FEDERAL RESERVE BANK OF ST. LOUIS

REVIEW

MARCH/APRIL 2001 | VOLUME 83, NUMBER 2

REVIEW

Review is published six times per year by the Research Division of the Federal Reserve Bank of St. Louis. Single-copy subscriptions are available free of charge. Send requests to: Federal Reserve Bank of St. Louis, Public Affairs Department, P.O. Box 442, St. Louis, Missouri 63166-0442, or call (314) 444-8808 or 8809.

Contents

Volume 83, Number 2

approach to modeling is not a panacea for recession forecasting. Although it does well explaining the data within the sample, its out-of-sample forecasts do not improve upon the benchmark probit specification.

Expectations

The Twenty-Second Henry Thornton Lecture, Department of Banking and Finance, City University Business School, London, England, November 28, 2000

William Poole

It is a great honor for me to be here tonight to present the Twenty-Second Henry Thornton lecture. In preparing this lecture, it has been fascinating to read parts of Thornton's great book, *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*, published in 1802, and F.A. Hayek's introduction to the 1962 reprint of *Paper Credit*. I recall reading Thornton years ago, but remember little of it. Rereading him today, I certainly appreciate Thornton's insights to a far greater extent than I did when I first read his book. I have also found it instructive to read several previous Thornton lectures. I'll refer to these lectures and to Thornton himself on several occasions this evening.

It is standard practice for Federal Reserve officials, with the exception of the Chairman, to begin every public presentation with a disclaimer. Thornton himself wrote a disclaimer in the introduction of his book, and I will adopt his disclaimer as my own for this lecture. Thornton wrote:

> That [this work's] leading doctrines are just, the writer feels a confident persuasion. That it may have imperfections, and some, perhaps, which greater care on his part might have corrected, he cannot doubt. But he trusts, that a man who is much occupied on the practical business of life, will be excused by the public, if he should present to them a treatise less elaborate, and, in many respects, more incomplete, than those on which he has found it necessary to remark. Future inquiries may possibly pursue, with advantage, some particular topics on which he has felt a certain degree of distrust.
>
> It may not be irrelevant or improper to observe, that the present work has been written by a person whose situation in life has supplied information on several of the topics under discussion . . .[1]

As one now pursuing the "practical business" of central banking, I can relate easily to Thornton's disclaimer. I would just add that I value the conversations on these subjects with my colleagues at the St. Louis Fed, especially Robert H. Rasche, but that I am responsible for the views expressed. These views do not necessarily reflect official positions of the Federal Reserve System.

Almost every aspect of human behavior is conditioned by expectations. Indeed, a distinguishing feature of humans among all living things is that humans, to an unmatched degree, calculate behavior in light of possible future outcomes. I cannot discuss the whole of human behavior in one lecture, or in one lifetime. Even the topic of expectations in a macroeconomics context is overly broad; I will concentrate rather unsystematically on aspects of this topic that are of special interest to me because of my current responsibilities. I will discuss issues from the perspective of central banking problems, but much of what I say applies to other areas of government policy.

By "rational expectations" I mean that market outcomes have characteristics as if economic agents are acting on the basis of the correct model of how the world works and that they use all available information in deciding on their actions. That information includes probable future monetary policy actions and, more generally, how monetary policy actions are likely to depend on various possible states of the economy. Expectations may be nonrational in an infinity of ways.[2] Almost every economist is familiar with the colorful language Keynes used to describe his view on how security values were determined. In one of his more succinct statements, Keynes (1936, p. 154) said that, "A conventional valuation which is established as the outcome of the mass psychology of a large number of ignorant individuals is liable to change

William Poole is the president of the Federal Reserve Bank of St. Louis. The author thanks colleagues at the Federal Reserve Bank of St. Louis for their comments, especially Robert H. Rasche, director of research. The views expressed do not necessarily reflect official positions of the Federal Reserve System.

[1] F.A. von Hayek (1962, p. 69): first published in 1939, this book contains Hayek's introduction, Thornton's *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain,* two of Thornton's speeches in the House of Commons in 1811, and other materials.

[2] I use the word "nonrational" rather than "irrational" because the latter sometimes carries connotations that I do not intend. Expectations may depart from full rationality without being "crazy," "silly," "emotional," or "stupid."

violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference to the prospective yield." Keynes and many others have viewed expectations as being driven by emotion and efforts to ride market trends without regard to underlying values. Popular commentary on bond, stock, commodity, and foreign exchange markets often focuses on presumed patterns in the data, such as resistance and support levels, that make no theoretical sense and are completely unsupported by careful empirical investigation.

In econometric models, economists have often used adaptive expectations, which are simple, and simple-minded, extrapolations of the past. Adaptive expectations are the antithesis of the emotional process Keynes emphasized. Adaptive expectations, as averages of recent observations, change relatively smoothly and continuously. They are unaffected by news items per se; if news moves the market, the adaptive expectation incorporates only a fraction of the unexpected price adjustment into expected future prices.

I will take up four topics. The first is what we can learn about expectations from banking panics and, more generally, from sharp disturbances in financial markets. The second is central bank credibility. The third is inflationary expectations. The fourth is the extent to which the market can predict central bank actions. I will connect these topics to produce what I hope will be a coherent account of certain expectations issues from the perspective of a practicing central banker.

I do not doubt that expectations are sometimes nonrational. My main theme, however, is that we central bankers should not be smug in assessing our presumably superior understanding of what expectations *ought* to prevail. We need to reflect on our possible role in creating and sustaining expectations that we regard as nonrational, and on the possibilities for pursuing policies that yield market outcomes closer to those reflecting rational expectations.

WHAT DO PANICS TELL US ABOUT EXPECTATIONS?

Sudden and unpredictable changes in market sentiment create problems for all sorts of businesses. Hayek, in his introduction to Thornton's *Paper Credit,* quotes from a contemporary account of an incident Thornton had to face in 1810.

> [Thornton] was on his road with his family to Scotland. It was a time of severe pressure upon banks and trading interests . . . The bank in which Mr. Thornton was a partner felt the pressure, and felt it severely, just after their most able partner had left London for the North. Had Mr. Thornton known what was impending, he would not have absented himself. The news reached him on his route to Scotland, and caused him some embarrassment. To return from a journey undertaken and generally known, would have spread rumors which might have brought on the very crisis that was to be feared. This course, therefore, could not be thought of. He decided to continue his journey, but he opened himself in confidence to one valued friend, and stated his wish that some thousands of pounds might be placed at the disposal of his partners in the bank. No sooner was the hint given than it was met by ample support. Funds poured in from all quarters—Wilberforce, with generous ardour, hastening to lead the way; and the money came in such a flood, that his bank saw itself lifted above the sands on which it was settling, and floated into deep waters with abundant resources. (p. 27)

This incident is interesting because it focuses on the problem of managing market expectations. From a central-banking perspective, the issues have been quite well understood since Walter Bagehot published *Lombard Street* in 1873. A central bank can resolve a banking panic by providing liquidity to solvent banks.

Let's look at the nature of the expectations issue when financial panic strikes. The place to start is with this question: Are the rumors sparking the crisis true? In the incident recounted above, the rumor was untrue. The bank was solvent and had access to ample sources of liquid funds; once it marshaled the funds, the problem was solved. In other cases, of course, rumors are true. In the fall of 1998, for example, Long Term Capital Management (LTCM) was severely overextended. The firm was indeed in danger of being unable to meet its obligations, and market participants were right to question its solvency. Moreover, the obligations outstanding were so large that significant market disruption might have occurred had the firm defaulted.

A central bank faces several issues in cases like LTCM. Without action, market prices may decline so much that a thinly capitalized firm goes under. But intervention may have the undesirable effect of propping up an institution that failed to meet the market test. This is the problem of moral hazard; other firms may bet on central bank intervention in similar cases in the future and thereby manage their affairs in a way that increases the probability of a crisis. Bagehot's solution to the moral hazard problem was for the central bank to lend at a penalty rate of interest. Marshaling private lenders who accept the risk works the same way.

What is the nature of expectations in a panic? Is the distinction between rational and nonrational expectations helpful here? I think we must look at two issues: One is that solvency may not be clear even to the best informed, most rationally calculating observer; the other is that the problem is sometimes just informed versus incompletely informed expectations.

Academic battles over rational expectations have often focused on rational expectations versus expectations driven by emotion or a failure to calculate sensibly. However, I think that panics large and small are sometimes driven by the lack of complete information, and in those cases the policy issue is relatively simple.

Consider an incident during the banking crisis in the state of Rhode Island in 1990-91. I was on the faculty of Brown University and lived in Providence, Rhode Island, at that time. A number of state-chartered credit unions and savings banks were insured by a private deposit insurance company. One of these credit unions, by the way, was the Brown University Employees' Credit Union. As I recall the chronology of events, in November 1990 one of the savings banks discovered a large embezzlement, which led to its failure. That failure nearly wiped out the assets of the deposit insurance company, which in turn led to widespread concern about the safety of deposits in other insured institutions. The crisis was reported in the Providence newspaper day after day. CNN sent a reporter to cover the story, and the reporter went on camera standing in front of a local bank—the Old Stone Bank. The next day, following the CNN report, there was a run on Old Stone Bank. Old Stone was federally insured and had nothing whatsoever to do with the crisis of the locally insured credit unions and savings banks.

Was it rational for Old Stone's depositors to pull their funds out of the bank? Those of us involved in banking and finance might easily say that such behavior was irrational because that bank was federally insured. But as I reflect on my own behavior in areas where I am less well informed, I am not so sure that judgment is sound. For example, when the recent publicity concerning Firestone tires hit the newspapers, I went out to my garage to look at the tires to figure out how my cars were equipped.

Information is costly, and our brains have only a finite number of cells to hold information. When an event or rumor brings an issue to public attention, many people will inevitably and appropriately react on the basis of highly imperfect information. The reactions may be perfectly sensible—rational, if you will—given the limited information at hand. Given incomplete information, I think it is completely rational for depositors to pull funds out of a suspect bank. Indeed, in the Eighth Henry Thornton Lecture, Karl Brunner argued that money itself exists because it helps to alleviate information problems. I agree with Brunner that the full-information version of the rational expectations hypothesis provides valuable insights for certain problems but is incapable of explaining some important phenomena.

Returning to the case in Rhode Island, the run on Old Stone Bank was quickly halted through the spread of accurate information. The bank itself and banking authorities emphasized to the public that Old Stone was federally insured and had no connection to the statewide banking crisis.

Many panic cases in practice reflect highly incomplete information. Given the costs of obtaining information, I think situations of this kind, which are not uncommon, provide compelling evidence against a pure, full-information version of the rational expectations hypothesis. Not only are some market participants poorly informed, which is obvious, but market outcomes can reflect poorly informed views. However, it is essential that we not equate expectations based on incomplete information with expectations that are hopelessly emotional and irrational; provision of information does have observable effects on market outcomes. From a policy perspective, that means that provision of accurate information is the first line of defense in cases of financial panics.

If this argument seems almost self-evident, we need to remember that from time to time central banks (and government authorities more generally) have contributed to the problem rather than alleviating it. Sometimes panics are driven by rumors that turn out to be substantially accurate. In such circumstances, those in authority may attempt to alleviate or avoid panic by glossing over the severity of the problem. Doing so may help to manage a particular incident, but at the cost of damaging the long-run credibility of the authorities.

A particularly clear, and expensive, example of this process was the U.S. savings and loan (S&L) industry. From the mid-1960s to the late 1980s, the U.S. government and regulatory bodies took numerous steps to deal with the institutional and financial weaknesses of numerous S&Ls. The process culminated in a $150 billion government bailout of the Federal Savings & Loan Insurance Corporation (FSLIC). Congress closed down FSLIC and the regulatory agency, the Federal Home Loan Bank Board. The political careers of several members of Congress were damaged or ended by the voters. I am convinced that the government could have avoided this entire mess if it had required market value accounting for S&Ls from the beginning.

Providing information prospectively, as with market value accounting, is perfectly feasible in many cases. In the Rhode Island banking crisis, and others, part of the problem has been that depositors genuinely believed that their deposits were perfectly safe—as safe as the currency in their wallets. The Rhode Island incident was not unique; the United States has a long history of failure of private and state deposit insurance funds.[3] If a government can standardize the definition of Scotch whiskey, why can't it standardize the definition of "deposit"? Given that depositors have so often been confused in the past, why not reserve the word "deposit" in the United States for a liability insured by the U.S. government?

Along the same lines, in the United States we need to clarify the extent of the federal guarantee for the liabilities of governmentally sponsored enterprises (GSEs). Although the legal situation differs from one enterprise to another, the liabilities of GSEs often carry no explicit guarantee, yet the market prices these obligations as if there were a federal guarantee. Based on past practice and continuing debate, market participants have every reason to assign a relatively high probability to a federal bailout should a GSE come close to defaulting on its obligations. Similar issues surround the "too big to fail" doctrine applied to large private financial institutions.

If a market crisis emerged one day because investors came to believe that the federal government was prepared to let one or more of these firms fail, would the crisis be the fault of nonrational expectations or of government policy that failed to clarify the issue?

The appropriate government role in guaranteeing financial obligations is a complex issue, and I don't intend to explore the merits of various positions here. But I do feel strongly that the government itself, not the market, is responsible if market expectations over a potential default seem emotionally driven and volatile. I hope I'm wrong, but I'm willing to speculate that the issue will remain unresolved in the United States until a threatened or actual default forces the issue. The United States did not address the S&L issue until it became too large to ignore. The political response is likely to depend heavily on the facts, or perceived facts, at the time, especially claims about who will be hurt by whatever decision is made and who is "at fault" and therefore deserves to be punished. Neither I nor market experts who know more about these matters can form confident expectations about outcomes in such cases. But I want to reiterate that the issues surrounding government guarantees can and should be addressed before a crisis strikes.

The rational expectations revolution in macroeconomics made clear that the distinction between policy and policy actions is critical. *Policy* reflects the general regularity of behavior of policymakers over time; *policy actions* are the individual responses case by case. Whenever policymakers believe that market expectations are irrational, policymakers ought first to look into the mirror and ask whether policy is coherent. Market expectations about policy cannot be coherent if policy is not coherent. I've suggested that U.S. policy toward federal guarantees is currently ill defined, and now I want to turn more explicitly to monetary policy.

I must say that there is amazingly little academic research providing solid guidance as to what I ought to do to help define a more coherent monetary policy. I am not implying, of course, that I believe that Fed policy is incoherent today. What I am saying is that research showing how we can do better, or even just characterizing more accurately the policy followed in recent years, is surprisingly thin. Research on monetary policy reaction functions seems quite unfruitful to date. Among those who have worked on this issue, I think the view is nearly unanimous that in recent years Federal Reserve policy has been better than

[3] See English (1993).

any proposed explicit policy rule. That means that no one has been able to write down a policy rule that accurately characterizes Fed policy.

This observation has a direct implication for research into the rationality of expectations. The key idea of the rational expectations hypothesis is that the market forms expectations based on estimates of model parameters that match the true model parameters. No one should be surprised if economists have difficulty confirming the rationality of market expectations about inflation, for example, if economists cannot even characterize Fed policy with much accuracy. Why should economists judge the market by standards they themselves, with all their knowledge of theory and econometrics, cannot meet? Indeed, this line of argument opens up the possibilities (i) that the market may behave *as if* it were able to characterize policy correctly and (ii) that economists' tests of rational expectations fail because economists fail rather than because markets fail. In the last section of this lecture, I'll describe some recent research at the St. Louis Fed suggesting that markets in fact understand recent Fed policy far better than economists do.

I've argued that market panics, and inexplicable changes in asset prices more generally, may not reflect the irrationality that many economists seem to assume. Panics may arise, at least in part, from the failure of policymakers to follow clear and coherent policies. Everyone agrees that, in general, asset prices ought to change when policy changes. If policy is ill defined, then no one should be surprised when asset prices change as market perceptions about prospective policy change. These perceptions will be weakly held and are therefore subject to change, perhaps even abrupt change, because it is not rational to have firm views about policy when policy is ill defined.

An objection to this view might be that it provides no explanation of the timing of panics and sharp changes in asset prices. But this objection is unconvincing. If an accurate empirical model—whether an economic or a psychological model—of timing existed, then the market would use that information to seek the profit implied. An uncontestable implication of rational expectations theory and evidence is that there are no easy profits to be had in asset markets. Panics and market crises *must* be unpredictable. To me, as a policymaker, the implication of inexplicable and unpredictable panics and asset price changes is not that we need a new, nonrational expectations approach to understanding expectations. Instead, we need to examine how policy bodies can more effectively transmit accurate information to the market and how policy can be made more coherent and reliable. In short, policymakers need to reallocate their thinking time more to looking inward at what they do and less to looking outward at what markets do.

WHERE DOES THIS CREDIBILITY COME FROM?

This discussion leads naturally to the broader subject of credibility. Markets should view economic policy in terms of a rule or regularity of behavior. Markets interpret individual policy actions in the context of their consistency with the policy, given the facts of the current situation. If authorities mislead the public in a particular situation, then public confusion or distrust will make it more difficult for policymakers to deal with the next crisis. It is important to emphasize the enormous benefit of central bank credibility in all areas in which it exercises its powers.

Central bankers have not always appreciated the importance of credibility. To relate a personal example from the 1970s, while on the faculty of Brown University I had many contacts with Federal Reserve officials. As inflation continued over the course of the decade, I became increasingly skeptical of the Federal Reserve's profession of allegiance to the goal of low inflation. I said, in effect, to some of my Fed friends, "I don't believe you." I think they were insulted by what I said, but the markets increasingly did not believe the Fed either. Although criticism from many different directions is a fact of life for central bankers, they should take such criticism seriously. At the same time, they should be careful not to assume that comments reflecting general esteem for those in office necessarily are a vote of confidence in the policies being pursued.

There is now an extensive literature on central bank credibility; I can hardly claim to be familiar with all of it. But what does strike me about this literature, as useful as it is, is that it does not go very far in providing specific advice to central banks about building credibility. The practical problem I face is in trying to decide how, if at all, to react to the latest release of employment data, inflation data, and the steady flow of other information of all kinds day by day. The problem is to make individual policy actions add up to a coherent policy. To be credible, the central bank must be successful in achieving its stated goals. To deliver on these goals, the central bank must know how to respond to the steady flow of information, and its responses to this information must make sense as policy. That is, every central bank needs a monetary policy strategy in which the goals are clear and the policy actions to achieve the goals are well defined.

Many market participants have great expertise in monetary matters, and they form reasoned judgments about the performance of central banks. We may call the view that emerges "reasoned credibility." But there is another aspect of credibility that arises from the fact that most of any individual's views and expectations come not from personal study and investigation but from acceptance of views of trusted authorities, or experts. No one has the time to be expert about everything. Reliance on experts is a consequence of the costliness of information. If a central banker is a trusted authority, his or her view on a wide range of economic issues, including many far removed from monetary policy, will carry great weight. Because trusted experts differ, and we all face the problem of picking which experts to believe, over time a central bank can develop special credibility among competing authorities. We may call this general trust of a central bank "institutional credibility."

Credibility in both its dimensions is earned, or lost, day in and day out, over big issues and small, and is not compartmentalized. In other words, a central bank cannot be distrusted in one area of its operations and retain high credibility in other areas.

The value of credibility is particularly clear in a crisis. When information is highly incomplete and the true state of affairs murky, it is extremely valuable for society if the markets can look to the central bank as a trusted authority and accept its judgments and actions. If the central bank is indeed well informed and competent, its credibility in the markets will obviously make its task far easier.

In 1985, Michael Parkin presented the Seventh Henry Thornton Lecture. His title was, "Inflation Expectations: From Adaptive to Rational to …?" As a part of his insightful review of expectations issues, Parkin discusses the failure of inflation expectations to fall promptly with the change in U.K. monetary policy in the early 1980s. He concluded that, given the history and the incentives to inflate, "it is not rational to expect, and act upon the basis of, a low rate of inflation" (p. 13). Both the United States and the United Kingdom bore heavy costs to reestablish expectations of low inflation and central bank credibility.

Central banks around the world today enjoy high credibility compared with the situation only 20 years ago. Just as there were observable market consequences—deep recession—of impaired credibility in the United States and United Kingdom in the early 1980s, there are observable market consequences today.

What are these observable consequences? I will speak only to the situation in the United States, where I know the history and data in detail. I think that there are many such observable consequences and that one of them is the sustained favorable surprise in the unemployment rate. Unemployment as low as the rate the United States has enjoyed in recent years could not have occurred without entrenched expectations of continuing low inflation. In the conventional Phillips curve, the rate of inflation depends on expected inflation and the gap between the actual and natural rates of unemployment. Anecdotal reports from employers and systematic information suggest that the U.S. labor market has been stretched abnormally tight for several years now. I think the best explanation of how these tight labor market conditions can continue is that expectations trump the gap. Firms are just not willing to bid aggressively for labor to fill empty positions because senior management does not believe that higher wages can be passed on in higher prices. Expectations of continuing low inflation dominate the outcome.

That is my tentative hypothesis anyway, but because I do not have research results to support it at this time I'll not pursue the matter further except to offer one more observation. Most economists believe, I think, that the rational expectations hypothesis is extremely valuable in understanding outcomes in auction markets—like those for equities, bonds, foreign exchange, and commodities—but is of limited application in the labor market. The labor market, so the argument goes, is dominated by institutional behavior, attitudes concerning equity, and slow adjustment to changing conditions. In econometric models of the labor market, adaptive expectations seem to work well enough. What I'm suggesting is that the U.S. unemployment rate has departed from the conventional estimate of the Phillips curve because that estimate failed to account adequately for the role of rational inflationary expectations in the labor market. The theory of rational expectations provides guidance in understanding economic behavior in all parts of the economy, not just in auction markets.

WHAT DO WE MAKE OF INFLATION EXPECTATIONS?

Thornton had a clear understanding of the distinction between the nominal and real rate of interest. In a speech before the House of Commons in 1811, he noted the following: "If, for example, a man borrowed of the bank a thousand pounds in 1800, and paid it back in 1810, having detained it by means of successive loans through that period, he paid back that which had become worth less by 20 or 30 percent than it was worth when he first received it. He would have paid an interest of 50 pounds per annum for the use of this money; but if from this interest were deducted the 20 pounds or 30 pounds per annum, which he had gained by the fall in the value of the money, he would find that he had borrowed at 2 or 3 percent, and not at 5 percent as he appeared to do" (Hayek, pp. 335-36).

A thorough understanding of the distinction between real and nominal interest rates is a great advance in central banking practice over the last 35 years. We've finally caught up with Thornton. In the United States, at least, in the mid-to-late 1960s, the practical importance of the distinction between real and nominal interest rates was not appreciated. Rising interest rates in the late 1960s were misinterpreted as evidence of a more restrictive monetary policy, when, in fact, nominal rates were not even keeping up with the increase in inflation expectations.

Compared with 35 years ago, the Federal Reserve today has access to far more data on expectations. With inflation-indexed bonds outstanding, we have day-by-day evidence on the behavior of the spread between conventional and indexed bonds. Survey information is widely available. I watch these data closely because they provide clear evidence of the central bank's success in maintaining credibility in achieving sustained low inflation.

The logic of the credibility argument, however, suggests that inflation expectations data do not provide definitive evidence about whether monetary policy itself is on track. Given that the Fed enjoys very high credibility today, the markets will not necessarily bid up inflation expectations when and if policy goes astray. High credibility means that the market trusts the Federal Reserve's policy judgments. That being the case, the Federal Reserve cannot reliably extract information from data on expectations about the appropriateness of current policy actions.

It is logically possible that policy actions are inconsistent with sustained low inflation at the same time that the market simply trusts the Fed and does not perform a separate analysis of policy actions. Why should any of us, on any matter, engage in a costly investigation when we can instead simply accept the judgment of a trusted authority? The answer is obvious: If the authority is completely trusted, and if separate confirmation of the information is costly, then the cost-efficient thing to do is simply to accept the authority's judgment.

Let me summarize this analysis. The expected rate of inflation over a five-year, or longer, horizon is a direct measure of central bank credibility regarding inflation. At any given time, monetary policy—policy, not policy actions—may or may not be consistent with long-term inflation expectations. Eventually, of course, expectations and policy must be consistent because one or the other will adjust.

Failure to understand this point could foster policy mistakes. When credibility is high, as it is in the United States today, inflation expectations will be slow to adjust. Actual inflation, influenced by expected inflation, may also be slow to adjust. Therefore, expected inflation, certainly, and actual inflation, probably, are poor guides as to the appropriateness of monetary policy in the short run. Similarly, when inflation expectations are high and credibility low, the central bank has the twin problems of getting policy turned around to be consistent with lower long-run inflation and of adjusting policy as credibility builds over time.

If the Fed cannot rely on actual and expected inflation to judge the appropriateness of current policy, because these measures are dominated by the market's assessment of Fed credibility, what can it rely on? We need to concentrate on the underlying determinants of inflation and early warning signs. The rate of money growth, spreads in financial markets, the supply-demand balance across industries in general, and the behavior of specific prices likely to lead overall inflation are relevant. The aim of policy should be to act *before* changes in inflation appear; clearly, once these changes do appear, the task of restoring credibility and reversing all the adjustments that firms and households have started to make becomes more difficult.

WHAT IS THE SIGNIFICANCE OF MARKET PREDICTIONS OF CENTRAL BANK POLICY?

I'm now going to bring the various strands of my discussion together. My colleague Robert Rasche and I have been pursuing a line of research on the predictability of monetary policy actions. The paper (Poole and Rasche, 2000) is available in the working papers section of the St. Louis Fed Web site; it will be published in the *Journal of Financial Services Research*. I'll outline the basic idea in that paper and then connect it to the argument of this lecture.

Consider a state of monetary policy nirvana in the world we actually live in. That is, if the central bank did as good a job as you can imagine in today's world—a world with many gaps in knowledge, data inaccuracies, and all the real problems real central banks face—what would we observe?

Let's suppose that you think a measured CPI inflation rate of 1 percent per year is optimal and that you believe the central bank can offset some financial and real disturbances to cushion fluctuations in output and employment without compromising the inflation objective. This is a short description of what I believe, but you can substitute your own specification for mine.

The market will, in due time, learn of the policy objective and the policy actions designed to achieve that objective. Real central banks almost without exception implement policy by setting a target for a short-term interest rate, usually an overnight bank rate. In the United States, that target rate is the federal funds rate. So, I'll assume that our real central bank implements policy actions that way.

I've given you a very simple description of what the central bank wants to do and its procedure for pursuing its objective. Given the assumed nirvana state of monetary policy, the central bank does its job efficiently. By that I mean that it responds sensibly to all the ambiguities and problems real central banks face. As new information arrives, the central bank efficiently processes its significance and adjusts its target for the overnight rate as required to achieve its policy goals. Given the inherent gaps in knowledge and data, sometimes the central bank will act too quickly or too slowly, by too much or too little. But my presumption is that the central bank can avoid cumulative errors and recover from policy missteps without missing its objectives.

Participants in financial markets will understand what the central bank is doing. To understand market outcomes in this setting, one other observation is needed. In the United States—I'm not sure about the situation elsewhere—the central bank has no significant informational advantage over the market. The Fed and the markets receive government statistical data at essentially the same time. The Fed does have an advantage over the market in that it has a very large staff and does obtain anecdotal information not generally available. However, individual firms have much more extensive information about their own markets than the Fed does. I think it is approximately correct, and certainly appropriate at the level of theoretical modeling, to assume that the markets and the Fed receive the same information at the same time.

Market participants have ample incentive to form accurate expectations about central bank policy actions. How accurate are those expectations likely to be? Given my assumptions, the market ought to be very accurate in predicting policy actions. The market and the central bank get the same data at the same time; the market understands the policy objectives and the policy actions appropriate to achieve the objectives. As new data arrive, the market should interpret the data the same way the central bank does, at least most of the time, and reach the same conclusion about the significance of the data.

Rasche and I have explored this hypothesis for the United States. Our research is ongoing, but at this time we can report that as of the last few years the market has been quite accurate in forecasting Fed policy actions. Since 1988, when trading opened in the federal funds futures market, we have had a very direct reading on market expectations about Fed policy. Since 1994, that market has predicted policy actions quite accurately on the whole.

It is instructive to note that 1994 was a watershed year. In February 1994, the Federal Open Market Committee (FOMC)—the Fed's main monetary policy body—first began to release a policy decision about its federal funds rate target immediately following the FOMC meeting. Before that time, the market learned of policy actions by observing open market operations conducted by the Open Market Desk at the New York Fed. Moreover, before 1994, policy actions occurred more often between regularly scheduled FOMC meetings than at the meetings. Since February 1994, almost all policy actions have been taken at regular FOMC meetings.

Although the FOMC adjusts the target federal funds rate most often by only 25 basis points, it sometimes has made larger adjustments. But these adjustments have been well predicted by the market.

This evidence shows conclusively that it is possible for a central bank to pursue a highly predictable policy, in the sense that, given the available information at the time of a policy meeting, the market can predict the policy action. Policy actions cannot be predicted far in advance because the information driving policy decisions cannot be predicted far in advance. But, as information accumulates before a policy meeting, the market and the central bank can converge on a common interpretation of the information.

Moreover, the market is well ahead of economists in understanding this process. I know of no econometric models that predict both the timing and the magnitude of Fed policy moves with anything close to the accuracy of the predictions in the federal funds futures market. There is an important research agenda implied by this observation. We need a deeper understanding of U.S. monetary policy to increase the probability of extending recent policy successes into the indefinite future.

This experience also shows that the central bank can change what it does to promote more accurate market expectations. By disclosing policy decisions quickly and by confining policy actions to regularly scheduled meetings, the FOMC has made possible improved market forecasts of monetary policy actions. The change in practice in February 1994 illustrates the point I emphasized earlier—that the central bank can improve the accuracy of information available to the market.

I believe that the simple step of prompt disclosure in February 1994 also imposed a valuable discipline on the FOMC itself. By confining most policy actions to days of FOMC meetings, the Committee made its own behavior more predictable. Now, everyone knows that a policy action at another time is special. The FOMC must think carefully about whether it wants to send a special message by changing policy between meetings and, if it does, what the message is. What the central bank does will shape expectations; for the central bank to be able to predict its effects on expectations, its own behavior must be as regular as possible.

WHAT SHOULD THE AGENDA FOR CENTRAL BANKS BE?

The rational expectations revolution in macroeconomics changed forever how we think about economic policy. We know that understanding markets requires that we understand market expectations about monetary policy. We know that the distinction between policy actions and policy itself is of central importance. We know that expectations are not always fully rational, but I have been at pains to argue that some of the problems caused by nonrational expectations are correctable.

I know of no policy models indicating that the economy works better when markets are kept guessing about monetary policy. The presumption must be that market participants make more efficient decisions—decisions that maximize economic growth by minimizing the wastage of resources from expectational errors—when markets can correctly predict central bank actions. That does not require that central bankers and market participants be able to forecast the unforecastable, but that they have a common understanding of the strategy governing policy actions.

I've suggested a large agenda—one that is indefinitely large—for central banks and governments. We need to focus on areas where market expectations are hazy because government policy itself is or may be ill defined. These include the nature of government guarantees, monetary policy objectives, and the strategy to reach those objectives. Some of the things we need to examine may appear terribly mundane. For example, I think that the FOMC probably meets more often than necessary. Market interest rates have ample room to fluctuate for any given federal funds rate, and it is rare that anything happens within the usual six weeks between FOMC meetings to require a reassessment of policy. If the markets and the central bank really do have a common understanding of monetary policy, it is hard for me to believe that outcomes for the 10-year bond rate, say, will depend on whether the policy meetings occur once a month or once a quarter. However, each meeting is an object of speculation; the market would be better served if traders would concentrate on the fundamentals behind policy decisions than on the meeting itself. My point is not actually to take a firm position on the minor issue of the meeting schedule but instead to point out that all sorts of things should be discussed as possible ways to improve the market's understanding of monetary policy.

I finish with a plea to both academics and central bankers. Of academics, I ask that research address this question: How, very explicitly, should policy instruments be adjusted? That is, what should central banks do and when should they do it? Of my central bank colleagues, I ask that we spend more time focused on defining general policy rules, or regularities, within which we will fit individual policy actions. Both enterprises promise significant improvements in the accuracy of market expectations and the stability of markets.

REFERENCES

Bagehot, Walter. *Lombard Street: A Description of the Money Market*. London: 1873.

English, William. "The Decline of Private Deposit Insurance in the United States." *Carnegie-Rochester Conference Series on Public Policy,* June 1993, pp. 57-128.

Hayek, F.A. von, ed. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain (1802)* [reprint]. New York: Augustus M. Kelley, [1939] 1962.

Keynes, John Maynard. *The General Theory of Employment, Interest and Money.* New York: Harcourt Brace, 1936.

Parkin, Michael. "Inflation Expectations: From Adaptive to Rational to ...?" Seventh Henry Thornton Lecture, City University Business School, London, 14 November 1985.

Poole, William and Rasche, Robert H. "Perfecting the Market's Knowledge of Monetary Policy." Working Paper 2000-010A, Federal Reserve Bank of St. Louis, April 2000; *Journal of Financial Services Research* (forthcoming).

The Expected Federal Budget Surplus: How Much Confidence Should the Public and Policymakers Place in the Projections?

Kevin L. Kliesen and Daniel L. Thornton

In 1998, the federal government recorded its first budget surplus in more than 25 years. Now, after an extended period of deficits and three consecutive years of surpluses, both the White House Office of Management and Budget (OMB) and the Congressional Budget Office (CBO) have projected annual budget surpluses for at least the next decade. The turnaround in the outlook for the U.S. government's finances is stunning. Under current policies, budget projections show that publicly held government debt, which is currently a little more than $3.5 trillion, will be eliminated by around 2010—perhaps earlier if the economy continues to grow faster than anticipated.[1] The political response to these projections has been rapid. Despite legislated budgetary "caps," fiscal policymakers have already increased spending and reduced taxes relative to previous agreements.

Projected large government surpluses have implications for the current political debate—possibly even the future implementation of monetary policy.[2] The purpose of this article is to assess the likelihood that the projected surpluses will materialize by comparing previous budget projections with actual outcomes. No one can say for sure whether these projected surpluses will materialize or whether publicly held government debt will be eliminated. If past experience is any guide, however, it seems likely that the market in default risk–free government debt will be with us for some time. Policymakers, accordingly, might be wise to consider this fact before deciding to ramp up spending or cut taxes solely on the basis of current projections.

The first section of the paper reviews the current budget projections of the CBO and the OMB; it also reviews government budget accounting practices and discusses whether government debt is rising or falling. As a means of ascertaining the probability that the current projected surpluses will materialize, the second section analyzes the CBO's projections since 1976 relative to actual outcomes. The third section of the paper analyzes the major sources of error in these projections. The conclusions are presented in the fourth section.

CURRENT BUDGET PROJECTIONS

Table 1 presents the 10-year budget projections and major economic assumptions published by the CBO and the OMB in 2000. The projection period is fiscal years 2001 to 2010. There are two sets of projections because each agency publishes a major report early in the year and then during the summer. One other fact worth noting is the approach taken by each agency. Generally speaking, CBO baseline budget projections follow the current services approach, which assumes that the current spending and tax programs remain in place throughout the projection period—though the CBO, at least in recent years, has typically published alternative projections using different spending assumptions. The OMB also publishes a current services (baseline) projection, as well as a projection that traces out the path of the surplus or deficit over time assuming that the Administration's specific policy proposals are enacted. We label the former OMB and the latter OMB*.

According to the January 2000 budget projections shown in Table 1, the CBO projected that the unified budget surplus (the sum of "on budget" and "off budget" revenues and expenses) will rise from $177 billion in fiscal year 2001 to just under $489 billion by fiscal year 2010.[3] Cumulatively, the 10-

Kevin L. Kliesen is an economist and Daniel L. Thornton is a vice president and economist at the Federal Reserve Bank of St. Louis. Stephen Majesky and Thomas A. Pollmann provided research assistance.

[1] Authors' note: This article was prepared prior to the publication of the projections for FY 2002-10 by the CBO and OMB. See Appendix.

[2] The prospect of eliminating publicly held federal debt poses a potential problem for the Fed. Historically, the Fed has implemented monetary policy by buying and selling government securities (open market operations). Because the public's demand for money tends to rise with nominal GDP, over time, the Fed buys more government securities than it sells to increase the supply of money. If the public debt were eliminated, the Fed would be unable to acquire additional government debt. Moreover, it would have to replace its holdings of government debt with something else. The question is what?

[3] Off-budget includes Social Security and the U.S. Post Office. On-budget is everything else—including other trust funds. This baseline—one of three published by CBO—assumes that *discretionary* spending grows at the rate of inflation after FY 2000. Because this projection assumes considerably greater discretionary spending levels, it produces the smallest cumulative surplus over the forecast horizon. Specifically, it presumes that the discretionary budget caps that are set to expire in 2002 will not be adhered to. Most budget analysts use this baseline projection.

Table 1

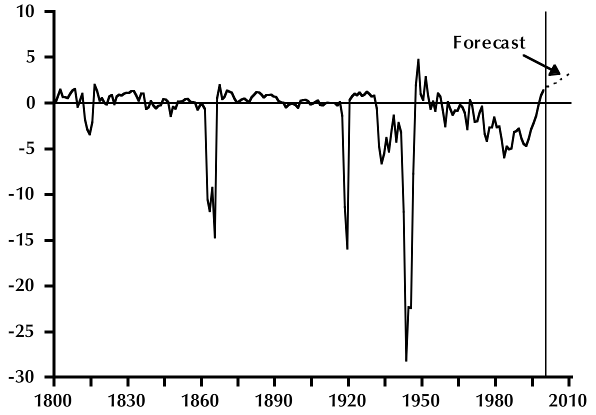
OMB/CBO Unified Surplus Baseline Projections and Economic Assumptions

	Surplus/deficit projections (billions of dollars)			Economic assumptions (fiscal year)		
	2001	2010	Cumulative	Real GDP	Inflation	10-Year bond
CBO (January 2000)	177	489	3,152	2.77	2.49	5.84
CBO (July 2000)	268	685	4,561	2.79	2.60	5.90
OMB* (February 2000)	184	363	2,519	2.71	2.59	6.10
OMB	171	457	2,919	2.71	2.59	6.10
OMB* (June 2000)	228	416	2,912	2.93	2.60	6.30
OMB	239	670	4,193	2.93	2.60	6.30

NOTE: Growth of real GDP, inflation, and the 10-year U.S. Treasury bond rate are averages of annual forecasts/projections. CBO and OMB projections use the CPI inflation rate. OMB is the Administration's current services projection.

Figure 1

The Federal Surplus/Deficit as a Percentage of GNP



SOURCE: *International Historical Statistics: The Americas 1750-1993*; and Congressional Budget Office (2000).

year unified budget surplus was expected to sum to just under $3.2 trillion. Although the Administration expected a somewhat smaller cumulative surplus because of the President's budget proposals, roughly $2.5 trillion (OMB*), its cumulative current services projection (OMB) was fairly close to the CBO's.

These projections are not only large in dollar terms, but, if realized, would be historically large as a percentage of gross domestic product (GDP). Figure 1 shows the U.S. annual surplus/deficit as a percent of GNP/GDP since 1800 and the projections to 2010. There is no prior multiyear period when the actual surplus was as persistently large as the CBO's current projections. Prior to 1970 the only multiyear periods of protracted deficits were associated with wars.

Since these projections were made, the U.S. fiscal outlook has improved even further according to the mid-year updates. As seen in Table 1, the CBO and OMB have increased substantially their cumulative 10-year budget surplus projections. The CBO's *Budget and Economic Update* released in July 2000 projects a cumulative $4.6 trillion surplus. In updated projections published in its *Mid-Session Review* released in late June 2000, the Administration's projected 10-year surplus (OMB*) was increased from $2.5 trillion to $2.9 trillion. The biggest surprise was the Administration's sharply higher projection of the 10-year current services surplus. Using slightly more optimistic economic assumptions, the OMB later estimated that the unified budget surplus will total $4.2 trillion over the 2001-10 period, which is more than 40 percent, or $1.3 trillion, greater than the February 2000 current services baseline. Revised economic assumptions account for $984 billion, or more than 75 percent, of this $1.3 trillion upward revision.[4]

Basic Budget Accounting: Is the Government Debt Rising or Falling?

One can think of the government debt as being the sum of all surpluses and deficits during U.S. history. In the very long run this is approximately true. In any given year, however, when the government

[4] Specifically, real GDP growth averages about one quarter of a percentage point more a year. Technical reestimates account for another $375 billion of the total upward revision, roughly split equally between increased receipts and lower expenditures. See OMB (2000) Tables 5, 8, and 9. Finally, legislation enacted since February 2000 is expected to decrease the projected surplus by $84 billion over this 10-year horizon.

Table 2

Government Accounting with Trust Fund Surpluses (Billions of Dollars)

	Fiscal year 1998	Fiscal year 1999	Change 1998-99
Debt outstanding			
Total government securities	5,478.7	5,606.1	127.4
Less:			
Securities held by government accounts	1,757.1	1,973.2	216.1
Equals:			
Government securities held by the public	3,721.6	3,632.9	–88.7*
Sources of financing			
Unified surplus, FY 1999			124.4
Borrowing from the public			–88.7*
Operating cash, change			–17.6
Other			–18.2
Surplus/deficit (–)			
Federal funds	–92.0	–88.3	3.7
Trust funds	161.2	212.7	51.5
Total unified surplus/deficit	69.2	124.4	55.2
Memorandum			
Treasury cash balance, end of FY 1999			56.5

NOTE: Includes debt issued by government agencies. Value of debt outstanding is the face value less the net unamortized premium and discount, otherwise known as the accrual amount. Totals may not sum because of rounding. *Identical values calculated by two separate methods.

SOURCE: Office of Management and Budget, *Historical Tables*, Tables 1.4 and 7.1; Federal Reserve Board of Governors, *Federal Reserve Bulletin,* November 2000, Table 1.38.

runs a deficit, it can borrow from the public or change its holdings of cash or other assets.[5] Conversely, when the government runs a surplus, it can retire its debt or increase its holdings of cash or other assets. Consequently, over any given period, government debt will not necessarily change dollar-for-dollar with the surplus or deficit.

To further complicate matters, whether the debt changes or not depends on where in the government the surplus or deficit is generated. Within the government there are a number of trust funds, the best known of which is the federal Old-Age and Survivors Disability Insurance (OASDI) trust fund—Social Security. When a surplus is generated within the trust fund, the government issues *nonmarketable* interest-bearing debt to the trust fund. This is essentially an IOU that one area of government issues to another. Likewise, when a trust fund runs a deficit, the government must buy back the nonmarketable debt using surpluses from elsewhere (should they exist), borrow from the public, or raise taxes.[6]

In the government accounts, revenues and expenditures are thus classified as either federal funds or trust funds. Trust fund revenues are earmarked for a specific program or purpose, such as OASDI or federal road, bridge, and highway construction (highway trust fund). If the government runs a surplus in both its federal and trust funds accounts, total debt will fall.[7] If the government runs a surplus in only one of its accounts, total debt need not fall. For example, in fiscal year (FY) 1999 the federal government's $124.4 billion budget surplus was the result of a $88.3 billion deficit in the federal

5 The other assets include such things as special drawing rights (SDRs), allocations of SDRs, the reserve position in the International Monetary Fund (IMF), loans to the IMF, changes in the gold stock, etc. See Rasche (1980) for more details.

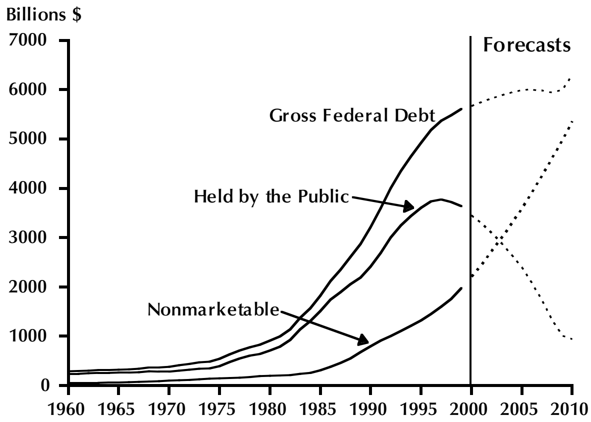
6 The accounting is more complicated because the interest income from the nonmarketable debt issued to trusts is accumulated in a separate account. Consequently, the government does not have to buy back the debt issued to the trusts until the accumulated interest income in these accounts is exhausted. This interest "income," however, is a mere intragovernment bookkeeping entry, since it cannot be used to offset expenditures elsewhere. In other words, because budget accounts measure transactions with the public, even though the trust funds may credit the government with interest income from nonmarketable debt, the reported unified budget is unaffected.

7 This is the general case. It is also possible to accumulate other assets with the surplus or decide to hold increased cash balances.

Figure 2

Gross Federal Debt and Federal Debt Held by the Public

Fiscal Years



SOURCE: Congressional Budget Office (2000)

funds account—mostly individual and corporate income tax revenues used to fund, for example, defense outlays or international aid programs—and a $212.7 billion surplus in the trust funds account.[8] As detailed in Table 2, the net government surplus was thus $124 billion, of which nearly $89 billion was used to retire government debt held by the public. Hence, publicly held government debt fell while total government debt rose.

Figure 2 shows the evolution of total (gross) federal debt, publicly held government debt, and nonmarketable debt since 1960, along with the CBO's projections through 2010. Despite cumulative budget surpluses of nearly $200 billion in fiscal years 1998 and 1999, gross federal indebtedness rose by nearly $130 billion. The reason, of course, is that the government issued nonmarketable debt to trust funds equal to their surpluses during this period. The largest of these is the OASDI fund, which "held" $855 billion in government debt at the end of FY 1999. Another $1.15 trillion was held by other trust funds, such as the federal employees' retirement fund ($474.7 billion) and the hospital insurance portion of Medicare ($153.8 billion).

The practice of issuing nonmarketable debt to trust funds was motivated by a desire to assure the public that the benefits that have been promised under specific programs will be forthcoming. Unlike private sector accounting rules, which mandate that firms fund their pension-plan benefits on an accrual basis, the federal government is not required to hold sufficient assets to compensate future beneficiaries their accrued benefits. The reason, of course, is that the government can, if necessary, raise taxes by the amount necessary to pay existing retirees—tax future workers to pay today's workers' retirement benefits. Alternatively, the government could borrow from the public when Social Security program payments exceeded payroll tax revenues, as it did frequently prior to 1983.[9]

Given that there was no explicit commitment to either raise taxes or borrow from the public, however, some feared that the government might instead renege on its commitment to Social Security recipients by reducing program benefits when the OASDI program began to run persistent deficits. Consequently, Congress attempted to tie its own hands by issuing nonmarketable, interest-bearing debt to the Social Security Administration in the amount of the Social Security surplus—the so-called Social Security "Lock Box."

The "Gray" Area of Future Budgets

Although publicly held debt is currently declining and may decline further, it is unlikely that public debt will stay at zero very long, should it ever get there. The reason is demographics. Recent generational accounting by Gokhale et al. (2000) suggests that the net tax liabilities of future generations will increase significantly because of the pending retirement of the baby boom generation. Likewise, according to the 2000 Social Security Trustees Annual Report, program expenses are projected to exceed payroll tax revenue beginning in 2016.

The federal government's unfunded liability of the Social Security program measured on an accrual basis was about $10.4 trillion in 1999 according to congressional testimony by CBO Director Dan Crippen. This is more than 12 times larger than the $855 billion in nonmarketable debt held by the OASDI trust fund in FY 1999. Consequently, it seems unlikely that the government will be able to meet its future obligations without borrowing from the public. In other words, at some point, publicly held debt

8 This accounting is a little misleading because the surpluses in some trusts are due to intragovernmental transfers from the general funds to these trust funds. Much of these transfers are the direct consequence of the trust funds holding interest-bearing government debt. See CBO (2000a, pp. 22-23) for additional details.

9 Since the 1983 Social Security reforms (Greenspan Commission), payroll tax revenues collected by the government to finance OASDI benefits have exceeded program payments to recipients by a little more than $50 billion a year, while the surplus for all federal trust funds averaged a little more than $100 billion a year.

will inexorably rise.

Indeed, projections published in the CBO's latest *Long-Term Budget Outlook* indicate such a scenario: Even if Congress and the Administration manage to wall off all of the projected (cumulative) $2.4 trillion *off-budget* surplus during the next decade in a "lock box," the debt held by the public will rise to a little more than 62 percent of GDP by 2040, compared with about 40 percent in 1999. Further simulations project that the share of the debt held by the public would rise to 184 percent of GDP by 2040 if there were no annual surpluses during the next 10 years. Accordingly, unless there is a surplus elsewhere in government to offset the Social Security deficit, the government will have to make up the shortfall by raising taxes, reducing benefits, and/or issuing marketable debt.[10] If the last option is used, publicly held debt will rise as the trust funds' holdings of debt decline.

PAST PROJECTIONS AND ACTUAL EXPERIENCE

Will publicly held government debt go to zero? Not if the current projected surpluses don't materialize.[11] Given the relative magnitude of the current surplus projections and the speed with which earlier CBO projections shifted from deficit to surplus, perhaps it is not surprising that nearly 60 percent of the respondents in a recent survey of business economists (conducted by the National Association for Business Economics) indicated that the CBO's budget projections are too optimistic. Only 31 percent of the respondents believed that the projections were reasonable.[12] It is not just economists who are skeptical of official budget projections. A survey by the Business Council suggests similar skepticism among chief executive officers of major U.S. corporations.[13]

Some insight into the accuracy of budget projections can be obtained by evaluating the CBO's past projections. We focus on the CBO's projections because the CBO is viewed as an impartial arbiter of the federal government's fiscal outlook. The budget projection process was not intended as a forecasting exercise. It was created to provide a method of analyzing alternative budget proposals. Nevertheless, the public and policymakers frequently treat these projections as forecasts, suggesting—or implementing—changes in spending or taxes based on them.

It is important to note, however, that this exercise is biased against the CBO for a couple of reasons. First, many of the rules employed to calculate the baseline budget projection are set by statute. Consequently, the CBO is forced to construct its baseline projections under these assumptions, regardless of whether CBO believes they are realistic. For example, the Deficit Control Act of 1985 mandates that the CBO project the future discretionary spending levels from current-year appropriations, whether the current-year appropriations are unusually high or low due to special appropriations (e.g., Desert Storm).

The second reason, which is related to the first, stems from the difference between budget projections and forecasts. Baseline budget projections are based on the current services approach, that is, the assumption that the government will leave its current tax and spending programs in place.[14] This is highly unlikely, so budget projections are almost assuredly going to be wrong. For example, the Omnibus Budget Reconciliation Act (OBRA) of 1993, which sought to produce markedly lower future deficits in relation to earlier projected deficits, produced budget projections markedly different from those of the previous baseline. As described later, the CBO estimates that about one third of its projection error is legislative, i.e., due to changes in tax and spending programs that were assumed to be constant for the purpose of making the projections. Despite the difficult nature of projecting future fiscal outcomes, it is nonetheless instructive to analyze the CBO's past budget projections.

The CBO began publishing five-year budget projections in January 1976 for FY 1977-81. Figures 3 through 5 present the actual one-year, five-year, and cumulative five-year projected surplus/deficit annually for the period 1976-99. If the CBO's past projections coincided with what subsequently transpired—that is, a zero projection error—then all of the points would lie on the 45-degree line. While few observations fall on the 45-degree line, at the one-

10 Social Security actuaries estimate that the Trust Fund will exhaust its holdings of nonmarketable government debt by about 2037. The extended period reflects the interest "income" that is annually credited to the trust fund. See footnote 7.

11 See Auerbach (2000) and the other papers published by the Federal Reserve Bank of New York (2000) for discussion of this and other issues.

12 See National Association for Business Economics (2000).

13 See Weill (2000).

14 In making their current services projections, the CBO and OMB each assume a path for several key economic variables and then calculate the path of outlays and receipts under the assumption that tax rates and spending programs will not change over the projection period. Tax rates and spending programs—other than automatic stabilizers such as income support programs for the unemployed—do not vary with forecasts of the economy, nor does a policy change endogenously alter the path of the key economic variables. As discussed by Rasche (1985), this is different from a model-based forecast, such as those used by most forecasters.

Figure 3

Actual and Projected Budget Deficit/Surplus at One-Year Horizon

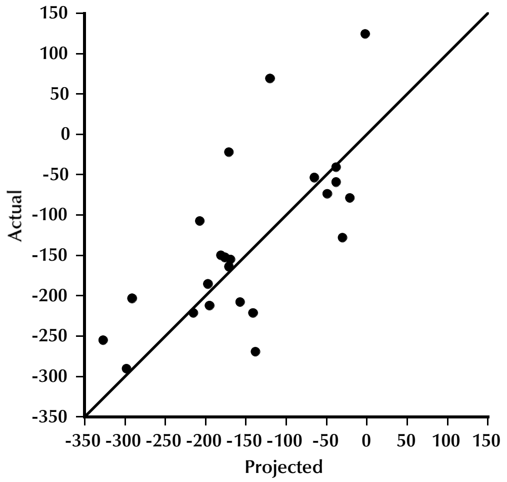


Figure 5

Actual and Projected Budget Deficit/Surplus: Cumulative Five-Year

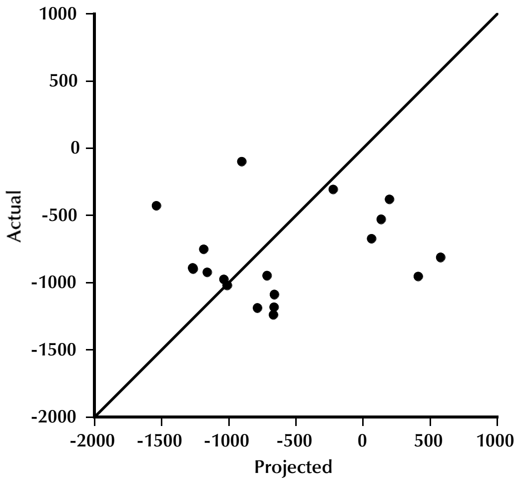
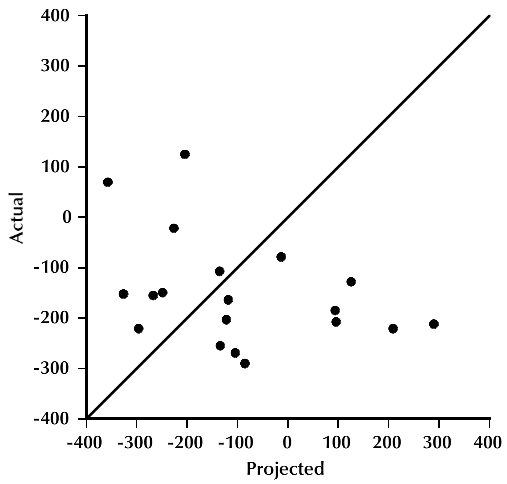


Figure 4

Actual and Projected Budget Deficit/Surplus at Five-Year Horizon



year horizon they are about equally spaced above and below the line, suggesting that the projections are approximately accurate *on average*. In other words, at the one-year horizon, the average projection error is not too far from zero. More important, at the one-year horizon there is a positive relationship between the projections and the actual outcome. Hence, if the CBO projected that the surplus would increase the next year, there was a tendency for the actual surplus to increase.

At the five-year horizon, outcomes are quite different. Figure 4 reveals that there is no relationship between the CBO's projections and the actual surplus/deficit. Not only do none of the observations fall on the 45-degree line, but a large deficit was as likely to occur when the CBO projected a surplus as it was when the CBO projected a deficit. At the five-year horizon, the CBO's projections have not been a useful indicator of what the future is likely to entail.

The CBO's cumulative five-year projections are more important for the question of whether publicly held government debt will be eliminated. Figure 5 shows that the CBO's cumulative projections have frequently missed their mark. In some cases, the projection errors are very large. For example, in 1980 the CBO projected a $578 billion cumulative surplus over the next five years. The actual outcome, however, was a deficit of a little more than $800 billion. If past experience is a guide to the future, the CBO's current projection of a cumulative surplus of a little more than $4.5 trillion dollars should not be treated as a forecast that can be relied on. The government is as likely to experience a deficit as it is to experience a surplus.

The general impression obtained from Figures 3 through 5 is summarized numerically in Table 3, which presents summary statistics for the CBO's projection errors (projected less actual) at the one-year and five-year horizons for the surplus/deficit, outlays, and receipts. To make the comparisons more useful, the projection errors are presented as a percentage of nominal GDP. At the one-year horizon, the mean absolute error (MAE) is about 1 percent of GDP. The

Table 3

CBO Projection Error Statistics: Deficit/Surplus, Total Outlays, and Total Receipts (Percent of GDP)

Series	Average error	MAE	RMS error	Minimum	Maximum
One-year-ahead projection errors					
Deficit/surplus	0.05	1.05	1.31	–2.19	3.03
Total receipts	0.03	0.75	0.99	–1.79	2.83
Individual income taxes	–0.08	0.49	0.63	–1.40	1.30
Corporate taxes	0.10	0.22	0.28	–0.37	0.64
Social insurance taxes	0.00	0.26	0.46	–1.93	0.52
Total outlays	0.25	0.56	0.69	–0.83	1.46
Mandatory	0.09	0.55	0.67	–1.32	1.27
Discretionary	0.07	0.17	0.20	–0.15	0.41
Five-year-ahead projection errors					
Deficit/surplus	1.97	4.10	5.28	–4.94	12.13
Total receipts	1.55	2.41	3.35	–2.78	8.86
Individual income taxes	1.16	1.65	2.28	–1.93	6.11
Corporate taxes	0.35	0.71	0.91	–0.79	2.00
Social insurance taxes	0.21	0.57	0.68	–1.60	1.13
Total outlays	–0.43	2.22	2.74	–6.31	3.67
Mandatory	–0.09	0.70	0.84	–1.66	1.30
Discretionary	0.81	0.83	1.13	–0.12	2.37

root-mean-squared error (RMSE) is slightly larger at 1.3 percent of GDP, and the range is about 5.2 percent of GDP.

How large are these numbers? One way to put these numbers into perspective is to note that the average deficit during the 1976-99 period was 3.0 percent of GDP. Hence, at the one-year horizon, the MAE is about one third of the average deficit. Moreover, actual surpluses and deficits ranged from –6.0 percent of GDP to 1.4 percent of GDP, so the range of projection errors was about 70 percent of the range of actual surpluses/deficits over the period (5.2 percent/7.4 percent, respectively). Hence, even at the one-year horizon, the projection errors are relatively large.

At the five-year horizon, the MAE is more than 4 percent of GDP and the RMSE is nearly 5.3 percent of GDP. As measured by the RMSE, the projection errors are nearly twice the actual average deficit over the period. This means that the probability is high that the error in the five-year projection would be larger than the largest single-year deficit as a percent of GDP, 6.0 percent. Furthermore, the range of projection errors is over 17 percent of GDP. This is more than double the range of the actual surplus/deficit as a percent of GDP during the 1976-99 period.

Table 3 also shows that the CBO's projections of future revenues are somewhat less accurate than its projections of outlays. Perhaps this is not surprising given the sensitivity of tax revenues to economic conditions. At the one-year horizon, the RMSE for total receipts is nearly 1.0 percent of GDP and is nearly 45 percent larger than the RMSE for total outlays. Not surprisingly, the projection errors by source are much larger at the five-year horizon. Nevertheless, it remains the case that receipts appear to be more difficult to project than outlays. The relative ability to project receipts and outlays appears to equalize somewhat as the projection horizon lengthens.

Table 3 also decomposes receipts and outlays by their major categories.[15] Individual income tax receipts are the most difficult component of receipts to project at both horizons. In terms of total outlays, discretionary spending projection errors are the smallest at the one-year horizon but are the largest at the five-year horizon.[16]

The reader is once again cautioned that these results should not be taken as criticism of the CBO per se. It is very doubtful that any other organization or government agency could do significantly better. Rather, these results point out how inherently difficult it is to project the future, especially given

[15] Caution must be used with respect to the components of total outlays because we only have a consistent breakdown of total outlays into mandatory and discretionary outlays back to FY 1983.

[16] See Auerbach (1999) for an analysis of CBO and OMB revenue forecasts.

that the no-change assumption with respect to tax and spending programs will eventually be violated.[17]

Are the CBO's Projections Biased?

The data in Table 3 show that the CBO's projections are unbiased at the one-year horizon. The average projection error is 0.05 percent of GDP, and it not significantly different from zero at the 5 percent significance level. At the five-year horizon, however, the average projection error is very large—nearly 2 percent of GDP. Moreover, the average error is significantly larger than zero at the 5 percent level. The average error is about half of the range in the actual surpluses and deficits: 1.4 percent to –6 percent of GDP. Because the CBO is projecting a surplus through 2005 that is equal to about 3.25 percent of GDP, making a correction for the average bias suggests a bias-corrected projected surplus that is only 1.25 percent of GDP.

THE SOURCES OF PROJECTION ERRORS

What accounts for the inaccuracies in the CBO's projections and why are they biased? Some insight into the answer to these questions can be obtained by looking at the possible sources of projection errors. Projection errors are typically classified into three types. The first is called *legislative* errors. As mentioned previously, budget projections are made under the assumption that current government program levels and the current tax structure will remain unchanged over the projection period. This assumption will almost certainly be false, especially at the five-year horizon. It is virtually certain that there will be legislative changes.

For example, the CBO estimated that the FY 2000 budget passed by Congress and signed into law by the President will reduce the estimated surplus by about $127 billion over the 2000-09 horizon. Much of this additional spending has been classified as "emergency" legislation. The combined "emergency" appropriations for FY 1999-2000 totaled $65.5 billion, roughly 60 percent of the previous eight years combined. But unlike the appropriations for Desert Storm, the 1993 Great Flood, Hurricane Andrew, and the 1994 Northridge (Los Angeles) earthquake, which were largely passed through traditional *supplemental* legislation, the bulk of the emergency appropriations in the FY 1999-2000 budgets were *regular* rather than supplemental. According to the CBO, regular appropriations classified as emergency spending have totaled $58.6 billion since 1990, with nearly 90 percent of that spending occurring in FY 1999-2000.[18] By contrast, supplemental emergency appropriations totaled just $14.7 billion in FY 1999-2000. In effect, then, Congress and the Administration have circumvented the discretionary budget caps that were put into place during the 1990 Budget Enforcement Act by increasingly classifying appropriations as emergency spending. Because the CBO could not have predicted this outcome, it should not be held responsible for projections that miss the mark because of budgetary chicanery. But this is exactly the point. In other words, although the budget projection process is to some extent biased against the CBO and thus potentially leads to large subsequent projection errors, legislative errors will occur because neither the CBO nor anyone else is able to predict how much of the public's money Congress and the Administration will choose to spend.

It is possible that policymakers create legislative errors by responding to budget projections. For example, policymakers could enact changes in tax rates and/or spending from current baseline levels in response to projections of a surplus or deficit. This is not far fetched: Policymakers are debating whether and by how much to increase government spending and/or reduce taxes in anticipation of the historically large projected surpluses.[19]

Furthermore, the interaction of the legislative changes and budget projections can bias the CBO's projections. For example, assume that policymakers, facing a projected surplus, decide to increase spending and/or reduce taxes. The result will be a smaller actual surplus than anticipated (or perhaps a deficit). As a consequence, the projection error (projection less actual) would be positive, i.e., the projections will be biased upward. On the other hand, if policymakers responded to projections of a deficit by reducing spending, raising taxes, or both, the average projection error would be negative.

---

[17] The CBO (2000a, p. 97) readily acknowledges "considerable uncertainty" in making its multiyear budget projections.

[18] These appropriations exclude spending associated with Desert Storm/Desert Shield operations.

[19] For example, the House of Representatives has voted to repeal the estate tax and the marriage penalty. In addition, the Congress passed, and the President signed, legislation repealing the earnings tax for senior citizens. On the spending side, there is considerable discussion of creating a new Medicare entitlement program to pay for prescription drugs for the elderly.

Does such behavior account for the positive average projection error in the CBO's projections at the five-year horizon? It seems unlikely: During most of the period, the CBO was projecting a deficit such that the bias created by the interaction of budget projections and legislative changes should have created a negative bias. Figure 5 shows, however, that there were five years (1977-81) when the CBO was projecting a surplus at the five-year horizon. The actual experience was that the government ran a deficit. Moreover, it is clear from Figure 5 that these observations account for the positive bias reported in Table 3 at the five-year horizon. Hence, it is possible that the bias occurred because the government increased spending or reduced taxes in response to projections of a budget surplus five years out. Another and perhaps more likely explanation is that, like most forecasters at the time, the CBO failed to forecast the rapid run-up in oil prices, the subsequent acceleration in inflation, and the recessions in 1980 and 1981-82.

Even if tax and spending programs are unchanged, the projections are not likely to materialize because they are tethered to the CBO's forecasts of major macroeconomic variables. The CBO reportedly does not base its economic forecast on a single model, but rather uses a series of models, ad hoc judgments, and input from private sector forecasters and economists.[20] For near-term budget projections, the CBO makes forecasts of the growth rate of real GDP, consumer price index (CPI) inflation, the unemployment rate, and the 3-month and 10-year Treasury rates. For longer-term projections (five years or more), however, the economy's underlying growth rate (which is largely determined by growth of labor productivity and population growth) plays the dominant role. These forecasts are referred to as *economic assumptions.* Accordingly, budget projection errors that arise because of erroneous economic forecasts are referred to as *economic assumption errors*. The magnitude of these errors depends on the sensitivity of the budget projections to these economic assumptions and whether assumption errors for different macroeconomic variables offset or reinforce each other.

Finally, even if there were no changes in spending and tax programs and the CBO's economic assumptions materialized exactly, projections could still be wrong because of *technical* errors. Government expenditures and tax revenues vary with their source. For example, tax revenue broadly depends on not only the growth rate of the economy but also the relative size of corporate profits, the growth rate of personal income, and the extent to which individuals turn paper profits into cash (capital gains), to mention but a few. Hence, the CBO makes forecasts not only of key macroeconomic variables but also of many variables upon which taxes and expenditures depend. Hence, even if the CBO's forecasts of key economic variables were 100 percent accurate, its budget projections could be in error because it got the details wrong.[21]

There were two important examples of this phenomenon in recent years. The first was the welfare reform legislation passed in 1996, which reduced spending on income security and food stamps. The second was the unexpected boom in equity prices which, in spite of the reductions in long-term capital gains taxes in 1998, led to a surge of individual income tax revenues.[22]

Another "detail" that may have contributed to a one-time projection error was the change in the construction of the CPI beginning in 1995. By 1999, the methodological changes in the construction of the CPI reduced the annual inflation rate by about 0.6 percentage points compared with the methodology employed to calculate the 1994 inflation rate. According to the CBO's analysis published in 1997, a reduction of 1 percentage point per year in the growth of the CPI would reduce the deficit by $653 billion over the 1998-2007 projection period.

Overall, because many of these components, such as capital gains taxes, are notoriously difficult to forecast, one would not be surprised to find that technical errors account for a large proportion of the projection errors. Of course, economic assumption errors and technical errors need not be independent. For example, if economic growth is stronger than anticipated, corporate profits or capital gains may be stronger than anticipated as well.

---

[20] See CBO (1998).

[21] Note that these are all economic variables and, to that extent, the distinction between technical errors and economic assumption errors is somewhat blurred. See Altig (2000) for a discussion of this point.

[22] Regarding the latter, see Kasten, Weiner, and Woodward (1999). The increased revenue from the surge in capital gains might be less important than the resulting change in the distribution of income—that is, if proportionately more people are shifted into a higher (marginal) tax bracket.

Table 4

Decomposition of CBO Deficit/Surplus Projection Errors: Five Years Ahead (Billions of Dollars)

	Forecast error due to:				
	Economic	Technical	Legislative	Total error	Total absolute error
1990 Baseline	–82.54	–163.38	199.70	–46.22	445.617
1991 Baseline	10.77	–151.60	89.92	–50.91	252.296
1992 Baseline	57.83	8.43	137.91	204.17	204.165
1993 Baseline	126.75	170.44	129.54	426.73	426.730
1994 Baseline	160.05	178.47	–9.68	328.84	348.199
Absolute error (sum)	437.94	672.32	566.75	862.61	1,677.01
Percentage of total absolute error	26.11	40.09	33.80		100.00

SOURCE: Congressional Budget Office, unpublished data.

The Relative Importance of Legislative, Economic, and Technical Errors

The CBO's analysis of the relative magnitude of legislative, economic, and technical errors is presented in Table 4. The CBO estimated that legislative errors account for about 35 percent of the average absolute projection error at the five-year horizon. They estimated that technical errors are the largest, accounting for over 40 percent of the average absolute projection error. Errors in forecasting the general economic outlook were the smallest at the five-year horizon, accounting for about 25 percent of the total projection error. Altig (2000) notes, however, that there is a close correspondence between technical errors and economic assumption errors because the former are the details related to the latter.

Projection errors arising from economic forecast errors have been relatively large in recent years. In 1999 and 2000, the CBO estimates that economic assumption errors accounted for 41 percent and 57 percent of the total absolute change in the projected 10-year surplus in 1999 and 2000, respectively.

What Caused the Change in the Budget Outlook?

Projections of relatively large unified budget surpluses over the next decade would have been viewed as highly implausible as recently as five years ago. Indeed, projections of large and rising federal budget deficits were the norm as late as January 1997. Figure 6 shows CBO's 10-year surplus/deficit projections for each year from 1994 through 2000, along with the revenue and outlay projections over the same periods. It was not until the federal government experienced its first surplus in 1998 that the CBO began projecting a persistent surplus. Prior to 1998, the CBO projected a persistent deficit. Not only did the projections shift from deficits to surpluses in 1998, but the slope of the projection line went from being negative to being positive (i.e., the CBO went from projecting larger deficits over time to projecting larger surpluses over time).

What accounts for this dramatic turnaround in the budget outlook? Some have suggested that the outlook changed because of the Omnibus Budget Reconciliation Act of 1993 (OBRA93).[23] This legislation ostensibly provided for $433 billion in deficit reduction over the five-year period spanning 1994-98. This was to be accomplished through a combination of tax increases, reductions in mandatory spending, and caps on discretionary spending. Altig (2000) argues that the improved budget outlook is not due to OBRA93 because similar legislation, the Omnibus Budget Reconciliation Act of 1990 (OBRA90), promised—but failed to deliver—a $482 billion reduction in the baseline deficit during the 1991-95 period. Despite OBRA90, Altig (2000) notes that the deficit increased by $6 billion relative to levels projected prior to the bill's passage.

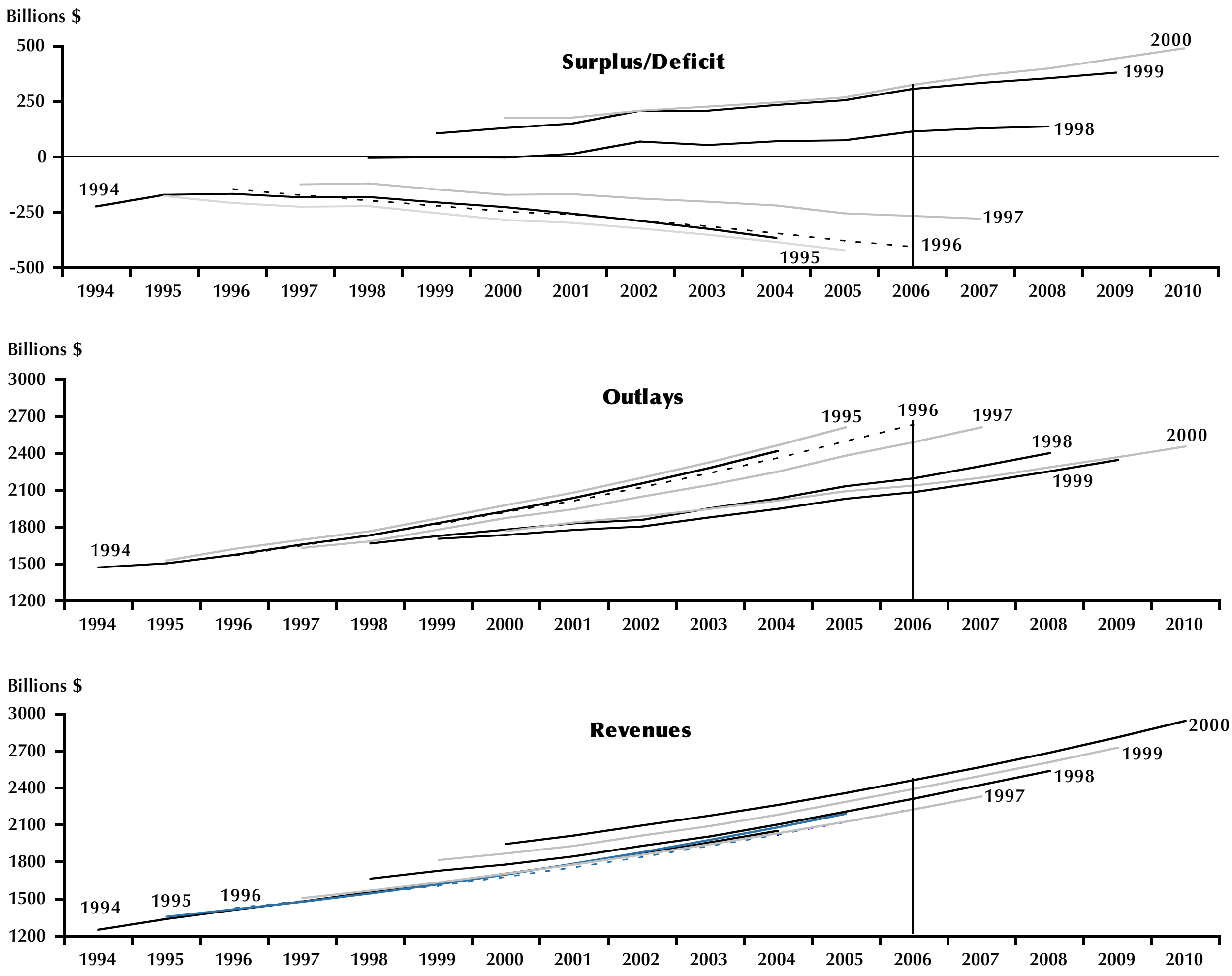
There is little question, however, that the marked improvement in the economy's performance between 1996 and 2000 has been a decisive factor in explaining the dramatic improvement in the federal government's budget.[24] Consider the analysis of the CBO's baseline projections between 1996 and 2000 for the period 2000-06 in Table 5. First, there was a

[23] The 2000 *Economic Report of the President* makes this claim. See p. 52.

[24] Altig (2000) also stresses this argument.

Figure 6

Baseline 10-Year Budget Projections, 1994-2000



NOTE: Year indicates the CBO's first-of-the-year baseline budget forecast (typically published in January or February).

marked change in the economic assumptions used by the CBO in their projections between 1996 and 2000. As other private forecasters did, the CBO boosted its forecast of actual and potential GDP growth; it also scaled back its assumption for CPI inflation and for the unemployment rate.

Faster economic growth means larger tax revenues, whereas a lower unemployment rate leads to reductions in mandatory spending for unemployment compensation, welfare outlays, and the like. As line 3 in Table 5 shows, nearly all of the improvement in the outlook for revenues stemmed from a greater than expected surge in individual income taxes—the area where the CBO's revenue projections have been the least accurate at the five-year horizon. Moreover, individual income tax revenue is projected to be higher by 1.2 percent of GDP in the 2000 projections compared with the 1996 projections.

On the outlay side, most of the improvement comes in mandatory expenditures, which are projected to be 2.3 percent lower as a percentage of GDP in 2000 compared with 1996. The other major improvement on the outlay side is net interest. This reflects the reductions in inflation (which, all else equal, reduces nominal interest rates) and the reduction in the public's holding of government debt that have already occurred and are projected to occur. Perhaps more important, the evidence in Table 5 suggests that the improved economic outlook may be largely responsible for the improved long-term outlook as well.

Table 5

Regime Shift? Change in Baseline Budget Projections and Economic Assumptions: 1996 vs. 2000 (Percent of GDP)

	Tax revenues				Outlays			
	Individual income	Corporate income	Social insurance	Total tax revenues	Discretionary	Mandatory	Net interest	Total outlays
Change in budget assumptions								
1996 Baseline projection, 1996-99 average	8.5	2.2	6.7	18.9	6.8	12.1	3.1	21.1
Actual outcome, 1996-99 average	9.2	2.2	6.6	19.5	6.6	10.9	2.9	19.4
Projection error	–0.7	0.0	0.1	–0.6	0.2	1.2	0.2	1.7
1996 Baseline projection, 2000-06 average	8.6	1.8	6.7	18.5	6.1	13.1	3.1	21.5
2000 Baseline projection, 2000-06 average	9.8	1.8	6.8	20.0	6.1	10.8	1.8	17.9
Difference: 1996 vs. 2000	1.2	0.0	0.1	1.5	0.0	–2.3	–1.3	–3.6
Memorandum (1970-95)								
Average	8.2	2.1	5.9	18.0	9.7	10.4	2.3	21.2
High	9.3 (1981)	3.2 (1970)	6.7 (1991)	19.6 (1981)	11.9 (1970)	11.9 (1983)	3.3 (1991)	23.5 (1983)
Low	7.6 (1976)	1.1 (1983)	4.4 (1970)	17.2 (1976)	7.5 (1995)	7.1 (1970)	1.3 (1973)	18.7 (1974)

	Growth of real GDP	Growth of potential GDP	CPI inflation	Unemployment rate
Change in economic assumptions				
Date of baseline projections				
May 1996 (1996 to 2006)	2.1	2.2	3.0	6.0
January 2000 (1999 to 2010)	2.9	3.1	2.5	4.8

NOTE: Averages of fiscal years indicated.

CONCLUSION

Projecting the path of federal government receipts and expenditures—and thus the unified budget surplus or deficit—several years into the future is a daunting task. This is why the CBO cautions users on the highly uncertain nature of their long-term budget projections. Moreover, budget projections provide a method of analyzing alternative budget proposals: they are not intended to be used as forecasts. Nevertheless, policymakers and the public often treat them as such. Indeed, there are many proposals to cut taxes or increase spending based on these projections. Our analysis of the CBO's record of projecting budget surpluses and deficits at the five-year horizon suggests that the public and policymakers should be wary of changing spending and tax programs based on projections of the surplus or deficit. Specifically, there is no statistically significant relationship between projections and experience. There is no reason to suspect that things are different now.

In addition, past projections were biased in the direction of under-projecting the size of the deficit. Moreover, the size of the bias is very large—nearly 2 percent of GDP. If the current projections are biased to a similar degree and policymakers choose to alter current tax and spending programs based on these projections, it is possible that the projected surpluses will never materialize.

Even if historical regularities fail to hold and legislators decide to forgo sizable future income tax reductions and expenditure increases, the budget surpluses that are currently projected to prevail over the medium term are likely to evaporate over the long term because of the retirement of the baby boom generation. The trust fund surpluses, which account for the bulk of the current projected unified budget surpluses, will not be available after about 2015. When this happens, the government will have to resume borrowing from the public, unless there are surpluses in other areas of the government. Given this reality and the likelihood that actual surpluses will be much smaller than projected, we believe that the concern expressed in some quarters—that there will not be a market for risk-free government debt—is significantly overstated. Even if the budget projections were reasonably accurate, the demise of this market likely would be only temporary.

REFERENCES

Altig, David. "Fiscal Policy and Fickle Fortunes: What's Luck Got to Do With It?" Federal Reserve Bank of Cleveland, *Economic Commentary,* 1 April 2000.

Auerbach, Alan J. "Formation of Fiscal Policy: The Experience of the Past Twenty-Five Years," in *Fiscal Policy in an Era of Surpluses: Economic and Financial Implications.* Proceedings of a Conference sponsored by the Federal Reserve Bank of New York, April 2000, pp. 9-23.

___________. "On the Performance and Use of Government Revenue Forecasts." *National Tax Journal*, December 1999, *52*(4), pp. 767-82.

Board of Governors of the Federal Reserve System. "Minutes of the Federal Open Market Committee Meeting, March 21, 2000." *Press Release,* 18 May 2000.

Cecchetti, Stephen G. "The Disappearance of U.S. Treasury Securities: Should We Care?" Occasional Essays on Current Policy Issues No. 7, 3 May 2000, <http://www.econ.ohiostate.edu/cecchetti/pdf/cpi7.pdf>.

Congress of the United States. *The Budget and Economic Outlook: Fiscal Years 2001-2010.* A Report to the Senate and House Committees on the Budget. Washington, DC: Congressional Budget Office, January 2000a.

___________. *The Budget and Economic Outlook: An Update.* A Report to the Senate and House Committees on the Budget. Washington, DC: Congressional Budget Office, July 2000b.

___________. *The Long-Term Budget Outlook*. Washington, DC: Congressional Budget Office, October 2000c.

___________. *Description of Economic Models*. CBO Paper. Washington, DC: Congressional Budget Office, November 1998.

Council of Economic Advisers. *Economic Report of the President.* Washington, DC: U.S. Government Printing Office, February 2000.

Dupont, Dominique and Sack, Brian. "The Treasury Securities Market: Overview and Recent Developments." *Federal Reserve Bulletin*, December 1999, pp. 785-806.

Federal Reserve Bank of New York. *Fiscal Policy in an Era of Surpluses: Economic and Financial Implications.* Proceedings of a Conference sponsored by the Federal Reserve Bank of New York, April 2000.

Frum, David. "The Amazing Vanishing Surplus." *New York Times,* 7 July 1999.

Gokhale, Jagadeesh; Page, Benjamin R.; Potter, Joan and Sturrock, John R. "Generational Accounts for the United States: An Update." *The American Economic Review, Papers and Proceedings,* May 2000, *90*(2), pp. 293-96.

Harris, John F. and Pianin, Eric. "Surplus Estimate is Doubled." *Washington Post,* 26 June 2000.

Kasten, Richard A.; Weiner, David J. and Woodward, G. Thomas. "What Made Receipts Boom and When Will They Go Bust?" *National Tax Journal*, September 1999, *52*(3), pp. 339-47.

Meulendyke, Ann-Marie. *U.S. Monetary Policy & Financial Markets*. Federal Reserve Bank of New York, 1998.

Meyer, Laurence H. "How Does a Surplus Affect the Formulation and Conduct of Monetary Policy?" Remarks before the 16th Annual Policy Conference of the National Association for Business Economics. Washington, DC: 23 February 2000.

National Association for Business Economics. "NABE Economic Policy Survey." *NABE News,* March/April 2000, pp. 8-9.

Office of Management and Budget. *Mid-Session Review.* Budget of the United States Government, Fiscal Year 2001. Washington, DC: U.S. Government Printing Office, 26 June 2000.

Rasche, Robert H. "Deficit Projections vs. Deficit Forecasts." Federal Reserve Bank of San Francisco *Weekly Letter,* 5 July 1985.

___________. "Financing the Government Deficit," in Warren J. Samuels and Larry L. Wade, eds., *Taxing and Spending Policy.* Lexington Books, 1980, pp. 77-83.

Reifschneider, David; Tetlow, Robert and Williams, John. "Aggregate Disturbances, Monetary Policy and the

Macroeconomy: The FRB/US Perspective." *Federal Reserve Bulletin,* January 1999, pp. 1-19.

Reischauer, Robert. "The Phantom Surplus." *New York Times*, 28 January 2000.

Stevenson, Richard W. "A Debate Over Dividing Tax Dollars Yet to Be Collected." *New York Times,* 10 August 1999.

Treasury News. "Undersecretary of the Treasury for Domestic Finance Gary Gensler Remarks at the May 2000 Treasury Quarterly Refunding." Washington, DC: Office of Public Affairs, 3 May 2000.

Weill, Sanford I. "Report to the Business Council on Economic Issues." Business Council Press Release, 10 May 2000.

Appendix

On January 31, 2001, the Congressional Budget Office released its baseline budget projections for fiscal years 2002 to 2011.[1] Consistent with the theme of recent reports, the Agency once again revised upward its projected cumulative budget surplus during the coming decade. According to the new projections, unified budget surpluses for the fiscal years 2002 to 2011 are expected to total just over $5.6 trillion. The outgoing Clinton Administration's final baseline (current services) projection is slightly less optimistic, but not by much, totaling roughly $5 trillion. These figures are substantially larger than a year earlier, when the CBO was projecting a cumulative $3.2 trillion surplus and the OMB (baseline) was projecting a $2.5 trillion surplus, and modestly greater than last year's mid-year estimates (see Table 1, page 12).

In part, the larger surpluses reflect the addition of an extra "out year" (2011) that is considerably larger than the surplus for the year that was dropped (2001). For example, the CBO's projected surplus for 2011, $889 billion, is $608 billion more than the projected 2001 surplus. The factor that is most responsible for the upward revision, however, is a more optimistic assessment of the economy's potential output growth. Both the CBO and the OMB now project that real GDP growth will average slightly more than 3 percent during the next 10 years, roughly a quarter of a percentage point more than last year's average projected growth. Faster economic growth, accordingly, is expected to boost projected revenues significantly. Comparing this year's 10-year projections for fiscal years 2001 to 2010 with last year's, the CBO estimates that cumulative revenues will be nearly 9 percent larger, while outlays are only expected to be 1.5 percent higher.

Since this article was prepared, the economy slowed appreciably during the second half of 2000, and most forecasters—including the CBO—expected its growth in 2001 to be about a percentage point below its potential growth. Since most forecasters assume the economy will bounce back quickly to a rate near or slightly above its trend growth, the effect on the cumulative surplus is expected to be minor. Moreover, the CBO ran an alternative simulation which assumes a recession in 2001 of approximately the same magnitude as the 1990-91 recession. In that scenario, the cumulative surplus for the period 2002 to 2011 is cut from $5.61 trillion to $5.48 trillion.

A more important development, which is in part a consequence of the projected surpluses, is the likelihood of expansionary fiscal policy. In early February 2001, the Bush Administration announced that it was sending to Congress a proposal that, among other items, reduces marginal tax rates, doubles the child tax credit, and reduces the marriage penalty. Though specifics of the proposal were not available as this article went to press, commentary from Administration officials suggests the total reduction in tax revenues would amount to $1.6 trillion over a 10-year horizon (the timing of the tax cuts was yet to be decided).

Generally speaking, an expansionary fiscal policy reduces tax revenues. The Bush package, however, includes reductions in marginal tax rates, which could enhance the economy's growth potential by increasing the incentives to save and invest—thereby boosting capital formation and productivity growth. Consequently, should it be adopted, the ultimate effect of the President's package on the budget is difficult to predict.

[1] Congress of the United States. *The Budget and Economic Outlook: Fiscal Years 2002-2011.* A Report to the Senate and House Committees on the Budget. Washington, DC: Congressional Budget Office, January 2001.

Are Small Rural Banks Vulnerable to Local Economic Downturns?

Andrew P. Meyer and Timothy J. Yeager

RISKINESS OF GEOGRAPHICALLY CONCENTRATED BANKS

In recent years, supervisory agencies have streamlined the bank examination process to focus attention on identified risk areas rather than the full scope of bank operations (Board of Governors, 1997). If supervisors can identify the potential risks, they can recommend preventive measures and respond more quickly to actual banking problems when they arise. Because of this shift to risk-focused supervision, off-site surveillance of banks has become much more important. Supervisory economists and staff who gather and interpret bank and economic data must direct examiners to areas of heightened risks.

Historical experience in the financial services industry demonstrates that institutions exposed to serious risk can run into trouble quickly. Many commercial banks were exposed to fluctuations in commercial real estate markets in the latter half of the 1980s. A sharp decline in real estate prices caused several hundred banks to fail (FDIC, 1997). In addition, agricultural bank failures represented a large share of the banks that failed from 1984 to 1987. These failures occurred within a few years after the peak in farmland prices, reflecting the inability of agricultural banks to absorb the losses accruing from falling farm incomes (Kliesen and Gilbert, 1996). Because these banks were not diversified in terms of geography or industry, the losses eventually overwhelmed the equity accumulated during the prosperous years. If exposure to these risks had been targeted as potential problems earlier, they might have been addressed sooner and their negative impacts dampened.

Because of the way that U.S. banking laws evolved, many U.S. banks have geographically concentrated offices and operations. Historically, national and state banking laws prevented banks from branching into other counties and states. Justification for such legislation was to promote sound and stable banking markets by limiting competitive pressures on existing banks and to prevent an excessive concentration of financial power (Spong, 2000, p. 146, and Jayaratne and Strahan, 1997). As we have noted, however, such laws left banks vulnerable to local economic downturns.

Over the last few decades, branching restrictions gradually have been lifted. By 1990, most states had granted banks permission to branch within state boundaries, and most states permitted some form of interstate banking (Berger et al., 1995, pp. 188-89). The Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 removed most remaining barriers to interstate branching. Bank holding companies (BHCs) were allowed to acquire a bank in any state and—as of June 1, 1997—merge it with an existing bank.[1] Hence, most of the legal barriers to geographical expansion have been removed.

Despite liberalized branching laws, thousands of small banks with geographically concentrated offices remain and are likely to exist for some time.[2] Some argue that small banks have a comparative advantage over large banks in small-business lending for which little public information about the borrower's creditworthiness is available. This advantage (and thus incentive to remain small) arises because small banks can originate and monitor relationship loans at a lower cost than larger banks.[3] Relationship lending requires that loan officers have autonomy to set underwriting standards and discretion to monitor and evaluate borrowers. Management at small banks can more easily monitor loan officers; consequently, small banks are better able to develop the community relationships necessary to underwrite small business loans. In addition, anecdotal evidence

Andrew P. Meyer is an economist, and Timothy J. Yeager is an economist and senior manager at the Federal Reserve Bank of St. Louis. We thank Boyd Anderson for his invaluable research assistance and Mark Vaughan, Alton Gilbert, Dusan Stojanovic, Tom King, David Wheelock, Ken Spong, Michele Gambera, Bob Rasche, Rick Sullivan, and Ron Feldman for their comments. Thomas A. Pollmann provided research assistance.

[1] Montana and Texas initially opted out of this system, but Texas reversed its decision in 1999.

[2] Such banks are more prevalent in non-western states that historically had severe intrastate branching restrictions (Gilbert, 2000).

[3] See Berger et al. (1998), Peek and Rosengren (1998), and Keeton (1995) for evidence supporting the small bank advantage hypothesis; see Jayaratne and Wolken (1999) for evidence against the hypothesis.

suggests that branch managers of large banks are rotated more often into and out of communities as they progress in their organizations; therefore, they do not develop the same personal relationships with customers that long-time community bankers develop. To contain loan origination and monitoring costs, larger banks often prefer to lend to customers for which credit information is more readily available. Consequently, community banks are likely to continue engaging in small-business lending where assessments of credit quality rely heavily on intangibles.

Two additional factors may limit community bank consolidation, at least in the short- to medium-term. First, most community banks are unit banks; larger banks seeking to expand often prefer to purchase banks with branching networks. Second, community banks that are family-owned and managed are put on the market infrequently. Hence, even some relatively unprofitable community banks could remain in the banking industry for some time.

Despite the continued presence of geographically concentrated banks, we find little evidence to support the hypothesis that such banks located in the Federal Reserve's Eighth District are particularly vulnerable to local economic downturns. Specifically, we find that county economic data are only weakly correlated with small rural bank performance. Our results provide little justification for bank supervisors to require geographically concentrated banks to take special measures to reduce their vulnerability to local economic downturns. In addition, we find little basis for systematically using county-level economic data in risk-focused supervision.

IS BANK PERFORMANCE CORRELATED WITH LOCAL ECONOMIC DATA?

In this section, we investigate the empirical relationship between the performance of geographically concentrated banks and local economic activity. Banks with geographically concentrated operations are potentially vulnerable to local economic contractions because of an inherent concentration of loan and deposit customers. Laderman, Schmidt, and Zimmerman (1991) find that a bank's location significantly influences its choice of borrowers because monitoring costs increase as the distance between lender and borrower increases. Hence, banks tend to make loans to the people and businesses that are geographically nearby. Should many firms in the area become distressed at the same time, the bank's credit quality will likely suffer more than credit quality at a bank with credit dispersed across economic markets. In addition, liquidity risk is likely to be higher at geographically concentrated banks because such banks often rely on deposits from fewer entities. In contrast, a more geographically diversified bank can attract deposits from a larger base of individual and business customers; therefore, large swings in deposits and withdrawals are less likely (Liang and Rhoades, 1988).

If local economic activity affects bank performance, this association is more likely to be evident in the data for small banks with offices in rural areas than for other banks. Smaller banks (as measured by assets) typically have more geographically concentrated operations and, due to their lower levels of capital, lend to smaller, less diversified businesses. Therefore, the performance of small banks may depend more heavily on local conditions. Bank performance also is more likely to be correlated with local economic data in rural rather than urban areas because rural banks tend to lend to a relatively high percentage of firms and residents in their own counties. If enough of those firms or residents are faring poorly, local economic data should reflect the poor performance. In contrast, banks located in metropolitan areas usually lend to a smaller fraction of all the firms and individuals in their area. Poor performance by individual small businesses and households will likely have less effect on measures of aggregate economic activity in urban areas than in rural areas.

METHODS

Data Collection

We obtained financial data on small rural banks—those with less than $300 million in assets located outside of a Metropolitan Statistical Area (MSA)—between 1990 and 1997 from the Reports of Condition and Income (call reports). We chose the $300 million size limit because it is commonly used in the banking industry to define small banks, yet the size cutoff is high enough that most banks in rural areas are included in the sample.[4] We excluded banks less than five years old

[4] Indeed, banks with assets less than $300 million file a call report that is different from the report filed by larger institutions.

Table 1

Summary Statistics: Bank Performance and Economic Variables, 1990-97

	Number	Mean	Standard deviation	Minimum	Maximum
Bank performance measures					
Adjusted ROA	6,741	1.75	0.65	–5.12	10.67
Nonperforming loans to total loans	6,741	1.23	1.63	0.00	78.35
Loan losses to total loans	6,741	0.33	0.70	–9.65	16.71
OREO to total assets	6,741	0.20	0.43	0.00	13.67
Economic variables					
County unemployment rate	6,740	7.55	2.61	2.20	22.20
State unemployment rate	6,741	5.98	1.00	3.50	8.76
County employment growth	5,781	1.30	3.92	–15.42	33.97
State employment growth	6,741	1.09	1.95	–3.59	7.44
County per capita income growth	6,034	2.41	3.09	–12.68	21.51
State per capita income growth	6,035	1.84	1.30	–0.74	4.86
County personal income growth	6,034	2.90	3.04	–12.73	23.65
State personal income growth	6,035	2.64	1.43	–0.11	6.28

NOTE: The bank sample includes Eighth District banks with less than $300 million in assets located outside an MSA, and the performance measures are standard ratios used in bank supervision to assess bank performance. County and state economic data are matched by year with the headquarters county and state of each bank in the sample. County economic data have much higher standard deviations than state economic data.

Adjusted ROA, net income plus provision expense as a percentage of total assets; *nonperforming loans,* loans past due 90 days or more plus nonaccrual loans as a percentage of total loans; *loan losses,* charge-offs minus recoveries as a percentage of total loans; and *OREO,* other real estate owned as a percentage of total assets.

because new banks have erratic ratios that could bias the results against finding a correlation of performance with economic data. We did not exclude banks involved in merger activity as long as the post-merger bank size did not exceed $300 million. Inclusion of these banks potentially biases the sample against finding correlation if the merger results in a more diverse operating market for the acquiring bank, but we deal explicitly with that issue below. Although banking data are readily available before 1990, county-level labor data were significantly revised in 1990, and a consistent time series before then is not available. Because labor data are lagging indicators of the business cycle, the sample period should pick up the effects of both the 1990-91 recession and the subsequent expansion.

To make our study directly relevant to the Federal Reserve's Eighth District risk-focused process and to limit county-level data collection to a reasonable size, we looked at the Eighth District only, which includes banks in Arkansas, eastern Missouri, southern Illinois, southern Indiana, western Kentucky, western Tennessee, and northern Mississippi.

We obtained four different pairs of county and state economic data from the Bureau of Labor Statistics and the Bureau of Economic Analysis. These data include unemployment rates, employment growth, personal income growth, and per capita personal income growth. The criteria for including a regional economic variable in the data set were that the information had to be published at least annually at both the state and county levels and enough observations had to be available in each county to make meaningful statistical inferences.[5] Because neither the county economic data nor the bank performance data were seasonally

[5] County bankruptcy data are available annually, but personal bankruptcies had little to do with economic conditions during the 1990s. Building permits are also available; however, rural counties often lack enough observations to make reliable inferences.

adjusted, we used year-end call report data and average annual economic data to smooth out monthly and quarterly fluctuations. Summary statistics for the sample banks and economic variables are reported in Table 1.

To measure bank performance, we used three asset quality ratios and one earnings ratio that are commonly used in bank supervision to monitor a bank's condition. Deterioration in asset quality and, hence, earnings are the primary reasons that small banks become distressed. The asset quality ratios chosen were nonperforming loans to total loans, net loan losses to total loans, and other real estate owned (OREO) to total assets. Nonperforming loans are loans 90 days or more past due or nonaccruing, net loan losses are charge-offs less recoveries, and OREO is the value of the temporary real estate assets on the bank's books as a result of debtor default. The earnings ratio chosen was adjusted return on assets (ROA) or net income plus provision expense, divided by assets. We added back provision expense because the asset quality measures already account for credit effects. Other things equal, we expected asset quality and earnings to improve with employment and income growth and to worsen with a rise in the unemployment rate.

Regression Model

We used ordinary least-squares (OLS) regression and tobit analysis to test whether rural bank performance depends on state and county economic variables. We used OLS for each regression involving adjusted ROA and net loan losses; however, we used a tobit procedure for each regression in which nonperforming loans or OREO was the dependent variable because these two variables are censored at zero.[6]

A simple methodology that regresses bank performance measures on local economic data suffers from omitted variable bias because several factors besides local economic conditions affect bank performance. We included bank-specific intercepts, lagged bank performance ratios, and state economic data as explanatory variables to reduce the bias. The intercept terms control for bank-specific effects on the levels of performance ratios. For example, some banks may have higher ROA than other banks because managers at the more profitable banks are more competent or take more risks than other banks' managers. Lagged bank performance ratios control for persistence in performance. For example, a bank that has high nonperforming loans one year is more likely to have high nonperforming loans the next year because it takes time for borrowers to improve their cash flows and for banks to sever their relationships with customers who are in poor financial condition.[7] Finally, research has demonstrated a strong link between bank performance and state economic data; therefore, we included state economic data in the initial regression equation to reduce omitted variable bias.[8]

We regressed the four bank performance ratios on lagged performance ratios and the four pairs of county and state economic variables between 1990 and 1997:

$$(1)\quad BP_{it} = \alpha_i + \gamma_0 \cdot BP_{i,t-1} + \gamma_1 \cdot CEcon_{it} + \gamma_2 \cdot SEcon_{it} + e_{it}.$$

In equation (1), BP_{it} represents bank i's performance at time t. The α_i coefficient is the bank-specific intercept term. The variables $CEcon_{it}$ and $SEcon_{it}$ represent, respectively, county and state economic data relevant to bank i at time t. We matched economic data with the county and state of the bank's headquarters. We tested lagged economic variables, but they contributed insignificant explanatory power to the model. We ran a separate regression for each pair of county and state economic variables to avoid multicollinearity across the different economic variables. For example, the first regression included the county and state unemployment rates of the bank's headquarters, the second regression included only county and state employment growth, and so on.

To properly assess the vulnerability of small rural banks to local economic activity, the primary focus was on the economic and statistical signifi-

---

6 Results from a straight OLS regression for nonperforming loans and OREO were quite similar to the tobit results because the non-negative constraint was nonbinding for a majority of banks.

7 Nickell (1981) has shown that, in a fixed-effects model, a lagged dependent variable coefficient will be biased unless the number of time periods is large. To a lesser extent, this bias can carry over into the other variables if they are correlated with the lagged dependent variable. This potential bias is not worrisome in our context, however, because our primary concern is with the explanatory power of each county economic variable relative to its state counterpart. In addition, we tested the effect of this bias by dropping the lagged dependent variable from the analysis. The results showing that state economic data influenced bank performance much more than county data remained unchanged.

8 See Neely and Wheelock (1997), FDIC (1997), and Laderman, Schmidt, and Zimmerman (1991).

Table 2

The "Large Change" Test as a Measure of Economic Significance

	2-Rated banks	Number of 2-ratings	3-Rated banks	Number of 3-ratings	Difference
Bank performance measures					
Adjusted ROA	1.61	636	1.12	150	0.49
Nonperforming loans to total loans	1.53	546	2.64	129	1.11
Loan losses to total loans	0.42	546	0.89	129	0.47
OREO to total assets	0.24	546	0.52	129	0.28

	Minimum	Maximum	Difference
Economic variables			
County unemployment rate	6.48	8.69	2.20
State unemployment rate	4.91	7.16	2.25
County employment growth	–0.39	3.36	3.76
State employment growth	–0.68	3.35	4.03
County personal income growth	1.10	4.65	3.55
State personal income growth	1.02	4.35	3.34
County per capita personal income growth	0.51	4.27	3.76
State per capita personal income growth	0.58	3.45	2.86

NOTE: The "large change" test measures the percentage of a large change in the bank performance variable explained by a large change in the economic variable. The ratio is computed as $\chi = \gamma\Delta X/\Delta Y$, where γ is the regression coefficient, ΔX is a large change in the economic variable, and ΔY is a large change in the bank performance ratio. A large change in the bank performance ratio is measured by calculating the average difference between a CAMELS 2-rated bank and a CAMELS 3-rated bank in the sample. For example, the average 2- and 3-rated banks had nonperforming loan-to–total loan ratios of 1.53 percent and 2.64 percent, respectively. The difference of 1.11 percent is a "large" change. The benchmark for a "large" change in the economic variable is the maximum change in the average annual values over the sample period. For example, the minimum average county unemployment rate between 1990 and 1997 was 6.48 percent, and the maximum county unemployment rate was 8.69 percent. The difference of 2.2 percentage points is considered a "large" change.

Adjusted ROA, net income plus provision expense as a percentage of total assets; *nonperforming loans*, loans past due 90 days or more plus nonaccrual loans as a percentage of total loans; *loan losses,* charge-offs minus recoveries as a percentage of total loans; and *OREO,* other real estate owned as a percentage of total assets.

cance of the county coefficient, γ_1. A significant coefficient indicated that economic changes at the county level had statistically important effects on small rural bank performance. We placed little emphasis on the overall model fit (R^2) because we were not attempting to fully explain bank performance.

Economic Significance Benchmarks

Economic significance is more difficult to assess than statistical significance. One cannot simply use the size of the regression coefficients to judge economic significance because of different variances among the regression variables. For example, state unemployment rates are typically less variable than county unemployment rates. An increase of 1 percentage point in the state unemployment rate represents a lower probability event than a 1 percentage point change in a county unemployment rate. To control for the differing variances of economic and bank data, we normalized the regression coefficients by calculating a "beta" coefficient, $\beta = \gamma\sigma_x/\sigma_y$, where γ is the regression coefficient, σ_x is the standard deviation of the independent variable, and σ_y is the standard deviation of the dependent variable. The ratio β—reported in the tables as a percentage—measures the effect of a one standard deviation increase in the economic variable on the bank performance variable relative to a one standard deviation change in the bank performance variable. Because previous research demonstrated a link between bank performance measures and state economic data, we used the size of the β coefficients for the state economic variables as benchmarks for the economic significance of the county β coefficients.

Although a β coefficient provides a measure of the relative importance of an independent vari-

Table 3

Correlation Between Small Rural Bank Performance and County and State Economic Data

Economic variable:	Unemployment rate				Employment growth				Personal income growth			
Dependent variable:	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO
Lagged dependent variable	0.251***	0.367***	0.099***	0.408***	0.188***	0.365***	0.118***	0.390***	0.198***	0.360***	0.105***	0.404***
	(0.012)	(0.011)	(0.012)	(0.010)	(0.013)	(0.012)	(0.012)	(0.011)	(0.013)	(0.012)	(0.012)	(0.011)
County economic variable	0.019***	0.000	–0.007	–0.007	–0.001	0.000	0.003	0.004**	–0.008***	0.011*	0.008***	0.006***
	(0.005)	(0.013)	(0.007)	(0.005)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)	(0.006)	(0.003)	(0.002)
State economic variable	–0.035***	0.175***	0.087***	0.116***	0.002	–0.074***	–0.036***	–0.031***	0.062***	–0.086***	–0.032***	–0.047***
	(0.007)	(0.021)	(0.011)	(0.008)	(0.003)	(0.010)	(0.005)	(0.004)	(0.004)	(0.012)	(0.006)	(0.004)
N	6,717	6,717	6,717	6,717	5,758	5,758	5,758	5,758	6,019	6,019	6,019	6,019
County economic significance												
β	6.6	0.0	–1.7	–3.1	–0.8	0.0	1.8	4.6	–5.2	2.5	3.9	5.2
χ	–8.5	0.0	–3.3	–5.4	–0.8	0.0	–2.4	–5.3	–5.8	–3.5	–6.0	–7.4
State economic significance												
β	–7.4	12.5	12.9	31.8	1.1	–10.8	–11.8	–18.5	19.7	–9.1	–6.9	–19.2
χ	16.1	35.6	41.6	91.2	1.6	27.0	30.8	43.7	42.2	26.0	22.7	54.8

Standard errors in parentheses; */**/*** significant at the .10/.05/.01 level.

NOTE: If banks with geographically concentrated operations are heavily influenced by local economic conditions, then the coefficients on the county economic variables should be statistically and economically significant even after controlling for the influences of state economic variables. The results show that county economic variables are neither statistically nor economically significant while state economic variables are almost always statistically significant and often economically significant. Absolute values of beta coefficients (β) for the county economic variables are much smaller than those for state economic variables, and values from the "large change" test—the percentage of large changes in bank performance ratios explained by large changes in county economic variables (χ)—are consistently small and negative, suggesting that small rural banks are not vulnerable to changes in local economic activity.

Adjusted ROA, net income plus provision expense as a percentage of total assets; *nonperforming loans,* loans past due 90 days or more plus nonaccrual loans as a percentage of total loans; *loan losses,* charge-offs minus recoveries as a percentage of total loans; and *OREO,* other real estate owned as a percentage of total assets. Because nonperforming loans and OREO can never be negative, we used a tobit censored regression for these dependent variables.

able, it does not provide an adequate measure of the overall importance of a given variable for explaining movements in bank performance ratios. For example, suppose that we regress nonperforming loans on county and state unemployment rates. The regression coefficients are 7 basis points and 17 basis points, respectively, and the β coefficients are 4 percent and 10 percent. The state unemployment rate clearly dominates the county unemployment rate using β coefficients to judge economic significance, but whether the 7 or 17 basis point coefficient values are economically large remains unclear. The answer depends on the definition of "large" changes in both nonperforming loans and unemployment rates.

We constructed a "large-change" test to determine whether a regression coefficient was economically significant. This test measures the percentage of a large change in the bank performance ratio explained by a large change in the economic variable. Bank examination ratings guided our assessments of large changes in bank performance ratios. Each time a bank is examined, regulators assign a composite rating and an individual rating to each of the CAMELS components. (CAMELS is an acronym that stands for Capital adequacy, Asset quality, Management, Earnings, Liquidity, and Sensitivity [to interest rate risk]). CAMELS ratings range from 1 (the safest banks) to 5 (the riskiest banks). Banks with composite ratings of 1 and 2 are considered to exhibit "strong" and "satisfactory" performances, respectively. Banks that fall below a 2 rating may prompt supervisory action, which could include a board resolution, a memorandum of understanding, a written agreement, or a cease and desist order. Hence, regulators consider a drop from a 2 rating to a 3 rating to be a significant change.

The differences in average bank performance ratios between 2- and 3-rated banks serve as our benchmarks for evaluating economic significance.[9] We used only bank performance ratios at the time of the bank examination instead of using all performance ratios for 2- and 3-rated banks to avoid endogeneity issues that might arise if supervisors required 3-rated banks to improve performance. Inclusion of all the ratios would potentially decrease the differences between 2- and 3-rated banks. Table 2 displays average measures of performance ratios for banks rated 2 and 3 between 1990 and 1997. For example, the average bank with a 2-rated earnings component had an adjusted ROA of 1.61 percent; the average bank with a 3-rated earnings component had an adjusted ROA of 1.12 percent. We consider the 49 basis point spread between the two sets of banks to be economically large.

The last step in measuring economic significance using the large-change test was to identify economically important changes in the economic variables. We identified large changes as the typical changes over the course of a business cycle. The 1990 to 1997 sample period includes the 1990-91 recession and the subsequent recovery and expansion. To identify large changes in the economic data over the business cycle, we calculated the yearly means of each county and state economic variable and computed the differences between the maximum and minimum values. For example, the maximum annual average county unemployment rate was 8.7 percent in 1991, and the minimum value was 6.5 percent in 1997, a difference of 2.2 percentage points. Large changes in economic variables are also displayed in Table 2.

Somewhat arbitrarily, we interpreted a regression coefficient as economically significant using the large-change test if the maximum change in the economic variable over the business cycle accounted for over one third of the difference in the average bank performance ratio between a 2- and 3-rated bank. The large-change economic significance ratio was calculated as $\chi = \gamma\Delta X/\Delta Y$, where γ is the regression coefficient and ΔX and ΔY are economically large changes in the independent and dependent variables, respectively. For example, suppose that a 1 percentage point change in the county unemployment rate increases the nonperforming loan-to–total loan ratio by 7 basis points. If ΔX is 2.2 percentage points and ΔY is 111 basis points, then $\chi = 13.9$ percent ($7 \times 2.2 / 111$) of the change in the bank performance ratio. We would conclude that the county unemployment rate coefficient is economically insignificant. Values of χ with negative signs imply that the regression coefficients had the theoretically unexpected signs.

RESULTS

Regression Results

The regression results indicate that county

[9] We also computed median differences in bank performance measures, but the differences between mean and median ratios were small.

Table 4

Correlation Between Small Rural Bank Performance and County Economic Data

Economic variable:	Unemployment rate				Employment growth				Personal income growth			
Dependent variable:	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO
Lagged dependent variable	0.259***	0.382***	0.111***	0.445***	0.189***	0.372***	0.124***	0.398***	0.212***	0.368***	0.109***	0.420***
	(0.012)	(0.011)	(0.011)	(0.010)	(0.013)	(0.012)	(0.012)	(0.011)	(0.013)	(0.012)	(0.012)	(0.011)
County economic variable	0.005	0.068***	0.027***	0.037***	–0.000	–0.018***	–0.006***	–0.004**	0.000	0.000	0.004	0.000
	(0.003)	(0.010)	(0.005)	(0.004)	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)	(0.002)
N	6,717	6,717	6,717	6,717	5,758	5,758	5,758	5,758	6,019	6,019	6,019	6,019
County economic significance												
β	1.8	7.8	6.4	16.2	–0.3	–5.1	–3.8	–4.6	0.3	0.0	2.0	0.0
χ	–2.2	13.5	12.6	28.5	0.0	6.1	4.8	5.3	0.0	0.0	–3.0	0.0

Standard errors in parentheses; */**/*** significant at the .10/.05/.01 level.

NOTE: This table reports the results from regressing small rural bank performance ratios on county economic variables, omitting state-level variables. Multicollinearity between state and county economic data may bias downward the correlation between county economic data and bank performance ratios. The results show that county economic data are sometimes significantly correlated with asset quality measures, but the economic significance is weak. For example, a large change in the county unemployment rate explains just 13.5 percent of a large change in nonperforming loans. In contrast, Table 3 shows that a large change in the state unemployment rate explains 35.6 percent of a large change in nonperforming loans. The implication is that multicollinearity between county and state economic data is not accounting for the weak correlation between county economic data and bank performance measures. The results suggest that small rural banks are not vulnerable to local economic downturns.

Adjusted ROA, net income plus provision expense as a percentage of total assets; *nonperforming loans,* loans past due 90 days or more plus nonaccrual loans as a percentage of total loans; *loan losses,* charge-offs minus recoveries as a percentage of total loans; and *OREO,* other real estate owned as a percentage of total assets. Because nonperforming loans and OREO can never be negative, we used a tobit censored regression for these dependent variables.

economic data are weakly correlated with bank performance measures. Results are reported in Table 3 along with both measures of economic significance (β and χ).[10] None of the county-level coefficients was both statistically significant and had the theoretically expected sign, suggesting that county economic activity did not have an important influence on bank performance. In contrast, state economic data were highly correlated with bank performance measures. In these regressions, 15 out of 16 state-level coefficients were statistically significant at the 1 percent level and had the expected sign. The β coefficients for state variables were typically several times higher than the county β coefficients. In addition, the large-change measure of economic significance showed that the state unemployment rate accounted for economically significant changes in nonperforming loans, loan losses, and OREO. State personal income growth rates were also economically significant, accounting for 42.2 percent of a large change in adjusted ROA.

Collinearity between each pair of county and state economic data may have reduced the significance of the county-level coefficients. Multicollinearity is likely because county data were derived explicitly from state data, and correlation coefficients between the county and state data were high. Although collinearity among explanatory variables does not bias OLS or tobit estimates, the state economic coefficient might have reduced the statistical importance of the county economic coefficient. To account for this potential effect, we reran equation (1) removing state economic variables from the equation. Results are reported in Table 4.

When state-level data were excluded from the regressions, several county-level economic variables were statistically significant. In particular, we found that changes in county unemployment rates and employment growth affect bank asset quality. Personal income growth, however, had no statistically significant effect on earnings or asset quality.

Although the county unemployment rate was statistically significant when regressed against asset quality ratios, its economic significance was low. The β coefficient from a change in the county unemployment rate was just 7.8 percent for nonperforming loans, about half the economic significance of the state unemployment rate coefficient reported in Table 3. Similar results held for loan losses and OREO. The large-change test confirmed the lack of economic significance.

Because the correlation between county economic data and bank performance was not economically significant, the implication is that small rural banks with concentrated operations are not particularly vulnerable to local economic downturns, as currently measured by county-level economic data. The results also suggest that economic data gathered at the county level are not useful to systematically assess community bank performance; however, state economic data may be useful for such a purpose.[11]

Sensitivity of Results to Sample Selection

The poor correlation between bank performance and county-level economic data could exist because the bank sample included a significant number of banks that were not sensitive to local economic conditions for one reason or another. By removing those banks from the sample, we may be able to uncover a significant relationship. If a significant correlation is found for a smaller subset of banks, then supervisors might wish to focus their attention on this set of banks.

Perhaps some banks had a significant portion of their branches, and hence loan activity, in other counties; therefore, including such banks in the sample weakened the correlation of bank performance with local economic performance. We matched each rural bank in the full sample with the county economic data in which the bank's headquarters resided, even though the bank may have had significant operations in other counties.

To reduce the bias from including banks with significant operations outside of their headquarters county, we re-ran the regressions including only those banks that had 100 percent of their deposits in the headquarters county of the community bank. At the end of each year, the Federal Deposit Insurance Corporation (FDIC) and Office of Thrift Supervision publish the *Summary of Deposits,* which contains branch deposit informa-

---

[10] Because the results for personal income and per capita personal income were similar, we report only the personal income results.

[11] Besides state data, we experimented with county, contiguous county, nearest MSA, and county data aggregated over only the Eighth District portion of each state. Although coefficients on economic data from the nearest MSA and the Eighth District portions of states are similar to the state-level results, the state economic data are preferred because they are available with a shorter time lag and are subject to less measurement error than the more disaggregated data.

Table 5

Correlation Between County Economic Data and Performance of Banks with All Deposits in a Single County

Economic variable:	**Unemployment rate**				**Employment growth**				**Personal income growth**			
Dependent variable:	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO
Lagged dependent variable	0.271***	0.364***	0.096***	0.421***	0.206***	0.352***	0.111***	0.378***	0.225***	0.349***	0.095***	0.393***
	(0.013)	(0.012)	(0.012)	(0.011)	(0.014)	(0.013)	(0.013)	(0.012)	(0.014)	(0.013)	(0.013)	(0.012)
County economic variable	0.006	0.069***	0.029***	0.038***	0.000	–0.019***	–0.007***	–0.004**	0.002	0.000	0.005*	0.000
	(0.004)	(0.011)	(0.006)	(0.005)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)	(0.006)	(0.003)	(0.002)
N	5,758	5,758	5,758	5,758	4,869	4,869	4,869	4,869	5,223	5,223	5,223	5,223
County economic significance												
β	2.0	7.5	6.5	16.0	0.3	–5.1	–4.2	–4.4	1.5	0.0	2.4	0.0
χ	–2.7	13.7	13.6	29.2	0.0	6.5	5.6	5.3	–0.7	0.0	–3.8	0.0

Standard errors in parentheses; */**/*** significant at the .10/.05/.01 level.

NOTE: This table reports the results of regressing bank performance ratios on county economic data, but the bank sample includes only those banks that have all deposits in a single county. This sample reduces the potential bias from matching a bank's performance with just one county's economic data when the bank may be operating in several counties. About half of the county economic coefficients are statistically significant, but all of the coefficients are economically insignificant as measured by beta coefficients (β) and the "large change" test (χ). The results suggest that even banks with deposits in a single county are not vulnerable to local economic downturns.

Adjusted ROA, net income plus provision expense as a percentage of total assets; *nonperforming loans,* loans past due 90 days or more plus nonaccrual loans as a percentage of total loans; *loan losses,* charge-offs minus recoveries as a percentage of total loans; and *OREO,* other real estate owned as a percentage of total assets. Because nonperforming loans and OREO can never be negative, we used a tobit censored regression for these dependent variables.

tion for FDIC-insured institutions as of June 30 of that year. Bank deposits inside and outside of a particular county can be computed from the *Summary of Deposits* data. Of the 6,717 observations in the full sample, 5,758 of them (86 percent) had all deposits in one county. Results are presented in Table 5.

The results from including only banks with all their deposits in the headquarters county were nearly identical to the full sample results. None of the 16 county economic coefficients was economically significant. Banks with all deposits within a single county are no more vulnerable to local economic downturns than banks with deposits in multiple counties.

Another possibility is that many of the community banks in our sample are controlled by larger bank holding companies. Community banks controlled by large BHCs may behave differently from the more isolated banks because they may have a higher share of syndicated loans and other interaction with entities outside their communities. To ensure that none of the banks in the sample was controlled by a large BHC, we included only banks in our sample in which the consolidated assets of the BHC were no more than twice the assets of the bank. For example, a rural bank with $50 million in assets owned by a BHC with more than $100 million in consolidated assets would have been excluded from the sample. Most banks in the Eighth District were controlled by BHCs smaller than this cutoff because the sample size fell from 6,717 to 4,837, a decline of only 28 percent. The results are reported in Table 6.

The statistical and economic significance of the sample with large BHCs excluded was similar to the full sample results. The economic significance was small as measured by the β coefficients and the large-change tests. The implication is that even the potentially most vulnerable banks—small rural banks not part of relatively large holding companies—are not strongly influenced by changes in available measures of county level economic activity.

An interesting question is why performance at rural banks is not significantly correlated with county economic data. Perhaps such banks were vulnerable to local economic conditions in the past, but changes in intrastate and interstate branching and acquisition laws, advances in transportation and communication technologies, and continued integration of domestic markets have reduced or eliminated this vulnerability. Indeed, Petersen and Rajan (2000) found that community banks have increased their lending to more distant borrowers over time. They found that the distance between small firms and lenders grew from an average of 51 miles in the 1970s to 161 miles in the 1990s. They attributed most of the gain to improvements in gathering and analyzing information. Banks reduced the importance of person-to-person contact by increasingly relying on financial statements and credit reports to evaluate potential borrowers. In addition, Gunther and Robinson (1999) found that banks faced less risk from variations in regional economic performance in 1996 than in 1985 due to industry diversification at the state level and geographical diversification by banks. Although it is theoretically possible to test the relationship between bank performance and county economic activity during a time period before branching restrictions were widely relaxed, county-level data limitations reduce the reliability of these estimates.

Regardless of whether banks were systematically vulnerable to local economic conditions in the past, the result that small rural bank performance is only weakly correlated with county economic data suggests that geographic concentration of a bank's offices may not be a significant risk factor for these banks today. The results also suggest that further intrastate branching in Eighth District states will not significantly affect a bank's vulnerability to local economic downturns because this risk factor is already low.[12]

Are Noisy Data to Blame?

One caveat in drawing implications from the lack of correlation between bank performance measures and county economic data is that the county data may not be of sufficiently high quality (see boxed insert). Perhaps rural banks are affected by county economic conditions but county labor and income data are so noisy that the statistical correlation is masked. Regardless of whether banks are insensitive to local economic downturns

[12] Because rural bank performance is significantly correlated with state-level income and labor measures, bank expansion across state lines is more likely to reduce the bank's risk. If interstate loan customers are influenced by a different set of economic events, then diversification benefits will materialize.

Table 6

Correlation Between County Economic Data and Performance of Banks Not Owned by Large Holding Companies

Economic variable:	Unemployment rate				Employment growth				Personal income growth			
Dependent variable:	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO	Adjusted ROA	Non-performing loans	Loan losses	OREO
Lagged dependent variable	0.269***	0.367***	0.059***	0.500***	0.185***	0.346***	0.062***	0.472***	0.230***	0.376***	0.053***	0.490***
	(0.014)	(0.008)	(0.012)	(0.016)	(0.015)	(0.008)	(0.014)	(0.019)	(0.015)	(0.009)	(0.013)	(0.017)
County economic variable	0.009**	0.070***	0.031***	0.028***	–0.002	–0.012***	–0.007***	–0.003*	–0.001	0.001	0.002	0.002
	(0.004)	(0.008)	(0.005)	(0.004)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.002)
N	4,837	4,837	4,837	4,837	4,126	4,126	4,126	4,126	4,340	4,340	4,340	4,340
County economic significance												
β	3.2	10.5	8.2	15.0	–1.5	–4.7	–4.5	–4.2	–0.4	0.3	1.1	2.2
χ	–4.0	13.9	14.5	21.5	–1.5	4.1	5.6	3.9	–0.7	–0.3	–1.5	–2.5

Standard errors in parentheses; */**/*** significant at the .10/.05/.01 level.

NOTE: This table reports the results from regressing bank performance ratios on county economic data, but the bank sample includes only those banks that are owned by a BHC with consolidated assets of no more than twice the bank's assets. This sample eliminates the potential bias from including banks that might operate differently from other small rural banks due to their holding company affiliations. About half of the county economic coefficients are statistically significant, but all of the coefficients are economically insignificant as measured by beta coefficients (β) and the "large change" test (χ). The results suggest that even banks without large holding company affiliations are not vulnerable to local economic downturns.

Adjusted ROA, net income plus provision expense as a percentage of total assets; *nonperforming loans,* loans past due 90 days or more plus nonaccrual loans as a percentage of total loans; *loan losses,* charge-offs minus recoveries as a percentage of total loans; and *OREO,* other real estate owned as a percentage of total assets. Because nonperforming loans and OREO can never be negative, we used a tobit censored regression for these dependent variables.

HOW RELIABLE ARE COUNTY-LEVEL ECONOMIC DATA?

To assess the reliability of county economic data, one must understand how the data are compiled. Because it is too costly to survey enough people directly, the Bureau of Labor Statistics (BLS) computes county unemployment rates by disaggregating state unemployment rates. When constructing local area unemployment statistics, the agency must estimate the number of people unemployed. The BLS constructs unemployment estimates by utilizing data on the number of people currently receiving unemployment insurance benefits in a county. A problem with this procedure, of course, is that insurance benefits may end before the worker finds a new job. To correct for this bias, the BLS uses historical trends to estimate the number of people who have exhausted their benefits. They also estimate the number of unemployed new entrants and reentrants into the labor force who are not eligible or have not filed for unemployment insurance. To be consistent with the more reliable state figures, the county unemployment estimates are adjusted to sum to the state unemployment rate. Clearly, several assumptions must be made to derive county labor estimates, increasing the potential for error.

County-level personal income data are also constructed from state personal income estimates, which in turn are constructed from national estimates.[1] Personal income of an area is defined as the income received by all the residents of the area. Because most of the source data are reported by the "place of work," the data must be adjusted to a "place of residence" basis. Therefore, the adjustments require making assumptions about cross-county commuting patterns. Data on intercounty commuting are available only every 10 years; interpolation must be used for other years by estimating changes in commuting patterns. Because commuting patterns are likely to change significantly over a 10-year period, measurement error is surely introduced and increases over the decade.

Source data used to construct local labor and income estimates are primarily collected for other purposes; therefore, the additional assumptions required to convert these data into county-level economic data increase the potential for error. However, as long as assumptions remain relatively consistent from year to year, then the *change* in county labor and income data should be relatively accurate. Therefore, it is not obvious that county economic data are so noisy that they are disguising the true correlation between small rural banks and local economic activity.

[1] Bureau of Economic Analysis (1994).

or the data are noisy, county economic data do not add systematic value to the risk-focused process. If the correlation is masked by noisy data, however, banks may indeed be vulnerable to local economic contractions.[13]

CONCLUSIONS

One aspect of risk-focused supervision involves identifying potential risk areas so that the risks can be addressed before a crisis erupts. Geographical concentration of bank offices and, hence, operations potentially leaves such banks vulnerable to local economic downturns. If this risk factor is important, then, given economic data of sufficient quality, bank performance of geographically concentrated banks should be significantly correlated with local economic data. Statistical analysis of small rural banks in the Federal Reserve's Eighth District suggests that geographical concentration is not a significant risk factor, that bank performance is not significantly correlated with county economic data, and that small rural banks in the Eighth District are not particularly vulnerable to local economic downturns.

Two policy implications arise from these results. First, a priori, little justification exists for imposing more stringent regulatory requirements on banks with geographically concentrated offices than on other banks. For example, all else equal, higher capital standards on geographically concentrated banks are not warranted. Second, coun-

[13] Although we cannot discern whether our results are driven by noisy county economic data or that bank performance truly is independent of local economic conditions, ongoing research supports the latter hypothesis. Meyer and Yeager (2000) find that small Eighth District banks with all deposits in a single county bear no more risk than similar banks with deposits across counties. In other words, intrastate diversification fails to reduce bank risk, which is consistent with the hypothesis that small rural banks are not vulnerable to local economic downturns.

ty economic data are not systematically useful in the risk-focused supervision process because the data are weakly correlated with bank performance.

REFERENCES

Berger, Allen N.; Kashyap, Anil K. and Scalise, Joseph M. "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been." *Brookings Papers on Economic Activity,* 1995, (2), pp. 55-218.

__________; Saunders, Anthony; Scalise, Joseph M. and Udell, Gregory F. "The Effects of Bank Mergers and Acquisitions on Small Business Lending." *Journal of Financial Economics,* November 1998, *50*(2), pp. 187-229.

Board of Governors. "Risk-Focused Framework for the Supervision of Community Banks." SR 97-25. Division of Banking Supervision and Regulation, 1 October 1997.

Bureau of Economic Analysis. "Local Area Personal Income, 1969-92." U.S. Department of Commerce, September 1994, pp. 1-56.

Federal Deposit Insurance Corporation. "An Examination of the Banking Crises of the 1980s and Early 1990s." *History of the Eighties: Lessons for the Future.* Vol. 1. 1997.

Gilbert, R. Alton. "Nationwide Branch Banking and the Presence of Large Banks in Rural Areas." Federal Reserve Bank of St. Louis *Review,* May/June 2000, *82*(3), pp. 13-28.

Gunther, Jeffery W. and Robinson, Kenneth J. "Industry Mix and Lending Environment Variability: What Does the Average Bank Face?" Federal Reserve Bank of Dallas *Economic and Financial Review,* Second Quarter 1999, pp. 24-31.

Jayaratne, Jith and Strahan, Philip E. "The Benefits of Branching Deregulation." Federal Reserve Bank of New York *Economic Policy Review*, December 1997, *3*(4), pp. 13-29.

__________ and Wolken, John. "How Important Are Small Banks to Small Business Lending? New Evidence from a Survey of Small Firms." *Journal of Banking and Finance,* February 1999, *23*(2-4), pp. 427-58.

Keeton, William R. "Multi-Office Bank Lending to Small Businesses: Some New Evidence." Federal Reserve Bank of Kansas City *Economic Review,* Second Quarter 1995, *80,* pp. 45-57.

Kliesen, Kevin L. and Gilbert, R. Alton. "Are Some Agricultural Banks Too Agricultural?" Federal Reserve Bank of St. Louis *Review,* January/February 1996, *78*(1), pp. 23-35.

Laderman, Elizabeth S.; Schmidt, Ronald H. and Zimmerman, Gary C. "Location, Branching, and Bank Portfolio Diversification: The Case of Agricultural Lending." Federal Reserve Bank of San Francisco *Economic Review,* Winter 1991, pp. 24-38.

Liang, Nellie and Rhoades, Stephen A. "Geographic Diversification and Risk in Banking." *Journal of Economics and Business,* November 1988, *40*(4), pp. 271-84.

Meyer, Andrew P. and Yeager, Timothy J. "Do Geographically Diversified Banks Operate With Lower Risk?" Supervisory Policy Analysis Working Paper, December 2000.

Neely, Michelle Clark and Wheelock, David C. "Why Does Bank Performance Vary Across States?" Federal Reserve Bank of St. Louis *Review,* March/April 1997, *79*(2), pp. 27-40.

Nickell, Stephen J. "Biases in Dynamic Models with Fixed Effects." *Econometrica,* November 1981, *49*(6), pp. 1417-26.

Peek, Joe, and Rosengren, Eric S. "The Evolution of Bank Lending to Small Business." Federal Reserve Bank of Boston *New England Economic Review,* March/April 1998, pp. 27-36.

Petersen, Mitchell A. and Rajan, Raghuram G. "Does Distance Still Matter? The Information Revolution in Small Business Lending." National Bureau of Economic Research, Working Paper 7685, May 2000.

Spong, Kenneth. *Banking Regulation.* 5th Ed. Federal Reserve Bank of Kansas City, 2000.

Zimmerman, Gary C. "Factors Influencing Community Bank Performance in California." Federal Reserve Bank of San Francisco *Economic Review,* 1996, (1), pp. 26-42.

Forecasting Recessions: Can We Do Better on MARS?

Peter Sephton

Macroeconomists spend much of their time developing theories and building models to demonstrate how shocks propagate and affect the overall level of economic activity. Both policymakers and the private sector maintain a keen interest in understanding the state of business affairs and the most likely path the economy will take over a planning horizon. Although there are a number of economic events that concern the authorities—including excessive inflation and unemployment—considerable attention is paid to the forecasting of recession. If policymakers can anticipate a recession, they take preemptive corrective action. The private sector uses this information to shelter itself from the vagaries of the business cycle and the most likely reaction of policymakers.

Recently a number of studies have examined the ability of financial variables to forecast recessions. Many analysts find that financial indicators contain information that can be used to increase forecast accuracy. Estrella and Mishkin (1998) found that the slope of the yield curve helped predict recessions beyond one quarter. Haubrich and Dombrosky (1996), Bernard and Gerlach (1996), Dueker (1997), and Atta-Mensah and Tkacz (1998) reported similar results.[1]

Many of these studies employed probit models to estimate the probability of recession. Probit models are sometimes used when economists model the behavior of a dependent variable which takes on two values, e.g., recession = 1, no recession = 0. The traditional approach to probit modeling requires the researcher to choose the variables that will be included in the equation, determine their level of interaction, and assume each variable plays the same role across all recessions in the sample period. These assumptions imply that the causal nature of recessions remains fixed over time, which we know to be at odds with the stylized facts of American business cycles in the twentieth century.[2] Consequently, probit models may not adequately capture the underlying processes related to recession.

The purpose of this paper is to revisit the information contained in financial variables using nonlinear, nonparametric methods, in particular, multivariate adaptive regression splines (MARS).[3] As with the probit specification, MARS models provide probability forecasts that lie between zero and one, yet they admit a much wider range of possible relationships in the data. The MARS approach allows the series to enter both individually and in combination. Given the idiosyncrasies of the American business cycle, this nonlinear, nonparametric approach may provide greater insight into the factors contributing to recession while avoiding some of the pitfalls associated with the probit specification.

MODELING WITH MARS

Data

The National Bureau of Economic Research (NBER) has identified six recessions from January 1960 through September 1999. The dates of these recessions are indicated in the list below. A dichotomous dependent variable that is equal to one if the economy is in recession and equal to zero otherwise will be used as the dependent variable to be forecast.

- April 1960 – February 1961
- December 1969 – November 1970
- November 1973 – March 1975
- January 1980 – July 1980
- July 1981 – November 1982
- July 1990 – March 1991

Recession dates are available at the NBER Web site at http://www.nber.org.

Peter Sephton is a professor of economics at the University of New Brunswick. This paper was written while visiting the Federal Reserve Bank of St. Louis. The author thanks Bob Rasche, Dan Thornton, Richard Anderson, and Michael Dueker for their helpful comments and the Research Department for their hospitality. Rachel Mandal provided research assistance.

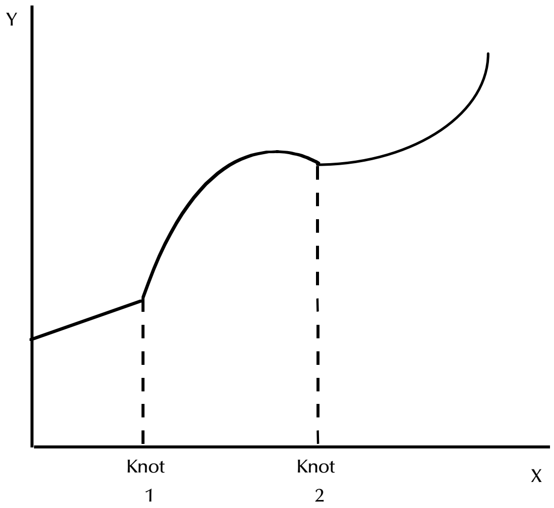
[1] Friedman and Kuttner (1998) report that the closely related paper-bill spread fared less well at predicting the 1990-91 recession. They argue that relative supply conditions in the commercial paper and Treasury Bill markets led to this result. It is worth remembering that although spreads and yield curves contain information on monetary policy, they are a function of returns on assets which are not always perfect substitutes.

[2] See Temin (1998) for an economic historiography of American recessions since 1890.

[3] The professional release of MARS is available from Salford Systems at <http://www.salford-systems.com>.

Figure 1

An Example of Thresholds



A wide variety of financial and real variables have been used as predictors of recession and output growth. The choice of which variables to include depends on whether the analysis is undertaken on monthly or quarterly data. Here the data frequency is monthly, and we employ six variables. The slope of the yield curve (measured by the difference between the 10-year constant maturity Treasury bond rate and the rate on 3-month Treasury bills [secondary market]) has been most prominent in previous studies. Changes in real factors will be captured by the change in the logarithm of the index of industrial production as well as the change in the civilian unemployment rate. Recessions are, after all, persistent declines in real output; thus, past changes in industrial production and the unemployment rate are natural candidates for use as predictors of recessions. The change in the logarithm of the S&P 500 Index has been shown to contain predictive content by Estrella and Mishkin (1998) and Dueker (1997), as have changes in the logarithm of real money, defined to be M2 deflated by the consumer price index. The change in the federal funds rate is also included in the model. These last three variables might capture the effects of both expected and unexpected monetary policy. All series are similar to those examined by others in the literature.

Nonlinear, Nonparametric Methods

The basic problem facing any forecaster is to determine the fundamental relationship between a dependent variable, Y, and a vector of predictors, expressed by X. The question is how best to specify the functional form $f\{.\}$ in equation (1):

$$Y = f\{X\} + \varepsilon, \tag{1}$$

where ε is the deviation of the dependent variable from the relationship linking X to Y. Equation (1) could involve time series on X and Y, or cross-sectional data on X and Y. The idea behind local nonparametric modeling is to allow for a potentially nonlinear relationship over different ranges of X.[4]

Friedman (1991a, 1991b) introduced the MARS approach of using smoothing splines to fit the relationship between a set of predictors and a dependent variable. A smoothing spline is similar to a cubic spline, in which a cubic regression[5] is fit to several pre-selected subsets of the data. By requiring the curve segments to be continuous (so that first and second derivatives are non-zero), one obtains a very smooth line that can capture "shifts" in the relationship between variables. These shifts occur at locations designated as "knots" and provide for a smooth transition between "regimes." The MARS algorithm searches over all possible knot locations, as well as across all variables and all interactions among all variables. It does so through the use of combinations of variables called "basis functions," which are similar to variable combinations created by using principal components analysis. Once MARS determines the optimal number of basis functions and knot locations, a final least-squares regression provides estimates of the fitted model on the selected basis functions.

As an example, Figure 1 presents the relationship between a single predictor and a dependent variable. This relationship changes at two knot locations—values of X_t at the points where the relationship between X_t and Y_t shifts. We can view

4 There are a number of attractive nonlinear, nonparametric regression models. Granger (1995) recognized that modeling nonlinear relationships between extended-memory variables holds promise and has recently devised a test for threshold unit roots with Enders (Enders and Granger, 1998). Granger and Terasvirta (1993) examined issues relating to modeling nonlinear relationships in economics. Qi (1999) has demonstrated that stock returns can be predicted accurately using financial and economic variables in a neural network model, whereas Cao and Soofi (1999) and Fernandez-Rodriguez et al. (1999) show that nonlinear methods can be successfully employed in exchange rate prediction.

5 In a cubic regression between a predictor X and a dependent variable Y, the regressors would include a constant, the level of X, the square of X, and the cube of X.

these knots as threshold effects, in that if X_t is below the first knot (threshold), the relationship appears to be linear. If X_t is between the two knots, the relationship appears curved; whereas if X_t is above the second knot, the relationship changes once again. If we label the X_t variable as time and the Y_t variable as the price level, Figure 1 tells us something about the behavior of the inflation rate over time. It changes at the knots. A smoothing spline provides a curved transition between the various thresholds exhibited in Figure 1.

When modeling the relationship between a single predictor X_t and the dependent variable Y_t, a general model might take the form

$$(2) \qquad Y_t = \sum_{k=1}^{M} a_k B_k(X_t) + \varepsilon_t$$

where $B_k(X_t)$ is the kth basis function of X_t. Basis functions can be highly nonlinear transformations of X_t, but note that Y_t is a linear (in the parameters) function of the basis functions. Estimates of the parameters a_k are chosen by minimizing the sum of squared residuals from equation (2). The advantage of MARS is in its ability to estimate the basis functions so that both the additive and the interactive effects of the predictors are allowed to determine the response variable.

An example will aid in understanding MARS modeling. Suppose the rate of inflation, π, money growth, μ, output growth, δ, and the rate of currency depreciation, γ, are related according to the following equation:

$$(3) \quad \pi = 1.25 + 0.1\max(0, \mu - 2.0) + 0.5\max(0, \delta - 5.0) + 0.8\max(0, \gamma - 2.5) + 0.25\max(0, \mu - 2.0)\max(0, \delta - 5.0).$$

The terms in parentheses have effects on inflation only if they are positive and are zero otherwise; max (0, μ – 2.0) is interpreted as the maximum value of the two elements, 0 and (μ – 2.0), and so on. When money growth, output growth, and the rate of currency depreciation are below their threshold values, the inflation rate is 1.25 percent. If money growth is above 2 percent (the value at which there is a knot), this has both direct and indirect (or joint) effects on inflation. The direct effect raises inflation by 0.1 times the difference between money growth and its knot. The joint effect depends on whether output growth is above 5 percent *at the same time that money growth is above its knot or threshold effect*. The rate of currency depreciation raises inflation by 0.8 times the difference between the rate of currency depreciation and 2.5. Below these knots, for each variable in this example, there are no effects on inflation.

MARS would take money growth, output growth, and the rate of currency depreciation as predictors and attempt to fit the best model for the inflation rate by placing knots and choosing additive and interactive effects to minimize the sum of squared errors. The basis functions would be interpreted as the additive and interactive effects of the variables relative to their knot locations. Thus, in this example, the first basis function would involve max (0, μ – 2.0); the second basis function would contain max (0, δ – 5.0); the third basis function would be max (0, γ – 2.5); and the final basis function would involve two variables and be nonlinear (in variables): max (0, μ – 2.0) max (0, δ – 5.0).[6]

MARS identifies the knot locations that most reduce the sum of squared residuals. For example, with a single predictor the sum of squared residuals would be

$$(4) \qquad \sum_{i=1}^{N} \left\{ Y_t - \sum_{j=0}^{Q} b_j X_t^{\,j} - \sum_{k=1}^{K} a_k (X_t - t_k)_+^{Q} \right\}^2$$

where b_j and a_k are multiple regression coefficients on cubic ($Q = 3$) splines of X_t, and X_t relative to knot location t_k. The notation $(X_t - t_k)_+^{Q}$ indicates that the cubic spline of X_t relative to knot location t_k is included if the difference is positive; otherwise it is zero.

From equation (4) it is clear that the addition of a knot can be viewed as adding the corresponding $(X_t - t_k)_+^{Q}$. A forward and backward stepwise search is incorporated in the MARS algorithm, with the forward step purposely overfitting the data. Insignificant terms are deleted on the backward step of the routine.

Model selection is based on the generalized cross-validation (GCV) criterion of Craven and Wahba (1979). The GCV can be expressed as

$$(5) \quad \text{GVC} = (1/N)\sum_{t=1}^{N} \left\{ \left[Y_t - f_M(X_t)\right]^2 / \left[1 - C(M)/N\right]^2 \right\}$$

where there are N observations, and the numera-

6 Note that values below these thresholds could be included in the final model if they add to the fit of the equation. For example, one might find a knot in money growth at 1.1, which has a different effect on inflation than that when money growth is above 2.0. Assuming the coefficient to be 0.04, equation (3) would become:

$$\pi = 1.25 + 0.1\max(0, \mu - 2.0) + 0.5\max(0, \delta - 5.0) + 0.8\max(0, \gamma - 2.5) + 0.25\max(0, \mu - 2.0)(0, \delta - 5.0) + 0.04\max(0, \mu - 1.1).$$

CAN MARS EXPLAIN M2 VELOCITY SHIFTS?

Because MARS is a relatively new tool in the econometrician's toolkit, an example will help illustrate its potential value. Orphanides and Porter (2000) recently demonstrated how regression trees can be used to explain shifts in M2 velocity, with a view to resurrecting the P* model of inflation. Regression trees can serve to identify breaks in the reduced-form velocity equation as changes in the coefficient on the opportunity cost of M2 and the time trend. Inflation forecasts based on their estimates of equilibrium velocity outperform those based on the simple Hallman, Porter, and Small (1991) P* model.

To demonstrate the advantages of using MARS, I constructed estimates of M2 velocity using data identical to those employed by Orphanides and Porter (2000), spanning a somewhat longer time frame, 1959:Q1 to 2000:Q1. Velocity is assumed to be a function of the opportunity cost of M2 balances (the difference between the three-month Treasury bill rate and the average rate paid on M2 balances) and a time trend. MARS allows threshold effects in the opportunity cost and time trend series to accommodate shifts in velocity resulting from financial innovation. Moreover, it allows both series to jointly affect velocity over the sample.

Table A1 provides the final fitted model, allowing as many as 40 basis functions and two variable interactions. The time trend series is most important, whereas the opportunity cost series is only 31.7 percent as important as the trend series. (These figures are constructed on the basis of what happens to the explanatory power of the model when each individual series is excluded from the equation.)

There appear to be threshold effects in the opportunity cost series at 0.27, 2.36, and 3.30 percent, while there are time trend thresholds at observations 20 (1964:Q1), 32 (1967:Q1), 47 (1970:Q4), 58 (1973:Q3), 63 (1974:Q4), 120 (1989:Q1), 130 (1991:Q3), 135 (1992:Q4), and 155 (1997:Q4). Orphanides and Porter (2000) identified time effects at 1960:Q3, 1962:Q3, 1978:Q1, 1988:Q4, and 1992:Q3, as well as a number of interest rate effects spanning from 1.643 percent to 2.034 percent.

MARS provides graphical information on the optimal fit of the data. The surface plot demonstrates that the optimal transformation and combination of both series in explaining M2 velocity is nonlinear.

The actual and fitted MARS model for velocity appear here. As you can see, the fitted MARS model captures velocity shifts in the post-1991 era very well. How this can be used to forecast inflation within a P* model will be the subject of further work. However, this example demonstrates the potential benefits to MARS modeling.

Figure A1

Surface 1: Maximum = 0.57952

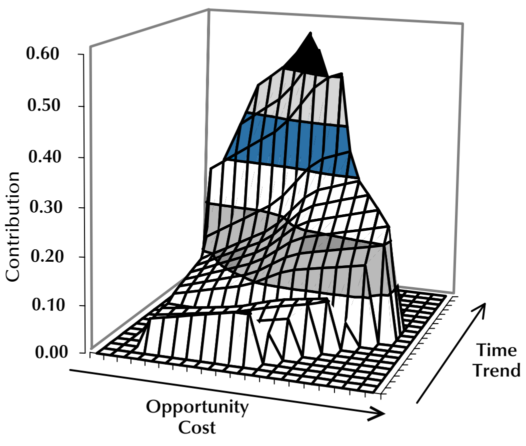


Figure A2

MARS Model for M2 Velocity

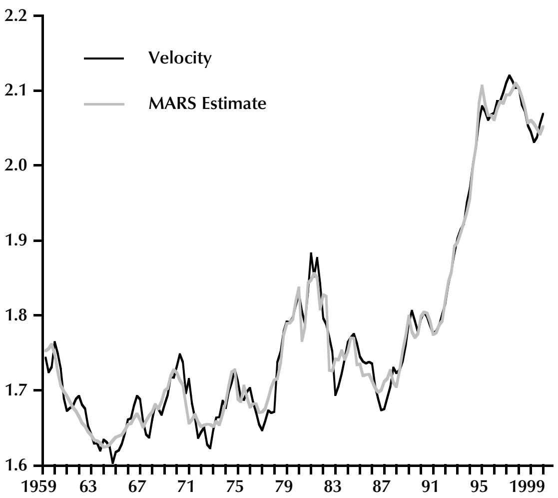


Table A1

MARS Velocity Estimates

Independent variable	Coefficient	t-Statistic	Variable
Constant	0.982	19.278	
Basis function 2 (BF2)	0.006	13.011	Max (0, 120–time)
Basis function 3 (BF3)	0.018	7.171	Max (0, oppcost–0.270)
Basis function 6 (BF6)	0.008	12.005	Max (0, time–20)
Basis function 8 (BF8)			Max (0, time–155)
Basis function 10 (BF10)	0.007	5.686	Max (0, time–58)* BF3
Basis function 14 (BF14)	–0.002	–5.940	Max (0, time–47)* BF3
Basis function 16 (BF16)	–0.004	–4.481	Max (0, time–63)* BF3
Basis function 18 (BF18)	0.017	7.566	Max (0, time–130)
Basis function 20 (BF20)	–0.017	–7.279	Max (0, time–135)
Basis function 25 (BF25)			Max (0, 3.301–oppcost)
Basis function 27 (BF27)	0.509×10^{-3}	3.187	Max (0, 2.364–oppcost)* BF6
Basis function 28 (BF28)	-0.809×10^{-3}	–4.168	Max (0, time–32)* BF25
Basis function 40 (BF40)	–0.009	–8.997	Max (0, oppcost–0.270)* BF8
R^2	0.983		
Number of observations	164		

NOTE: This table provides the final fitted model, allowing as many as 40 basis functions (of which only 11 are retained) and two variable interactions. The time trend series is most important, whereas the opportunity cost series is only 31.7 percent as important as the trend series. These figures are constructed on the basis of what happens to the explanatory power of the model when each individual series is excluded from the equation (*time* denotes the time trend, *oppcost* the opportunity cost series). Also, there appear to be threshold effects in the opportunity cost series at 0.27, 2.364, and 3.301 percent, whereas there are time trend thresholds at observations 20 (1964:Q1), 32 (1967:Q1), 47 (1970:Q4), 58 (1973:Q3), 63 (1974:Q4), 120 (1989:Q1), 130 (1991:Q3), 135 (1992:Q4), and 155 (1997:Q4). Orphanides and Porter (2000) identified time effects at 1960:Q3, 1962:Q3, 1978:Q1, 1988:Q4, and 1992:Q3, as well as a number of interest rate effects spanning from 1.643 percent to 2.034 percent.

tor measures the lack of fit on the M basis function model $f_M(X_t)$. This term corresponds to the sum of squared residuals from the fitted model. The denominator contains a penalty for model complexity, $C(M)$, which is related to the number of parameters estimated in the model.

Interpretation

MARS estimates can most readily be interpreted from an analysis of variance (ANOVA) representation of the model, where the fitted function is expressed as a linear combination of additive basis functions in single variables and interactions between variables. MARS provides graphical plots which illustrate the optimal transformation of the variables chosen by the algorithm, much like the alternating conditional expectations (ACE) algorithm of Breiman and Friedman (1985). The ACE approach to modeling finds the nonlinear transformation of the predictors which maximizes the correlation between the dependent variable and the transformed predictors. A plot of the transformed series against the dependent variable is sometimes helpful in identifying a functional form to be used in parametric modeling. Hallman (1990) and Granger and Hallman (1991) employed ACE to examine nonlinear cointegration. The accompanying box provides a simple example of the MARS algorithm applied to estimates of M2 velocity.

In MARS, a comparison of the low- and high-order models assists in determining whether to allow variables to enter individually or in combination. Friedman (1991a) suggests a comparison of a measure analogous to an "adjusted R-squared," with a model involving interaction terms chosen over an additive model only if its adjusted R-squared is "substantially" larger. As part of the MARS output, the relative contribution of each variable is determined, as are estimates of the model's adjusted R-squared given that a particular ANOVA function (variable) has been omitted from the model. This assists in interpreting the significance of each ANOVA function.

MARS has been extended to incorporate categorical variables, logit regression, and missing data.

It has been successfully applied to the Wolf sunspot data by Lewis and Stevens (1991), to cointegration testing by Sephton (1994), to forecasting exchange rates by Sephton (1993) and De Gooijer et al. (1998), and to nonlinear causality testing by Sephton (1995), as well as in describing large cross-sectional data sets by Steinberg and Colla (1999). The objective here is to examine the extent to which the logit specification provides useful information on the probability of recession.

FORECASTING RECESSIONS USING MARS

There are two interesting questions to consider. The first relates to in-sample forecasts of the probability of recession based on information that is available at time $(t-k)$. That is, how well does MARS fit the historical data? Given the flexibility of the algorithm, one might expect to see MARS perform very well in capturing the probability of recession. The second, more interesting question examines out-of-sample forecasts to determine whether information on financial variables can predict the probability of recession k periods ahead. This is the type of question one might ask of an "operational forecasting" model: Given data at time $(t-k)$ how likely is recession within the next few months?

A number of previous studies have examined the ability of probit models to capture recession probabilities.[7] The probit specification examines the probability of recession, Prob $(Y_t = 1)$, using the cumulative standard normal distribution, $\Phi(\,.\,)$ and a set of regressors, X_{t-k}:

$$(6)\qquad \text{Prob}\left(Y_t = 1\right) = \Phi\left(a + \beta X_{t-k}\right).$$

Given data up to period $(t-k)$ these models are estimated and used to generate recession forecasts at time t. Statistics on pseudo-R-squared, root-mean-squared error, mean absolute error, and quadratic probability scores are used to gauge forecast accuracy.[8] In-sample forecasts are generally more accurate than out-of-sample forecasts in which an estimated model is used to forecast beyond the estimation period. In the probit model, the parameters are assumed to be temporally stable: that is, a and β are assumed to be constant. The effects of X at time $(t-k)$ are assumed to have the same influence on the probability of recession at every point in the sample. This assumption ignores Temin's (1998) historiography of American recessions over the past 100 years. Temin concluded that it was difficult to assign a unique correspondence between an economic variable and the likelihood of recession. The interesting question here is whether MARS results outperform those based on this simple probit specification.

In-Sample Estimates

For present purposes, recession forecasts are examined at the three-, six-, nine-, and twelve-month horizons. Information available at times $(t-3)$, $(t-6)$, $(t-9)$, and $(t-12)$ is used to model the probability of recession at time t. The in-sample evidence is based on estimating a MARS model over the entire sample period. Actual dates of recession are compared with forecasted probabilities to measure the information content of the MARS models.

For example, the six-month horizon model examines the following specification:

$$(7)\qquad R_t = f\left\{Y_{t-6}, \Delta IP_{t-6}, \Delta UR_{t-6}, \Delta RM_{t-6}, \Delta SP_{t-6}, \Delta FF_{t-6}\right\} + \varepsilon_t$$

where R_t is 1 if the economy is in recession in period t and 0 otherwise; Y is the yield spread, ΔIP is the change in the logarithm of industrial production, ΔUR is the change in the unemployment rate, ΔRM is the change in the logarithm of the CPI deflated value of M2, ΔSP denotes the change in the logarithm of the S&P 500 Index, and ΔFF denotes the change in the federal funds rate. The error term is given by ε_t, with the nonlinear nonparametric functional form given by $f\{.\}$. The fitted value can be used to obtain an estimate of the probability of recession given data at time $(t-6)$. Information previous to $(t-6)$ and subsequent to $(t-6)$ is not included in the model.

The MARS algorithm involves setting a number of parameters used in model selection. The most important are the maximum number of basis functions allowed and the highest order of interaction possible. Because there are six predictors, up to six variable interactions are allowed. The MARS algorithm will fit as many interactions as help

[7] These include Estrella and Mishkin (1998), Haubrich and Dombrosky (1996), Bernard and Gerlach (1996), Dueker (1997), and Atta-Mensah and Tkacz (1998).

[8] The quadratic probability score is simply the average of twice the squared errors. For present purposes the root-mean-squared error and mean absolute error rates will be used to gauge forecast accuracy. Currently there is no measure analogous to the pseudo-R-squared used in probit models.

describe the data, with up to 40 basis functions allowed in the forward search strategy.

Reporting the results of each model would be of little merit because a large volume of output is generated by each estimation. More important is a comparison of the forecast of recession with the actual data. The upper half of Table 1 contains summary statistics on how well each MARS model fit the historical data, as well as those derived from a probit specification using the same explanatory variables. MARS recession probability estimates are superior to those derived from the probit specification, with the root-mean-squared error for the three-month forecasting horizon 16.7 percent relative to 28.9 percent for the probit model. The MARS root-mean-squared error is lowest at the three-month horizon and is highest at the twelve-month horizon, at almost 24 percent. At all horizons the MARS models appear to dominate those based on the probit specification. Figure 2 presents a plot of the MARS probability forecasts for the four different forecasting horizons. The algorithm provides a very good in-sample fit in the short-term, yet exhibits a number of false signals beyond three months.

These results appear to suggest that there are benefits to the modeling of the dichotomous variable at the monthly frequency using MARS. This is to be expected given that nonlinear nonparametric models fare well at explaining relationships in-sample. They are designed to be sufficiently flexible to capture historical data, as are neural network models of Kuan and White (1994). The interesting issue is whether they perform well in an out-of-sample forecasting exercise.

Before turning to that question, it is useful to consider results presented by Dueker (1997). He found that adding a lagged recession variable to the probit framework improved forecast accuracy, arguing that the probability of recession could be affected by duration effects associated with different "states of the world." Does adding a lagged recession variable affect the in-sample results of both the probit and MARS frameworks?

The bottom half of Table 1 contains information on this augmented model. A recession variable dated at the same time as the other explanatory series was added to the predictor space and MARS models were re-estimated. Forecast accuracy improves at the three-month horizon, with a reduction in the root-mean-squared error from 16.7 to 11.7 percent but remains relatively unchanged at the other time horizons. The probit specification benefits from the addition of the lagged dependent variable at the three-month horizon, but continues to underperform relative to MARS.

Table 1

In-Sample Forecasting Statistics

	Root-mean-squared error		Mean absolute deviation	
Lag	MARS	Probit	MARS	Probit
Six predictors				
3	.167	.289	.057	.170
6	.197	.298	.078	.177
9	.189	.299	.071	.177
12	.244	.311	.117	.193
Six predictors and lagged dependent variable				
3	.117	.219	.074	.094
6	.199	.280	.158	.154
9	.193	.299	.187	.178
12	.239	.307	.199	.191

NOTE: Root-mean-squared error is calculated by summing the squared differences between the actual and forecast probabilities of recession, dividing by the number of periods in the sample, and taking the square root of the result. The mean absolute deviation is the average absolute value of the prediction less the true state of the recession variable.

Table 2

Out-of Sample Forecasting Statistics

	Root-mean-squared error		Mean absolute deviation	
Lag	MARS	Probit	MARS	Probit
Six predictors				
3	.317	.292	.152	.181
6	.311	.293	.145	.160
9	.317	.303	.135	.167
12	.319	.290	.169	.168
Six predictors and lagged dependent variable				
3	.289	.229	.111	.094
6	.299	.281	.129	.135
9	.319	.305	.137	.167
12	.341	.292	.183	.169

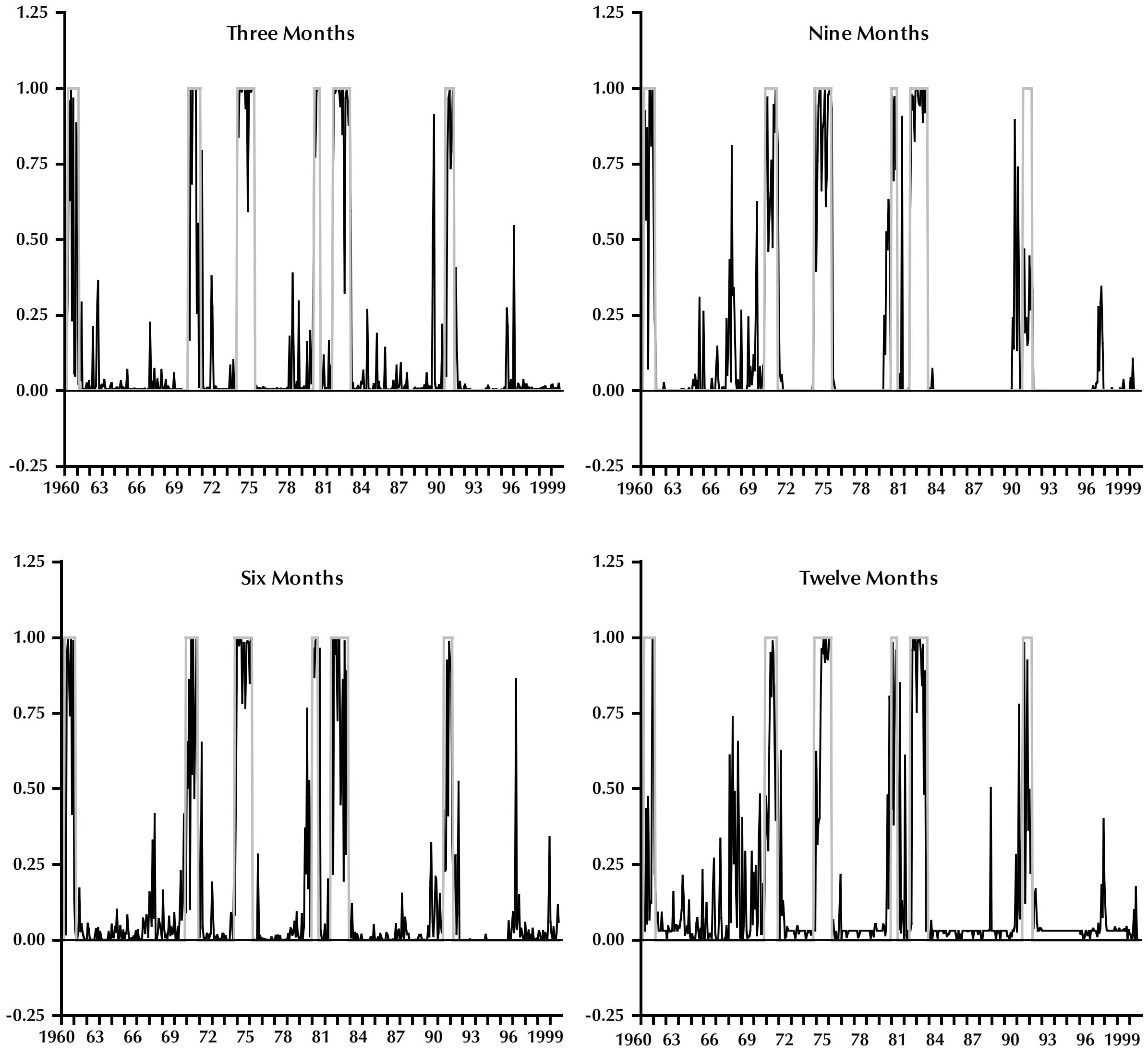
NOTE: Root-mean-squared error is calculated by summing the squared differences between the actual and forecast probabilities of recession, dividing by the number of periods in the forecast horizon, and taking the square root of the result. The mean absolute deviation is the average absolute value of the prediction less the true state of the recession variable.

Out-of-Sample Estimates

Although neural network and nonparametric regression models frequently fit well in-sample, their out-of-sample performance is not as impressive. This is in part a result of the large data samples which are required to fit the models. As well, the models are constructed to provide an optimal

Figure 2

In-Sample Forecasts of Recession, MARS Models



in-sample fit, and, as in traditional linear parametric methods, there is no guarantee they will provide a good fit out-of-sample. A realistic out-of-sample exercise is required to determine whether there are true benefits to modeling recession probabilities using a data-mining procedure such as MARS.

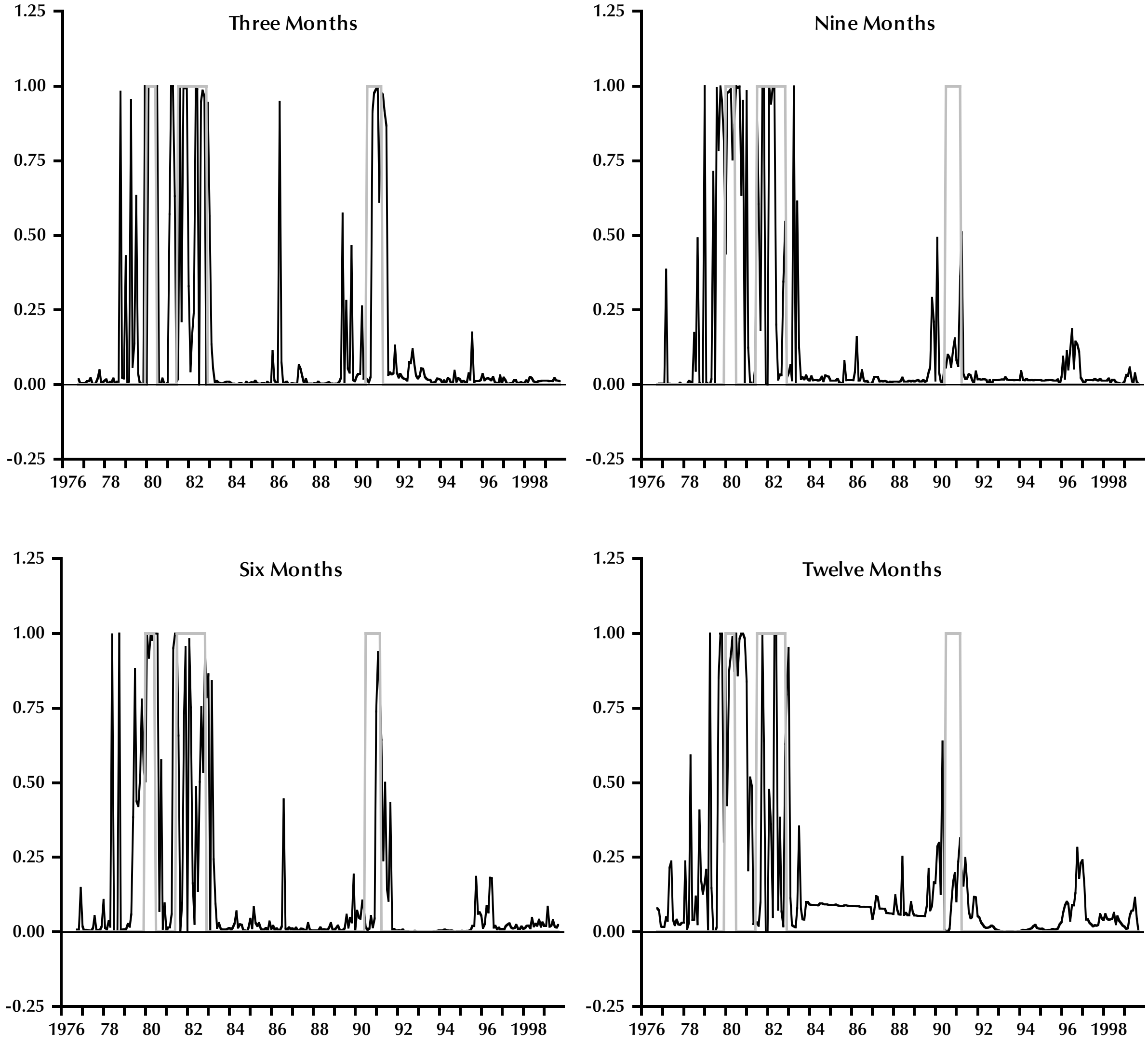
Toward this end, the first 200 observations of the data were used to fit MARS and probit models, which were subsequently used to forecast the probability of recession k periods hence, with k = 3, 6, 9, 12, as before.[9] Each forecast was compared with the state of the economy to determine forecast accuracy. The sample was then extended by one observation, and the process continued until the entire sample was used to forecast the probability of recession. This process is similar to a rolling regression forecast with model updating.

Table 2 contains summary statistics for the original and the augmented (lagged recession variable) predictor space. The MARS specification does not fare as well at predicting recessions in

[9] The sample used to estimate the first MARS model spans February 1960 through September 1976 and expands by one month until all the data through September 1999 are included.

Figure 3

Out-of-Sample Forecasts of Recession, MARS Models



out-of-sample forecasting, with root-mean-squared errors around 31 percent using the four different forecast horizons. Adding a lagged recession variable reduces the error rates by less than 3 percent at the three- and six-month horizons, with prediction errors at roughly 29 percent and 30 percent, respectively. The estimated MARS models perform nearly as well as the probit approach to estimating recession probabilities.

Figure 3 presents a plot of the actual recession dates and the MARS forecasts. At the three-month horizon, MARS appears to forecast the 1990-91 recession fairly well, but the large number of false signals across all time frames suggests that the adoption of nonlinear nonparametric methods is not a panacea for recession forecasting.

CONCLUSION

For in-sample recession forecasting, the application of multivariate adaptive regression splines to financial predictors of recession shows great promise. The out-of-sample evidence indicates that the MARS models considered here contain helpful, but not entirely accurate, predictions of recession.

There are a number of areas in which the present analysis can be extended. The first is to include a broader set of financial variables to determine whether they contain information in addition to that already contained in the six series included in the present analysis. Similarly, it may be reasonable to examine these questions using quarterly data rather than monthly data, since the latter may be characterized by a high noise-to-signal ratio.

Finally, the construction of a leading index using the MARS modeling strategy may provide useful forecasts against which to compare other leading indicators maintained by the Conference Board and others. The logit specification may be too difficult for the algorithm to fit effectively; a dynamic model examining economic growth which allows for variable interactions and duration-dependence may offer significant advantages over the present analysis.

REFERENCES

Atta-Mensah, Joseph and Tkacz, Greg. "Predicting Canadian Recessions Using Financial Variables: A Probit Approach." Working Paper No. 98-5, Bank of Canada, 1998.

Bernard, Henri and Gerlach, Stefan. "Does the Term Structure Predict Recessions? The International Evidence." Working Paper No. 37, Bank for International Settlements, 1996.

Breiman, Leo and Friedman, Jerome H. "Estimating Optimal Transformations for Multiple Regression Correlation." *Journal of the American Statistical Association,* September 1985, *80*(391), pp. 580-98.

Cao, Liangyue and Soofi, Abdol. "Nonlinear Deterministic Forecasting of Daily Dollar Exchange Rates." *International Journal of Forecasting,* October 1999, *15*(4), pp. 421-30.

Craven, P. and Wahba, G. "Smoothing Noisy Data with Spline Functions—Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation." *Numerische Mathematik,* 1979, *31*(4), pp. 317-403.

De Gooijer, Jan G.; Ray, Bonnie K. and Krager, Horst. "Forecasting Exchange Rates Using TSMARS." *Journal of International Money and Finance,* June 1998, *17*(3), pp. 513-34.

Dueker, Michael. "Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions." Federal Reserve Bank of St. Louis *Review,* March/April 1997, *79*(2), pp. 41-51.

Enders, Walter and Granger, Clive W.J. "Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates." *Journal of Business and Economic Statistics*, July 1998, *16*(3), pp. 304-11.

Estrella, Arturo and Mishkin, Frederic S. "Predicting U.S. Recessions: Financial Variables as Leading Indicators." *Review of Economics and Statistics*, February 1998, *80*(1), pp. 45-61.

Fernandez-Rodriguez, Fernando; Sosvilla-Rivero, Simon and Andrada-Felix, Julian. "Exchange-Rate Forecasts with Simultaneous Nearest-Neighbor Methods: Evidence from the EMS." *International Journal of Forecasting*, October 1999, *15*(4), pp. 383-92.

Friedman, Benjamin and Kuttner, Kenneth. "Indicator Properties of the Paper-Bill Spread: Lessons from Recent Experience." *Review of Economics and Statistics,* February 1998, *80*(1), pp. 34-44.

Friedman, Jerome H. "Multivariate Adaptive Regression Splines." *Annals of Statistics*, March 1991a, *19*(1), pp. 1-67.

__________. "Estimating Functions of Mixed Ordinal and Categorical Variables Using Adaptive Splines." Laboratory for Computational Statistics Technical Report No. 108, Stanford University, 1991b.

Granger, Clive W.J. "Modelling Nonlinear Relationships Between Extended-Memory Variables." *Econometrica,* March 1995, *63*(2), pp. 265-79.

__________ and Hallman, Jeffrey J. "Long Memory Series with Attractors." *Oxford Bulletin of Economics and Statistics,* February 1991, *53*(1), pp. 11-26.

__________ and Terasvirta, Timo. *Modelling Nonlinear Economic Relationships.* New York: Oxford University Press, 1993.

Hallman, Jeffrey J. "Nonlinear Integrated Series, Cointegration and an Application." Ph.D. Dissertation, University of California, San Diego, 1990.

__________; Porter, Richard D. and Small, David H. "Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run?" *American Economic Review,* September 1991, *81*(4), pp. 841-58.

Haubrich, Joseph G. and Dombrosky, Ann M. "Predicting Real Growth Using the Yield Curve." Federal Reserve Bank of Cleveland *Economic Review*, First Quarter 1996, *32*(1), pp. 26-35.

Lewis, P.A.W. and Stevens, J.G. "Nonlinear Modeling of Time Series Using Multivariate Adaptive Regression Splines (MARS)." *Journal of the American Statistical Association,* December 1991, *86*(416), pp. 864-77.

Kuan, Chung-Ming and White, Halbert. "Artificial Neural Networks: An Econometric Perspective." *Econometric Review*, March 1994, *13*(1), pp. 1-91.

Orphanides, Athanasios and Porter, Richard D. "P* Revisited: Money-Based Inflation Forecasts with a Changing Equilibrium Velocity." *Journal of Economics and Business,* January-April 2000, *52*(1), pp. 87-100.

Qi, Min. "Nonlinear Predictability of Stock Returns Using Financial and Economic Variables." *Journal of Business and Economic Statistics,* October 1999, *17*(4), pp. 419-29.

Sephton, Peter S. "Nonlinear Exchange Rate Prediction on MARS." Unpublished manuscript, 1993.

__________. "Cointegration Tests on MARS." *Computational Economics,* February 1994, *7*(1), pp. 23-35.

__________. "A Nonparametric View of Money/Income Causality." *Applied Financial Economics*, April 1995, *5*(2), pp. 79-84.

Steinberg, D. and Colla, P. *MARS: An Introduction.* San Diego: Salford Systems, 1999.

Temin, Peter. "The Causes of American Business Cycles: An Essay in Economic Historiography," in J. Fuhrer and S. Scott, eds., *Beyond Shocks: What Causes Business Cycles.* Federal Reserve Bank of Boston Conference Series No. 42, June 1998, pp. 37-50.

All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* are available to our readers and can be obtained from three sources:

1. **FRED (Federal Reserve Economic Data), an electronic bulletin board service.** You can access FRED either through the Internet's World Wide Web or by dialing (314) 621-1824 through a modem-equipped personal computer. FRED's Internet address is **www.stls.frb.org/fred.** For a free brochure on FRED, please call (314) 444-8809.

2. **Research Department,** Federal Reserve Bank of St. Louis, Post Office Box 442, St. Louis, MO 63166-0442. You can request data and programs on either disk or hard copy. Please include the author, title, issue date, and page numbers with your request.

3. **Inter-university Consortium for Political and Social Research (ICPSR).** Member institutions can request data through the CDNet Order facility. Nonmembers should write to: ICPSR, Institute for Social Research, P.O. Box 1248, Ann Arbor, Michigan 48106, or call (734) 998-9900.

ON THE INTERNET

www.stls.frb.org

THE FEDERAL RESERVE BANK of ST. LOUIS

Federal Reserve Bank of St. Louis
Post Office Box 442
St. Louis, Missouri 63166-0442