, but I’ll hazard a guess that once it starts issuing regular financial statements the company will increase its dividend rather than build capital further.

I began this speech noting that the Federal Reserve has a responsibility to maintain financial stability. That responsibility includes increasing awareness of threats to stability and formation of recommendations for structural reform. I do not believe that a GSE crisis is imminent. However, for those who believe that a GSE crisis is unthinkable in the future, I suggest a course in economic history.

REFERENCES


Lehnert, Andreas; Passmore, Wayne and Sherlund, Shane M. “GSEs, Mortgage Rates, and Secondary


APPENDIX

Summaries of Recent Studies on GSE Issues

Jaffee (2003) Study of GSE Interest Rate Risk

Dwight Jaffee was one of the first to “peer through” the public disclosures provided by the GSEs about the interest rate risks they incurred and how they managed them. Jaffee concluded that the GSEs actually incurred significant interest rate and liquidity risks, despite their own characterization of such risks as being minimal. Subsequent events and analysis have proven Jaffee correct.

OFHEO (2003) Study of Potential Systemic Risks Posed by GSEs

Even before the GSE accounting scandals broke, the GSEs’ safety-and-soundness supervisor had prepared a study comprising scenarios in which the GSEs might contribute to systemic risk. Although OFHEO concluded that the likelihood of one or both GSEs contributing to financial-system instability was very small, the agency recommended to Congress that its (OFHEO’s) supervisory powers should be enhanced to further safeguard the GSEs and the financial system.
Public Statements by Chairman Alan Greenspan (2005a,b,c)

Federal Reserve Chairman Greenspan (2005a) rejected the idea of stronger GSE regulation in favor of portfolio limits, stating that,

World-class regulation, by itself, may not be sufficient and, indeed, might even worsen the potential for systemic risk if market participants inferred from such regulation that the government would be more likely to back GSE debt in the event of financial stress...We at the Federal Reserve believe this dilemma would be resolved by placing limits on the GSEs’ portfolios of assets.

Chairman Greenspan also drew attention to the strains the GSEs could place on the over-the-counter interest rate derivatives markets due to their portfolio-hedging activities.

Research Papers by Federal Reserve Staff6

One of these papers estimated the pass-through by Fannie Mae and Freddie Mac of their funding-cost advantage into primary mortgage rates, finding a mere 7 basis points of pass-through. Another paper provided evidence against the GSEs’ claims that their purchasing behavior stabilizes mortgage rates during periods of market turbulence. Other papers discuss (i) likely competitive interactions between the GSEs and large banks that will be subject to Basel II capital regulation and (ii) the ill-structured incentives the GSEs face to increase the size of their portfolios.

Ad Hoc Federal Reserve Study Group Examining GSE Impacts on Interest Rate Derivatives Markets (Board of Governors, 2005)

The study group identified potential channels through which disruptions at the GSEs could flow through to other market participants in the over-the-counter markets for interest rate derivatives, like swaps, interest rate options, and swaptions (options on swaps). The study group reported that market participants felt current risk-management practices were sufficient to contain risks posed by the GSEs.

Economic-Capital Analysis of GSEs by Morgan Stanley (Posner, 2005)

Kenneth Posner, an equity analyst at Morgan Stanley, isolated the distinct economic risks faced by the GSEs and estimated how much capital the firms would need to provide adequate protection to debtholders to justify an AA senior-unsecured bond rating. This analysis assumed that there would be no support forthcoming (or expected by financial-market participants) from the federal government. His estimate of the required equity-to-assets capital ratio was in the range of 4 to 7 percent, about twice as high as the current GSE ratios of closer to 3 percent. Thus, the GSEs would be significantly under-capitalized today if there were no expectation of government support of their liabilities.

---

6 These include Passmore (2005), Passmore, Sherlund, and Burgess (2005), Lehnert, Passmore, and Sherlund (2006), Hancock et al. (2005), Frame and White (2005), and Emmons and Sierra (2004).
Milton Friedman and U.S. Monetary History: 1961-2006

Edward Nelson

This paper, using extensive archival material from several countries, brings together scattered information about Milton Friedman’s views and predictions regarding U.S. monetary policy developments after 1960 (i.e., the period beyond that covered by his and Anna Schwartz’s Monetary History of the United States). The author evaluates these interpretations and predictions in light of subsequent events. (JEL E31, E51, E52, E58)


Milton Friedman and Anna J. Schwartz’s (1963) A Monetary History of the United States covered a 93-year period: 1867 to 1960. Milton Friedman lived until November 2006—46 years beyond the period covered by the Monetary History and a length of time equal to nearly half that covered in that book. Throughout 1961-2006, Friedman commented publicly on U.S. monetary policy developments. This paper attempts to provide a perspective on Friedman’s account of 1961-2006 U.S. monetary history by studying the observations he provided over that period.

Friedman provided no single detailed assessment of post-1960 U.S. monetary developments. Though Paul Samuelson once speculated of a time when “Milton Friedman and Anna Schwartz come to write their history of the crimes of 1974 and 1975” (NYT, 02/26/75),1 the Monetary History was the single volume of monetary history produced by the two authors; Friedman and Schwartz’s subsequent collaborations were not a continuation of their historical analysis.2 Friedman (1984a) did provide a capsule history of monetary developments from 1960 to 1983; but not only did this study ultimately cover only one half of the post-1960 period through which Friedman lived, but Friedman’s own views of 1979-83 developments underwent major revision in the years after 1984. There is thus no single definitive record by Friedman available of his version of events. But major elements of it can be recovered, as there are many forums to which he contributed over this period that contain his observations on contemporary monetary policy. I therefore study Friedman’s account of 1961-2006 using a variety of sources: not only Friedman’s post-

1 In this paper, newspaper and magazine articles are cited in the text with their abbreviation and date. A key to the abbreviations is given in Appendix A, and Appendix B gives the newspaper articles referenced in chronological order.

2 In particular, Friedman and Schwartz’s (1982) Monetary Trends has “1867-1975” in its title but is not an update of the monetary history to 1975; rather, the authors offered a “statistical and theoretical analysis” to complement the “chronological and largely qualitative analysis” of the Monetary History (Friedman and Schwartz, 1982, p. xxviii). Consistent with this, Friedman and Schwartz (1982) mention Arthur Burns several times, but it is always with reference to Burns’s scholarly work on business fluctuations and never with reference to his position as Federal Reserve Chairman from 1970 onward.
1960 writings on monetary affairs, but other public statements he gave in media interviews, speeches, and congressional testimony. This broad source base includes—over and above Friedman’s contributions to books, journals, and media—reporting on Friedman’s public statements that appeared in newspapers and magazines in the United States, the United Kingdom, Australia, Canada, Germany, Hong Kong, Ireland, Israel, New Zealand, Singapore, and South Africa. ³

My discussion moves between material drawn from Friedman’s statements in the media and material from his contributions to specialist monetary economics outlets. This approach might be questioned. Friedman’s critics have sometimes argued that Friedman’s economic analysis was not consistent across his technical and popular writings. In particular, they alleged that Friedman used a harder-line and more mechanical version of monetarism in his journalism. This criticism, first made by Paul Samuelson and James Tobin, has recently been revived by Krugman (2007a). ⁴ In my view, however, the criticism is without merit. Friedman (1970a, 1972a) provided detailed examples showing that the monetary analysis in his popular writings was consistent with that in his technical work, and this judgment has been affirmed by Gordon (1976) and Nelson (2004). More recently, Schwartz and Nelson (2007) provide a rebuttal to Krugman (2007a) on this issue.

**CRITIC OF MONETARY POLICY: 1961-68**

The London *Financial Times* in late 1962 referred to “Americans’ surprising addiction to pre-Keynesian economics” (*FT*, 12/29/62), a remark undoubtedly inspired by the U.S. economic policy record of the 1950s. That period was characterized, in addition to fiscal conservatism, by a monetary policy directed at restraint on aggregate demand, as Romer and Romer (2002) show. Romer and Romer provide evidence, based on documentary material and policy-rule estimates, that U.S. monetary policy had a modern inflation-oriented outlook—one acknowledged, albeit pejoratively, by Paul Samuelson in 1965 when he referred to the Federal Reserve’s “anti-inflation paranoia of the 1950s” (*FT*, 12/31/65).

Romer and Romer (2002, p. 121) argue that Milton Friedman did not judge 1950s monetary policy favorably. This understates Friedman’s praise of 1950s monetary policy, both in the *Monetary History* with Schwartz and subsequently. Though a critic of the Federal Reserve’s overall record since its inception, Friedman praised the mid- and late-1950s policy of “restraint...[which] eliminated inflation by 1960.” ⁵ Friedman and Schwartz (1963, p. 628) had praised the “near-revolutionary change” in official statements in 1952-54 toward acknowledging the importance of the money supply, and in 1965 Friedman suggested that the Federal Reserve’s attention to money supply data during 1959-60 helped stop the recession of 1960-61 from being worse. ⁶ Friedman argued that a large part of the credit for the anti-inflationary monetary policy of the 1950s was due to the Eisenhower Administration, for showing solidarity with the Federal Reserve and resisting the “temper of the time” that favored aggressive stimulation of aggregate demand (*NW*, 12/06/76). The burying of inflationary expectations by the early 1960s, Friedman contended, resulted from “the [economic] slowdown that Mr. Eisenhower was willing to accept at the end of the 1950s” (*CHA*, 11/73).

While applauding the price stability that the monetary policy of the late 1950s had produced, Friedman felt that a steadier policy, involving fewer fluctuations of money growth around its downward trend, could have generated the same result. Actual policy, he believed, had worsened business cycle fluctuations by ill-timed fine-tuning. Thus the 1950s were not excluded from

---

³ Previous studies of Friedman’s contribution to policy debates include Frazer (1988), Hammond (1996), Leeson (2000), and Nelson (2004). Due to my expanded base of source material and my focus on U.S. monetary history, the overlap here with those studies is minor.

⁴ The earlier articulations of this argument appeared in Tobin (1970) and a 1968 commentary by Samuelson (*STE*, 12/15/68).

⁵ Friedman (1984a, p. 26).

⁶ See Friedman (1968a, p. 146).
(i) his generalization in 1967 that “[t]hroughout the post-war period...the Fed has tended first to delay action and then, when it did act, to go too far” (NW, 10/30/67) or (ii) his statement in 1964 that “we ought to convert monetary policy from being a destabilizing force into at least being a neutral factor.”

Monetary policy, already expansionary from mid-1960, became more so from 1961, with both the Kennedy Administration and the Federal Reserve supporting stimulation of the economy. This monetary expansion is evident in an increase in money growth (using today’s M1 and M2 definitions) during the first half of the 1960s. Most of the shift to a higher growth rate was completed in 1961, with money growth in the succeeding years quite smooth. That smoothness, Friedman suggested, was why the United States avoided a recession after 1960. Friedman developed this theme in an October 1965 consultant’s memorandum to the Board of Governors, which noted that in the prior three years “monetary growth has been relatively stable by past standards,” an “excellent” monetary policy performance. Consistent with his 1965 judgment, Friedman came to regard the early-to-mid 1960s as testament to the value of stable monetary growth; in 1977, for example, he calculated that the standard deviation of M2 growth for 1963:Q3–1966:Q4 was close to the lowest observed in the postwar era for any period of comparable length.

Friedman served as economic advisor to Republican candidate Barry Goldwater during the 1964 presidential election campaign. Goldwater eventually suffered a landslide defeat to President Lyndon Johnson. Friedman’s monetary policy proposals were not likely to be implemented even if Goldwater had won, as Goldwater had already said that a constant money growth rule “would require too fundamental a change in our monetary system. We must continue to work within the framework of the independent Federal Reserve System” (CSM, 09/28/64).

Friedman’s 1965 memorandum applauded the increased average money growth in 1962-65 as an improvement over the inadequate rates of 1958-60, but judged that M2 growth had now overshot its desirable rate. He acknowledged that the upward trend of money growth had so far led to greater output growth with very little increase in inflation, but attributed this to the impact on expectations of the late 1950s period of austerity. Inflationary expectations, he felt, were now returning, and so the time was now overdue for a “moderate reduction” in M2 growth.

At the same time, Friedman spoke out against the Johnson Administration’s increasing emphasis on wage-price guidelines to restrain inflation. “Price control by exhortation and threat and use of extra-legal powers never has worked and never will, except to disrupt the economy. The cure is worse than the disease. Holding down a few prices here and there only diverts inflationary pressure elsewhere, just as squeezing the corner of a balloon pushes the air into the rest of it” (SLGD, 12/09/65).

Friedman was equally scornful of more general controls: “Direct control of prices and wages does not eliminate inflationary pressure. It simply shifts the pressure elsewhere and suppresses some of its manifestations. The only way to stop inflation is to restrain the rate of growth of the quantity of money” (CT, 05/08/66).

By late 1966, Friedman was worried that the Federal Reserve had moved too far in the direc-

---

7 Friedman (1964, p. 1156).
8 See Friedman (1984a, p. 27).
9 See Friedman (1968a, p. 147).
10 See Friedman and Modigliani (1977, p. 16). This calculation roughly agrees with what emerges from examining present-day data on the modern definition of M2. Using this series, the lowest value of the 14-quarter standard deviation of money growth takes place for the period 1962:Q4–1966:Q1. See Figure 1.
11 In fact, Friedman and Goldwater, while sharing similar limited-government aspirations, had little agreement on economic management issues. Friedman supported the Kennedy-Johnson tax cut, while Goldwater voted against it in the Senate. Friedman viewed inflation in 1970-71 as monetary in origin and opposed wage-price controls; Goldwater blamed cost-push factors and supported price controls.
12 Friedman (1968a, p. 152).
13 Here and below, where newspaper coverage of Friedman’s public appearances are quoted in the text, the quotations correspond to the direct quotations attributed to Friedman in the articles.
tion of restraint and said the U.S. economy was “headed straight for recession in the next two or three months” (CSM, 12/19/66). Similarly, in Newsweek in early 1967 Friedman wrote that it was “almost surely too late to prevent a recession” (NW, 01/09/67). Actual data now show a sharp slowdown in real GDP growth in late 1966 and the first half of 1967, but no decline in output. There is also no recession in 1967 according to the official National Bureau of Economic Research (NBER) chronology. Nevertheless, Friedman and Schwartz (1982, p. 74), in their own business cycle dating, amended the NBER chronology to categorize 1966-67 as an economic contraction. Partial justification for this choice is that their study was concerned with relations between money stock movements and nominal income movements, and 1966-67 is an important episode in this respect. Nominal income slowed down much more than output, in what Friedman later judged was an unusual instance of inflation responding as rapidly as total spending to a tightening of monetary policy.¹⁴ In fact, four-quarter CPI inflation actually came back below 3 percent in the first half of 1967 after crossing the 3 percent barrier in 1966. It then rose about 3.5 percentage points over the rest of the 1960s.

In 1968, the joint House-Senate economic committee made a nonbinding recommendation that the Federal Reserve be judged by the criterion of meeting a 2 to 6 percent growth rule for the money supply. Friedman applauded the proposal, suggesting that the most important change would be “for the Fed to be aware of the attitude of Congress” regarding the desirability of steady money growth (CT, 07/05/68). The 1968 proposal in itself came to nothing, but was a precursor to the congressional resolution that established monetary targets in 1975.

**WAGE-PRICE CONTROLS AND MONETARY TARGETS: 1969-75**

In January 1969 Friedman called for Federal Reserve Chairman William McChesney Martin Jr. to step down. Though Martin was due to retire in January 1970, Friedman told Time magazine, “It would be a very good thing if he went early” (TIME, 01/10/69). Despite the public acrimony, Friedman and Martin continued to correspond privately, and in early 1969 Friedman wrote to Martin urging the Fed to adopt a money stock target, with Friedman offering to help design procedures to ensure better control of monetary aggregates. Martin replied that it was “quite true” that better control of monetary aggregates was feasible, but expressed doubt that stabilization of monetary growth was actually desirable.¹⁵

Friedman welcomed the appointment of Arthur Burns as Federal Reserve Chairman in January 1970, declaring Burns as “the first person ever named Chairman of the Board who has the right qualifications for that post,” and expressing the hope that Burns would deliver money growth “high enough to encourage recovery…but low enough to avoid renewed inflation” (NW, 02/02/70). Friedman rapidly became disillusioned when Burns began a series of speeches calling for an incomes policy to help in fighting inflation. And over the course of 1970, Burns adopted a progressively stricter cost-push analysis of inflation. To Friedman, such an analysis, along with the corresponding policy recommendation of direct government intervention in price and wage formation, was outdated and discredited by empirical evidence. Friedman himself had embraced cost-push views in his pre-monetarist days (see Despres et al., 1950), but it had been superseded in the previous 20 years by his monetarist view of inflation, according to which inflation depended on monetary forces through excessive aggregate spending. As he put it in July 1970, “To each businessman separately it looks as if he has to raise prices because costs have gone up. But then, we must ask, ‘Why did his costs go up? Why is it that [for example] from 1960 to 1964 he didn’t find that he had to pay so much more for labor he had to raise prices, but that suddenly from 1964 to 1969 he did?’ The answer is, because, in the second period, total demand all over was increasing” (CDN, 07/29/70).

---

¹⁴ See Friedman (1972b, p. 14).

¹⁵ William McChesney Martin Jr. letter to Friedman of April 7, 1969, quoted in Friedman (1982a, p. 106).
The disagreement led Friedman in May 1970 to send Burns a lengthy handwritten letter critical of Burns’s statements on incomes policy. The Philadelphia Inquirer reported that Burns was shaken by the letter and suggested that relations between Burns and Friedman had deteriorated (PHI, 05/29/70). A permanent rift between the two occurred. Nevertheless, claims that Friedman and Burns broke off communications completely in 1970 (or 1971) are false. The two met and spoke at the Federal Reserve Board during Friedman’s appearances (until 1974) as an academic consultant; and they continued to correspond over 1970-71. In 1985 Friedman even described Burns as “still among my closest friends,” but this was certainly an exaggeration.

In an interview with the Chicago Daily News, Friedman specifically disputed the notion that strong labor unions had, as claimed by Burns, become important as a source of inflation. Friedman noted: “You can look around the world and find countries that have had very strong trade unions and no inflation...The fact is that there is little relation between trade unions and inflation” (CDN, 07/29/70). Wage growth, he stressed on a later occasion, was only a symptom of inflation: “It isn’t the wet street that caused the rain” (OKL, 05/19/81).

Nevertheless, belief in union monopoly power as a cause of inflation remained prevalent. President Nixon, with Burns’s support, introduced wage-price controls on August 15, 1971, beginning with a three-month freeze. Friedman’s reaction was that the measures were “purely cosmetic in nature, not therapeutic” (NYP, 08/16/71). Friedman also criticized the controls in a meeting with Nixon in September 1971.

Friedman did react favorably to another of the measures Nixon announced in August 1971: the end of the U.S. commitment to a pegged gold price. A longtime advocate of floating exchange rates, Friedman welcomed the end of this mainstay of the postwar fixed exchange rate system. He proclaimed in September 1971, “The Bretton Woods system is dead” (JT, 09/24/71). This judgment was borne out by the failure, over the following 18 months, of attempts to restore international cooperation on fixed-rate arrangements.

At the end of 1971, Friedman renewed his attack both on Burns’s diagnosis of cost-push inflation and Nixon’s attempted cure of controls. “We have been driven into a widespread system of arbitrary and tyrannical control over our economic life, not because ‘economic laws are not working the way they used to,’” Friedman began, alluding to a phrase Burns had used, “not because the classical medicine cannot, if properly applied, halt inflation, but because the public at large has been led to expect standards of performance that as economists we do not know how to achieve.” Friedman said that Burns’s embrace of cost-push views reflected “the propensity of economists to appeal to a change in our economic structure whenever they are puzzled” and that Burns had revived old fallacies about inflation.

In 1973 Friedman went on to declare that controls were the “worst mistake in American economic policy that has been made by an American president in the last 40 years” (OAK, 06/15/73) and added that U.S. business leaders had shown “ignorance and shortsightedness” in endorsing the control measures (JT, 10/15/73). Friedman’s consistent opposition to guidelines and controls contrasted with the position of leading Keynesian economists during the 1960s and 1970s. For example, Arthur Okun, who had helped shape the Kennedy and Johnson administrations’ wage/price guideposts, applauded President Nixon’s adoption of wage/price controls. In December 1972 Okun said that some form of wage-price controls would be in effect “for the rest of our lives,” adding, “I’ll bet a nickel that we never hear another U.S. President say what President Nixon said in 1969 and 1970, that the government has no role over private wage and price decisions” (KCS, 12/07/72). This example

16 See Friedman (1982a, pp. 106-10).
17 Friedman (1986, p. 81).
18 See Friedman and Friedman (1998, p. 387). On a visit to the National Archives in May 2002, Friedman listened to the recording of his September 1971 meeting with Nixon, but found that substantial portions of the recording were unintelligible.
19 Friedman (1972b, p. 17).
20 Friedman (1972b, p. 11).
and many similar ones show that Krugman (2007b) is incorrect to suggest that Friedman’s opposition to the Nixon wage/price controls was widely shared by nonmonetarists.

Though opposed to the changes in inflation analysis that were taking place over 1970-71, Friedman did modify his own position on inflation in one major respect during this period. Before 1971, he had believed that output and inflation reacted simultaneously to prior monetary policy actions.21 This led him to predict in August 1969 that inflation would begin declining by the fourth quarter of 1969 (DOM, 08/24/69) and to admit in June 1970 that the impact of the 1969 monetary policy tightening on inflation was “coming later than many of us hoped or expected” (CT, 06/29/70). In late 1971 Friedman reexamined postwar evidence and found an 11- to 31-month lag from monetary growth to inflation,22 leading to his later summary of the evidence that there was a two-year lag from money growth to inflation and “output responds more quickly than prices…”23 The two-year rule of thumb from monetary policy actions to inflation, which entered Friedman’s framework in 1971, has since become standard.24

If Friedman’s confidence about timing relations between money growth and the components of nominal income growth increased in the 1970s, he remained as skeptical as ever about estimating structural aggregate supply relationships. His proposition that the appropriate Phillips curve specification was an expectational version (Friedman, 1968b) was gaining acceptance in academia, but Friedman continued to doubt that a durable specification could be implemented empirically. This was especially so with those Phillips curves that used unemployment-based measures of economic slack; Friedman thought that Okun’s law relationships were unstable. His doubt, expressed in 1968, that the natural rate of unemployment could be measured reliably was reinforced by evidence that in the 1970s the relationship between the rates of employment and unemployment had loosened (NW, 02/07/77).

Asked in 1974 how much unemployment was needed to cure inflation, Friedman gave the non-answer, “To answer that question is like answering the question, ‘When did you stop beating your wife?’” (IT, 09/18/74). Nevertheless, he confirmed that “There is no way of slowing down inflation that will not involve a transitory increase in unemployment, and a transitory reduction in the rate of growth of output. But these costs are far less than the costs that will be incurred by permitting the disease of inflation to rage unchecked” (TG, 09/16/74).

In explaining why monetary conditions had been allowed to become so relaxed in 1972, Friedman judged that “the Fed became worried about rising interest rates” (LAT, 02/18/73), adding more generally in 1978 that pressures “to keep interest rates low are a major reason for high monetary growth” (NW, 04/24/78). Friedman emphasized that the failure to allow nominal interest rates to exceed inflation was inhibiting price stability and that only a period of disinflation could permanently deliver low nominal interest rates. Hence his recommendation: “The best way to hold rates down in the long run is for the Fed to raise them temporarily” (LAT, 02/18/73).

In addition to criticizing official efforts to hold down nominal interest rates in securities markets, Friedman condemned the U.S. government for providing inadequate returns to holders of nonmarketable debt. Savings bonds, he declared in early 1976, were the “greatest ripoff in modern history” and “anyone who has bought savings bonds in the past ten years has been taken to the cleaners” (OAK, 02/12/76). He was no less acerbic when he returned to the theme in 1987: “Savings bonds were the biggest steal of all. The government was able to get suckers to buy savings bonds yielding much less than the inflation rate” (SFC,
In 1974, Friedman had called for government-indexed savings bonds to provide a vehicle for small savers to protect the purchasing power of their assets (DC, 08/16/74).

With Friedman’s criticisms of U.S. monetary policy gaining ground, Chairman Arthur Burns wrote a letter to Senator William Proxmire in November 1973 (subsequently published in Federal Reserve outlets: Burns, 1973) that apparently denied Federal Reserve responsibility for inflation and, in part, for money growth as well. This prompted Friedman himself to write a bristling rebuttal, also in the form of a letter to Proxmire, in early 1974 (Friedman, 1974). Friedman was particularly scornful of Burns’s statement, “The severe rate of inflation that we have experienced in 1973 cannot responsibly be attributed to monetary management.” “As written,” Friedman began (1974, p. 20), “this sentence is unexceptionable. Delete the word ‘severe,’ and the sentence is indefensible.” Commodity price increases and other one-time events, Friedman said, might explain why U.S. inflation in 1973 reached 8 percent instead of 6 percent, but not why inflation could reach 6 percent in the first place; moreover, the more years one considered, the less valid it was to invoke nonmonetary events to explain inflation. Thus “the Fed’s long-run policies have played a major role in producing our present inflation.”

The Federal Reserve did tighten monetary policy over 1973, and in October 1973 Friedman predicted a recession of which already “we might be in the early stages” (PHB, 10/12/73). The National Bureau of Economic Research later dated the 1973-75 recession as having begun in November 1973.

October 1973 also saw the OPEC oil embargo and the announcement of permanent oil price increases. Friedman would have a checkered record in predicting the course of oil prices. In September 1973 he had said that food and commodity prices had undergone a “tremendous bubble...[which] cannot last” (NW, 09/17/73), thereby implying that their likely direction was down, not up. When OPEC then announced its large price increases, Friedman predicted that in four to five years “oil will be coming out of our ears” because “cartels inevitably break up sooner or later” (PHB, 11/09/73). He later cited the Iran revolution and counterproductive energy policies in the United States as the source of resilience in the oil price for the rest of the 1970s. Weaker oil prices in the early 1980s eventually provided some support for Friedman’s initial skepticism about OPEC, and he accurately predicted further weakness to come. “Betting on continued high oil prices is equivalent to betting on a disruption in the Middle East,” he said in Norway in September 1982. “I am not talking about a reduction from $32 to $30 a barrel, but a drop to somewhere around $15 to $20” (SCMP, 09/17/82). His further conjecture that, longer-term, the real oil price would return to its pre-1973-shock value was borne out during some of the 1990s.

Friedman consistently stressed, however, that success in controlling inflation was not contingent on oil prices beyond the very short run. “If you spend more on oil, doesn’t that leave you less to spend on something else? Why don’t other prices come down, or not rise as rapidly? It’s a complete fallacy to suppose that the rise in the price of oil, or of other commodities, has had any significant effect on inflation” (TG, 09/16/74).

Wage and price controls came to an end in 1974, but Friedman remained irritated by the proliferation of false cures for inflation. After attending a two-day summit on inflation hosted by President Ford, Friedman deplored the “bewildering variety of proposals for governmental action” (NW, 10/14/74). “You can’t stop inflation without unemployment and stagnation,” he added in 1975. “All this business about consumers saving more and reusing paper plates to control inflation—that’s nonsense” (PHB, 03/02/75). Friedman also criticized Chairman Burns for still failing to acknowledge sufficiently the Fed’s role, in the early 1970s, in producing the inflation surge of 1973-74 (WSJ, 08/21/75). On a more positive note, Friedman was pleased that floating exchange rates had become the norm and judged that they had been “effective shock absorbers” against recent years’ events, such as the OPEC shock (WSJ, 06/30/75).

Another development that Friedman initially saw as very positive was the Federal Reserve’s adoption, after a congressional resolution, of monetary targets in 1975. But within a few years he would conclude that the targets had been implemented in a way that had not led to the desired changes in monetary policy. Among Friedman’s criticisms of the implementation were that the Federal Reserve avoided accountability by introducing target bands for both M1 and M2 growth instead of a point target for a single aggregate (Friedman, 1984a, p. 27). The Federal Reserve also continued to use a federal funds rate instead of the monetary base instrument, which Friedman preferred. While Friedman acknowledged that monetary targeting with a funds instrument was feasible, he contended that inertia in adjusting the interest rate meant that monetary target misses would be serially correlated when a funds instrument was used (NW, 12/08/75).

NEW PRESIDENTS, NEW CHAIRMEN, NEW OPERATING PROCEDURES: 1976-80

With money growth having come down from its peaks in 1972, Friedman early in 1976 predicted accurately that U.S. inflation would fall to 4 to 5 percent by the end of 1976 (JOH, 04/05/76). When this decline came about, however, Friedman voiced unhappiness that achieving 5 percent inflation was now thought of “as doing a good job” (PHB, 11/18/76).

Friedman warned in April 1976 that, while money growth had recently picked up, slow money growth observed in 1975 could interrupt the U.S. economic recovery (JOH, 04/05/76). Indeed, Friedman’s later verdict on 1976 was that “the [monetary] slowdown in late 1975 produced the economic pause in the second half of 1976 that played such a prominent role in the Ford-Carter election battle” (NW, 10/03/77). Ford’s former press secretary later said that a rise in unemployment before the election, associated with the economic pause, was a decisive factor in President Ford’s defeat (USAT, 12/27/06).26

Election day 1976 coincided with Friedman’s announcement of a change of location. He would leave Chicago at the end of November and move to San Francisco to become a research fellow at the Hoover Institution, Stanford University, following a six-month spell in late 1976 and early 1977 as a visiting scholar at the Federal Reserve Bank of San Francisco (CT, 11/02/76; NYT, 11/03/76).

Shortly after the election, Friedman said he hoped that once in office President Carter would “rise above his advisors” and not try to stimulate aggregate demand aggressively (PHB, 11/15/76). He also provided a scenario, in a December 1976 Newsweek column, regarding the likely result of following expansionary policies in 1977, which proved to be prophetic:

If Mr. Carter tries to put his advisors’ policies into effect and succeeds in doing so—including getting the Federal Reserve System to speed up substantially the rate of monetary growth—there might be a sudden spurt in the economy and a quick reduction in unemployment. However, these good results would be temporary. By 1978 or 1979, inflation would be back in double digits and wage and price controls would be in place or in contemplation. By 1980 at the latest, unemployment would be rising sharply. As Machiavelli might say: what a way to face the 1980 election! (NW, 12/06/76)

It was probably this commentary that President-elect Carter had in mind when in December 1976 he asked to be put through to Milton Friedman and began, “I’ve wanted to talk to you, but first let me congratulate you on the [Nobel] prize.” Unfortunately, Carter’s staff had connected him not to Milton Friedman but to Mr. Milton (Milt) Friedman, a member of President Ford’s staff (DFP, 12/25/76).27 Carter’s staff finally connected him to the Milton Friedman, who later described his conversation with Carter as a “pleasant talk” but provided no details (Friedman and Friedman, 1998, p. 459).

---

26 The unemployment rate in the third quarter of 1976 was slightly higher than in the previous quarter.

27 Ford’s staff member Milton Friedman had served as senior speechwriter and special assistant to the President (TDN, 02/05/76).
The pleasantries did not endure into 1977 and, shortly after Carter took office, Friedman spoke dismissively of the new administration’s proposed tax rebate. Reaffirming his skepticism about the effectiveness of fiscal policy, Friedman asked an interviewer, “How can the government stimulate the economy by taking money out of one pocket of the public and putting it into another pocket?” (USNR, 03/07/77). Carter withdrew the tax rebate proposal in April 1977.

In an interview with a St. Louis reporter in late 1977, Friedman returned to the themes covered in his Nobel lecture of a year earlier.28 “[I]nflation is a lot like alcoholism,” he said. “When you drink, the good effects come first, and the hangover comes the next morning. When an inflationary period starts, spending goes up and employment rises along with it.” By an “inflationary period,” Friedman explained he meant the initiation of an easy monetary policy, which usually preceded the actual upturn in inflation. “By printing money at a faster rate you may be able temporarily to create an appearance of prosperity, but only so long as you fool the people. Once the public comes to realize what is going on, higher inflation means higher unemployment! Just look at the example of the U.S., Great Britain, and every other country” (SLGD, 12/16/77).

Friedman was asked whether he thought Chairman Burns would receive another term from President Carter. Friedman said it would not make much difference: Despite an inflation-fighting reputation, Burns in practice had produced “a Fed which is promoting inflation.” Monetary targets had not made a difference to this, said Friedman, because “while the targets are going down, actual monetary growth is going up” (SLGD, 12/07/77). His own opinion was that the “chances are good” that Burns would be reappointed (SLGD, 12/16/77). In late December 1977, Carter actually announced a new Chairman, G. William Miller, who took office in March 1978 and about whom Friedman ultimately wrote very little.

In April 1978, double-digit M2 growth in 1977 led Friedman to predict that “inflation from February 1977 to October 1979 will average something like 7 to 10 per cent...[and] no sustained reduction in inflation can be expected before mid- or late 1979” (NW, 04/24/78). The following month Friedman made more specific predictions: Inflation “probably will be eight or nine per cent by the end of the year, and may be in double digits in 1979” (MHER, 05/22/78). These predictions were realized.

The appointment of Paul Volcker as Federal Reserve Chairman in 1979 gave Friedman an opportunity to reflect on the record of money growth and inflation in the past two decades. He noted that inflation exhibited a steeper upward trend since 1960 than did M2 growth. This was “no mystery...It reflects the widening recognition that inflation is the way of the future. Those inflationary expectations make it prudent for all of us to reduce the fraction of our assets in the form of money.” This pattern, in reverse, would become important in interpreting 1980s monetary developments. Friedman continued, “The problem is not, as President Carter asserts, a lack of confidence. The problem is rather that the public is very confident that the government will produce inflation and will mismanage the economy. We do not need more confidence in bad policies. We need better policies” (NW, 08/20/79).

In October 1979, the Federal Reserve inaugurated its new operating procedures, intended to deliver greater control of monetary growth and permit larger transitory fluctuations in the federal funds rate. Speaking to a local reporter shortly after the procedures were announced, Friedman called them a “long overdue change,” but added, “I remain skeptical until I see proof” that a material change in policy had taken place. The need for a change in regime was urgent, Friedman said, because “sooner or later, the adverse effect of inflation will destroy the country” (SFC, 10/18/79).

In July 1980, Friedman pronounced a negative verdict, declaring that the erratic money growth of the preceding months was a “disgraceful performance” that “confirms the doubts rather than the hopes” about the new regime (NW, 07/14/80). In May 1981 he further declared, “After studying the Fed for [its] 67 years, I have no doubt that the United States would be better off if the Federal

28 The 1976 Nobel lecture was published as Friedman (1977).
Reserve had never been established” (NYDN, 05/22/81). In these and other critiques, Friedman cited the Fed’s failure to reform its arrangements with commercial banks as a major reason for the instability in money growth. In particular, throughout 1979-82 (and into 1984), the required amount of reserves that commercial banks had to keep with the Federal Reserve continued to be determined on the basis of prior deposit levels. This made aggregate reserves predetermined in the short run and so inhibited short-run control of the money stock.

Besides the new operating procedures, Friedman was highly critical of direct controls on credit introduced by President Carter and the Federal Reserve in March 1980. Friedman (1982a, p. 103) said the rationale for the controls was the fallacious real-bills doctrine, according to which certain types of spending are particularly inflationary, so that control of inflation requires discouraging these spending categories. This contrasted with Friedman’s own view that inflation responded to total spending. Friedman also argued that the controls would have adverse effects on saving and investment (NW, 04/14/80). While ostensibly the controls applied to credit rather than money, the controls had a noticeable downward effect on money growth, and output collapsed in the quarter following the controls’ imposition. The credit controls were withdrawn in July 1980.

On November 4, 1980, Ronald Reagan, about whom Friedman had written supportively both during 1980 and in Reagan’s 1976 campaign for the Republican presidential nomination, was elected president. Three days after the election, the Soviet Union’s state-controlled press commented on the result, in the form of an article in Pravda. The commentary noted, “In Ronald Reagan’s entourage are found such experienced public figures as former Treasury Secretary William Simon, economists Alan Greenspan and Milton Friedman, and other well-known people.”29 Ulam (1983) suggests that this respectful reference to Friedman exemplified the Soviet authorities’ regard for opponents who exhibited ideological consistency. If this assessment is accurate, the respect for consistency presumably offset the Soviets’ opposition to much of Friedman’s record, including his advocacy of free markets and his recent meeting with Chinese government officials. On the other hand, Friedman’s views on monetary policy were probably not among his most objectionable opinions for Soviet officials; after all, as Friedman wryly observed on several occasions, Karl Marx was a monetarist.30

THE YEARS OF CONFUSION: 1981-85

The years 1981 to 1985 saw several major blemishes appear on Friedman’s forecasting record, and many changes in the details of his analysis of recent U.S. monetary policy developments. The outcome was a confused record of public statements on Friedman’s part. By far, the period 1981-85 marks Friedman’s most inconsistent period since he became a monetarist in 1950-51.31 It was 1986 before Friedman settled on positions in describing 1980s developments that he maintained for the remaining 20 years of his life.

In September 1981, Friedman commented in Newsweek that “Institutional change, notably the explosion in money-market mutual funds, has rendered narrow monetary aggregates misleading.” He accordingly used M2 to judge monetary policy settings (NW, 09/21/81). If Friedman had stuck with this judgment throughout the following years, he would probably have been able to avoid some of his worst—and most-publicized—macroeconomic forecasts. Certainly in 1981-82 Friedman’s record on inflation forecasting continued to be good; in October 1981 President Reagan told the press that inflation was “in single digits now, and I was interested to see that our Nobel economics prize winner, Milton Friedman, has just been quoted as saying that he believes it’ll be down


30 See for example Friedman’s article “Marx and Money” (NW, 10/27/80).

31 David Laidler, an admirer and former student of Friedman, nevertheless certainly had Friedman in mind when he referred (Laidler, 1990, p. 59) to “careless monetarist predictions” that were made in the early stages of the economic recovery that began in 1982.
to six percent next year” (WCPD, 10/16/81). This forecast, presumably based on the lower M2 growth after 1977, was borne out in U.S. inflation outcomes in 1982.\footnote{In February 1982, Friedman went on to forecast 5 to 6 percent CPI inflation by the end of the year (STR, 02/19/82).}

But in early 1982, Friedman abruptly switched to using M1 on a near-exclusive basis (e.g., NW, 02/15/82). In October 1982, the Federal Reserve, which had already shown signs since late summer of switching to a federal funds rate target, announced that it was “de-emphasizing” its M1 target on account of continuing financial innovations. Friedman’s reaction to the double-digit M1 growth that followed in 1982-83 was to predict a revival of inflation.

Indeed, from 1982 to 1985, Friedman repeatedly predicted a major revival of inflation that never occurred. In 1982 he predicted 8 percent inflation for 1983; the outcome was around 4 percent (FORT, 03/19/84). In July 1983, Friedman wrote, “We shall be fortunate indeed if we escape either a return to double-digit inflation or renewed recession in 1984” (NW, 07/25/83). In August 1983, he said, “U.S. inflation rates will rise appreciably in 1984, although it’s not yet determined where they’ll go from there” (TSN, 08/30/83). In April 1984, Friedman said, “I believe [the CPI] will be rising in the neighborhood of 8 to 10 percent in 1985.”\footnote{Even in November 1985, Friedman said that “Inflation is not dead. It will emerge once again and will be higher next year than it is this year. We almost surely are currently at the bottom of this inflationary episode and are likely to be starting up again” (NYDN, 11/13/85). Defying these predictions, inflation was consistently below 5 percent in every month from 1983 to 1986; moreover, apart from a brief uptick in early 1984, inflation continued to decline after 1982, and was lower in 1986 than it was in 1985.}

Questions that arise are why Friedman increased his reliance on M1 in 1982 even as the Fed was deemphasizing it; and why he in 1982-85 attached little weight to the position that financial innovations were distorting M1, even though he had taken this position himself in 1981. To answer these questions, it is worth going back to Friedman’s past preferences in defining money. As Friedman recalled in 1984, “In the past, I always found M2 to be a more reliable guide to economic events than the earlier M1.”\footnote{From his remarks in Heller et al. (1984, p. 51).} His and Schwartz’s Monetary History had used M2 as their money series, and they had defended this choice at length in their 1970 Monetary Statistics.\footnote{See Friedman and Schwartz (1963, 1970).} This position hardened in 1978 when Friedman saw the prospect of sweeps programs making M1 “a nearly useless aggregate” (NW, 10/30/78), though in practice sweeps did not become a pervasive distortion until the 1980s and 1990s.

In 1980, the Federal Reserve made M1 and M2 much broader definitions than they had previously been. The new definitions of money were based on the type of deposit, irrespective of whether the deposit was a liability of commercial banks or of nonbank financial institutions. (So, for example, demand deposits of nonbanks, previously included in neither M1 nor M2, were now included in both M1 and M2.\footnote{See Anderson and Kavajecz (1994).}) Friedman was initially inclined to rely on the new M2 and, as we have seen, argued in 1981 that M2 was more immune to distortions from financial innovation. But again, why did he come to rely on M1?

One major reason was the close relationship that M1 growth and nominal income growth enjoyed in the early 1980s—a “hot streak” of M1, discussed here later. In addition, M1 gave a more unambiguous picture of tight money over 1981 and into 1982 than did M2, and so gave an accurate signal both of the severe 1981-82 recession and the 1982-83 decline in inflation.

Friedman’s discounting of the Fed’s rationale for its deemphasis on M1 arose partly from experience; it was not the first time the claim had been made that financial innovations were disturbing velocity behavior. In January 1977, Friedman had said, “I have observed over a long period of time that whenever anything goes wrong with monetary policy, the favorite excuse of the monetary authorities is that there has been an exogenous
shift in the demand for money." \(^{37}\) Similarly, Friedman wrote in August 1982 that "the talk about changes in the demand for money is simply a red herring introduced by the Federal Reserve... In each case it has turned out that there has been no change in the demand for money..." \(^{38}\) Friedman was also chastened by his own lapse in 1972, when he had briefly persuaded himself that the high money growth of 1971-72 had been permanently absorbed by a velocity shift and so did not signal future inflation (\(NW,\ 10/16/72\)).

An article Friedman presented at the December 1983 American Economic Association meetings, and which was published in 1984, represented the high point of his confidence in M1 and so combined a number of judgments that he would retract within a few years. He stated (Friedman, 1984b, p. 398) that "Few if any monetarists ever recommended the use of such broad aggregates as the current M2 or M3 as monetary targets—certainly, this one did not." (This was not accurate—he had advocated M2 targeting in Friedman [1982a, p. 117] and would again from 1986 on.) He also stated that "The current M1 is conceptually...closer to the aggregate we [Friedman and Schwartz] labeled M2 rather than to our M1..." (This claim was also inaccurate: Federal Reserve data showed that old M2 and new M2 were more correlated than old M2 and new M1 before 1979. \(^{39}\)

Also, the new M1 shared with the old M1 the "stylized fact" of an upward velocity trend in the 1960s and 1970s, while the new M2 shared with the old M2 a basically trendless velocity over those decades.)

In retrospect, Friedman underestimated the impact of financial innovations on M1, as well as the extent to which high M1 growth reflected a recovery of real balances after a disinflation. He did acknowledge the latter as a factor in reducing the level of velocity in 1982-83 (\(WSJ,\ 09/01/83\)), but underestimated the extent of the increase in money demand. Part of his error here reflected his viewing M1, which is highly interest-elastic, as similar to old M2, whose demand function was less interest-inelastic. Consequently, he underestimated the impact that the shift to lower nominal interest rates from 1982 had in reviving the demand for real balances. In addition, while aware that much of the 3 percent per year pre-1981 trend growth rate of M1 velocity was due to the inflation-ary monetary regime, Friedman initially conjectured that even under price stability velocity would grow at about 1.5 percent per year because "technological improvements in cash management" were not yet exhausted. \(^{40}\) It soon became clear that disinflation since 1981 had instead eliminated the M1 velocity trend altogether.

In May 1986, Friedman undertook a fundamental review of the monetary data in light of developments in recent years, including his own poor forecasting record on inflation, and settled on M2 (with a one-time adjustment for the introduction of money market deposit accounts in the four months to March 1983) as the appropriate money series for U.S. monetary analysis. \(^{41}\) This became the position he stuck to. If, as Friedman thus concluded, monetary policy developments in the early 1980s should have been analyzed using M2 rather than M1, how does the picture of this period change? The basic assessment that monetary policy variability increased in the early 1980s is supported by using series other than M1. Friedman’s (1984a, p. 29) tabulation of money growth showed that the ups and downs of M1...


\(^{38}\) Friedman (1982b). Other critics of the Federal Reserve took the same perspective but expressed it with more temperate language. For example, Maxwell Newton, a monetarist financial columnist, wrote in October 1982: “The justification offered to an evidently credulous world for the Fed’s decision to allow the money stock M1 to float way above target in the coming weeks is the hoary old lie brought out by the Fed from time to time—namely that ‘financial innovations’ have made M1 less relevant or less accurate as a measure of ‘money.’ This lie has been proffered by the Fed for at least 40 years” (\(NYP,\ 10/12/82\)). By 1987, Newton had accepted that M1 actually was distorted by financial innovations and had switched to focusing on M2 in his analysis (\(TT,\ 07/20/87\)).

\(^{39}\) Correlations computed from the data for 1973:Q1–1979:Q4 tabulated in the Federal Reserve Bulletin (Simpson, 1980, p. 112) indicate that old M2 quarterly growth had a correlation coefficient of 0.54 with the new M1 growth series and 0.67 with the new M2 growth series. At various points, Friedman (e.g., 1984b, p. 398) claimed that the new M1 was similar to the old M2 because the new M1 series included interest-bearing deposits. But this claim did not really hold up. The velocity of M1 remained much more sensitive to variations in market interest rates than did M2, reflecting the fact that interest-bearing, variable-rate accounts were much more important for M2 than for M1.

\(^{40}\) Friedman (1984a, p. 58).

\(^{41}\) See Friedman (1988, pp. 229, 239-40).
and M2 growth over 1979-83 were basically similar, while Erceg and Levin (2003) estimate that the monetary policy swings in the early 1980s produced considerable variability in private sector estimates of the long-run inflation rate. Also consistent with these authors’ findings is Friedman’s statement that a principal channel through which money growth variability was manifesting itself in economic variability was interest-rate variability: As he put it, “What accounts for this unprecedentedly erratic behavior of the U.S. economy? The answer that leaps to mind is the correspondingly erratic behavior of interest rates” (NW, 02/15/82). Friedman also judged that the high variability had itself contributed to the downward pressure on aggregate demand, making the recession worse and producing a larger decline in inflation than would be expected from the decline in money growth alone (ESN, 10/21/82; Friedman, 1984b).

There are, however, some differences in an account of 1979-82 that uses M2 rather than M1. M1 growth unambiguously fell in the period from late 1979 to mid-1982. M2 on the other hand, as Friedman noted in 1981, was “flat” over 1978-81 in its annual average growth rates (NW, 09/21/81), at about 8.5 percent, and continued at around this rate in 1982. Seen from the perspective of M2 growth, the 1979-82 period amounted to a monetary policy sufficiently tight to hold the line on the reduction in M2 growth achieved in 1977-78, frustrating the strong upward pressure on the nominal quantity of money demanded that was coming from high rates of nominal spending growth.

In addition, while M2 velocity was roughly trendless since the mid-1950s, it nevertheless exhibited sustained movements over shorter periods, rising in the 1970s and falling about 7 percent over 1981 and 1982. In line with his 1979 discussion of why inflation had risen more steeply than M2 growth, Friedman attributed the high money growth relative to inflation to agents’ flight back into money once they were confident inflation had peaked (WSJ, 02/12/87). Studies by Friedman and others found that M2 velocity behavior in the postwar period could be accounted for by opportunity-cost variables, leading Friedman to conclude that velocity in the 1980s was explicable “despite all the talk about how the relation between money and other variables has shifted drastically in recent years.”42 M2 velocity settled down after 1983, by which time most of the larger opportunity-cost movements had taken place.43 Therefore, Friedman was able to observe in June 1988 that from 1983:Q1 to 1988:Q1, nominal GNP had grown 7.4 percent per year and M2 7.5 percent: “How much closer can you get for a five-year period?” (WSJ, 06/22/88).

Friedman’s reliance on M1 in the mid-1980s led him not only to successive errors in forecasting inflation, but also a well-publicized erroneous short-term forecast for 1984. The sharp decline in M1 growth during the second half of 1983 led him to predict a recession starting in 1984:Q1. “Mr. Reagan may well face a very difficult situation by next fall,” Friedman said in late 1983. “There is a real threat of a recession in the first half of 1984” (NYT, 12/31/83). Instead, real growth continued throughout 1984 and was exceptionally strong in 1984:Q1. “I have no easy explanation of what went wrong,” Friedman said after the strong output numbers were released (NYT, 04/30/84).

Friedman was, however, not as misguided in his recession forecast as he had been in his inflation predictions. While he was proven wrong in his forecast of a recession in 1984:Q1, Friedman was on the right track when he continued to maintain that “there is not enough monetary fuel going into the economic engine” and that this had to show up before long in slower real and nominal income growth (NYT, 04/30/84). The Fed had, indeed, withdrawn a lot of monetary stimulus in the second half of 1983 and continued to do so in 1984. The move to restriction was evident not only in the questionable M1 data but in monetary base growth and rising real interest rates. What happened, in line with Friedman’s expectation, was a substantial slowdown in real and nominal income growth during 1984—but, contrary to his expectation, it commenced in 1984:Q2 and was not associated with a recession.

---

42 Friedman (1988, p. 229).
Rather, the United States had a “soft landing,” instead of an overreaction to the unsustainably high real growth of 1984:Q1. In fact, the early-to-mid 1984 period is now regarded as the pivotal period in which the United States transitioned to the “Great Moderation” and so a much more stable business cycle (see, e.g., Bernanke, 2004; Blanchard and Simon, 2001; McConnell and Perez-Quiros, 2000; and Stock and Watson, 2002). Based on the prior experience that Friedman had studied, big withdrawals of monetary stimulus implied a recession; but the 1984 episode produced an example of a better transition of the economy from over-rapid growth.44

On top of his very public errors in forecasting inflation and recession, 1984 proved a dismal year for Friedman on several other fronts. His and Rose Friedman’s television program The Tyranny of the Status Quo (a follow-up to their 1980 success, Free to Choose) flopped, Newsweek dropped him as a columnist, and in October he suffered a heart attack. In December he was able to record success by putting things into a longer perspective, telling the Wall Street Journal, “It’s widely accepted that you can’t hold inflation down unless you hold down the money supply. Thirty years ago, people didn’t agree with that” (WSJ, 12/10/84).

The Battle of the Friedmans

In light of Milton Friedman’s awful forecasts for 1984, finding fault with his monetary analysis during this period should be like shooting fish in a barrel. It happens, however, that two such critiques, by Benjamin Friedman and Alan Blinder, that appeared during the 1980s contain oversights that, when corrected, lead to support for some aspects of Milton Friedman’s analysis during the early 1980s.

Benjamin Friedman (1988, p. 61) critically considers Milton Friedman’s (1984b) statement that the relation between nominal GNP and M1 growth was “unusually close” and, using 1987-vintage data, states that “The GNP-to-lagged-M1 correlation was not ‘unusually close’ during 1979:Q4–1983:Q4 compared with the past...The correlation of 0.45...is essentially identical to that for the previous 79 quarters.”

On closer inspection, this claim of Benjamin Friedman’s does not stand up as a refutation of Milton Friedman, as it overlooks key elements of what was said in the 1984 article. What Milton Friedman (1984b, p. 399) said of M1 growth and nominal GNP growth was “From 1981 on, the relation is extraordinarily close—indeed, instead of being less close than during the earlier [pre-1979] years, is considerably closer...Two things are notable...first, the lag is both shorter on the average...second, the relation is unusually close” (emphasis added).

In evaluating Milton Friedman’s position, Benjamin Friedman overlooked the statement that the claimed close relationship was from “1981 on.” Once this aspect of Milton Friedman’s statement is taken into account, Benjamin Friedman’s challenge to Milton Friedman’s contention regarding the M1/nominal income relation under the new operating procedures is refuted. As Table 1 shows, the peak M1/future income correlation is unusually high from 1981 on (whether 1981-83 or 1981-84 is considered), compared with pre-1979 correlations. Moreover, Milton Friedman’s statement, ignored by Benjamin Friedman in his calculation of correlations, that the lag was shorter after 1979 is also supported: The peak correlation is with a two-quarter lead for M1 growth up to 1979, but with a one-quarter lead for both 1979-83 and 1981-83.45 Thus, when the sample considered is restricted to begin with the period to which Milton Friedman explicitly indicated that his comments pertained, Milton Friedman’s contention that the relationship is unusually close is confirmed, while Benjamin Friedman’s counterclaim is rejected.

On a similar note, Blinder (1984, p. 267) offered as evidence of a “shift in the money-income relationship” a table which “show[ed]...
that during the nine-quarter period 1981 Q1 to 1983 Q1 there was actually a remarkably strong negative correlation between money growth and nominal GNP growth.” But as Table 1 shows, the data given in Blinder (1984) again confirm the strong positive correlation between M1 growth and next-quarter nominal GNP growth claimed by Friedman (1984b), for samples ending in both 1983:Q1 and 1983:Q4. The negative correlation noted by Blinder is evident only in the contemporaneous relationship. The specific evidence offered by Blinder in 1984 is therefore less a demonstration of the absence of a money-income relationship than it is confirmation of Friedman’s (1980, p. 59) warning that “failure to allow for lags in reaction is a major source of misunderstanding” and of the 1997 observation of one former policymaker that “lags tend to be trivialized or ignored in academia...Failure to take proper account of lags is, I believe, one of the main sources of central bank error.”

The breakdown of M1’s tight early-1980s relationship with nominal income led Friedman back to M2, but he conjectured that the same break had led critics of monetarism to overemphasize the extent to which money/income relationships had changed in the 1980s. Friedman came to see the closeness of M1 to GNP from 1981 as a “hot streak” that ended in 1984 and judged that M1 was, in fact, more distorted by financial innovation than other aggregates—i.e., the monetary base and M2 (WSJ, 09/18/86; Friedman 1988, p. 224). More generally, Friedman reaffirmed, “The relationship between the money supply and the state of the economy has always been a loose one” (SFC, 07/04/86).

Friedman also stressed that, while financial deregulation had distorted monetary aggregates such as M1, he supported deregulation. “Deregulation is good. That is the one good thing—the only good thing, in my opinion—that has come out of this [i.e., the inflation/disinflation experience].”

Indeed, Friedman (1960, 1970c) had called for the deregulation of commercial banks’ interest rates that ultimately took place in the 1980s. He did argue that the inflation and interest-rate peaks in the 1970s and early 1980s had exacerbated the distortions arising from financial regulation and thus increased the distortion to monetary aggregates when deregulation took place. He did not feel that financial regulation was a cornerstone for control of monetary aggregates or of aggregate demand: Open market operations were sufficient for monetary control (Friedman, 1974; FT, 02/23/87), and lack of reg-

---

Table 1

**Correlations of Nominal Income Growth and Prior M1 Growth**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In same quarter</td>
<td>0.480</td>
<td>0.098</td>
<td>-0.196</td>
<td>-0.152</td>
<td>-0.425</td>
</tr>
<tr>
<td>One quarter earlier</td>
<td>0.449</td>
<td>0.543</td>
<td>0.660</td>
<td>0.604</td>
<td>0.515*</td>
</tr>
<tr>
<td>Two quarters earlier</td>
<td>0.515</td>
<td>0.227</td>
<td>0.595</td>
<td>0.570</td>
<td>0.024†</td>
</tr>
<tr>
<td>Three quarters earlier</td>
<td>0.383</td>
<td>-0.217</td>
<td>-0.260</td>
<td>-0.155</td>
<td>-0.741‡</td>
</tr>
</tbody>
</table>


---

46 Blinder (1997, p. 8).
ulation of financial institutions “in no way enables them to escape the discipline of the Fed’s monetary actions” (WSJ, 06/30/75).

As shown above, the main critiques of Friedman’s 1980s positions have their own weaknesses. Nevertheless, Benjamin Friedman (1988) was on solid ground in his demonstration that Milton Friedman (1984b) forecast higher inflation for 1983-85 from higher M1 growth and that this proved very erroneous. “I’ve gone back to using M2 and only M2,” Milton Friedman told the present author much later. “I slid away from the true and narrow path in the 1980s, and was sorry for it!… I think those predictions in the ’80s were bad, but I think on the whole my Newsweek predictions stand up pretty well."

FROM VOLCKER TO GREENSPAN: 1986-92

Friedman remained a persistent critic of Chairman Volcker, criticizing Volcker’s reappointment by President Reagan in July 1983. Speaking in October 1983, Friedman acknowledged that Volcker had “brought inflation down,” but maintained that it was still “not a good monetary policy” because it was so erratic. In early 1987, Friedman returned to this theme, criticizing Volcker’s current “seat of the pants” policy and claiming of 1979-82 that “If somebody had wanted deliberately to discredit monetarism they would have done what Volcker did” (FT, 02/23/87).

Friedman’s relations with Volcker’s successor, Alan Greenspan, would be vastly better. Though Friedman had rarely mentioned Greenspan in his writings, the two were old friends. When Greenspan was appointed Chairman of the Council of Economic Advisors in 1974, Friedman had praised him as a “very good man” who would resign on principle if there were another 1971-style lurch to expansionary policies (CT, 07/17/74). In turn, Greenspan in late 1986 had praised Friedman as one of the 20th century’s major intellects (SFC, 12/26/86). After Greenspan was nominated as Chairman, Friedman said, “I’m a good friend of Alan Greenspan. He’s an able person” (MINN, 06/03/87), though Friedman maintained he would still rather have the Federal Reserve replaced by a computer. “I send my congratulations to Alan. And my sympathy” (SFC, 06/08/87).

In a 1981 appearance at a Brookings Institution meeting, Greenspan had stressed the value of M2 as a monetary indicator, and M2 did figure quite heavily in monetary policy discussions in Greenspan’s first four and a half years as Chairman. Aside from a loosening after the 1987 stock market crash, Greenspan undertook a moderately restrictive policy in his first two years that saw annual M2 growth fall below 5 percent and the federal funds rate briefly reach double digits. Friedman approved, offering his opinion that Greenspan “on the whole has been doing very well” (TST, 03/11/89). As for inflation, Friedman said that it was “likely to decline sharply over the next several years” due to the fall in M2 growth (WSJ, 07/05/89). At the end of the decade Friedman’s assessment was very optimistic: “There’s no reason why the ’90s shouldn’t be as good as the ’80s, or better. There’s no reason we shouldn’t have a decade of rapid growth and relatively low inflation” (TIME, 01/01/90).

Friedman’s optimism about the 1990s was eventually vindicated in spades. But in the interim, there was a spike in inflation in 1990 and what Friedman in May 1992 called “three years of stagnation” from mid-1989, including the 1990-91 recession. By late 1992, while noting the achievement of low inflation, Friedman felt that U.S. monetary policy had begun to exhibit “perverse fine-tuning”—an inadvertently restrictive policy due to giving too little weight to low money growth and too much weight to low interest rates in judging conditions. Friedman called for a more expansive monetary policy consistent with bringing M2 growth back inside its announced target range. Doing so, Friedman wrote, would mean that inflation would converge...
on its ideal value—“a level that would make it irrelevant to individual and business decisions” (WSJ, 10/23/92). This form of words describing the ideal inflation rate was similar to those that became prevalent among policymakers later in the 1990s.53

ENTERING THE NEW ECONOMY: 1993-2001

In October 1993 Friedman wrote that “the opportunities are there for the making of a second industrial revolution” (FEER, 10/28/93)54 from the previous decades’ breakthroughs in information and telecommunications technology. In August 1995, however, he added his voice to those economists expressing puzzlement that U.S. productivity growth did not seem to have increased in the information age (WSJ, 08/01/95). It transpired that the U.S. productivity trend was already undergoing an upward shift as he wrote those words.

The mid-1990s also gave Friedman opportunities for introspection, including an appearance on CSPAN’s program Booknotes in 1994. The question of how many books he had written prompted Friedman’s reflection, “Oh, I don’t know—fifteen…[T]here’s no question that the most influential book I’ve written is not Free to Choose, but a book that sold probably one-twentieth as many, five percent as many copies, namely A Monetary History of the United States, which I wrote jointly with Anna Schwartz…A Theory of the Consumption Function is, in my mind, the best thing I ever did as piece of science, [while] Monetary History is undoubtedly the most influential, and Free to Choose is the best selling, so they are not similarly characterized” (BOOKN, 11/20/94).55

Friedman had predicted in early 1991 that inflation would fall to 2 percent in coming years, based on the low M2 growth maintained under Greenspan (FUT, 03/01/91). In January 1998, by which time inflation had stood at about this level for five years, Friedman said in a television interview that “Right now, inflation is relatively low. Alan Greenspan and the Federal Reserve have done a good job of keeping monetary growth low and fairly steady” (CNN, 01/23/98).

Shortly afterwards, Friedman gave a more detailed evaluation of the Greenspan regime in an interview for Barrons (BAR, 08/24/98). He again praised the outcome of “the lowest and least variable rate of monetary growth” compared with comparable periods. But he also acknowledged some weaknesses of M2 as an indicator during the 1990s: “I think there is no doubt that in the Nineties, from ’92 to ’95, around there, there was a very sharp uptick in the velocity of M2 and that targeting money supply at that time in a rigid fashion would not have been a good thing to do… I only say in retrospect that Greenspan did the right thing in abandoning primary reliance on M2 during that period.” While reaffirming that “I don’t like operating through interest rates” and restating his confidence in longer-term money-income relationships, Friedman said that Greenspan had “done a very good job of threading his way through those [velocity] difficulties.” This observation would be fleshed out in the thermostat hypothesis that Friedman championed in the 2000s.

The Thermostat Hypothesis

In 1966 Alan Walters, a U.K. monetarist, had observed that “If the [monetary] authority was perfectly successful then we should observe variations in the rate of change of the stock of money but not variations in the rate of change of

53 Friedman had given a related formulation in 1981: “I think the best of all worlds is one where you have no inflation and no indexation. Indexation is not a good thing in and of itself” (NZH, 04/18/81). In his 1992 article Friedman wrote as though the definition of price stability he gave was also that stated by the Federal Reserve, which is consistent with Orphanides’s (2006) tracing of this definition to Paul Volcker.

54 Friedman attributed the post-1973 slowdown in economic growth to greater government intervention in the economy since the 1960s (NW, 08/23/82). While he believed that inflation (and false cures for inflation) had damaged potential output, Friedman did not believe that the restoration of low inflation would in itself restore high growth. He said in 1982 that low growth and high inflation were “two distinct phenomena, and monetarism is concerned only with the inflation problem” (DS, 01/18/82; my translation) and reaffirmed in 1987 that “it’s a great mistake to connect too closely the level of inflation and the rate of economic growth” (JP, 11/11/87).

55 Besides Monetary History, the books Friedman was referring to here were Friedman (1957) and Friedman and Friedman (1980).
income...[a]ssuming that the authority’s objective is to stabilize the growth of income.” The basic idea behind this statement is that if a variable Y has a stochastic relationship with a variable X that monetary policy can affect, a policy that stabilizes Y will lead to fluctuations in X to offset other potential sources of variation in Y. The result will be a low empirical correlation between X and Y, despite the two enjoying a structural relationship.

A student of Friedman’s, Levis Kochin, worked in the 1970s on the relevance of this idea for money-income relations,57 while Poole (1995) argued that roughly optimal policy since 1982 had produced low correlations between inflation and money growth.

Friedman initially showed little interest in ideas along these lines. He acknowledged in 1984 that successful stabilization should mean that “high monetary variability...[is] associated with low economic variability,” but believed that the relevance of this was contradicted by the positive relationship between monetary variability and output (and nominal income) variability in the data.58 This was, of course, in line with his long-held position that stabilization policy had often been tried, but had proved counterproductive. Thus in 1967 Friedman had described the Federal Reserve’s stated aim of responding to emerging economic conditions as “a formula guaranteed to produce bad policy” (NW, 01/09/67) and had written that forward-looking formulations of this approach were likely to be destabilizing, with policymakers “vainly trying to lean against next year’s wind which, in the process, they are themselves stirring up” (WP, 11/05/67).

But by early 2000, however, in light of the much-reduced degree of output variability, Friedman had reached the view that stabilization policy had been executed successfully by the Federal Reserve in recent years and was a major reason for high M2 growth variability relative to output variability since 1992 (see Taylor, 2001). To articulate the idea, Friedman made an analogy between successful stabilization policy and the behavior of a thermostat. This was a revival of an analogy prevalent in the 1950s; for example, a commentator on U.K. economic affairs, Anthony Crosland, wrote in 1956 that economic policy in practice would never achieve the ideal where “demand could be exactly set, as though by a thermostat control, at just the right point.”59 U.S. monetary policy in the 1990s, as Friedman saw it, had come closer to achieving the “thermostatic” ideal than many, including himself, had thought practicable.

Friedman was thus able to identify a period in which money growth targeting was inferior to the monetary policy actually pursued. He had always argued that this was possible in principle: Friedman (1960, p. 98) had said that there was “little to be said in theory” for a constant money growth rule and “persuasive theoretical grounds” for varying money growth “to offset other factors.” Rather, the money growth rule recommendation emerged from the uncertainty underlying the policy responses required for successful stabilization, which meant that “deviations from the simple rule have been destabilizing rather than the reverse.” Nor in later expositions did Friedman claim that money growth targeting was optimal in theory.60

There are some weaknesses in the thermostat hypothesis as an explanation for shifting money-income relations in the data.61 For his part, Friedman advanced the thermostat story only to account for some of the looseness in money-

57 Kochin (1973); see also Kishor and Kochin (2004).
58 Friedman (1984a, p. 34).
59 Crosland (1956, p. 398). Similarly, Warburton (1953, p. 10), in commenting on activist rules that Friedman had advanced in the 1940s, said that they were designed to be “a self-regulating mechanism, such as a thermostat...”
60 For example, in Friedman (1970b, pp. 18-19) the money growth rule was acknowledged as inferior in principle to “deliberate changes in the rate of monetary growth...to offset other forces,” this alternative being rejected by Friedman “until such time as we demonstrably know enough to limit discretion by [using] more sophisticated rules.” While usually made with nominal income stabilization as the implicit criterion, this conclusion was also applied to direct targeting of prices (or inflation) by Friedman (1966b, 1982b).
61 For example, it does not distinguish between weak correlations that emerge because policymakers make use of a stable structural relationship between money and the economy and the case where much of the weakness reflects the fact that the structural relationship is itself undergoing change.
income relations on quarterly data; he continued to believe that longer-period stability in nominal income required moderate rather than high variability in money.

**21st Century Developments**

The U.S. stock market peaked in early 2000. Friedman attributed some of the stock market’s strength to higher M2 growth (WSJ, 01/22/02; Friedman, 2005). Nevertheless, he concurred that stock prices had far exceeded economically justified values (FORB, 05/03/99). Thus, Friedman noted in October 2001, the Federal Reserve faced the issue of what to do in the aftermath of an “excessively bullish stock market” (UPI, 10/09/01).

“The monetary policy Greenspan is following I think really has no precedent,” Friedman remarked on this occasion, adding that he had a “great deal of sympathy” with the interest-rate cuts that Greenspan had initiated in January 2001. Whether these cuts were sufficient, Friedman said, depended on “what their effect has been on monetary growth,” though Friedman cautioned that money growth would now be difficult to interpret because “there’s no doubt one of the effects of September 11 will be to increase the demand for cash balances” (UPI, 10/09/01).62 Four years later, Friedman (2005) reached a favorable verdict on the period. He maintained that monetary policy behavior, by keeping M2 growth high in the wake of the stockmarket decline, had stopped the equity price collapse from leading into anything more than a shallow recession. The monetary stimulus had also been withdrawn at “just about the rate required for a rapidly growing non-inflationary economy” (WSJ, 04/07/05).

**STILL A MONETARIST: 2002-06**

In June 2003 a profile appeared in a Financial Times magazine supplement that included brief remarks from Friedman on monetary targeting:

> “The use of quantity of money as a target has not been a success. I’m not sure I would as of today push it as hard as I once did” (FT, 06/07/03). To readers of Friedman’s interviews in 1998-2000 with Barrons and John Taylor, the 2003 quotations were not a great surprise. Friedman had previously made it clear that he had been unhappy with how monetary targeting had typically been implemented in practice and had been impressed by how U.S. monetary policy in recent years had done better on economic stabilization than he had thought practicable and outperformed what he would expect from a money growth rule. The 2003 remarks were not a break with his position in these earlier interviews.

Some commentators, however, took the remarks in isolation and interpreted the 2003 interview as repudiation by Friedman of monetarism.63 William Keegan, the economics editor of the London Observer and a longtime critic of Friedman’s, used the quote as the basis for an article entitled “So Now Friedman Says He Was Wrong” and asserted that “Friedman, now 91 [sic; 90], … feel[s] it is time to own up. It is ‘true confession’ time” (OBS, 06/22/03). In light of this and similar “spins” on his remarks, Friedman decided to produce a lengthy exposition of his interpretation of the monetary policy developments in the preceding 15 years. The resulting article, in the Wall Street Journal, would disappoint anyone looking for confirmation that he had turned his back on monetarism. On the contrary, while again praising recent policy for doing better than a constant money growth rule, Friedman maintained the following position: “Velocity is ordinarily very stable, fluctuating

---

62 The powerful effect of national emergencies in increasing the demand for real money balances (equivalently, reducing velocity) was a long-standing Friedman theme (e.g., Friedman and Schwartz, 1963, pp. 673-75; Friedman and Schwartz, 1982, pp. 4, 228; Friedman, 1984b, p. 399).

63 Over the years, discussions of Friedman often seized on new public statements of his and alleged that these amounted to recanting of previous positions. For example, a commentator reacting to a talk Friedman gave in 1981 wrote, “Milton Friedman has made a significant change in his doctrine since his previous visit to Australia in 1975” (SMH, 04/19/81). And in 1987 John Kenneth Galbraith said in an interview, “I think Professor Friedman, who was an interesting man, has disappeared into the shadows very much. He has accused the Reagan Administration of prime incompetence in the managing of economic policy” (SMH, 02/24/87). These claims of reversals in position, like those made in 2003, were unfounded—consisting either of incorrect readings of Friedman’s earlier statements or illegitimate extrapolations from his later ones. Nevertheless, as we have seen, Friedman did change his position on certain issues, such as the relative merits of M1 and M2, and on the lag between monetary actions and inflation.
only mildly... The $MV = Py$ key to a good thermostat was there all along” *WSJ, 08/19/03.* Indeed, at the time that he was working on this article, Friedman noted to the present author:

In my original support for a straight money target, I always emphasized that it was partly a case based on ignorance, based on the fact that we really did not understand sufficiently well the detailed relationship between money, income, interest rates, and the like to be able to fine-tune, that our goal should be to develop a detailed enough understanding so that we could do better than a simple constant monetary growth target.

However, I believe still, as I did then, that constant monetary growth would produce a highly satisfactory price path, and, if it enabled you to get rid of the Federal Reserve System, that gain would compensate for sacrificing the further improvement that a more sophisticated rule could produce.64

Friedman also took the opportunity of the *Wall Street Journal* article to offer a story about why monetary policy had changed since the inflationary 1960s and 1970s. Over the years—in interviews he gave in the late 1980s (e.g., *FT, 02/23/87; SFC, 08/18/87*) and even as late as 2000 (in the Taylor interview)—Friedman had appealed to far-fetched and unconvincing explanations for why policymakers had shifted to low inflation. These stories centered on the political profitability of inflation having been reduced by indexation of tax brackets and government securities. Accounts such as this attributed too much knowledge to pre-1980s policymakers regarding how to control inflation.

Now, instead, Friedman returned to themes that he had put forward during the 1970s and early 1980s, which emphasized the errors in policymakers’ basic theories of inflation. In 2002 he noted that policymakers in the United States and United Kingdom “used to say inflation was caused by trade union power, cost pressures, and so on; nobody talks like that any more.”65 His 2003 article elaborated on this “shift in the theoretical paradigm.” The major change since the 1970s, Friedman now affirmed, was the shift by policymakers from adherence to cost-push views to acceptance that inflation was a monetary phenomenon (*WSJ, 08/19/03*). While pleased with this development, Friedman was unhappy with the fact that focus on monetary aggregates had receded. Always ready with an analogy, Friedman had written in 1966 that discussing monetary policy without mentioning money was like writing a book about love without mentioning sex.66

As far as current economic developments were concerned, in 2003 Friedman dismissed “all the talk about deflation.”67 A necessary condition for deflation, monetary contraction, had not occurred. As for whether falling high-tech goods prices could produce deflation, Friedman had dismissed this back in 1978: “The high price of cars doesn’t cause inflation any more than a drop in the price of hand calculators causes deflation” (*TDN, 05/19/78,*).

In December 1986 Friedman, while acknowledging some discomfort at being better known for the likes of *Free to Choose* than for his monetary economics, had foreshadowed that his research output would dry up—“people’s capacities to do scientific work decline with age” (*SFC, 12/26/86,*). As things turned out, he never could stop completely, and almost 20 years later, in August 2006, Friedman (now age 94) was putting the finishing touches on his latest research paper. Friedman delivered the paper to the London, Ontario, conference by video link from San Francisco on August 18. The subject was tradeoffs in monetary policy, and the paper was a contribution to a Festschrift for Friedman’s former student David Laidler. As noted above, Friedman had used the thermostat hypothesis to help explain recent money/income patterns. That hypothesis was based on the observation that M2 variability was high relative to output variability during the 1990s. But in reviewing the longer picture using annual data, Friedman now stressed that the larger event was the large reduction in the absolute level of M2 growth variability in recent decades.

64 Milton Friedman, email to author, July 21, 2003.


(This is also evident in quarterly data: Figure 1 displays the 14-quarter standard deviation of money growth, a key statistic in the Friedman-Modigliani [1977] debate.) “The collapse of the variability of output,” Friedman wrote, “is clearly an effect of the collapse of monetary variability. In my opinion, the same results could have been obtained at any earlier time and can continue to be achieved in the future.”

68 Friedman (2006).

**CONCLUSION**

Despite his prolific writing, Milton Friedman did not provide a single, detailed record of his views on post-1960 U.S. monetary history. He instead relayed those views in a number of forums, including regular and semi-regular newspaper and magazine columns, books and articles, congressional testimony, and interviews with economists and the national and world media. The objective of this paper has been to bring together, using archival material from several countries, information about Friedman’s views and predictions regarding post-1960 U.S. monetary developments and to evaluate his interpretations in light of subsequent events. The 1960s and 1970s were notable for policymakers resistance to, then acceptance of, Friedman’s consistently held views on exchange rates and on the control of inflation. Probably the most important changes in Friedman’s outlook over the 1980s, 1990s, and 2000s were a firming of his preference for M2 as the definition of money, following several years of forecast errors using M1, and his concession that stabilization policy in the United States since the mid-1980s had been more successful than he had thought possible. Friedman’s outlook nevertheless remained monetarist, and he continued to advocate a fixed money growth rule, arguing that such a constraint on policymaking was desirable.

In addition to giving an impression of Friedman’s perspective on monetary policy
practice, this analysis must shed some light on Friedman’s theoretical framework, since, as he once said, there is always a “general economic model which each of us must have, underlying our specific temporal predictions.”

REFERENCES


Milton Friedman, interview with the author, Stanford University, January 22, 1992.


Nelson


APPENDIX A

Abbreviations for Periodicals Cited in Text

BAR—Barrons (U.S.); BOOKN—Booknotes (CSPAN television program; transcript); CHA—Challenge (U.S.); CNN—Moneyline (CNN television program; transcript); CSM—Christian Science Monitor (Boston); CT—Chicago Tribune; DC—Daily Courier (Connellsville, PA); DFP—Detroit Free Press; DOM—Dominion Post (Morgantown, WV); DS—Der Spiegel (Hamburg, Germany); ESN—Edmonton Sun (Alberta, Canada); FEER—Far Eastern Economic Review (Hong Kong); FORB—Forbes magazine (U.S.); FORT—Fortune magazine (U.S.); FT—Financial Times (London); FUT—Futures (U.S.); IT—Irish Times (Dublin); JOH—The Star (Johannesburg, South Africa); JP—Jerusalem Post (Israel); JT—Japan Times (Tokyo); KCS—Kansas City Star (Kansas City, MO); LAT—Los Angeles Times; MHER—The Herald (Melbourne, Australia); MINN—Minneapolis Star-Tribune; NW—Newsweek (U.S.); NYDN—Daily News (New York); NYT—New York Times; NZH—New Zealand Herald (Wellington); OAK—Oakland Tribune (Oakland, CA); OBS—The Observer (London); OKL—The Daily Oklahoman; PHB—Philadelphia Bulletin (Philadelphia, PA); PHI—Philadelphia Inquirer; SCMP—South China Morning Post (Hong Kong); SFC—San Francisco Chronicle; SLGD—St. Louis Globe-Democrat; SMH—Sydney Morning Herald; SML—Sunday Mail (Brisbane, Australia); STE—Sunday Telegraph (London); STR—Straits Times (Singapore); TDN—The Detroit News; TG—The Guardian (London and Manchester, U.K.); TIME—Time magazine (U.S.); TSN—Toronto Sun (Ontario, Canada); TST—Toronto Star (Ontario, Canada); TT—The Times (London); UPI—UPI news wire; USAT—USA Today; USNR—U.S. News and World Report; WCPD—Weekly Compilation of Presidential Documents (Washington, DC); WP—Washington Post; WSJ—Wall Street Journal (New York).
APPENDIX B

Chronological List of Media Articles Referred to in Text


Albert L. Kraus, “‘Mini-Recession’ Seen,” Dominion Post, August 24, 1969, p. 3D.


“Friedman Predicts Fall in Inflation Rate to 6pc,” *Straits Times*, February 19, 1982.


Bill Nichols and Tom Vanden Brook, “Tributes Pour In for Ford,” *USA Today*, December 27, 2006, pp. 1A-2A.
The Lower and Upper Bounds of the Federal Open Market Committee’s Long-Run Inflation Objective

Daniel L. Thornton

It is widely acknowledged that the Fed can control the average inflation rate over a period of time reasonably well. Because of this and the Federal Open Market Committee’s (FOMC’s) long-standing commitment to price stability, the author argues that the FOMC has an implicit long-run inflation objective (LIO)—lower and upper bounds to the long-run inflation rate. He shows that the statements made by the FOMC in 2003 clarified the lower bound of its LIO and that the average of long-run inflation expectations responded by rising about 80 basis points. Moreover, consistent with reducing the market’s uncertainty about the FOMC’s LIO, long-run inflation expectations became more stable. The FOMC has recently been more specific about the upper bound of its LIO as well. The FOMC could eliminate the remaining uncertainty by establishing an explicit, numerical inflation objective. (JEL E50, E52, E58)

Currently there are at least 21 countries with inflation targets. Inflation targeting is marked by a numeric inflation objective that the central bank attempts to achieve over a reasonably well-specified time horizon. The numeric target is most often a range in which inflation is permitted to vary over a multiyear horizon. Among the industrialized nations of the world, the central banks of the United States and Japan stand out in not adopting a formal, numeric inflation target.

Economists and policymakers have raised a number of objections to establishing an explicit numeric inflation target in the United States. The most important of these objections are discussed in Federal Reserve Bank of St. Louis (2004) and, hence, will not be discussed here. When he was a Federal Reserve System Governor, Bernanke (2004) suggested that the Fed take an “incremental move toward inflation targeting, in the form of the announcement of a long-run inflation objective” (LIO), which I take to be congruent with Bernanke’s optimal long-run inflation rate, which he defined as “the long-run (or steady-state) inflation rate that achieves the best average economic performance over time with respect to both the inflation and output objectives.” I argue that (i) the Federal Open Market Committee (FOMC) already has what can be reasonably characterized as an implicit LIO and (ii) having an implicit LIO is a consequence of the conventional wisdom that the Fed (indeed, all central banks) can control the long-run inflation rate. I present evidence

1 See Rasche and Williams (2005) for a list of these countries and the details of their inflation targets.


that the market believes that the FOMC has an implicit LIO by showing the market’s reaction to statements made by the FOMC in 2003.

The adoption of inflation targeting by many of the world’s central banks is directly linked to changing views about the central bank’s ability to control inflation. The now conventional wisdom that the Fed and other central banks control long-run or steady-state inflation is the primary reason central banks have an explicit or implicit LIO. Hence, the paper begins with a brief review of the evolution of thinking about the ability of central banks to control inflation during the nearly 100 years of the Federal Reserve System.

CENTRAL BANKS AND INFLATION

Although it is now widely believed that the FOMC has an implicit LIO, this has not always been the case. Indeed, there has been an ebb and flow in economists’ and policymakers’ thinking about the ability of the central bank to control inflation over the near 100 years of the Federal Reserve System. The quantity theory of money was the dominant theoretical paradigm in the economics profession at the time of the Fed’s founding. The quantity theory was embodied in the gold standard, where long-run price stability was maintained through an automatic equilibrating process known as the price-specie-flow mechanism. In essence, the quantity theory asserts a strong causal link between money growth and inflation. In fiat monetary systems, where the currency is not backed by gold or other precious metals, the quantity theory asserts that central banks control inflation by controlling the rate of growth of the money supply. Permitting the money supply to grow more rapidly than is warranted by labor force and productivity growth causes inflation. Too slow money growth results in deflation. The quantity theory hypothesizes that inflation is always and everywhere a monetary phenomenon that central banks can control.

The early Fed was skeptical of the quantity theory and followed what is called the real bills doctrine. The real bills doctrine hypothesized that reserves supplied by the Fed for “productive purposes”—to enhance the production of goods and services—would not be inflationary.

Economists’ thinking about what central banks could and could not do about inflation changed markedly following the publication of Keynes’s The General Theory of Employment, Interest and Money in 1936 and the subsequent Keynesian Revolution. For various reasons, perhaps the most important of which were concerns that the central banks had limited ability to control the supply of money and the belief that the demand for money was not stable, economists and policymakers came to question central banks’ ability to control (or even substantially influence) inflation. The conventional wisdom was that inflation was not the consequence of an excess supply of money, but of an excess of aggregate demand—the demand for all goods and services—which could occur independent of the relative supply of money. Moreover, it was believed that monetary policy had relatively little impact on aggregate demand. A small effect on aggregate demand translates into a small effect on inflation. Consequently, it was thought that the Fed could do relatively little to control inflation.

A high point of the belief that inflation is not a monetary phenomenon in the United States occurred in 1974 with the federal government’s WIN (Whip Inflation Now) campaign. The WIN campaign was an attempt to spur a grassroots movement to reduce inflation through a combination of public and private measures to reduce aggregate demand and, consequently, “demand-pull” inflation. Fiscal policy, increased saving, and other factors—not monetary policy—were thought to be the keys to curing the nation’s inflation woes.

The monetarist counterrevolution and the success with anti-inflation monetary policy in

---

4 See Nelson (2005a,b) and Nelson and Nikolov (2004) for a discussion of how neglect of monetary factors and belief that other factors caused inflation played a role in several of the great inflations of the 1970s.

5 See Bernanke (1993) for an excellent and concise discussion of some of these arguments.

6 On this note it is interesting to observe that inflation has been trending lower despite a corresponding negative trend in the United States saving rate.
the late 1970s and early 1980s dramatically changed economists’ and policymakers’ thinking about what central banks could do about inflation. Despite the fact that the Fed and nearly all other central banks now use a short-term nominal interest rate instrument and not monetary aggregates to conduct policy, it is widely acknowledged that central banks can control inflation on average over a period of time.

If central banks can control the average long-run inflation rate, it follows that they are responsible for it, whether they, or the governmental agencies to whom they report, establish a specific numerical objective for it or not or whether, by their actions, they effectively abdicate their responsibility: The Fed was responsible for the great inflation of the 1970s, even though it was not intended. For most central banks, the realization that they control the long-run inflation rate has led to the adoption of formal inflation objectives or inflation targets. Although it has never adopted a specific inflation target, the Fed has had a long-standing objective of “price stability.” Former Chairman Greenspan repeatedly noted that the primary role of monetary policy is to promote sustainable economic growth over time by fostering price stability. Chairman Bernanke has reiterated the causal link between price stability and economic growth, noting that “price stability is essential for strong and stable growth of output and employment.” Given the belief that the Fed can control the long-run inflation rate and its stated objective of price stability, it follows that the FOMC has an implicit LIO. That is, the FOMC has an implicit range of inflation that is consistent with its objective of “price stability.”

Despite his repeated assertion that sustainable economic growth is inextricably linked to price stability, former Chairman Greenspan opposed adopting a specific numerical LIO on the grounds that no price index adequately reflected price stability and because of political concerns (Greenspan, 2002; FOMC, 1995). In contrast, Chairman Bernanke has been a steadfast proponent of inflation targeting (e.g., Bernanke, 2002, 2003, 2004). Several other FOMC participants also have expressed a preference for announcing an explicit LIO recently (e.g., Poole, 2006; Lacker, 2005; Stern, 2005; and Yellen, 2006).

DOES THE MARKET BELIEVE THE FOMC HAS A LIO?

Given the belief that the Fed is responsible for the average long-run inflation rate and the FOMC’s commitment to price stability, it is reasonable to assume that market participants have a perception of the FOMC’s implicit LIO. For example, it is doubtful that anyone believes that the FOMC would be content with long-run inflation of 5 percent or with persistent and protracted deflation. Hence, it seems safe to assert that most market analysts believe that the FOMC’s implicit LIO is somewhere between zero and, say, a maximum of 5 percent. Of course, some may believe the implicit target to be narrower. Assuming that market participants expect the FOMC to behave consistently with its implicit LIO, the average of these expectations provides a point estimate of the FOMC’s implicit LIO.

The only market-based measure of inflation expectations is the spread between Treasury inflation-indexed securities (TIIS) and the corresponding non-indexed Treasury issue. TIIS are indexed to inflation as measured by the consumer price index (CPI). Nominal long-term bond yields reflect both the market’s expectation for the real yield and expectations for inflation, whereas the corresponding TIIS reflects only the real yield; accordingly, the TIIS spread—the difference between the nominal yield and the corresponding TIIS yield—is, in principle, a measure of the market’s expectation of inflation over the holding period of the long-term asset. The spread is not a pure measure of inflation expectations because...

---

7 See Federal Reserve Bank of St. Louis (2005) for a detailed discussion and analysis of this period in U.S. monetary policy.
8 See Bernanke (2003) for a discussion of the role of money and monetary aggregates in the inflation process.
9 See Rasche and Thornton (2006) for more details.
11 For example, Yellen (2005) notes that “inflation…is what the Fed can undeniably control in the long run.”
the spread also may reflect risk and liquidity premiums.\footnote{Also, it is impossible to match the maturities of TIIS and nominal government securities exactly. Even if the maturities were an exact match, the two securities have different payment flows. See Sack and Elsasser (2004) and Kwan (2005) for additional details.} That is, the spread is equal to

\[ TIIS_{t}^{sp} = \pi^{e} + rp - lp, \]

where \( TIIS_{t}^{sp} \) denotes the spread between nominal and inflation-indexed Treasury securities, \( \pi^{e} \) denotes expected inflation, \( rp \) denotes the inflation risk premium, and \( lp \) denotes the liquidity premium.

The possibility of a non-zero inflation risk premium arises from the fact that investors in nominal debt are uncertain about the inflation rate that will occur over the holding period of the assets. This is due in part to uncertainty about the FOMC’s LIO. Of course, no one expects the FOMC to achieve its LIO exactly at each point in time. Consequently, the inflation risk premium might exist because of stochastic variation in inflation around a known LIO. Either way, the greater the uncertainty, the more investors will have to be compensated—the larger the risk premium—assuming that investors are risk averse.\footnote{Even if investors were risk neutral, the risk premium would not vanish for technical reasons; see Sack (2000) and Sack and Elsasser (2004) for details.}

The liquidity premium stems from the fact that the market for TIIS is less liquid than the market for nominal Treasury securities. Consequently, TIIS investors likely receive a liquidity premium in the form of a higher return for holding TIIS rather than more-liquid conventional securities.

The risk and liquidity premiums have opposing effects on the spread as a measure of inflation expectations. The existence of a risk premium causes \( TIIS_{t}^{sp} \) to overestimate the market’s expectation for inflation. The existence of a liquidity premium causes \( TIIS_{t}^{sp} \) to underestimate inflation expectations. Consequently, \( TIIS_{t}^{sp} \) only approximates CPI inflation expectations. Nevertheless, if one is willing to assume that the sum of these premiums is relatively stable over time, marked changes in the TIIS spread should reflect changes in market participants’ expectations of inflation.

Inflation-indexed securities were first issued in January 1997. Figure 1 presents the monthly on-the-run \( TIIS_{t}^{sp} \) for 10-year government securities since January 1997.\footnote{On-the-run securities are the most recent issue. The on-the-run spread reported here is obtained by subtracting the yield on the most recently issued TIIS from the constant maturity yield on the most recently issued nominal Treasury security that most closely matches the maturity of the on-the-run TIIS.} \( TIIS_{t}^{sp} \) fluctuated considerably during the early years of the market.
but settled down by early 2001. From January 2001 to June 2003, TIPS\textsuperscript{16} averaged 1.72 percent. In the summer of 2003 the spread began to widen, but by January 2004 it appeared to have settled down again. From January 2004 to March 2006, the spread averaged 2.49 percent, 77 basis points higher than from January 2001 to June 2003. What is responsible for the widening of the spread?

**The 2003 Experience**

At the conclusion of the May 2003 meeting, the FOMC stated in its press release that “the probability of an unwelcome substantial fall in inflation, though minor, exceeds that of a pickup in inflation from its already low level.” This statement was widely analyzed in the press as suggesting the possibility that inflation was at the “lower bound” of the rate acceptable to the Committee. This interpretation was reinforced by the release of the minutes of the May meeting on June 25 and by the FOMC’s June 2003 press release. The minutes indicated that “substantial additional disinflation would be unwelcome because of the likely negative effects on economic activity and the functioning of financial institutions and markets, and the increased difficulty of conducting an effective monetary policy, at least potentially in the event the economy was subjected to adverse shocks.”\textsuperscript{15} Members agreed, however, that “there was only a remote possibility that the process of disinflation would cumulate to the point of a decline for an extended period in the general price level.”\textsuperscript{16} The June press release contained a statement identical to the May statement, but went on to indicate that, “On balance, the Committee believes that the latter concern [an unwelcome fall in inflation] is likely to predominate for the foreseeable future.”

Then-Governor Bernanke detailed his views in a speech entitled “An Unwelcome Fall in Inflation?” on July 23, 2003. Noting that “the May 6 statement was more than a procedural innovation,” Bernanke suggested that it “broke new ground as the first occasion in which the FOMC expressed the concern that inflation might actually fall too low.”\textsuperscript{17} Bernanke’s suggestion that the FOMC had a non-zero lower bound beyond which it did not want inflation to fall was made clear in the FOMC’s August 12 statement, which read, “The Committee judges that, on balance, the risk of inflation becoming undesirably low is likely to be the predominant concern for the foreseeable future” (emphasis added). This statement was repeated in the September 16 and October 28 statements.

That the 2003 experience effectively established a lower bound for the FOMC’s LIO is also suggested by Jeffrey Lacker, president of the Federal Reserve Bank of Richmond, who was the Director of Research at the Richmond Fed in 2003. Lacker (2005, p. 6) notes that

The statement issued following the May 2003 FOMC meeting asserted that a fall in inflation—then about 1 percent—would be “unwelcome.” This came as something of a surprise to markets and caused a sharp reaction in long-term rates. If an inflation target range had been in place in 2003 with a lower bound of 1 percent, the public could have inferred the Fed’s growing concern about disinflation as the inflation rate drifted down toward that bound... If the May 2003 statement is interpreted as the revelation of the lower bound of an inflation target range, then half of an inflation target range has been announced. And if revealing a dislike of inflation below 1 percent was useful in May 2003, is it not likely that revealing an upper bound will prove useful in some future circumstance?

The FOMC’s concern about disinflation began to abate in the fall of 2003. At the conclusion of the December meeting the FOMC indicated that “the probability of an unwelcome fall in inflation has diminished in recent months and now appears almost equal to that of a rise in inflation.” Refer-

\textsuperscript{15} Minutes of the Federal Open Market Committee, May 2003. The last concern is reference to the so-called zero bound problem. The zero bound problem arises because zero is the theoretical lower bound to nominal interest rates. Because the policy instrument is the nominal federal funds rate, the zero bound on nominal interest rates is thought by some to set a lower bound to the FOMC’s ability to conduct expansionary monetary policy. For a more detailed discussion of the zero bound problem, see Bernanke (2002).

\textsuperscript{16} Minutes of the Federal Open Market Committee, May 2003.

\textsuperscript{17} Bernanke (2003).
ence to an unwelcome fall in inflation was not mentioned in subsequent statements.

The Effect of the 2003 Experience on Market Participants’ Beliefs about the FOMC’s LIO

Market participants formulate their beliefs about the FOMC’s implicit LIO from verbatim FOMC transcripts, minutes of FOMC meetings, press releases made at the conclusion of each FOMC meeting, speeches and other statements made by the Chairman and other FOMC participants, and FOMC policy actions. It is not surprising that there was a sharp change in the TIIS spread following the FOMC’s revelations concerning the lower bound of its implicit inflation objective.

If the observed rise in the 10-year TIIS spread were due to the market having better information about the FOMC’s LIO, we might anticipate that expectations would be revised at a horizon that might reasonably be thought of as the horizon consistent with the FOMC’s LIO. Consequently, $\text{TIIS}_{10}\sp{sp}$ is also calculated for the 5-year horizon. Figure 2 presents $\text{TIIS}_{5}\sp{sp}$ using the 5-year TIIS beginning in October 2004 and the previous 10-year issues that are closest to having a 5-year remaining term for the previous period. Because the first 10-year TIIS was first issued in 1997, the sample period is January 2001 through August 2006. The behavior of the 5-year $\text{TIIS}_{5}\sp{sp}$ after the 2003 experience was similar to that of the 10-year $\text{TIIS}_{10}\sp{sp}$. Specifically, the average 5-year $\text{TIIS}_{5}\sp{sp}$ increased by about 100 basis points, from 1.46 percent for the period January 2002 through June 2003 to 2.48 percent for the period January 2004 through April 2006.

If the marked rise in long-run inflation expectations shown in Figures 1 and 2 is due to the FOMC eliminating some uncertainty about the lower bound of its long-run inflation objective, we might also expect to see a reduction in the variability of the TIIS spread. That this occurred is verified in Figure 3, which shows the intra-month standard deviation of the 10-year (solid line) and 5-year (dashed line) TIIS spreads using daily data. Despite the fact that the average level rose, consistent with long-run inflation expectations being more precisely held, there is a marked drop in the intra-month standard deviation in early 2004 for both the 5-year and 10-year TIIS spreads. For the 10-year $\text{TIIS}_{10}\sp{sp}$, the standard deviation declined from 7 basis points from January 2001 through June 2003 to 5 basis points from January 2004 through August 2006. For the 5-year $\text{TIIS}_{5}\sp{sp}$, the standard deviation dropped by a third, from 9 basis points for the period January 2002 through June 2003 to 6 basis points from January 2004 through August 2006.

Survey Measures of Long-Run Inflation

The Blue Chip and Michigan survey also poll their survey participants for inflation forecasts over longer-run horizons. The Blue Chip survey is biannual, while the Michigan survey is monthly. The Michigan survey asks respondents what they believe will be the average CPI inflation rate over the next 5 to 10 years, whereas the Blue Chip forecasts are for 5-year and 10-year horizons. The monthly average of the mean forecast of Michigan survey for the period 1997-2006 are presented in Figure 4. The qualitative implications from the Blue Chip survey are identical, so only the Michigan survey is presented here. Survey expectations drifted down slightly from early 1997 to early 2001; however, unlike $\text{TIIS}_{10}\sp{sp}$ there is no marked change in the survey forecasts of the long-term inflation rate after mid-2003. Survey forecasts for long-run inflation averaged 3.3 percent for both the period January 2001 through June
Figure 3
Intra-Month Standard Deviation of the 5- and 10-Year TIS Spreads

Figure 4
The Survey of Inflation Over the Next 5 to 10 Years
2003 and January 2004 though July 2006. Hence, the indication from \( TIIS_{t}^{up} \)—that there was a marked change in inflation expectations in the wake of the FOMC being more explicit about its LIO—is not reflected in survey expectations measures.

The lack of response of the survey forecasts after May 2003 is difficult to reconcile with the marked and sustained increase in the \( TIIS \) inflation expectations measure. It could be that there was a marked increase in the liquidity premium in the \( TIIS \) market that just happened to coincide with the FOMC’s statements. The liquidity premium between on-the-run nominal Treasuries and the less liquid on-the-run \( TIIS \) is not directly observable. However, comparisons of yields of on-the-run and off-the-run \( TIIS \) indicate that the liquidity premium for on-the-run \( TIIS \) is small and does not decline markedly in mid-2003. Estimates of the liquidity premium in the nominal Treasury market by Gurkaynak, Sack, and Wright (2006) also suggest that the liquidity premium is rather small, about 10 basis points. Hence, it is unlikely that the approximately 80-basis-point widening of the spread could be attributed solely or largely to a decline in the liquidity premium.

Alternatively, statements made at the May and subsequent meetings could have caused the inflation risk premium to rise. For example, some market participants who thought that the FOMC’s inflation target range was say 0 to 2 percent may have become more uncertain about the FOMC’s LIO. More generally, indications that the FOMC had a non-zero lower bound to acceptable inflation could have shaken some participants’ belief about the FOMC’s commitment to price stability. That the entire 80-basis-point increase can be attributed to a larger inflation risk premium seems unlikely, however.

Despite the fact that there is no corresponding rise in survey measures of inflation expectations, the fact that \( TIIS_{t}^{up} \) widened following the FOMC’s statements and leveled off when the FOMC indicated that its concerns about disinflation had waned strongly suggests that the sharp rise in \( TIIS_{t}^{up} \) reflects the market’s reevaluation of the FOMC’s implicit LIO. One possibility is that the FOMC’s statements effectively truncated the lower end of the probability distribution of inflation expectations. That is, consistent with Lacker’s (2005) interpretation, individuals who may have thought the Fed’s inflation objective was less than, say, about 1 percent revised their expectation upward. This explanation alone would not seem to account for the 80-basis-point rise in the spread. To see why, assume that prior to the summer of 2003 all market participants thought that the FOMC’s LIO was 0 to 3 percent, with a mean of 1.5 percent (which is close to the average spread from January 2001 through June 2003). If the statements caused all market participants simply to truncate their estimate of the lower bound of the FOMC’s LIO at 1 percent, this would cause the average spread to increase by only 50 basis points, from 1.5 percent to 2 percent. It is, however, difficult to assess the impact of truncation on the average level of inflation expectations because the FOMC was vague about the lower bound for inflation. Some market participants may have thought that the effective lower bound was now higher than the 1 percent rate suggested by Lacker. Nevertheless, if the nearly 80-basis-point rise in the spread is entirely due to a rise in inflation expectations among market participants, it would seem that the FOMC’s statements must have caused some market participants to raise their estimate of the upper bound of the implicit LIO as well. For example, participants who previously thought the FOMC’s target range was 0 to 2 percent could have raised the estimated range to 1 to 3 percent.

**The Upper Bound of the FOMC’s LIO**

Statements made in 2003 suggest that the FOMC has a lower bound of acceptable long-run inflation in the neighborhood of 1 percent. Given its commitment to price stability, there can be no doubt that the FOMC has an upper bound as well. However, as Bernanke (2004) has noted, the publicly expressed preferences for the LIO by various members of the FOMC range “from less than 1 percent to 2.5 percent or more.”

---

rise in inflation, the FOMC has provided more information about the upper bound of its implicit LIO. The minutes of the August 9, 2005, FOMC meeting indicate that

While recent monthly readings indicated that core inflation had been subdued, a number of participants noted that underlying core inflation appeared to be running at a pace around the upper end of the range they viewed as consistent with price stability—an assessment that was reinforced by the recent upward revisions to historical data on core PCE inflation. Participants commented that an increase in inflation from recent rates could have especially adverse effects on longer-run economic performance. (emphasis added)

A similar statement appeared in the minutes of the March 27-28, 2006, FOMC meeting, where

Some participants held that core inflation and inflation expectations were already toward the upper end of the range that they viewed as consistent with price stability, making them particularly vigilant about upside risks to inflation, especially given how costly it might be to bring inflation expectations back down if they were to rise. (emphasis added)

A similar statement appeared in the minutes of the May 10, 2006, meeting.

The minutes of June 28-29, 2006, meeting are more specific in that they indicate that inflation may have already reached the Committee’s upper bound. The minutes note that

All participants found the elevated readings on core inflation of recent months to be of concern and, if sustained, inconsistent with the maintenance of price stability. (emphasis added)

During the three months prior to this meeting, annual core CPI inflation averaged about 3.75 percent, while core PCE inflation averaged about 3.1 percent. The lack of specificity about the core measure to which the FOMC was referring creates uncertainty about the upper bound of the FOMC’s inflation objective. Nevertheless, if these numbers reflect “core inflation of recent months,” the upper bound of the FOMC’s LIO is somewhere in the neighborhood of 3.1 to 3.75 percent.

If one assumes that Lacker’s suggestion that the 2003 experience established the lower bound of the FOMC’s implicit LIO at 1 percent, one can conjecture that the FOMC’s implied LIO is, say, 1 to 3.5 percent. The midpoint of this range, 2.25 percent, is very close to Chairman Bernanke’s (2004, p. 166) suggestion that “something in the vicinity of 2 percent is the optimal long-run average inflation rate for a variety of assumptions about the costs of inflation, the structure of the economy, the distribution of shocks, etc.” Bernanke noted, however, that many details would have to be decided before such a number could be embraced by the FOMC and suggested that additional research would be worthwhile before the FOMC could decide on the optimal long-run inflation rate. The estimate of 2.25 percent is below the TIIS spread of 2.5 percent, but the difference can easily be accounted for by a non-zero inflation risk premium.

CONCLUSIONS

Because the Fed can control the average long-run inflation rate and because of the FOMC’s long-standing commitment to price stability, it is reasonable to assume that the FOMC has an implicit LIO. The 2003 experience clarified the lower bound of the FOMC’s LIO. The minutes of June 2006 have clarified the FOMC’s upper bound. That this recent guidance does not appear to have affected the TIIS spread significantly suggests that the information merely reinforced the market’s belief in the upper bound of the FOMC’s LIO. The FOMC could alleviate the remaining uncertainty by following Bernanke’s (2002) suggestion and formally announcing a LIO. Until it does, the market will have to rely on FOMC statements, actions, and other information to pin it down more tightly.

REFERENCES


Bernanke, Ben S. “An Unwelcome Fall in Inflation?” Remarks before the Economics Roundtable, University of California at San Diego, La Jolla, California, July 23, 2003.


Lacker, Jeffery M. “Inflation Targeting and the Conduct of Monetary Policy.” Speech at the University of Richmond, Richmond, Virginia, March 1, 2005.


Granger Causality and Equilibrium Business Cycle Theory

Yi Wen

Postwar U.S. data show that consumption growth “Granger-causes” output and investment growth, which is puzzling if technology is the driving force of the business cycle. The author asks whether general equilibrium models with information frictions and non-technology shocks can rationalize the observed causal relationships. His conclusion is they cannot. (JEL E13, E32)


There is a “causal” relationship among consumption, output, and investment. Postwar U.S. data show that consumption growth “Granger-causes” gross domestic product (GDP) growth but not vice versa and that GDP growth in turn Granger-causes business investment growth but not vice versa.1 This unidirectional causal chain suggests that consumption contains better information about the source of shocks hitting the economy than does output, and output in turn contains better information about such shocks than does investment.

This causal relationship cannot be explained by standard real business cycle (RBC) models. For example, under technology shocks, output contains the best information possible for the source of shocks; hence, it will not appear to be “Granger-caused” by consumption once the history of output is taken into account. To rationalize the causal relationship found in the U.S. data, it seems natural to consider demand shocks and to add a richer information structure into standard models such that demand shocks can affect consumption before affecting output and investment.2

I investigate whether existing equilibrium business cycle models driven by demand shocks (in particular, government spending shocks) can rationalize the observed causal relationship when the following information structure is embedded: (i) Employment and output cannot respond to demand shocks immediately; they can do so only with a lag behind consumption. And (ii) investment cannot respond to demand shocks immediately; it can do so only with a lag behind output.3

Under these ad hoc assumptions, I first show that standard general equilibrium business cycle models do predict the existence of a causal chain from consumption to output and investment, but with the wrong sign. Namely, consumption growth negatively causes output growth, and output

---

1 The concept of causality is defined according to Granger (1969).

2 Hall (1978) first points out that consumption appears to be exogenous with respect to output and investment. More recently, Cochrane (1994) argues that the predictive power of consumption on output indicates that consumption shocks are important for the business cycle.

3 The rationale could be time-to-build or adjustment costs in employment and investment. Notice that preference shocks do not resolve the problem because under preference shocks consumption is not able to respond at the impact period when neither output nor investment can respond to the shocks. Consequently, consumption would contain the same information as output and would fail to Granger-cause output.
growth in turn negatively causes investment growth. In the U.S. data, however, these causal relationships are strictly positive. The negative causal chain emerges from standard models because of the crowding-out effect among components of aggregate demand in general equilibrium.

I then choose to mitigate the crowding-out problem by allowing for variable capacity utilization and production externalities in standard models, following Baxter and King (1990), Benhabib and Farmer (1994), Burnside and Eichenbaum (1996), Wen (1998), and Benhabib and Wen (2004). Variable capacity utilization and mild production externalities mitigate the crowding-out problem by creating short-run increasing returns to labor, which permit the expansion of output to meet aggregate demand with little increase in marginal costs in the short run. These modifications, however, bring about only limited success. The model now predicts that output growth positively causes investment growth, but it fails to predict that consumption growth positively causes output growth. The source of failure is still the crowding-out effect: Demand shocks crowd out consumption at the impact period during which neither output nor investment is able to respond.

There seem to be no simple remedies for the problems identified. More fundamental modifications to existing models are required to fully explain the causal aspects of the business cycle in general equilibrium. One possible remedy is to allow for inventory accumulation, so as to further mitigate the crowding-out effect on consumption. Because general equilibrium business cycle models with inventories are still at an early stage of development, this channel is left as a future research topic.

THE CAUSAL RELATIONSHIPS

To document the causal relationships among aggregate consumption, output, and investment,

\[ \Delta y_t = f(\Delta y_{t-1}, \Delta y_{t-2}), \]

\[ \Delta y_t = f(\Delta y_{t-1}, \Delta y_{t-2}, \Delta i_{t-1}), \]

\[ \Delta y_t = f(\Delta y_{t-1}, \Delta y_{t-2}, \Delta c_{t-1}), \]

where \( \Delta y \) is growth in real GDP, \( \Delta i \) is growth in business fixed investment, and \( \Delta c \) is growth in real consumption of nondurable goods and services. A variable \( x \) is said to Granger-cause a variable \( y \) when a prediction of \( y \) on the basis of its history can be improved by further taking into account the previous period’s \( x \). Estimating (1), (2), and (3) gives the following results (\( t \)-values are in parentheses, the 5 percent significance level [denoted throughout by an asterisk] is \( \pm 1.96 \)):

\[ \Delta y_t = 0.005 + 0.29\Delta y_{t-1} + 0.12\Delta y_{t-2}, \]

\[ (0.18) \quad (4.36) \quad (1.80) \]

\[ \Delta y_t = 0.005 + 0.37\Delta y_{t-1} + 0.16\Delta y_{t-2} - 0.05\Delta i_{t-1}, \]

\[ (5.20) \quad (4.59) \quad (2.27) \quad (-1.79) \]

\[ \Delta y_t = 0.003 + 0.14\Delta y_{t-1} + 0.12\Delta y_{t-2} + 0.39\Delta c_{t-1}, \]

\[ (2.84) \quad (1.83) \quad (1.81) \quad (3.13) \]

These results lead to the following conclusions: First, based on regressions (4) and (5), I cannot reject the null hypothesis that investment growth in the preceding period has no explanatory power with respect to output growth in the current period, given the history of output growth.

\[ \Delta y_t = f(\Delta y_{t-1}, \Delta y_{t-2}, \Delta c_{t-1}), \]

where \( \Delta y \) is growth in real GDP, \( \Delta i \) is growth in business fixed investment, and \( \Delta c \) is growth in real consumption of nondurable goods and services. A variable \( x \) is said to Granger-cause a variable \( y \) when a prediction of \( y \) on the basis of its history can be improved by further taking into account the previous period’s \( x \). Estimating (1), (2), and (3) gives the following results (\( t \)-values are in parentheses, the 5 percent significance level [denoted throughout by an asterisk] is \( \pm 1.96 \)):

\[ \Delta y_t = 0.005 + 0.29\Delta y_{t-1} + 0.12\Delta y_{t-2}, \]

\[ (0.18) \quad (4.36) \quad (1.80) \]

\[ \Delta y_t = 0.005 + 0.37\Delta y_{t-1} + 0.16\Delta y_{t-2} - 0.05\Delta i_{t-1}, \]

\[ (5.20) \quad (4.59) \quad (2.27) \quad (-1.79) \]

\[ \Delta y_t = 0.003 + 0.14\Delta y_{t-1} + 0.12\Delta y_{t-2} + 0.39\Delta c_{t-1}, \]

\[ (2.84) \quad (1.83) \quad (1.81) \quad (3.13) \]

These results lead to the following conclusions: First, based on regressions (4) and (5), I cannot reject the null hypothesis that investment growth in the preceding period has no explanatory power with respect to output growth in the current period, given the history of output growth.
Second, regressions (4) and (6) suggest that past growth in consumption has a significant effect on current output growth even after the history of output growth is taken into account. In fact, consumption growth is such an important factor for determining future output growth that none of the dependent variables in regression (4) remain significant after past consumption growth is taken into account in regression (6). This result suggests consumption growth explains the bulk of future output growth.

For the reverse questions, whether past output growth has an effect on current investment growth given the history of investment growth and whether it also has an effect on current consumption growth given the history of consumption growth, I obtain the following results:

Regression (11) suggests that investment growth in the preceding period has no explanatory power for consumption growth in the current period, given the history of consumption growth. This establishes the one-way causal chain. Regression (12) simply confirms that the causal relationships are transitive; namely, if past consumption growth causes current output growth and past output growth causes current investment growth, then past consumption growth must also be significant in predicting current investment growth.

Regressions (7) and (8) suggest that past output growth has a significant effect on current investment growth. On the other hand, regressions (9) and (10) suggest that consumption growth in the preceding period is the best predictor of consumption growth in the current period. Taking into account past output growth does not improve the prediction statistically and economically. This is consistent with Hall’s (1978) empirical analysis that the history of consumption is the best predictor of future consumption, except that consumption does not follow a pure random walk.

These results suggest the existence of a one-way “causal” linkage among consumption, output, and investment growth. Namely, consumption growth in the preceding period Granger-causes output growth in the current period; and output growth in the current period in turn Granger-causes investment growth in the next period. To conclude that the causal chains are truly unidirectional, however, I must run two more regressions to eliminate the possibility of feedback from investment growth to consumption growth. I obtain the following results:

\[ \Delta c_t = 0.006 + 0.20 \Delta c_{t-1} + 0.08 \Delta c_{t-2} + 0.01 \Delta i_{t-1}, \]
\[ (7.84)^* \quad (2.89)^* \quad (1.16) \quad (0.79) \]

\[ \Delta i_t = -0.0015 + 0.33 \Delta i_{t-1} + 0.03 \Delta i_{t-2} + 1.05 \Delta c_{t-1}. \]
\[ (-0.49) \quad (4.95)^* \quad (0.51) \quad (3.11)^* \]

Regression (11) suggests that investment growth in the preceding period has no explanatory power for consumption growth in the current period, given the history of consumption growth. This establishes the one-way causal chain. Regression (12) simply confirms that the causal relationships are transitive; namely, if past consumption growth causes current output growth and past output growth causes current investment growth, then past consumption growth must also be significant in predicting current investment growth.

**ROBUSTNESS**

The standard Granger causality test gets into trouble when a time series has a moving average component that is not invertible. In that case, finite history of that time series can never be sufficient for predicting its current behavior, rendering other variables significant in improving the prediction. For example, let

\[ x_t = \varepsilon_t - \varepsilon_{t-1}, \]
\[ z_t = 0.9 z_{t-1} + \varepsilon_t, \]

where \( \varepsilon_t \) is an i.i.d. white noise innovation. If one defines the current information set as \( \Omega_t = [\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] \), then the prediction, \( P[x_t | \Omega_{t-1}] \), cannot be improved by further taking into account the history of \( z_t, [z_{t-1}, z_{t-2}, \ldots] \).

Strictly speaking, therefore, these two series, \( x_t \),
and \( z_t \), do not cause or predict one another. History of \( z_t \), however, can appear to be significant in predicting the current movement of \( x_t \) in the linear regression:

\[
    x_t = \alpha + \sum_{j=1}^{k} \gamma_j x_{t-j} + \beta z_{t-1}, \quad 0 < k < \infty.
\]

This is so because \( z_{t-1} \) contains the entire history of innovations \( \{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\} \) useful for predicting \( \{x_{t-k-1}, x_{t-k-2}, \ldots\} \), which are useful for predicting \( x_t \) when only the finite history, \( \{x_{t-1}, \ldots, x_{t-k}\} \), is included in the information set of the regression.

As a demonstration, a Monte Carlo experiment of the above series gives the following estimation results:

\[
    (13) \quad \begin{align*}
        x_t &= 0.0003 - 0.79x_{t-1} - 0.59x_{t-2} - 0.40x_{t-3} - 0.19x_{t-4} \\
            &\quad (0.03) \quad (-80.5) \quad (-49.0) \quad (-33.3) \quad (-19.5) \\
        x_t &= -0.0008 - 0.76x_{t-1} - 0.54x_{t-2} - 0.36x_{t-3} - 0.17x_{t-4} - 0.16z_{t-1} \\
            &\quad (-0.07) \quad (-82.0) \quad (-48.0) \quad (-31.5) \quad (-17.8) \quad (-34.8)
    \end{align*}
\]

Although \( \text{cor}(x_t, x_{t-j}) = 0 \) for \( j \geq 2 \), the first regression in (13) shows \( x_{t-j} \) are highly significant in predicting \( x_t \) even for \( j > 2 \). This happens because \( x_t \) does not have a finite autoregressive representation when its moving average component is not invertible. Failing to take into account the non-invertible moving average component can render other variables such as \( z_{t-1} \) significant in predicting \( x_t \), although the variable \( z_t \) contains no better information than what is in \( x_t \) regarding \( \varepsilon_t \). The second regression in (13) confirms that \( z_{t-1} \) is highly significant in predicting \( x_t \). Even though the history of \( x_t \) predicts \( x_t \) reasonably well (\( R^2 = 0.39 \)), past \( z_t \) improve the prediction (\( R^2 = 0.46 \)).

A sensible solution for this pitfall is to use a two-stage regression: Fit an optimal ARMA\((p,q)\) model to a stationary time series, and then regress the estimated residual from the ARMA\((p,q)\) model against the history of other variables that are of interest. If these other variables appear to be significant in predicting movements in the estimated residual series, then there is said to exist Granger causality between these other variables and the first time series.

If we apply this idea to the above example, regressions based on Monte Carlo simulations give the following results (\( t \)-values are in parentheses):

\[
    (14) \quad \begin{align*}
        x_t &= \varepsilon_t - 0.999\varepsilon_{t-1} + u_t; \\
             &\quad \text{(-2146.9)} \\
        u_t &= -0.02 - 0.006z_{t-1}; \\
             &\quad \text{(0.03) (0.48)}
    \end{align*}
\]

As expected, the results show that past \( z_t \) are not significant in predicting current \( x_t \) after the moving average component of \( x_t \) is taken into account.

This point is relevant to my analyses of the U.S. data because the first differences of output, consumption, and investment could contain moving average components that are not invertible when the log levels of these variables are not exactly random walk series. In such cases, consumption growth in the previous period can appear to be significant in predicting output growth in the current period even when in fact it does not contain any information superior to that in output. This point is also relevant to my theoretical analysis in what follows because the growth rates of output and other variables in the models have a moving average component.

With this extended notion of Granger causality in mind, I reexamine the identified causal relationship by estimating an ARMA\((4,1)\) model for the growth rate of each of the three macro variables. I find that the moving average coefficients for all three variables are highly significant and are all close to 1 in absolute value. I then use the estimated residuals obtained from each ARMA estimation in a second-stage regression with respect to a constant and the lagged growth rate of another variable. For the case of output growth, I obtain results, as shown in Table 1, in the second-stage estimation.

The second-stage regression shows that the estimated residual of output growth obtained from the ARMA model is not exogenous with respect to consumption growth in the preceding period. Namely, consumption growth in the preceding...
period helps predict current output growth even after the history of output growth and the moving average bias are taken into account. This is consistent with the earlier results obtained above: Consumption growth causes output growth.

Similarly, I cannot reject the null hypothesis that investment growth in the preceding period has no explanatory power with respect to output growth in the current period (see the bottom row in Table 1). This is also consistent with the results obtained earlier: Investment growth does not cause output growth.

With respect to consumption growth, I obtain the results shown in Table 2 in the second-stage estimation. It is also consistent with earlier results that neither output growth nor investment growth in the preceding period has explanatory power for consumption growth. The second-stage regression of investment growth gives the results in Table 3.

The table shows that investment growth in the current period is predictable by consumption growth in the preceding period. This is also consistent with the earlier result. Output growth in the preceding period, however, lost its significance in predicting current investment growth at the 5 percent significance level. It is, however, still significant at the 10 percent significance level. In addition, judged by the economic significance, past output growth still helps predict current investment growth very well. The coefficient of ∆yt–1 in the regression is 0.30 with a standard error of 0.18.

In sum, taking into account the potential bias caused by non-invertible moving average components in the growth rates does not change the conclusions I drew earlier: Postwar U.S. aggregate data exhibit a “causal” chain among consumption, output, and investment. That “causality” runs in only one direction: from consumption growth to output growth and from output growth to capital formation. Within this causal chain, the impact of consumption growth on both output and investment growth appears to be very powerful and highly robust.

These results reinforce the empirical findings by Hall (1978) and Cochrane (1994). They suggest that there exist certain types of shocks in the U.S. economy that affect consumption before having any impact on output and investment. These shocks cannot be total factor productivity shocks, as output would react immediately to productivity shocks and it is unlikely that consumers are better informed of these shocks than firms. For this reason, in what follows, I try to rationalize the documented empirical regularity by introducing information frictions and demand shocks. However, to be convincing, I first present results obtained in standard models under technology shocks.

### Table 1

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient</th>
<th>t-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆ct–1</td>
<td>0.58</td>
<td>3.83*</td>
</tr>
<tr>
<td>∆it–1</td>
<td>0.03</td>
<td>1.08</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient</th>
<th>t-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆ct–1</td>
<td>0.001</td>
<td>0.04</td>
</tr>
<tr>
<td>∆it–1</td>
<td>0.02</td>
<td>1.55</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient</th>
<th>t-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆ct–1</td>
<td>0.95</td>
<td>2.59*</td>
</tr>
<tr>
<td>∆yt–1</td>
<td>0.30</td>
<td>1.61</td>
</tr>
</tbody>
</table>

**Predictions of a Standard RBC Model**

A simple RBC model can be cast in the following form in which a representative agent solves
subject to

\[ k_{t+1} = \rho k_t + \varepsilon_t \]

subject to

\[ c_{t+1} + k_{t+1} - (1 - \delta)k_{t+1} \leq A_{t+1} k_{t+1}^{\alpha(1-\alpha)} \]

and \( \lambda > 0 \). In equilibrium, consumption, output, and investment in a RBC model should follow these decision rules near the steady state:

\[
\begin{align*}
    c_t &= \pi_{ck} k_t + \pi_{ca} A_t, \\
    y_t &= \pi_{yk} k_t + \pi_{ya} A_t, \\
    i_t &= \pi_{ik} k_t + \pi_{ia} A_t, \\
    k_{t+1} &= \pi_{kk} k_t + \pi_{ka} A_t,
\end{align*}
\]

where \( k \) is the capital stock and \( A \) is technology. Utilizing the law of motion for the capital stock, the above equilibrium decision rules can be further expressed as

\[
\begin{align*}
    c_t &= \pi_{ck} c_{t-1} + \pi_{ca} A_t + (\pi_{ck} \pi_{ka} - \pi_{ca} \pi_{kk}) A_{t-1}, \\
    y_t &= \pi_{yk} y_{t-1} + \pi_{ya} A_t + (\pi_{yk} \pi_{ka} - \pi_{ya} \pi_{kk}) A_{t-1}, \\
    i_t &= \pi_{ik} i_{t-1} + \pi_{ia} A_t + (\pi_{ik} \pi_{ka} - \pi_{ia} \pi_{kk}) A_{t-1}.
\end{align*}
\]

Clearly, these equilibrium laws of motion imply that consumption, output, and investment all contain the same information about the history of technology shocks, hence neither variable should appear to Granger-cause another variable. For example, once the history of output is taken into account, consumption in the previous period should have no additional explanatory power on current output.

To confirm this, I simulate the model to obtain artificial data series for consumption, output, and investment (sample size = 10,000), under the assumption that technology shocks follow the process \( A_t = \rho A_{t-1} + \varepsilon_t \). Because the equilibrium law of motion for \( \{c, y, i\} \) suggests that the growth rates of all variables follow an ARMA(1,1) process if \( \rho = 1 \) and an ARMA(2,1) process if \( \rho = 0.9 \), they all contain moving average components. I therefore apply the two-stage regression procedure discussed in the previous section to estimate causal relationships among the growth rates of the three variables. In the first stage, I apply an ARMA(2,1) model to obtain the residual series. The estimated residuals from the ARMA(2,1) model are then used in the second-stage regression to determine the presence of Granger causality. The results obtained are noted in Table 4.7

The table shows that none of the variables Granger-cause each other once its own history is taken into account. The coefficients are all statistically insignificantly different from zero. This is expected because all variables in the model share the same information about technology shocks. Hence, adding other variables to the regressions does not improve the prediction.

### Table 4

<table>
<thead>
<tr>
<th>Equation for Δyt</th>
<th>Equation for Δit</th>
<th>Equation for Δct</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta c_{t-1} ) 0.04 (0.73)</td>
<td>( \Delta c_{t-1} ) 0.24 (1.13)</td>
<td>( \Delta c_{t-1} ) 0.00 (0.09)</td>
</tr>
<tr>
<td>( \Delta i_{t-1} ) 0.00 (0.03)</td>
<td>( \Delta y_{t-1} ) 0.01 (0.30)</td>
<td>( \Delta i_{t-1} ) 0.00 (0.10)</td>
</tr>
</tbody>
</table>

**Predictions with Sequential Information Structure**

This section embeds a sequential information structure into the benchmark model so as to create information differentials among output, consumption, and investment. To create an information differential between output and consumption so that consumption contains better information

---

7 The time period is one quarter and the model’s parameters are calibrated as follows: the time discounting factor \( \beta = 0.99 \), the capital’s share \( \alpha = 0.3 \), the rate of capital depreciation \( \delta = 0.025 \), the inverse labor supply elasticity \( \gamma = 0 \) (Hansen’s [1985] indivisible labor), and the persistence parameter \( \rho = 0.9 \). The results are not sensitive to these parameter values.
than output, we need shocks that can affect consumption without affecting output in the initial period. This suggests we need to consider demand shocks instead of technology shocks.\(^8\) Hence I make the following assumptions: (i) The source of the business cycle is from aggregate demand, and demand shocks can affect consumption instantaneously. (ii) Decisions about employment must be made one period in advance; this implies that output cannot respond to demand shocks immediately, but only with a lag behind consumption. (iii) Investment decisions must be made two periods in advance; this implies that firms’ investment cannot respond to demand shocks immediately, but only with a lag behind output.

Under these assumptions, the representative agent’s problem is to solve

\[
\begin{align*}
\max_{k_{t-1}} E_{t-2} \left[ \max_{\ell_{t-1}} E_{t-1} \left[ \max_{\ell_{t-1}; f_{t-1}} \left( \sum_{j=0}^{\infty} \beta^j \log(c_{t+j}) - \alpha n_{t+j}^{1+\gamma} \right) \right] \right],
\end{align*}
\]

subject to

\[
c_{t+j} + g_{t+j} + k_{t+j} - (1-\delta)k_{t+j} \leq A_{t+j} k_{t+j}^\alpha n_{t+j}^{1-\alpha},
\]

and \(k_0 > 0\). I also assume that government spending follows an AR(1) stochastic process in logs, \(\log g_t = 0.9 \log g_{t-1} + \epsilon_t\), where the innovation \(\epsilon_t\) is i.i.d. white noise.

The first-order conditions with respect to choices in time periods \(t \geq 0\) are given by

\[
\begin{align*}
\frac{1}{c_t} - \lambda_t &= 0, \\
E_{t-1} \left\{ an_t^\gamma - (1-\alpha) \lambda_t A_t k_t^\alpha n_t^{1-\alpha} \right\} &= 0, \\
E_{t-2} \left\{ \lambda_t - \beta \lambda_{t+1} \left[ A_t k_{t+1}^\alpha n_{t+1}^{1-\alpha} + 1 - \delta \right] \right\} &= 0, \\
c_t + g_t + k_{t+1} - (1-\delta)k_{t+1} &= A_t k_{t+1}^\alpha n_{t+1}^{1-\alpha},
\end{align*}
\]

where the first equation equates the marginal utility of consumption to its shadow price; the second equation equates the expected marginal cost and benefit of hours based on time \(t-1\) information; the third equation equates the expected cost and benefit of investment based on time \(t-2\) information; and the last equation is the period-by-period resource constraint.

In equilibrium, consumption, output, and investment in the model should follow the following rules\(^9\):

\[
\begin{align*}
c_t &= c(k_t, A_t, A_{t-1}, A_{t-2}, g_t, g_{t-1}, g_{t-2}) \\
y_t &= y(k_t, A_t, A_{t-1}, A_{t-2}, g_t, g_{t-1}, g_{t-2}) \\
i_t &= i(k_t, A_{t-2}, g_{t-2}).
\end{align*}
\]

These equilibrium policy rules imply that consumption in the preceding period \((c_{t-1})\) helps predict output in the current period \((y_t)\) even after the history of past output, \(\{y_{t-1}, y_{t-2}, \ldots\}\), is taken into account. This is so because \(c_{t-1}\) has information about the demand shock \(g_{t-1}\) that is useful for predicting \(y_t\) but is missing in the history of \(y_t\).

They also imply that output in the preceding period \((y_{t-1})\) helps predict investment in the current period \((i_t)\) even after the history of investment, \(\{i_{t-1}, i_{t-2}, \ldots\}\), has been taken into account, because \(y_{t-1}\) has information about the demand shock \(g_{t-2}\) that is useful for predicting \(i_t\) but is missing in the history of \(i_t\). Notice that this information structure cannot be obtained under technology shocks because by definition technology shocks affect output directly before they can affect consumption. Therefore, in the following simulations only government shocks are used.

Using similar estimation procedures above, I obtain the results shown in Table 5 from the second-stage regressions.\(^{10}\)

The first column of the table shows that consumption growth in the preceding period has significant explanatory power for the residual of output growth in the current period. The middle column of the table shows that output growth (as well as consumption growth) in the preceding period has significant explanatory power for the residual of investment growth in the current period. The last column of the table shows that neither output growth nor investment growth in the preceding period has significant effects on

---

\(^8\) Technology shocks directly impact output. Hence, output always contains the best information possible under technology shocks, and, consequently, it is impossible for consumption to Granger-cause output under technology shocks.

\(^9\) See Wen (2005b) for details.

\(^{10}\) The steady-state government spending-to-output ratio is set at \(g/y = 0.15\).
the residual of consumption growth in the current period. (Although the coefficient on $\Delta i_{t-1}$ is statistically significant, it is economically insignificant.)

Hence, introducing the sequential information structure and demand shocks brings the standard RBC model into closer conformity with the data’s causal structure. However, the model fails on two grounds: (i) The causal relationships among consumption, output, and investment are of the wrong sign—they are all negative in the model. (ii) The order of the relative volatilities of consumption, output, and investment is exactly opposite to that of the data: In the model, consumption is more volatile than output, which in turn is more volatile than investment. Both failures are due to the well-known crowding-out effect of government shocks, which renders consumption and output to be negatively correlated and prevents consumption from smoothing when government expenditure fluctuates. This crowding-out problem cannot be resolved by introducing different forms of demand shocks, such as preference shocks.

**PREDICTIONS OF A SCALE-ECONOMY MODEL**

Because allowing for demand shocks in the standard models creates the well-known problem of negative comovement among components of aggregate demand, I introduce further modifications into the model to mitigate the crowding-out problem, following the ideas of Wen (1998) and Benhabib and Wen (2004). In the model, output is produced according to the technology

$$y_t = \left( e_t k_t \right)^{\alpha(1 + \eta)} n_t^{-(1-\alpha)(1+\eta)},$$

where $e$ is the rate of capital utilization and $\eta > 0$ measures the degree of externalities taken as parametric by representative agents. The rate of capital depreciation is linked to the rate of capital utilization in the preceding period according to

$$\delta_t = \frac{1}{\delta} e_t^\theta (\theta > 1),$$

implying that capital depreciates faster when used more intensively. Thus, the law of motion for capital accumulation is given by

$$k_{t+1} = i_t + \left(1 - \frac{1}{\delta} e_t^\theta \right) k_t.$$

Under these assumptions, the representative agent’s problem is to solve

$$\max_{\{k_{t+1}\}} E_{t+2} \left[ \max_{\{c_{t+1}, \alpha_{t+1}\}} E_{t+1} \left[ \max_{\{c_t\}} E_t \left[ \sum_{j=0}^\infty \beta^j \log(c_{t+j}) - \frac{\alpha_{t+j}^1}{1 + \gamma} \right] \right] \right],$$

subject to

$$c_{t+j} + g_{t+j} + k_{t+j} - (1 - \delta_{t+j})k_{t+j} \leq (e_{t+j}k_{t+j})^{\alpha(1+\eta)} n_{t+j}^{(1-\alpha)(1+\eta)},$$

$$\delta_{t+j} = \frac{1}{\delta} e_{t+j}^\theta, \quad \theta > 1;$$

and $k_0 > 0, 1 > e_{-1} > 0.11$

Variable capacity utilization and mild externalities can mitigate the crowding-out effect, as shown by Benhabib and Wen (2004). The scale economy model therefore improves the previous models substantially in explaining the observed Granger causalities. Applying the two-stage estimation procedures to the model gives the results shown in Table 6.12

11 See Wen (2003b) for details of how to solve this model.

12 The steady-state rate of capital depreciation is set at $\delta = 0.025$, which implies the utilization elasticity of depreciation $\theta = 1.4$. Following Wen (1998) and Benhabib and Wen (2004), I chose the externality parameter $\eta = 0.15$. The results are not sensitive to these parameter values.
The scale economy model improves the performance of the previous models along several dimensions. First, the middle column of Table 6 shows that both consumption growth and output growth in the scale economy model positively cause investment growth. Secondly, the first column of Table 6 shows that the negative causal relationship found between consumption growth and output growth in the previous models is no longer economically significant in the scale economy model, although it is still non-positive. Another significant improvement of the current model is that the relative volatilities among consumption, output, and investment are restored to the right order; namely, consumption is now the least volatile and investment the most volatile in the scale economy model. This smoothing effect is explained by Wen (1998). Capacity utilization and production externalities help smooth consumption because they render the real wage relatively smooth compared with employment.

What has prevented the model from generating a positive causal relationship between consumption growth and output growth? Figure 1 shows the impulse responses of the model to a positive government shock. On impact, consumption decreases significantly, whereas the other variables remain intact. Although consumption starts to rise above the steady state once employment and output are able to adjust in the following periods, the magnitude is relatively small because of consumption smoothing. Thus, the growth rate of consumption is significantly negatively autocorrelated in artificial time series of consumption generated by government shocks, whereas the growth rates of the other variables are all positively autocorrelated. This negative autocorrelation of consumption growth is caused by the crowding-out effect of government shocks at the impact period, during which output and investment are both fixed. As long as output is not allowed to respond to shocks at the impact period, such a crowding-out effect is unavoidable.

**Remarks**

It is important to reiterate that adding technology shocks into the model does not help resolve the problem, because the causal relationships found in the data are conditional predictions. What matters is the information differential between consumption and output. Technology (or any other) shocks will have no effect on the causal chain unless they can change the information differential. Even though technology shocks can create positive correlation between consumption and output, this has nothing to do with conditional predictions. Just as one time series leading another does not imply it also Granger-causes the other time series, a positive correlation between consumption and output does not imply consumption Granger-causes output. This is why simulations with mixed shocks are not considered. Also, adding other forms of demand shocks (such as taste shocks and sunspots shocks) does not change the fundamental picture because, given our framework, these shocks all have a crowding-out effect either on consumption or on investment at the impact period when output in the resource constraint is not able to react to shocks. In fact, if investment is also fixed in the initial period, then consumption is not going to be responsive at all to preference (or sunspot) shocks on impact; consequently, consumption will fail to Granger-cause output. In this case, the model is able to generate positive correlations between consumption growth and output growth.

Table 6

Scale Economy Model with Sequential Information

<table>
<thead>
<tr>
<th>Equation for $\Delta y_t$</th>
<th>Equation for $\Delta i_t$</th>
<th>Equation for $\Delta c_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta c_{t-1}$ $-0.02^*$ $(-2.74)$</td>
<td>$\Delta c_{t-1}$ $3.08^*$ $(100.3)$</td>
<td>$\Delta y_{t-1}$ $0.01$ $(1.29)$</td>
</tr>
<tr>
<td>$\Delta i_{t-1}$ $-0.03^*$ $(-21.1)$</td>
<td>$\Delta y_{t-1}$ $1.15^*$ $(35.5)$</td>
<td>$\Delta i_{t-1}$ $-0.00$ $(-0.70)$</td>
</tr>
</tbody>
</table>
under preference or sunspot shocks, but it still cannot make consumption Granger-cause output. This suggests that the concept of Granger causality adds additional restrictions on economic theory and is thus a powerful litmus test for equilibrium business cycle models.

**CONCLUSION**

The U.S. data suggest a “causal” relationship among consumption, output, and investment. This causal relationship may be surprising to some economists, but not to a businessman. According to a businessman’s intuition, production would not rise until consumption demand rises; and investment would not rise until profit rises along with the rise in production. The key elements missing in the businessman’s intuition, however, are the aggregate resource constraint and the price mechanism. Without changes in production possibilities or prices, what would enable consumption to rise in the first place without crowding out? General equilibrium business cycle models embodying the resource constraint and price mechanism, nevertheless, have trouble conforming to the data. There must be something fundamental missing in standard models, too. One possible missing element is inventory investment. Inventories provide a perfect buffer for consump-

---

**Figure 1**

**Impulse Responses of the Scale Economy Model to Demand Shocks**

- **Impulse Response of Y**
- **Impulse Response of C**
- **Impulse Response of I**
- **Impulse Response of N**
tion when output cannot react immediately to demand shocks. To model inventory behavior in general equilibrium, however, is itself a challenge and is therefore beyond the scope of this paper. My conjecture is that even with inventories introduced, the sequential information structure and the source of shocks (aggregate demand) are still crucial for the model to succeed. Hence, the Granger causality concept and the empirical regularities documented in this paper can prove to be a new litmus test for equilibrium business cycle models.

REFERENCES


13 For more recent work on this issue, see Kahn and Thomas (2004) and Wen (2005a).