Vol. 32, No. 1

ECONOMIC REVIEW

1996 Quarter 1

FEDERAL RESERVE BANK
OF CLEVELAND

http://clevelandfed.org/research/review
Economic Review 1996 Q1

ECONOMIC REVIEW

1996 Quarter 1
Vol. 32, No. 1

This paper outlines the considerable information requirements faced by monetary policymakers and looks at the data to see what we actually know and how well we know it. The author's main conclusion is that our forecasting ability is very poor, which creates uncertainty that leads to cautious policymaking. At a more practical level, he finds that nominal-income targeting rules are more robust than price-targeting rules in the sense that someone who cares about the aggregate price path loses little by targeting nominal income, but someone who cares about nominal income is made much worse off by moving to a price-level target, which substantially destabilizes real output.

The reduced-form empirical strategy has been used for more than 30 years to test the Modigliani–Miller model of corporate financial structure. Curiously, the early tests almost always accepted the model, whereas subsequent tests almost always reject it. This paper considers the limitations of the reduced-form strategy that led to the early, spurious results, and demonstrates why an empirical strategy that is not closely tied to an underlying economic theory of behavior will usually yield estimates that are too imprecise or too unreliable to form a basis for policy.

The yield curve, which relates interest rates to notes and bonds of various maturities, is often used by economists and business analysts to predict future economic growth. But how reliable is it? This article uses out-of-sample regressions to determine how well the 10-year, three-month yield spread predicts future real GDP growth. The authors show that although the yield curve is a good predictor over the entire 30-year sample period, it has become much less accurate over the last decade.

Economic Review is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland. Copies of the *Review* are available through our Corporate Communications & Community Affairs Department. Call 1-800-543-3489 (OH, PA, WV) or 216-579-2001, then immediately key in 1-5-3 on your touch-tone phone to reach the publication request option. If you prefer to fax your order, the number is 216-579-2477.

Economic Review is also available electronically through our home page on the World Wide Web: http:// www.clev.frb.org.

ISSN 0013-0281

Practical Issues in Monetary Policy Targeting

by Stephen G. Cecchetti

Stephen G. Cecchetti is a professor of economics at Ohio State University and a research associate at the National Bureau of Economic Research and the Federal Reserve Bank of Cleveland. The author thanks Charles T. Carlstrom for his helpful comments on this article.

Introduction

What do monetary policymakers need to know and when do they need to know it? Textbook descriptions and academic discussions of policymaking usually ignore the practical problems faced by those who make the decisions and take the actions. While most economists would agree that monetary policy has real short-run effects and is most likely neutral in the long run, they could provide no more than informed speculation in helping decide at what level to set the target for a policy instrument and when to change it.

This paper's purpose is to outline the type of information monetary policymakers need in practice, and to examine the data to see what we actually know.[1] Any policy rule must be formulated in several clearly defined steps. First, one must identify an operational instrument, best thought of as something policymakers can control precisely, like the federal funds rate or the monetary base. Next, there must be a target. Many central banks have stated that price stability is their goal, but an obvious alternative to targeting the aggregate price level is targeting nominal income.[2] In addition to choosing the target variable itself, formulating policy necessitates specifying a loss function: What is the relative importance of large and small, or positive and negative, deviations of aggregate prices from their target path? One might also assign a cost to large movements in the target variable. For example, it might be important for the Federal Reserve to have a reputation for changing the federal funds rate target smoothly, without large movements or sudden reversals, to avoid creating uncertainty in financial markets.

The next stage in devising a monetary rule is to link the operating instrument with the target. This requires specification and estimation of a macroeconomic model. One needs quantitative answers to questions of the form "If the federal funds rate is moved by one percentage point, what will be the path of the aggregate price level and real output over the following three years?" Not only do we require a point estimate of this response function, but it is also crucial that we know how precise our knowledge is in a statistical sense.

1 This work is based on Cecchetti (1995).

2 I will not discuss the difference between price-level and inflation targeting. While this is a potentially important practical distinction, it is beyond the scope of this paper.

Finally, policymakers need a timely estimate of their target variable's future path *in the absence of* any policy actions. In other words, they must know when external shocks hit the economy, how large they are, and what their impact on the time path of aggregate prices and real output will be.

The next section offers a detailed discussion of the modeling issue: How do we formulate and estimate the necessary simple, dynamic, empirical macroeconomic model? The section's first major part looks at econometric identification. What must we assume in order to disentangle the fluctuations in output and prices into their various components? How might we actually estimate the impact of monetary policy on macroeconomic quantities of interest?

The section's second major part discusses the issue of structural stability. Monetary policymakers change their emphasis fairly frequently, focusing on one indicator one year and another the next. How does this affect our attempt to estimate the relationship between the variables that policymakers control (like the federal funds rate) and the things we care about? Can we estimate a stable relationship between output, prices, and interest rates over any reasonable period?

The methodological discussion of modeling issues is followed by section II, in which I present a particular model and examine its properties. Several different types of results are included. First, I look at the impact of different sources of shocks on the variables of interest. Besides allowing answers to questions like "If the federal funds rate were to rise by 100 basis points, how much would output change over the next three years?" this approach makes it possible to examine the sources of fluctuations in output and prices. For example, has monetary policy been responsible for a significant share of output variation over the past decade?

Section III discusses how a policy rule can be formulated. The first step is to specify an objective function: What do policymakers actually want to stabilize? This discussion emphasizes the need for taking account of imprecision when forming a policy rule. We are uncertain how changes in the interest rate affect the size and timing of output and price movements. This means we cannot confidently predict policy actions' impact on target variables, so that policy actions differ from what they would be if we had perfect knowledge. From the theoretical discussion, I move on to examine several possible objective functions of policy and the interest rate paths implied by the combination of each rule and the estimated model. I focus throughout on the importance of employing rules that recognize the imprecision of our knowledge regarding the size of the linkages we need to estimate.

I reach three significant conclusions: First, since prices take time to respond to all types of economic shocks, the objective of price stability implies raising the federal funds rate immediately after a shock, instead of waiting for prices to rise. Second, and more important, comparing the results of price-level targeting with those of nominal-income targeting implies that the difficulties inherent in forecasting and controlling the former provide an argument for concentrating on the latter. Finally, it is possible to use policy rules to see how closely recent movements in the federal funds rate conform to those implied by either price-level or nominal-income targeting rules. The results show that the policy that is optimal in this limited sense involves faster, bigger movements than those shown by the actual federal funds rate path. This suggests that policymakers' actions have been based on something akin to nominal-income targeting, but with costs attached to interest rate movements.

I. Modeling Issues

The single biggest problem in formulating monetary policy rules is how to construct an empirical macroeconomic model that describes the critical aspects of the economy. It is important that the model be dynamic, summarizing the impacts of shocks to the economy—as well as those of intended policy actions—over time. The standard response to this challenge has been to construct various forms of vector autoregressions (VAR). A VAR can answer a question of the following type: "If the federal funds rate moves, when and by how much does the price level change?" Policymakers require quantitative answers to exactly these kinds of questions.

To construct any usable empirical model, a researcher must make a number of choices. I will describe four of these: 1) Which variables should be included in the model? 2) What is the appropriate measure of monetary policy? 3) How can the model be identified econometrically? and 4) Over what sample period should the model be estimated?

Variable Inclusion

When trying to discern the relationship between inflation, output, and monetary policy, should we include other variables in the model? Our answer is guided by the findings of Sims (1992), who estimates a model with prices, output, and an interest rate for several countries. His robust overall conclusion is that with this specification, increases in the interest rate (which should signal policy contractions) lead to prices that are higher than otherwise expected, not lower. This problem, which came to be known as the "price puzzle," can be eliminated by including commodity prices in the model. The reasoning is that the policymaker has additional knowledge about prices' future path that the three-variable model does not adequately summarize. Policy contractions, being based on this omitted information, signal that these other indicators are pointing toward higher prices.

More recent research, like that of Christiano, Eichenbaum, and Evans (1996a, 1996b), has shown that including commodity prices eliminates the puzzle. They suggest that higher commodity prices mean higher future overall prices, and that policymakers respond to this. In other words, an upward move in commodity prices precedes both a rise in the price level and a tightening of policy in the form of an increase in the federal funds rate. The omission of this information from the original Sims formulation led to a bias in which contractionary policy predicts higher aggregate prices. This is not a policy change, but simply a reaction to external events. The models of Christiano, Eichenbaum, and Evans do have the following property: Moving toward a more contractionary monetary policy drives prices down (relative to the trajectory they would follow without the policy change).

Choice of Policy Instrument

Beyond the question of which variables the model should include, it is necessary to specify a monetary policy instrument. Should one assume that policymakers are focusing on the federal funds rate itself (or behaving as if they were), or would it be more realistic to use nonborrowed reserves as the instrument? The literature takes up this issue in some detail.[3] Because events of the past 15 years suggest that the primary focus has been on the federal funds rate, I will assume that it contains the information necessary to gauge policy actions.[4]

Identification

A model builder's most complex decision is formulating a set of "identifying assumptions." This is also the subtlest issue and the one that has generated the most discussion in the literature. It is like the textbook question about estimating supply and demand curves: There, if data on the price and quantity of a good in a market both move, we cannot tell whether the root cause of the change was a shift in supply or a shift in demand. Here, things are a bit less transparent, because there are no clearly defined supply and demand curves in the standard microeconomic sense. Instead, it is necessary to distinguish whether prices, output, and interest rates moved as a result of policy shifts, or because of factors like changes in the price of oil (an aggregate supply shock) or in the demand for money (an aggregate demand shock).

To understand the problem and its solution more fully, we can begin by writing down a dynamic structural model in its moving-average form:

$$(1) \quad p_t = A_{11}(L)\varepsilon_{pt} + A_{12}(L)\varepsilon_{ct} + A_{13}(L)\varepsilon_{yt} + A_{14}(L)u_t$$

$$(2) \quad p_t^c = A_{21}(L)\varepsilon_{pt} + A_{22}(L)\varepsilon_{ct} + A_{23}(L)\varepsilon_{yt} + A_{24}(L)u_t$$

$$(3) \quad y_t = A_{31}(L)\varepsilon_{pt} + A_{32}(L)\varepsilon_{ct} + A_{33}(L)\varepsilon_{yt} + A_{34}(L)u_t$$

$$(4) \quad r_t = A_{41}(L)\varepsilon_{pt} + A_{42}(L)\varepsilon_{ct} + A_{43}(L)\varepsilon_{yt} + A_{44}(L)u_t\,,$$

where p_t, p_t^c, and y_t are the logs of the aggregate price level, commodity prices, and output, respectively, r_t is the policy indicator, the ε's are exogenous shocks, and u is the policy innovation. Equations (1)–(4) summarize the impact of all the shocks to the economy. The $A_{ij}(L)$'s are lag polynomials in the lag operator L. For example,

$$A_{11}(L)\varepsilon_{pt} = \sum_{i=0}^{\infty} a_{11i} L^i \varepsilon_{pt}$$
$$= a_{110}\varepsilon_{pt} + a_{111}\varepsilon_{pt-1} + \ldots .$$

3 See, for example, discussions in Christiano, Eichenbaum, and Evans (1996a, 1996b) and Bernanke and Mihov (1995).

4 Most results are unaffected by the substitution of nonborrowed reserves, suggesting that the funds rate elasticity of reserve demand is relatively stable.

Because we do not observe the shocks, it is not possible to estimate the model (1)–(4) directly. Instead, we estimate the more familiar VAR form and place restrictions on the coefficients (the a_{ijk}'s) in order to recover estimates of the shocks.

Identification entails determining the errors in this four-equation system, that is, the actual sources of disturbances that lead to variation in prices, output, and interest rates. As the appendix to this paper describes, when there are four endogenous variables, six restrictions are required for complete identification.

All identification schemes involve assumptions about how these sources of variation are correlated. Researchers use two types of restrictions for this purpose. The first, based on the pioneering work of Sims (1980), is what I will call a "triangular identification," which assumes that a shock does not affect a variable contemporaneously, and so one or more of the a_{ij0}'s are zero. For example, it is commonly assumed that no variable other than policy itself responds to monetary shocks immediately, and so $a_{140} = a_{240} = a_{340} = 0$.

A more formal description of a triangular identification begins by writing the matrix $A(0)$ that is composed of all the coefficients of the $A_{ij}(0)$'s—that is, all the a_{ij0}'s. Triangular identification means assuming that six of these a_{ij0}'s are zero, and so

$$(5) \qquad A(0) = \begin{bmatrix} a_{110} & 0 & 0 & 0 \\ a_{210} & a_{220} & 0 & 0 \\ a_{310} & a_{320} & a_{330} & 0 \\ a_{410} & a_{420} & a_{430} & a_{440} \end{bmatrix}.$$

In other words, triangular identification means that the monetary policy shock u_t is identified by assuming that no variable other than the federal funds rate responds to it contemporaneously. The output shock, ε_{yt}, is identified by assuming that it is the portion of the error in the output equation that is orthogonal to the policy shock, while the commodity price shock, ε_{ct}, is the portion of the error in the commodity price equation that is orthogonal to these. The final part of the residual in the aggregate price equation that is orthogonal to all three of these is the aggregate price shock, ε_{pt}.

There are many other ways to constrain the four-variable VAR and achieve identification. One, based on the work of Galí (1992), combines two types of restrictions. The first are contemporaneous and resemble those used in the triangular method. The second, following Blanchard and Quah (1989), assume that some shocks have temporary, but not permanent, effects on some variables. For example, we might claim that monetary shocks have no long-run effects on real output, and so the impact of u_t on y_t dies out. Formally, this involves assuming that the a_{34k}'s sum to zero: $\sum_{k=0}^{\infty} a_{34k} = 0$.

Recalling that we need six restrictions, the Galí-style procedure begins with two contemporaneous restrictions based on the logic of data availability and the time people in the economy take to act. The first constraint is that monetary policy does not affect real output contemporaneously (within the month). In the notation used above, the assumption is that $a_{340} = 0$. This seems sensible, since production planning is unlikely to change suddenly after a policy innovation. The second constraint is that the aggregate price level does not enter the money supply rule. This also seems sensible, because the Bureau of Labor Statistics does not publicly release the Consumer Price Index (CPI) until the month following its price survey.

The Galí-style, long-run restrictions, based on Blanchard and Quah (1989), amount to assumptions that neither monetary policy nor aggregate price (other aggregate demand) shocks permanently affect real output or the real commodity price level.

Together, the two contemporaneous and four long-run restrictions allow us to estimate the impact of monetary policy shocks on prices and output.

Structural Stability

Variable inclusion and identification are related. The way in which we name various estimated shocks in a model obviously depends on the quantities being modeled in the first place. While connected to the other choices, the final, more general issue concerns the period over which the empirical model is estimated. The problem is that the reduced-form relationships in the data are unlikely to be stable over any reasonable sample.[5] The problem, known widely as the Lucas (1976) critique, is that policy rule changes alter the relationship among endogenous variables in the economy.

It is easy to see why this might happen. For the sake of discussion, assume that inflation is actually determined by the following structural model:

$$(6) \qquad \pi_{t+1} = \alpha r_t + \beta_1 X_{1t} + \beta_2 X_{2t} + \omega_{t+1},$$

5 For a more detailed discussion, see section 4 of Cecchetti (1995).

where r_t is policy, w_{t+1} is a random variable, and X_{1t} and X_{2t} are measures of general economic conditions, like things that influence aggregate supply and money demand.

Next, assume that we can write down a reaction function whereby policymakers automatically change their policy control variable when economic conditions change:

$$(7) \qquad r_t = g_1 X_{1t} + g_2 X_{2t} + n_t.$$

The policymaker's role is to choose g_1 and g_2, the reaction of r_t to X_{1t} and X_{2t}. Since the g's can be zero, a policy regime need not react to the X's. The term n_t is a measure of the random component in the policy rule.

Now, consider the reduced-form regression:

$$(8) \qquad p_{t+1} = f_1 X_{1t} + f_2 X_{2t} + x_t.$$

Since $f_i = a g_i + b_i$, changes in policy, which are changes in the g's, will alter the correlation between the X's and p. In effect, the reduced-form inflation regression subsumes the monetary-policy reaction function (7), so that a change in the monetary authorities' policy rule—which may be a change in the relative weight placed on various indicators—will cause changes in (8).

As a practical matter, there are several ways to deal with the instability that may be caused by changes in monetary policy rules. First, one can use institutional information to restrict the data to a period when there were no large changes in policy procedure. Second, one can try to estimate the timing of structural breaks.[6] Alternatively, one can use time-varying parameter models, as Sims (1992) suggests. It is also possible to simply ignore the problem and use all of the available data.

Following my earlier work, I use only the past decade's data, beginning in 1984. Excepting the truncated sample period, I will ignore the problems created by the Lucas critique in all of the calculations that follow. This is an unfortunate necessity if any progress is to be made.

II. Results from Estimating the Model

Impulse Response Functions

Using monthly industrial production data for January 1984–November 1995, the CPI for urban wage earners (CPI-U), the *Journal of Commerce* index of industrial materials prices, and the federal funds rate, along with the triangular identification in equation (5), straightforward procedures yield estimates of the a_{ijk}'s, as well as a covariance matrix for these estimates. These are the time path of the impact of innovations on the model's endogenous variables. They tell us how any one of the four shocks will affect any of the four variables initially—and after several months.

It is easiest to present these results in a series of figures. Figure 1 shows estimates of 16 impulse response functions, plotted with two standard-error bands.[7] These are the response of output, aggregate prices, commodity prices, and the federal funds rate to a unit innovation to each of the four shocks.

The impulse response functions are straightforward and easy to understand. Taking the policy innovation as an example, the last column of figure 1 shows the result of an unanticipated 100-basis-point change in the federal funds rate for one month on y_t, p_t, p^c_t, and r_t over the next three years. For example, the fourth plot in the third row shows the impact of monetary policy shocks (u_t) on the aggregate price level (p_t). The estimates suggest that a one-time policy tightening—an increase in the federal funds rate—causes prices to rise slightly initially, then to fall below their original level after about six months. Over the next 30 months, the price level continues to fall. The standard-error bands on this figure imply that we are actually *very* unsure of the response. The data indicate a strong possibility that the policy tightening will result in a price-level increase.

Several additional features of figure 1 are worth noting. First, in all cases, commodity prices (second row) respond more quickly and in the same direction as aggregate prices (third row). Second, for the three e shocks, the output response seems to be more precisely estimated

6 This is the technique used in Cecchetti (1995).

7 The standard-error bands in the figure are constructed using the simple Taylor-series approximation:

$$F(\hat{b}) \approx F(b) + \frac{dF(b)}{db}\bigg|_{b=\hat{b}} (\hat{b} - b),$$

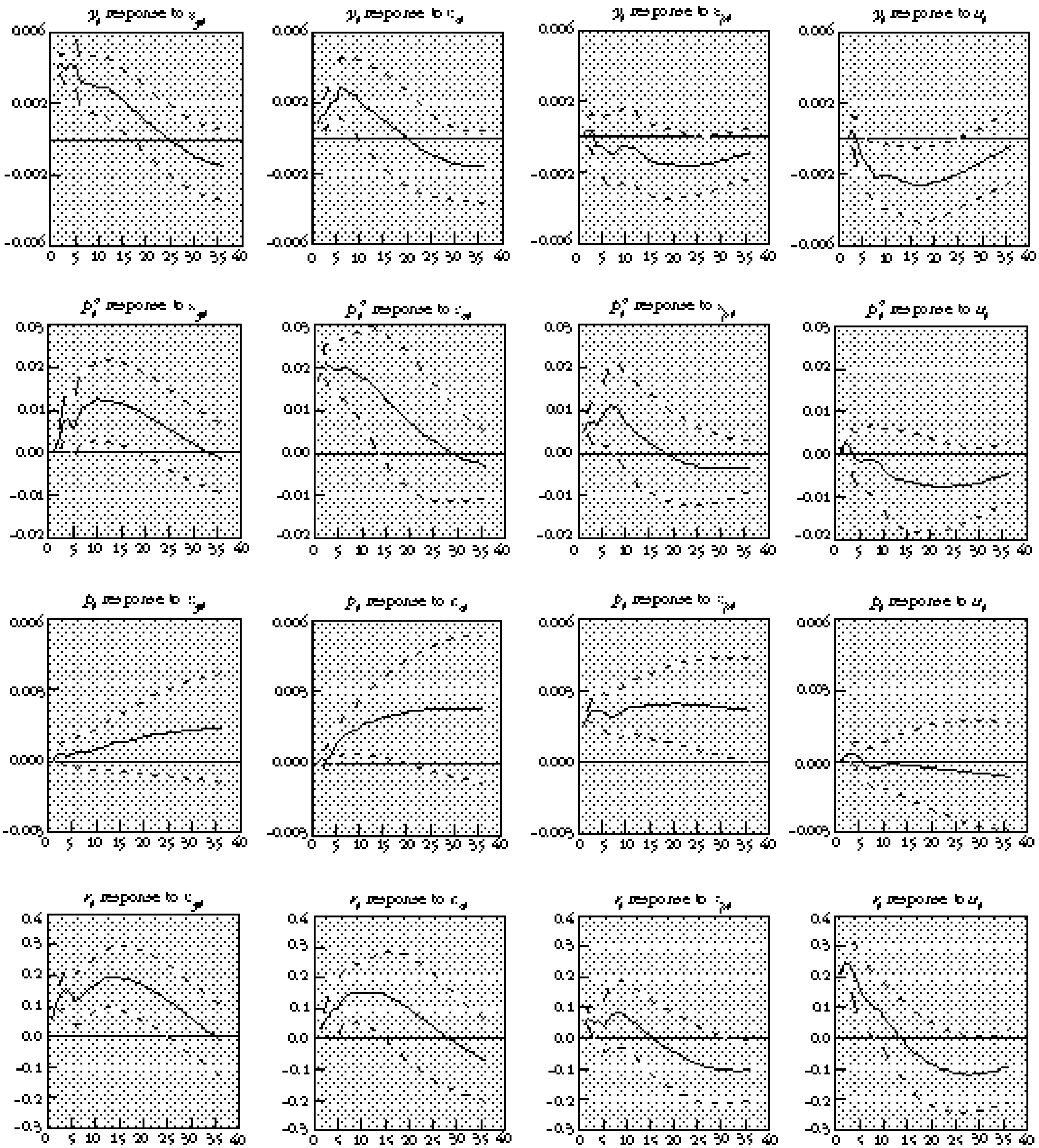
where F is any differentiable function. The variance of $F(\hat{b})$ follows immediately as

$$E\left[F(\hat{b}) - F(b)\right]^2 \approx \left[\frac{dF(b)}{db}\bigg|_{b=\hat{b}}\right]^2 Var(\hat{b}).$$

Here, we can think of the estimated impulse response functions, the $\hat{A}_{ij}$'s, as functions of the estimated reduced-form VAR coefficients, the elements of $\hat{R}(L)$. Given the estimated variance of these coefficient estimates, the variance of the $\hat{A}_{ij}$'s can be computed by numerical differentiation.

FIGURE 1

Impulse Response Functions: Triangular Identification[a]



a. Estimated response, with two standard-error bands.
NOTE: Horizontal axes are in months; vertical axes are in the change in the log per month.
SOURCE: Author's calculations.

than the aggregate price response. This second conclusion is consistent with Cochrane's (1994) observation that real output is forecastable with high R^2 at horizons of several years, and with my finding (see Cecchetti [1995]) that inflation is difficult to forecast at horizons longer than a single quarter.

It is very tempting to seek a correspondence between the shocks in this four-variable VAR and those discussed in macroeconomics textbooks. In a simple model, the basic result is that aggregate supply shocks move prices and output in opposite directions, while aggregate demand shocks move them in the same direction. With this categorization, the impulse responses shown in figure 1 suggest that all of the shocks in this model come from the demand side. While this makes intuitive sense for the monetary policy shock, it renders the other classifications unsatisfactory.

One can either accept this at face value or ask whether it might result from the identification used to generate the estimates. Taking the second possibility seriously leads to examination of an alternative identification—the one proposed by Galí being a natural choice. Figure 2 plots the impulse response functions from such a model, estimated using exactly the same data. Because of the technical difficulty associated with their construction, I do not include standard-error bands. Here, the results differ markedly. It now appears that the output shock, ϵ_{yt}, behaves like an aggregate supply shock, while the three remaining shocks, representing the aggregate price-level shock, the raw material price shock, and the monetary policy shock, lead to reactions consistent with those expected from aggregate demand shocks.

However, we can draw an important positive conclusion by comparing these two sets of identifying restrictions: The impulse response functions of the monetary policy shock are robust to changes in the identification procedure. Policy's impact on output and prices seems fairly robust to the exact methods used in estimation. Since they are easier to compute, I will now proceed using only the estimates obtained with the simpler triangular identification.

Historical Decompositions

While the ultimate goal is to use the estimated dynamic model to construct policy rules, the impulse response functions and structural innovations also allow us to compute the quantities known as "historical forecast decompositions." These allocate output and price movements into the portions accounted for by each of the structural shocks. It is easy to understand how these estimates are constructed from the structural model's equations (1)–(4):

Define the impact of the monetary policy shock on output as $H_{yu}(t)$. From equation (3), this is just

$$(9) \qquad H_{yu}(t) = \sum_{i}^{\infty} a_{34i} u_{t-i},$$

and analogously for the other shocks. Its estimated value, constructed from the parameter estimates, is

$$(10) \qquad \hat{H}_{yu}(t) = \sum_{i}^{\infty} \hat{a}_{34i} \hat{u}_{t-i},$$

where $\hat{u}_t$ is the estimated monetary policy innovation.

Figure 3 plots the decomposition of the movements in real output and aggregate prices into the components attributable to monetary and nonmonetary shocks. In constructing these, I have truncated the sum in (9) at 60 months. Because of the difficulty in identifying innovations from nonmonetary sources, it seems prudent to simply sum them together. That is, I plot the fluctuations in y_t and p_t attributable to u_t, $\hat{H}_{yu}(t)$, and $\hat{H}_{pu}(t)$, and the portion not attributable to policy, $[y_t - \hat{H}_{yu}(t)]$ and $[p_t - \hat{H}_{pu}(t)]$.

The results show that, for the past seven years, important movements in both output and prices are largely accounted for by innovations other than those coming from monetary policy. The blue line representing $\hat{H}_{yu}(t)$ and $\hat{H}_{pu}(t)$ in the figure's two panels has much less variation than the green line representing the fluctuations in y_t and p_t that are attributable to nonmonetary policy shocks. This result is particularly striking for prices, where variation seems to be driven by innovations to output, raw materials prices, and the aggregate price level itself. Aggregate supply shocks and nonmonetary aggregate demand shocks account for most of the recent movements in key macroeconomic variables.

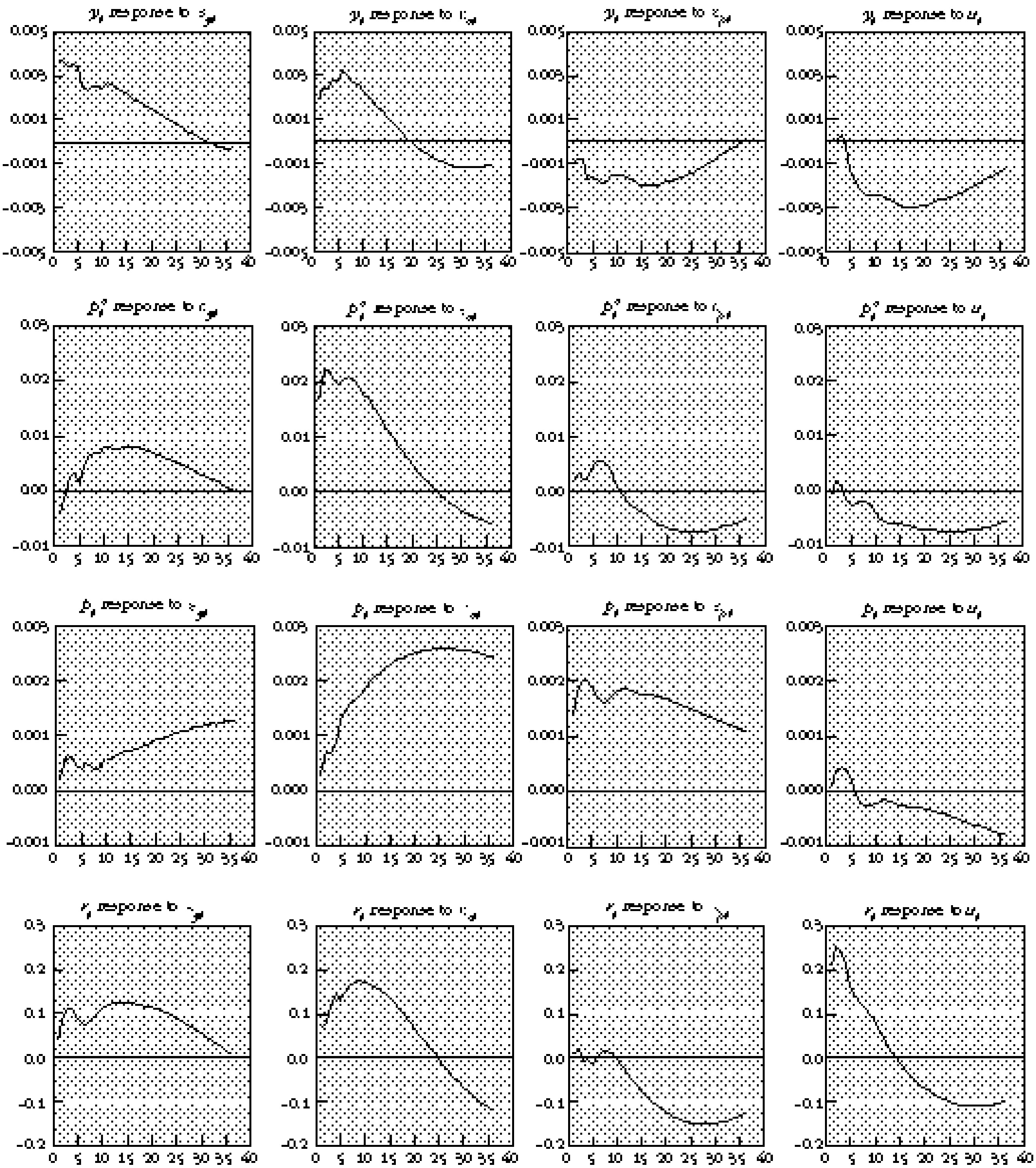
III. Formulating a Policy Rule

Issues

The main use of the empirical model described in section I and estimated in section II is to provide quantitative answers to the questions required for implementing a policy rule. To see

FIGURE 2

Impulse Response Functions: Galí Identification[a]



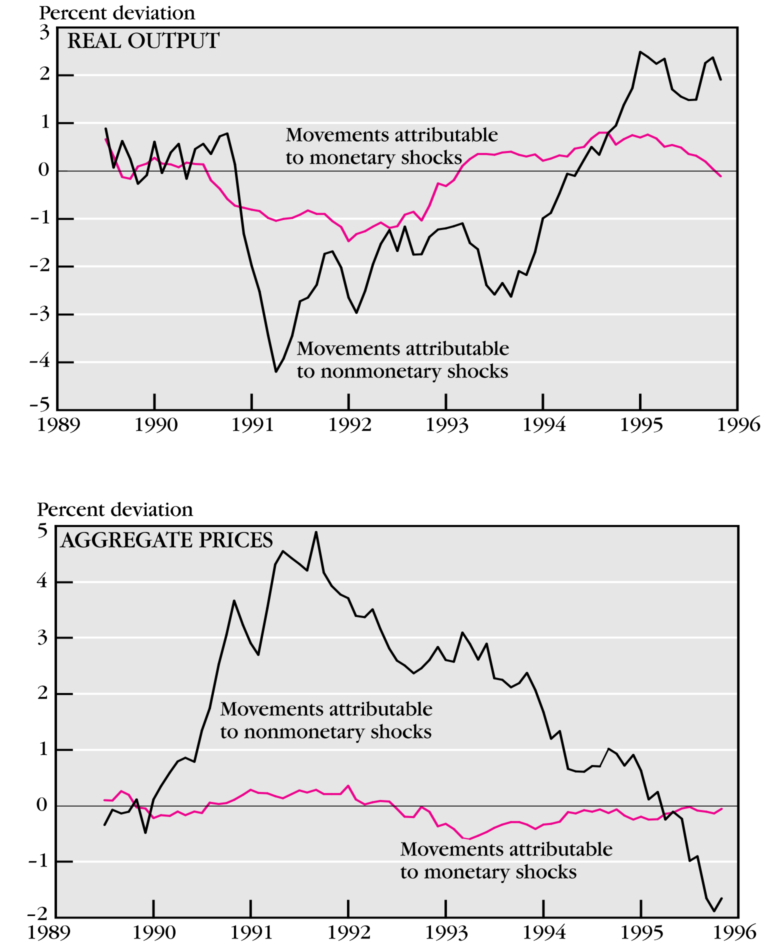
a. Estimated response.

NOTE: Horizontal axes are in months; vertical axes are in the change in the log per month.

SOURCE: Author's calculations.

FIGURE 3

Forecast Error Attributable to Various Innovations



SOURCE: Author's calculations.

how this is done, first note that the model implies estimated values for the aggregate price level and real output:

$$(11)\qquad \hat{p}_t = \hat{A}_{11}(L)\,\hat{\epsilon}_{pt} + \hat{A}_{12}(L)\,\hat{\epsilon}_{ct} + \hat{A}_{13}(L)\,\hat{\epsilon}_{yt} + \hat{A}_{14}(L)u_t,$$

$$(12)\qquad \hat{y}_t = \hat{A}_{31}(L)\,\hat{\epsilon}_{pt} + \hat{A}_{32}(L)\,\hat{\epsilon}_{ct} + \hat{A}_{33}(L)\,\hat{\epsilon}_{yt} + \hat{A}_{34}(L)u_t.$$

A policy rule is a sequence of u_t's that is constructed to meet some objective. In other words, the policymaker is allowed to pick the path of the federal funds rate to meet a particular objective.[8]

The monetary policy literature includes many discussions of the efficacy of various objective functions. Mankiw (1994) includes several papers that deal with this topic explicitly. There are two primary candidates: price-level targets and nominal-income targets. One version of these involves setting the policy instrument—the u_t's in the model—to minimize the average expected mean square error (MSE) of either inflation or nominal-income growth over some future horizon. In the inflation case, the objective function can be written as

$$(13)\qquad \min_{\{u_i\}} \frac{1}{h} \sum_{i=1}^{h} E(\hat{p}_i - p_o)^2,$$

where p_o is the log of the base-period price level and h is the policymaker's horizon. The expectation in (13) is over the sampling distribution of $\hat{p}$, which is related to the covariance matrix of the estimated coefficients in equation (11). Nominal-income targeting simply replaces the log price level in (13) with the sum of p_t and y_t.

One important distinction between the objective function (13) and more standard formulations is the treatment of parameter uncertainty. As the results in figure 1 clearly show, we are very unsure about the size and timing of price movements following innovations to the federal funds rate. When constructing a policy rule, it seems prudent to account for this lack of knowledge.

As Brainard (1967) originally pointed out, the presence of uncertainty has important implications. This is easily demonstrated in the present context. Consider a simplified version of the structural price and interest rate equations

$$(14)\qquad p_t = \epsilon_{pt} + \gamma u_t$$

$$(15)\qquad r_t = u_t,$$

where γ is a parameter. Next, take the horizon in (13) to be one period (h = 1), and the initial log price level to be zero, p_o = 0. The policy control problem then reduces to

$$(16)\qquad \min_{\{u_i\}} E[\hat{p}_i^2].$$

Substituting in the expression for p_t, this is simply

$$(17)\qquad \min_{\{u_i\}} E[\epsilon_{pi} + \hat{\gamma} u_i]^2.$$

■ 8 Feldstein and Stock (1994) examine an identical experiment, but without parameter uncertainty.

If we ignore that γ is estimated, then it is trivial to generate the policy rule. It is just

$$(18)\quad u_i^* = -\frac{1}{\hat{\gamma}}\varepsilon_{pi}.$$

Taking account of uncertainty in the estimate of γ, but continuing to assume that ε_{pt} is known, the minimization problem yields

$$(19)\quad u_i^* = -\frac{\hat{\gamma}}{[\hat{\gamma}^2 + \mathrm{Var}(\hat{\gamma})]}\varepsilon_{pi}.$$

For a given ε_{pt}, this leads to an unambiguously smaller response. In other words, imprecision creates caution, with policy reactions being muted in the face of uncertainty.

Reactions are further attenuated if policymakers attach a cost to the movement in instrument. Taking the same simple setup, imagine the modified objective function

$$(20)\quad \min_{\{u_i\}} E[\hat{p}_i^2 + \alpha \hat{r}_t^2].$$

This produces the reaction function

$$(21)\quad u_i^* = -\frac{\hat{\gamma}}{[\hat{\gamma}^2 + \mathrm{Var}(\hat{\gamma}) + \alpha]}\varepsilon_{pi},$$

which will yield an even smoother path for the interest rate than does (19).

Results

I examine results based on several policy objectives. It is worth noting that the exercise described here appears to be a gross violation of the Lucas critique. That is to say, contrary to the implications of the discussion in section I, I assume that the reduced-form correlations among output, prices, and interest rates described by equations (11) and (12) are unaffected by the change in the policymaker's reaction function.

There are two ways to defend the procedure. The first is to take the view of Sims (1982)—that parameters in these models evolve slowly enough to make Lucas-critique considerations quantitatively unimportant. The second defense is to reinterpret the exercise as an attempt to recover the objective function that policymakers were implicitly using, by trying to match the actual federal funds rate path with that implied by an optimal rule.

I report results for three different policy rules. The first, which might be termed passive, holds the federal funds rate fixed in the face of the shock. (The model makes it clear that this is not really a passive policy, since it involves shocks to overcome the estimated reaction function.) The other two, which I will call active, minimize the average MSE of either the log of the price level or the log of nominal output over a 36-month horizon ($h = 36$).[9] For each rule, I examine three experiments—one for each structural shock. In each of the nine resulting cases, $\varepsilon_{jo} = 1$ and $\varepsilon_{lk} = 0$ for $l \neq j$ and $k \neq 0$. In other words, there is a unit innovation to one of the structural disturbances in the base period, and that is all. I then construct individual estimates for the optimal response of interest rates to each of the shocks.

Figure 4 reports the implied path of the federal funds rate, aggregate prices, and industrial production for each policy objective in response to each of the three structural shocks. The fixed federal funds rate policy results in consistently higher output and prices than does either of the other two polices. The activist policies both have the same profile, whatever the source of the shock. Output and prices both rise initially, and then fall, with output dropping more than prices.

Interestingly, both of the activist policies involve raising the funds rate immediately and then lowering it slowly. This follows directly from the fact that prices respond slowly to policy innovations (see the third row of figure 1). The implication is that a policymaker who wishes to stabilize prices must respond to exogenous shocks quickly, in order to ensure that future price movements are minimized. That is the argument for the Federal Reserve's tightening up at the first sign of upward price pressure.

Comparing Targeting Objectives

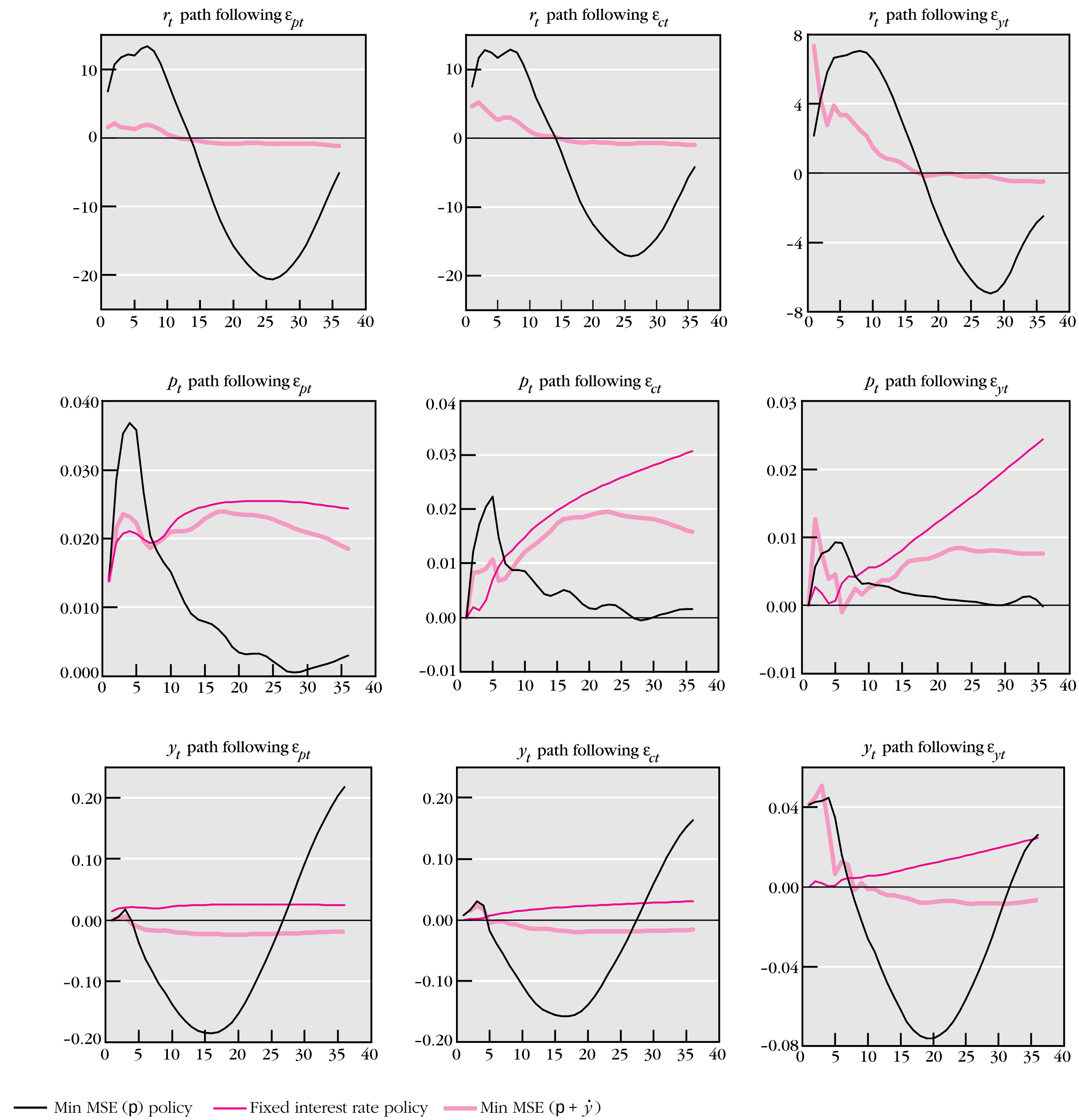
These calculations have direct implications for the debate between advocates of price-level targeting and those who favor targeting nominal GDP. To see why, I have computed the implied root-mean-square error (RMSE) for inflation and nominal income for each policy. For the price-targeting case, these are the square root of the minimized objective function (13).

Table 1 shows the results. The computations suggest that nominal-income targeting has a certain robustness, since inclusion of real output in the objective function increases the RMSE for inflation only slightly. For the case of an output shock, the increase is from 0.24 to 0.61. However, when the output shock is the

■ **9** Because the model is estimated in logs, the minimum MSE of the nominal-income policy minimized the MSE of the sum of the log of industrial production and the log of the CPI.

FIGURE 4

Interest Rate, Output, and Price Paths following Shocks, and the Policy Response



NOTE: Horizontal axes are in months; vertical axes are in the change in the log per month.
SOURCE: Author's calculations.

TABLE 1

Comparison of Policy Responses

Average RMSE of Inflation over a 36-Month Horizon

	Source of Shock		
Policy Rule	Aggregate Price	Commodity Price	Output
Fixed interest rate	2.35	1.98	1.14
Min MSE ($p + \dot{y}$)	2.15	1.50	0.61
Min MSE (p)	0.99	0.51	0.24

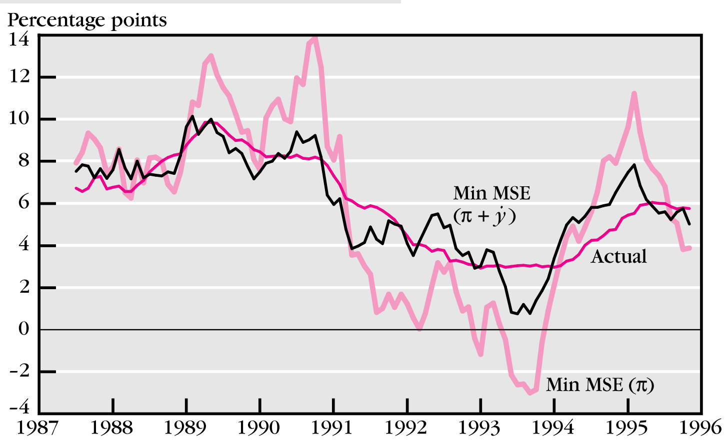
Average RMSE of Nominal Income over a 36-Month Horizon

	Source of Shock		
Policy Rule	Aggregate Price	Commodity Price	Output
Fixed interest rate	1.86	4.89	6.19
Min MSE ($p + \dot{y}$)	0.32	0.35	0.69
Min MSE (p)	0.99	10.85	4.12

SOURCE: Author's calculations.

FIGURE 5

Comparison of Optimal and Actual Federal Funds Rate Paths[a]



a. Monthly data, June 1987 to November 1995.
SOURCES: Author's calculations; and Board of Governors of the Federal Reserve System.

source of the instability, the move from price-level targeting to nominal-income targeting decreases the RMSE of nominal income substantially—from 4.12 to 0.69. In other words, the inability to estimate precisely either the impact of shocks on prices or prices' response to policy innovations argues strongly for including real variables in the objective function.

Comparing Actual and Implied Interest Rate Paths

Finally, one might ask how closely recent policy conforms to what would have been implied by either the price-level or nominal-income targeting rules plotted in figure 5. A simulated interest-rate path can be calculated by taking the estimated structural innovations, the $\hat{e}_{jt}$'s, and then computing the optimal policy responses implied by each rule before substituting the result into the equation for the federal funds rate, which is the equivalent of (11).[10]

Figure 5 compares the actual path of the federal funds rate with that implied by the estimated price-level and nominal-income targeting policies. When we examine the figure, several findings emerge. First, targeting the price level alone yields larger swings, as the funds rate reaches both higher and lower extremes. The actual funds rate is the least variable, looking like a smoothed version of the two simulated paths, but the general character of the plot suggests that the optimal policy response simply involves faster, bigger movements than those on the actual path.[11]

Figure 5, however, allows an even more interesting conclusion. From its results, it is possible to infer something about the procedures policymakers were actually following. Such a calculation does not violate the Lucas critique, since it is an attempt to recover the loss function implicit in the policy actions we actually observed.

The estimates imply that the actual funds-rate path was very similar to one that would

■ **10** Performing the calculations in this way ignores a number of elements. In particular, there is no guarantee that the policy rules generated from the artificial experiment of one unit shock in one e_{jk} at a time will be robust to sequences of shocks in all the e_{jk}'s simultaneously. One clear reason for this is that it ignores the covariance of estimated coefficients both within and across the elements of the $\hat{A}_{ij}(L)$'s.

■ **11** As one would expect, these large policy innovations result in less stable real output, highlighting that the ultimate issue in policymaking is still the relative weight of prices and output in the objective function.

have been implied by a nominal-income targeting procedure, only smoother. It is as if, over the past decade or so, the federal funds rate had been set to conform to a nominal-income targeting regime, but with policymakers attaching a cost to actually moving the funds rate. That is, the objective function that we can construct from the actual path of interest rates would minimize the sum of squared deviations in nominal income from a target path and squared movements in the federal funds rate, over a horizon of about three years.

IV. Summary

The information requirements for any policy rule are daunting. Not only do policymakers need timely information about current economic conditions, they also need forecasts of the future path of the variables they wish to control (aggregate prices and real output) and quantitative estimates of how their actions will affect these objectives.

This paper's purpose is to suggest that much of our knowledge is very inexact, and that our inability to precisely forecast the results of policy changes should make us cautious. Even more important, the fact that we have a much better understanding of the impact of our policies on real output than on prices suggests that nominal-income targeting rules are more robust than price targeting rules. From a purely pragmatic viewpoint, someone who cares about nominal income is made substantially worse off by moving to a price-level target, which destabilizes real output considerably. Thus, practical issues make a strong argument for nominal-income targeting.

In addition, we have seen that the actual path of interest rates over the past decade is very similar to that implied by a nominal-income targeting rule, albeit one in which interest rate movements are viewed as costly. By comparing the actual interest-rate path with the path implied by the nominal-income targeting rule, we see that policymakers have smoothed interest rate movements more than the rule would have dictated, but not by much.

Appendix: Identification

To understand the more general issues of identification, it is useful to rewrite the four-equation model [(1)–(4)] in a more compact form:

$$\text{(A1)} \qquad x_t = A(L)e_t,$$

where x_t and e_t are now vectors, and $A(L)$ is a matrix of lag polynomials. We can also write the model in its more familiar VAR reduced form as

$$\text{(A2)} \qquad R(L)x_t = \eta_t,$$

where $R(0) = I$, the η_t's are i.i.d. (implying that they are orthogonal to the lagged x_t's), and $E(\eta\eta') = \Sigma$. It immediately follows that $A(L)e_t = R(L)^{-1}\eta_t$. This allows us to write $A(0)e_t = \eta_t$, and $A(L) = R(L)^{-1}A(0)$. As a result, given estimates of $A(0)$, $R(L)$, and η, we can recover estimates of both the structural innovations—the e_t's—and the structural parameters—the components of $A(L)$.

The issue of identification is the problem of estimating $A(0)$. To show how this is done, note that $A(0)E(ee')A(0)' = \Sigma$, where $E(ee')$ is diagonal by construction. Normalizing $E(ee') = I$, we obtain the result that $A(0)A(0)' = \Sigma$. In a system with n variables, Σ has $[\frac{n(n+1)}{2}]$ unique elements, and so complete identification requires an additional $[\frac{n(n-1)}{2}]$ restrictions. In a four-variable model, six more restrictions are needed. This is a necessary but not sufficient condition for identification. Sufficiency can be established by proving that the restrictions lead to construction of an $A(0)$ matrix that is invertible.

The long-run restrictions of the Galí-style identification can be understood by defining $A(1)$ as the matrix of long-run effects computed by summing the coefficients in $A(L)$. That is, the (i,j) element of $A(1)$ is

$$A_{ij}(1) = \sum_{k=0}^{\infty} a_{ijk}.$$

There are two long-run restrictions. The first is that the impact of ε_{pt} and u_t on y_t is transitory, and so $A_{31}(1) = A_{34}(1) = 0$. The second is that ε_{pt} and u_t have no permanent impact on the relative price of commodities, $(p_{ct} - p_t)$, that is, $A_{11}(1) - A_{21}(1) = A_{14}(1) - A_{24}(1) = 0$.

References

Bernanke, B. S., and I. Mihov. "Measuring Monetary Policy," Federal Reserve Bank of San Francisco, Working Paper No. 95-09, March 1995.

Blanchard, O. J., and D. Quah. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review,* vol. 79, no. 4 (September 1989), pp. 655–73.

Brainard, W. C. "Uncertainty and the Effectiveness of Policy," *American Economic Review,* vol. 57, no. 2 (May 1967), pp. 411–25.

Cecchetti, S. G. "Inflation Indicators and Inflation Policy," in B.S. Bernanke and J.J. Rotemberg, eds., *NBER Macroeconomics Annual 1995.* Cambridge, Mass.: MIT Press, 1995, pp. 189–219.

Christiano, L. J., M. Eichenbaum, and C. L. Evans. "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics,* vol. 78, no. 1 (February 1996a), pp. 16–34.

______, ______, **and** ______. "Identification and the Effects of Monetary Policy Shocks," in M.I. Blejer et al., eds., *Financial Factors in Economic Stabilization and Growth.* New York: Cambridge University Press, 1996b, pp. 36–74.

Cochrane, J. H. "Shocks," *Carnegie–Rochester Conference Series on Public Policy,* vol. 41 (December 1994), pp. 295–364.

Feldstein, M., and J. H. Stock. "The Use of a Monetary Aggregate to Target Nominal GDP," in N.G. Mankiw, ed., *Monetary Policy.* Chicago: University of Chicago Press, 1994, pp. 7–70.

Galí, J. "How Well Does the IS-LM Model Fit Postwar U.S. Data?" *Quarterly Journal of Economics,* vol. 107, no. 2 (May 1992), pp. 709–38.

Lucas, R. E., Jr. "Econometric Policy Evaluation: A Critique," *Journal of Monetary Economics,* vol. 1, no. 2 (Supplementary Series 1976), pp. 19–46.

Mankiw, N. G., ed. *Monetary Policy.* Chicago: University of Chicago Press, 1994.

Sims, C. A. "Macroeconomics and Reality," *Econometrica,* vol. 48, no. 1 (January 1980), pp. 1–48.

______. "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity,* no. 1 (1982), pp. 107–52.

______. "A Nine Variable Probabilistic Macroeconomic Forecasting Model," Cowles Foundation for Research in Economics at Yale University, Discussion Paper No. 1034, June 1992.

The Reduced Form as an Empirical Tool: A Cautionary Tale from the Financial Veil

by Ben Craig and Christopher A. Richardson

Ben Craig is an economist at the Federal Reserve Bank of Cleveland, and Christopher A. Richardson is a graduate student of economics at Indiana University. The authors have benefited from the generous help and advice of Charles Carlstrom.

Introduction

Economic data usually influence policy through a reduced-form analysis. Using such an analysis, the researcher generally poses an empirical relationship between an outcome variable, such as a firm's total investment, and a policy variable, such as the design of a particular tax. This relationship serves as a point of departure in the analysis. Explicit assumptions about behavior that underlie the relationship are not emphasized; rather, the researcher asserts that the "data do the talking." Policy implications, where they exist, are directly observed in the pattern estimated in the data. Most empirical analyses of policy questions follow a reduced-form strategy.[1]

It is easy to understand why a reduced-form approach might, at first glance, appear to be the best way to analyze policy. It is a simple methodology, and thus can more easily keep track of what is happening during the complicated process of analyzing data. One does not need to specify a sophisticated and consistent model of behavior to use this approach. Further, the answers embodied in the model estimates may accord with a wide variety of behaviors that could be true of the firm.

A different approach to estimating the effect of taxes would be to specify a model of optimizing behavior on the part of the firm and to model the tax policy as a set of constraints on this optimizing behavior. A simplistic reason for preferring the reduced-form approach is that economists are interested only in the overall effect of a proposed tax policy on investment. Why should we care about the intermediate steps by which a tax will affect the firm?

How successful is the reduced-form approach at testing a behavior or measuring a policy effect? Given that we are never shown the truth behind the mystery, this paper will examine the history of an economic question that has been subjected to 35 years of intense scrutiny: Does corporate financial structure affect real investment? The empirical answer to this question, which lies at the heart of corporate financial economics, has heavily influenced every tax reform bill since the 1960s.

■ 1 *Reduced form* has a different meaning here than in simultaneous equations estimation, where a reduced form is estimated by regressing an endogenous variable on all of the exogenous variables in a system of equations. We use the term in a wider context, where the pattern in the data—not an assumed behavioral structure—forms the point of departure for estimation.

The initial econometric strategy was to follow a pure reduced-form approach. How well have the results of this research program held up to further scrutiny?

Modigliani and Miller (1958) provide the first theoretical model showing the influence of corporate debt structure on investment. In the world they portray, perfect capital markets, coupled with symmetric information about the investment prospects of the firm, the investors, and the lenders, mean that the firm's debt level is irrelevant to the amount of investment it undertakes.

One reduced-form approach would be to directly examine the empirical relevance of the Modigliani–Miller (hereafter "MM") hypothesis that with perfect capital markets, no taxes, and a given investment policy, capital structure is irrelevant to firm value. As a consequence, neither capital structure nor dividends should affect investment behavior. The MM propositions provide the following broad empirical prediction: In a properly specified regression of investment on the debt/equity ratio, dividends, and other covariates, the coefficients of debt/equity and dividends should equal zero. A reduced-form estimating strategy uses this prediction as the point of departure.

The following two sections discuss the history of tests of this hypothesis from both a cross-section and a time-series perspective. Next, we step back and explore the reasons for the pattern in the early reduced-form estimates through a simple structural model. We then look at what we have learned about whether a tax policy can affect investment through its influence on a company's financial structure. We conclude with the object lessons that accompany 35 years of intensive research on this topic—lessons that could be applied to other situations where the reduced form is used to help shape policy. What did we first believe the data were telling us, and how did these beliefs change under close scrutiny? After all this time spent researching a single hypothesis, what limitations in our knowledge may be embedded in the reduced-form approach?

I. Cross-Section Regression Tests of the MM Hypothesis

The clear and simple MM hypothesis that there is no relationship between financing and real capital investment seems to lend itself easily to cross-section regressions. The early reduced-form models assume away the importance of differential corporate and personal income taxes, which are a clear violation of the original MM statement. Thus, they jointly test the MM model and the hypothesis that the income tax structure is irrelevant to the effect of financial structure on real investment. We will treat the two tests separately later in this paper. The test of the joint hypothesis measures the statistical significance of financial variables in an investment equation where the dependent variable is capital investment and the independent variables are measures of a firm's financial position, which may include its debt/equity ratio, cash flow, and dividends. The hypothesis of no relationship between financing and investment is rejected if the coefficients of the debt/equity ratio and dividends are statistically close to zero. A simple regression is not adequate here because both dividend payments and the firm's debt are endogenous. Thus, absence of a correlation between the debt/equity ratio and investment is not necessarily evidence that the MM hypothesis holds.

To alleviate this problem, early cross-section studies specified instruments in a system that estimated investment (I), dividends (D), and new debt (ND) equations of the general form

$$
\begin{aligned}
I_{it} &= \alpha_0 + \alpha_1 D_{it} + \alpha_2 ND_{it} + \alpha_3 X_{it} + \varepsilon_I \\
D_{it} &= \beta_0 + \beta_1 I_{it} + \beta_2 ND_{it} + \beta_3 Y_{it} + \varepsilon_D \\
ND_{it} &= \gamma_0 + \gamma_1 I_{it} + \gamma_2 D_{it} + \gamma_3 Z_{it} + \varepsilon_{ND},
\end{aligned}
$$

where i and t are firm and time subscripts, the ε's are statistical error terms, and X, Y, and Z are vectors of exogenous explanatory variables. For the investment equation to be identified (so that we are estimating a separate equation for investment, not a hodgepodge of all three equations), the vectors Y and Z must contain variables that are not included in X. It is this process of identification that proved so problematic in the early reduced-form studies. What exogenous variable affects dividends and debt levels but does not influence investment behavior?

It is important to note here that the MM hypothesis is not a theory of investment, but of why a firm's financial structure does not influence investment. The estimating system of equations that test the MM hypothesis must include a theory of investment (even if it is implicit) to control for its endogeneity. Thus, the reduced form is a joint test of both the MM hypothesis and an underlying theory of investment. For example, if the researcher holds investment opportunities constant through using a measure of Tobin's q, then the test of the MM hypothesis also tests whether Tobin's q is an empirically useful model of investment behavior.[2] Hence, the test is only as good as the theory of investment.

The early studies used identifying instruments[3] that included profits, proxies for firm size and taxes, and firm and industry dummy variables to allow for fixed firm and industry effects.[4] (See, for example, Dhrymes and Kurz [1967], McDonald, Jacquillat, and Nussenbaum [1975], and McCabe [1979].) These studies, like so many modern consulting reports, argue by assertion—for example, the profit level should affect dividends but not investment. Unfortunately, a researcher's assertion that a variable is an instrument does not necessarily make it so. A reduced form offers few checks as to whether the assertion reflects reality.

Empirical tests of the MM irrelevance hypothesis during the 1970s and early 1980s, though more advanced econometrically, still came up short in modeling differences in the financial environment firms face. Although the studies varied in their conclusions, all suffered from the lack of a convincing instrument that would control for a firm's investment opportunities. Again, the studies could not adequately address the fact that firms with better investment opportunities might choose higher levels of debt. McDonald, Jacquillat, and Nussenbaum estimate cross-section models using ordinary least squares (OLS) and two-stage least squares (2SLS), as does McCabe, but their conclusions differ. McDonald et al. find support for the MM propositions, while McCabe does not. Because investment opportunities surely vary across firms and certainly affect investment independently of financial structure, these early studies were never conclusive tests of the MM irrelevance hypothesis.

Another reason for conflicting conclusions among the early empirical studies seems to lie in the differences in equation specification. McDonald et al., like many other researchers before McCabe, estimate investment as a function of contemporaneous variables only. Because it is likely that the decision to invest today will depend in part on financial decisions made previously, excluding lagged financing and dividend variables from an investment equation results in a misspecification.

How was one to choose between these early studies? If they had been structural, a researcher could affirm that a particular study was the most convincing if it had more believable parameter estimates (for example, if it generated rate-of-return estimates of the same general magnitude as the interest rate). A classic indication that a system is identified improperly (that is, by false assumptions) is that estimates of the individual equations yield parameters that make little sense economically. One reason the earliest tests of the MM hypothesis seemed, on balance, to support the theory was that the estimates which rejected MM had the "wrong" expected sign for the dividend equation. This seemed to indicate that the studies which did not reject the hypothesis used more convincing instruments. One pitfall of a simple reduced-form strategy is that it yields so few checks of whether an estimated parameter makes economic sense.

Subsequent cross-section studies made significant improvements over previous work. For example, Peterson and Benesh (1983) estimate a system of three equations similar to that used in earlier studies (adding a lagged profit variable to the investment equation and a lagged dividend variable to the dividend equation), but in addition to estimating the standard OLS, 2SLS, and 3SLS models, they also conduct MM hypothesis tests on the reduced-form equations by estimating a seemingly unrelated regressions (SUR) model. Their SUR results corroborate the 2SLS and 3SLS findings, which reject the null hypothesis that financing and investment decisions are independent. The lagged profit variable serves as a proxy (albeit an imperfect one) for investment opportunities, which makes the rejection of the MM irrelevance hypothesis somewhat more convincing.

The use of lagged profits suffers from a problem common to all studies that rely on lagged variables for identifying instruments. Although it is true that lagged profits are approximate measures of investment opportunities, they may also affect both dividends and debt in the same ways that these variables were correlated with the original contemporaneous error term. It is not clear that simply including the lagged profit term will correct the original statistical bias.

Most recent reduced-form cross-section models reject the MM hypothesis. (See, for example, Gilchrist and Himmelberg [1995].) However,

2 Tobin's q is defined as the ratio of the market value of capital to the replacement cost of capital.

3 An *instrument* is a variable that is correlated with a variable on the right-hand side of the equation (in this case, corporate debt or dividends) without being correlated with the statistical error term. An *identifying instrument* in this case is one that is correlated with the right-hand variables without being included as a right-hand variable. Thus, it may be included in the equation where dividends or debt are left-side variables, but it must be excluded from the original investment equation.

4 In some cases, X, Y, and Z contain the same variables.

these models often suffer a distressing lack of robustness to econometric specification. This makes precise determination of the estimated reduced-form parameters problematic. Further, the reduced-form approach does not present us with an easy point of departure for determining the correct econometric specification through convincing tests. We also lack a consensus on parameter estimates that are specific and precise enough to be more useful in a policy context than are cross-section reduced-form models. Can we glean additional evidence on the empirical validity of the MM irrelevance hypothesis from time-series patterns in the data?

II. Granger Causality Tests

Given the difficulty of pinning down a convincing set of instruments to tease out that part of the correlation of debt and investment stemming from debt's possible impact on investment, some researchers have tried to determine the causality by studying the timing of debt and investment. Thus, if investment precedes debt, the correlation may be spurious because the firm, seizing its more potent investment opportunities, creates more debt, whereas the less fortunate firms do not have as much debt. This would be the case when the higher debt level was due to more investment opportunities for the high-debt firm. The test of a causal relationship between the variables proposed by Granger (1969) says that if a variable or event X (a change in a financial variable) causes another variable or event Y (a change in investment), then X should precede Y. The test involves measuring the power of lagged values of X in predicting Y. A test of whether debt affects investment is connected to whether corporate debt "Granger causes" investment.

Smirlock and Marshall (1983) perform Granger causality tests on a sample of 194 firms from 1958 to 1977. Using annual data on dividends and investment, they fail to reject the null hypothesis of no Granger causality for the aggregate sample of firms. Causality tests on each of the 194 firms' series do not reject the null any more often than would be expected by chance, leading the authors to conclude that their results support the MM irrelevance hypothesis.

True to the pattern of cross-section tests of the MM hypothesis, the early test did not hold up to later scrutiny. It is imperative that enough variables be included in a Granger causality study so that nearly identical firms are being compared. For example, Smirlock and Marshall omit a financing variable, so that the analysis compares noncomparable firms that differ in *precisely* that dimension which the causality test requires to be the same. Mougoue and Mukherjee (1994) address this issue by including a long-term debt-financing variable in their causality tests. They find that dividend and investment growth rates Granger cause each other negatively, long-term debt and investment growth rates Granger cause each other positively, and debt and dividends Granger cause each other positively, thus rejecting the MM irrelevance principle.[5] If the reduced-form test is to be appropriate, some sort of implicit structure must underlie it. In this case, the researchers had to have an idea about which financing variables were important so that the Granger causality could test comparable firms.

Although Mougoue and Mukherjee's Granger causality tests can detect significant interactions among investment, debt, and dividend variables, they do not tell us much else. That dividends and investment Granger cause each other simply means that a motion in one precedes a motion in the other. Which comes first, the investment chicken or the dividend egg?

Moreover, it is somewhat ironic that Mougoue and Mukherjee's causality tests may also suffer from a misspecification bias due to the omission of a proxy for investment opportunities, such as cash flow. If internal funds are omitted from the system of equations, the observed negative causality from dividends to investment may actually stem from a negative causality from dividends to retained earnings and a negative causality from retained earnings to investment. The MM irrelevance hypothesis would still be rejected, but for different reasons. Although more properly specified equation systems may be useful in illustrating the existence of causal relationships, it appears that Granger causality tests have only limited utility in distinguishing among the different hypotheses of how, why, and to what degree financing and real investment decisions interact.

In addition, Granger causality tests suffer from a difficulty related to the forming of expectations. If debt Granger causes investment, the interpretation is that the corporate structure effects a change in investment behavior. However, expectations about investment

5 It is assumed here that firms use borrowed funds to finance future investment or to increase dividend payments. Because the variables are expressed as changes in logarithms (growth rates), positive bidirectional causality between debt and investment does not support or refute the presence of financing constraints, as it might if debt and investment were expressed in levels.

opportunities could just as well influence both investment and corporate debt levels, but affect debt sooner because debt levels adjust more quickly. Tests that center on the timing of debt and investment thus provide weak evidence on the relevance of the MM hypothesis.

Interestingly, the pattern of evidence in the Granger causality tests is the same as the pattern in the cross-section regression results. Initially, the evidence seemed to support the MM hypothesis. However, closer scrutiny and clearer identifying assumptions tend to reject the hypothesis. Even current studies are unable to provide more information than a crude rejection of the hypothesis. Precise parameter estimates needed for policy prediction seem to require a different estimation strategy. Why do we observe this pattern in the reduced-form estimates? It is not clear which part of the joint hypothesis—perfect capital markets or the empirical irrelevance of the personal and corporate income tax structure—is being rejected by the above tests. To further define the two hypotheses, a simple heuristic model is needed.

III. A MM Structural Model

In this section, we explore a model in which the underlying behavior of firms generates the data. A simple statement of the model will clarify the measurement problems inherent in testing the MM hypothesis with cross-section or time-series data. A MM firm chooses the levels of investment, I_0, in a project that will pay $F(I_0)$ dollars for each period in the future, $F'(I_0) > 0$, $F''(I_0) < 0$, so that the firm receives

$$(1) \quad \sum_{t=1}^{\infty} \frac{F(I_0)}{(1+\rho)^t} = \frac{F(I_0)}{\rho}$$

from the investment, where ρ represents the interest rate.[6] The firm starts with a predetermined amount of cash, C, and must decide how much of this cash to pay out in dividends, C_d, and how much to reinvest, C_I $(C = C_d + C_I)$. The firm can also issue debt, D, to finance the investment. A tax rate is imposed on a corporation at rate τ_c and on individuals at rate τ_p.

MM's first observation is that the market value of the shares, S, is just a tax-adjusted value of the investment payoffs (including the corporate cash paid out today, C_d) minus the value of debt, or

$$(2) \quad S = (1-\tau_p)C_d + (1-\tau_c)\left[\frac{F(I_0)}{\rho} - D\right].$$

The firm maximizes S with respect to the amount of investment subject to the financing constraints $I_0 = C_I + D$ and $C = C_d + C_I$. A simple substitution gives

$$(3) \quad S = (1-\tau_p)(C - C_I) + (1-\tau_c)\left[\frac{F(C_I + D)}{\rho} - D\right],$$

with first-order conditions

$$(4) \quad F' = \rho$$

and

$$(5) \quad F' = \rho\,\frac{(1-\tau_p)}{(1-\tau_c)},$$

corresponding to investment financed out of debt and cash, respectively.

If the personal tax rate is equal to the corporate tax rate (or $\tau_p = \tau_c$, which nests the special case of MM's no-tax scenario), the first-order conditions make it apparent that the expense of an additional dollar of investment is the same however it is financed, and that the firm finances from either debt or retained earnings until the marginal benefit from investment is equal to the outside rate of return (or $F'[I_0] = \rho$).

It is important to note that the first-order condition in this case simultaneously says two things about the firm's behavior. First, a firm's decision to finance a given level of investment out of debt or retained earnings is irrelevant: The firm is indifferent between the two. Second, the investment level is determined by the rule that the firm invests until the marginal benefit of investment is equal to the interest rate.[7] The level of debt says nothing about the value of the firm except that it has traded debt for dividends at a rate of one for one. Investors who are paying for a share of the company and who might prefer a higher level of debt in their portfolio can continue holding shares of this firm, but elect to borrow more on the outside market to increase the debt level within their own portfolios. This and other similar possible arbitrages force the share value to treat debt and retained earnings symmetrically.

In a world where the corporate tax rate is higher than the personal tax rate, $\tau_c > \tau_p$, the firm's rule is to finance until the return on

6 The MM results do not depend on risk neutrality, a constant stream of benefits, or a constant discount rate. These are assumed here for expositional convenience.

7 In the MM exposition, the firm invests until the marginal benefit equals the rate of return for the firm's risk class. The theory also shows that investment financed out of new equity issues is equivalent to investment financed out of retained earnings or debt.

investment is equal to ρ. The investment is financed entirely out of debt and all of the cash appears as a dividend. Financing out of debt rather than retained earnings costs less because debt payments are fully deductible. MM (1963) makes the point that the tax advantages of debt financing (interest payments are deductible as a cost in calculating corporate taxes) imply that investment should never be financed out of retained earnings in a pure example of their model. Indeed, because of the tax advantages, the firm prefers to borrow more than its investment amount to finance a larger dividend. Clearly, a more complicated model is needed to explain why a corporation chooses one method of financing over another. However, the simpler model may still be adequate to explain the level of investment if the data show no relationship between that level and corporate financial structure.

If the corporate tax rate is less than the personal tax rate, then cash is the relatively less expensive investment source. The firm will use only cash to finance investment unless the marginal return on investment after all of the cash is used (and the dividend is zero) exceeds the cost of financing additional investment out of debt (or $F'[C] > \rho$). At this point, debt becomes the marginal investment source, with the first-order condition given by equation (4). If $F'(C) < \rho$, then debt will equal zero, and only cash will be used to finance investment, until equation (5) is satisfied.

The top federal individual tax rate and the top federal corporate tax rate are currently about the same (39.6 percent versus 38 percent). However, in many states, the top corporate rate is higher than the top individual rate (in Ohio, the respective figures are 8.1 percent and 4.1 percent). Based on the above analysis, one would expect to find corporations financing investment entirely out of debt and never using retained earnings for this purpose. Clearly, this is not the case, as firms use both debt and equity financing. One reason companies do not rely solely on debt is that outside credit constraints (or the costs of bankruptcy) may make the marginal cost of debt rise as the total level of debt increases. In other words, the market value of debt may decrease the firm's value faster at higher debt levels because high debt may alert the capital markets that the firm is less likely to survive, or because lenders become less willing to lend to firms that could be hit with bankruptcy costs. In this world, the value of the shares might be written as

$$(6) \qquad S = (1-\tau_p)(C-C_I) + (1-\tau_c)\left[\frac{F(I_0)}{\rho} - [D+\delta(D)]\right].$$

The parameter δ is the increasing cost of debt not captured in the interest rate, where $\delta(0) = 0$, $\delta' > 0$, and $\delta'' > 0$. The new rule for debt-financed investment is

$$(7) \qquad F'(I_0) = \rho\,[1 + \delta'(D)].$$

The first-order condition for investment financed out of equity is the same as in equation (5). The rule for investment explicitly makes the amount of investment a decreasing function of the firm's debt level if the firm finances out of debt. Similar to the MM model, the first-order conditions can generate corner solutions in which the firm finances investment either completely out of debt or completely out of cash. For example, if $\tau_c \le \tau_p$, the firm will finance up to its total cash holdings out of equity, then finance out of debt only if the marginal return on investment at that point is greater than or equal to ρ. If the corporate tax rate is greater than the personal tax rate (as it is for most U.S. corporations), then the investment rule is more complicated. The firm will finance out of debt only if $[1 + \delta'(D)] < (1 - \tau_p)/(1 - \tau_c)$; that is, if the marginal cost of increasing the firm's debt burden is small enough. If this is not the case, firms will use both debt and cash to finance their investment projects. First-order conditions for investment financed out of cash remain the same (equation [5]), so that the equation determining the debt level, if both debt and cash are used to finance investment, is

$$(8) \qquad [1 + \delta'(D)] = \frac{(1 - \tau_p)}{(1 - \tau_c)}.$$

It is important for policymakers to know whether a world represented by MM or an environment of substantially increasing marginal cost of debt, crudely represented by equation (6), best reflects investment behavior. One easy reduced-form test of the MM assumptions in the earlier cross-section studies was to examine whether investment was negatively correlated with debt. If the study detected no negative relationship, then the conclusion might be that equation (6) did not make empirical sense. However, as the following example illustrates, lack of correlation between debt and equity might occur in a world that is very non-MM.

Suppose our sample consists of two types of firms that differ only in their set of investment opportunities. Each type faces an investment payoff of $a_i F(I)$, where $a_1 > a_2$. The empirical researcher observes only the debt and investment outcomes of the two firms. The firms face a very non-MM world, one in which increasing debt discourages investment, represented by equation (6). We further assume that both debt and equity are used to finance investment, so that equation (8) holds. First-order conditions for the two firms are $F'(I_0) = \frac{\rho}{a_i}[1 + \delta'(D)] = \rho(1 - \tau_p)/a_i(1 - \tau_c)$. The behavior rule gives the following outcome: $I_1 > I_2$ and, if both firms finance investment out of some of their retained earnings, $D_1 = D_2$.[8] A simple regression of investment on the firm's debt level would lend support to the MM hypothesis of the irrelevance of debt for the level of investment, even though the data are generated by a behavior where the debt level, ceteris paribus, discourages investment. Clearly, lack of a simple correlation between debt and investment is not enough to test the appropriateness of the theory. If the potential cash available to the individual firms, C, is unobserved by the empirical researcher (as is likely), then the dividend amount may also be uncorrelated with the investment level. The problem is that the corporate financial instruments are behavior variables chosen by the agents, not experimental variables applied by the researcher.

Any estimation must take into account that both investment and corporate financial structure are caused by the environment facing the firm, and that the available data contain very little of the information needed to reconstruct the decision process for each firm's investment and corporate financial structure, even if all of the correct variables are included. The key to the earlier studies lay in finding sufficient instruments to control for the different investment opportunities represented by a_i and for the fact that debt was a behavioral variable chosen by the firm. Poor instrument choice was bound to lead to poor estimates. In this estimating context, the underlying structure of behavior and a clear notion of what is generating differences across observations are needed to formulate a useful reduced form.

The problems with reduced-form analysis are clear from this example, yet researchers do not necessarily have the data to conduct a full-blown structural estimation. Despite these limitations, we can profit from structural models by using them to devise a test that can help uncover some of the important factors driving firms' investment decisions.

IV. Tax Effects and the Investment/Financing Relationship

Early reduced-form empirical work on the tax effects of the investment/financing relationship, such as Long and Malitz (1985), Titman and Wessels (1988), and Fischer, Heinkel, and Zechner (1989), failed to find economically or statistically significant effects, just as early reduced-form studies failed to find a link between corporate financial structure and investment. These early nonstructural studies had an important influence on the policy debate, particularly when federal tax reform was discussed. For example, when the Economic Recovery Tax Act of 1981 (ERTA) was being debated, it was theoretically understood that in a credit-constrained world, the investment tax credit might yield a strong substitution effect as firms changed their investment funding from debt financing to retained earnings. This was not considered important because the early reduced-form estimates indicated that the effect of taxes on corporate financial structure was negligible.[9]

Subsequent, more careful work has generally found evidence of a significant tax effect. For example, MacKie-Mason (1990) states that earlier studies suffered from a failure to fully consider the impact of a firm's tax shields (tax deductions or investment tax credits) on its effective marginal tax rate. He notes that if a firm has no taxable income, any additional tax shields it receives will have no impact on its marginal tax rate. The marginal rate will be affected only if tax shields lower taxable income to zero. By taking this point into account and investigating incremental financing decisions using discrete-choice models instead of debt/equity ratios, MacKie-Mason finds evidence that firms with high tax-loss carryforwards are less likely to rely on debt. This is certainly consistent with both the theoretical models of MM and the debt-constraint model, which predict that as the corporate tax rate decreases, debt should shrink.

This more recent finding of a significant tax effect forces the reduced-form research to be more careful in defining its hypotheses. How do taxes influence investment? They could

■ 8 This follows directly from the relation $\frac{(1 - \tau_p)}{(1 - \tau_c)} = 1 + \delta'(D)$.

■ 9 See Trezevant (1994), which discusses the contemporary debate surrounding ERTA. The author finds a significant substitution effect in taxes.

affect it directly through a change in the post-tax price of real investment, or indirectly through a change in corporate financial structure, as demonstrated in the previous section. The indirect influence is the one of interest to corporate finance. Separation of the indirect financial effect from the direct real-price effect requires a clarity that makes reduced-form estimation look more like structural estimation.

This clarity is seen in more recent research that concentrates on the impact of taxes on corporate financial structure. Givoly et al. (1992) and Cummins, Hassett, and Hubbard (1994) find evidence of a relationship between debt and taxes. The Givoly study empirically examines the response of business firms to the Tax Reform Act of 1986 (TRA). The authors find evidence of a substitution effect between debt and nondebt tax shields, as hypothesized by DeAngelo and Masulis (1980). In addition, both corporate and personal tax rates appear to affect leverage decisions.

Givoly et al. provide a good example of how forming an implicit structure about the effect of taxes provides a precise hypothesis for testing with a reduced-form estimation strategy. Consider how they use their structure and their knowledge of TRA specifics to develop simple statistical hypotheses. For example, they use only 1987 data to describe TRA's effect, because they assert that the Act was surrounded by uncertainty until its actual passage by the Senate. Their test year was 1987, and their control years were 1984 and 1985, before any tax reform legislation was introduced. Although firms might have behaved in anticipation of a new tax structure, it is unlikely that this Lucas effect would be of overriding importance in the statistical results.

Givoly et al. test their hypotheses involving tax code changes by estimating standard cross-section OLS regressions of the change in leverage on the firm's effective tax rate, nondebt tax shields, dividend yield, Tobin's *q*, firm size, business risk, and changes in depreciation and investment tax credits.[10] The authors are able to reach definitive conclusions about the effect of the TRA using cross-section analysis because they state their hypotheses carefully. For example, the TRA greatly reduced the statutory corporate tax rates, so that firms faced more similar rates. Thus, their Hypothesis 1 states that in response to the decline in the statutory corporate tax rate, firms with a high marginal effective rate will decrease their leverage more than will firms with a low marginal effective rate. In other words, the decline in the effective corporate tax rate will have a greater impact on firms with high marginal tax burdens. Low effective tax rates imply a low tax advantage of debt and result in a smaller decrease in leverage stemming from a cut in the statutory tax rate. Hence, the relationship between the effective tax rate and changes in leverage should be inverse.

Notice how the hypothesis embodies a solid underlying structure of how the firm reacts to a tax change. This structure gives the hypothesis a clarity and specificity that provide the necessary power for a reduced-form test. The hypothesis also illustrates a clear understanding of the workings of the tax code. Although tax laws specify the marginal statutory corporate tax rate, corporate decisions are based on the marginal effective tax rate, which is the present value of future tax payments resulting from an additional dollar of taxable income. The statutory rate is the same for all firms, but the effective rate differs from firm to firm above a certain dollar amount. Tax shields such as investment tax credits, tax-loss carryforwards, depreciation allowances, and interest expenses lower the effective tax rate.

The Givoly results support all of their hypotheses: Each of the relevant coefficients is statistically significant and has the proper sign for 1987, the first year the TRA was in effect. The hypotheses appear to be of moderate economic significance. For example, a firm with an effective tax rate 10 percent above that of another firm would have lowered its debt/equity ratio 1.15 percent more in response to the TRA.

Givoly and others provide only part of the answer regarding the effect of financial structure on real input decisions. By showing that taxes influence financial structure, the studies have shown that capital markets are imperfect, thus providing an important clue as to why the MM propositions are not supported by the data. However, they have not clearly shown how these financial decisions impact real input decisions. Furthermore, the magnitude of the effect is far from certain. Thus, the latest findings point out one link of the chain of indirect tax effects through capital structure by showing how taxes influence corporate finance. However, the complete change still suffers from lack of information on the magnitude of the effect of corporate financial structure on real investment decisions.

■ **10** In Givoly et al., Tobin's *q* proxies for bankruptcy costs and the collateral value of the firm when bankrupt. Business risk is proxied by the coefficient of variation in operating income (minus depreciation) over the firm's last 10 years.

Although earlier empirical work found no significant tax effects, more recent studies have addressed the inherent empirical problems and have produced evidence that supports the importance of taxes on financing decisions. Hence, one link between policy and real investment behavior has been established, but only after very clear statements about the firm's underlying structural behavior were used to define relevant variables and identifying assumptions. These are formulations that require a careful, informed analysis. Even so, only a fully structural model can provide policymakers with an accurate measure of how changes in tax policy influence real investment. Without parameter estimates from such a model, the short- and long-term effects of tax policy changes on real investment remain uncertain.

V. Conclusion

What have we learned from examining 35 years of research? In each case—the direct test of the MM hypothesis through cross-section regressions, the test of the timing of investment through Granger causality, and the test of whether taxes should matter to corporate financial structure—the findings exhibit the same pattern. Early research often failed to reveal statistical significance in the relationship between corporate real investment and the explanatory variable, be it financial structure or taxes. This seemed to provide prima facie support for the empirical relevance of MM's assertions.

Our present knowledge of corporate financial structure through reduced-form estimation is typical of what a more careful reduced-form strategy can do. The weight of the current evidence seems to reject the MM neutrality hypothesis. Financial structure does matter to a firm's investment decisions, and taxes do influence these decisions through their effect on financial structure. These are important statements to bear in mind both when deciding on policy and when formulating new theory with which to guide policy.

Our cautionary tale does not say that the reduced form is an unwise estimation strategy. Rather, it notes what conditions are necessary if a reduced form is to yield accurate results. In all cases, an underlying structure of behavior (even when not used explicitly in a structural estimation model) guided the research through the crucial steps of data definition and formulating the correct econometric test. It is also important to note that a reduced-form analysis is a critical step in any empirical study of a policy question. Simply estimating a structural model without first determining and reporting general directions in the data is a recipe for disaster.

However, the reduced-form strategy, when used without accompanying structural estimates, is distinguished by what it has *not* done. We do not have precise estimates of the magnitude of the effects. Because the estimating equations are formulated without an explicit structure, the resulting parameters are subject to fewer "reality checks" to determine whether they make economic sense. In addition, fewer comparisons can be made to related empirical literatures to determine the appropriateness of the estimating equations' specifications. Is this or that estimated parameter comparable to a risk-aversion parameter estimated in the portfolio-balance literature? We cannot tell from a reduced-form estimate because the reduced form resists a structural interpretation that will allow comparison.

References

Cummins, J.G., K.A. Hassett, and R.G. Hubbard. "A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments," *Brookings Papers on Economic Activity,* vol. 2 (1994), pp. 1–59.

DeAngelo, H., and R. Masulis. "Optimal Capital Structure under Corporate and Personal Taxation," *Journal of Financial Economics,* vol. 8, no. 1 (March 1980), pp. 3–29.

Dhrymes, P., and M. Kurz. "Investment, Dividend, and External Finance Behavior of Firms," in *Determinants of Investment Behavior.* New York: National Bureau of Economic Research, 1967.

Fischer, E., R. Heinkel, and J. Zechner. "Dynamic Capital Structure Choice: Theory and Tests," *Journal of Finance,* vol. 44, no. 1 (March 1989), pp. 19–40.

Gilchrist, S. and C.P. Himmelberg. "Evidence on the Role of Cash Flow for Investment," New York University, Stein School of Business, Working Paper, 1995.

Givoly, D., C. Hayn, A. Ofer, and O. Sarig. "Taxes and Capital Structure: Evidence from Firms' Response to the Tax Reform Act of 1986," *Review of Financial Studies,* vol. 5, no. 2 (1992), pp. 331–55.

Granger, C. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," Econometrica, vol. 37, no. 3 (July 1969), pp. 424–38.

Long, M.S., and I.B. Malitz. "Investment Patterns and Financial Leverage," in B.M. Friedman, ed., *Corporate Capital Structures in the United States.* Chicago: University of Chicago Press, 1985, pp. 325–48.

MacKie-Mason, J. "Do Taxes Affect Corporate Financing Decisions?" *Journal of Finance,* vol. 45, no. 5 (December 1990), pp. 1471–93.

McCabe, G.M. "The Empirical Relationship between Investment and Financing: A New Look," *Journal of Financial and Quantitative Analysis,* vol. 14, no. 1 (March 1979), pp. 119–35.

McDonald, J.G., Jacquillat, B., and M. Nussenbaum. "Dividend, Investment, and Financing Decisions: Empirical Evidence on French Firms," *Journal of Financial and Quantitative Analysis,* vol. 10, no. 5 (December 1975), pp. 741–55.

Modigliani, F., and M. Miller. "The Cost of Capital, Corporate Finance, and the Theory of Investment," *American Economic Review,* vol. 48, no. 3 (June 1958), pp. 261–97.

_________ and ________. "Corporate Income Taxes and the Cost of Capital: A Correction," American Economic Review, vol. 53, no. 3 (June 1963), pp. 433–43.

Mougoue, M., and T.K. Mukherjee. "An Investigation into the Causality among Firms' Dividend, Investment, and Financing Decisions," *Journal of Financial Research,* vol. 17, no. 4 (Winter 1994), pp. 517–30.

Peterson, P.P., and G.A. Benesh. "A Reexamination of the Empirical Relationship between Investment and Financing Decisions," *Journal of Financial and Quantitative Analysis,* vol. 18, no. 4 (December 1983), pp. 439–53.

Smirlock, M., and W. Marshall. "An Examination of the Empirical Relationship between the Dividend and Investment Decisions: A Note," Journal of Finance, vol. 38, no. 5 (December 1983), pp. 1659–67.

Titman, S., and R. Wessels. "The Determinants of Capital Structure Choice," *Journal of Finance,* vol. 43, no. 1 (March 1988), pp. 1–19.

Trezevant, R. "How Did Firms Adjust Their Tax-Deductible Activities in Response to the Economic Recovery Tax Act of 1981?" *National Tax Journal,* vol. 47, no. 2 (June 1994), pp. 253–71.

Predicting Real Growth Using the Yield Curve

by Joseph G. Haubrich and Ann M. Dombrosky

Joseph G. Haubrich is a consultant and economist and Ann M. Dombrosky is a financial reports analyst at the Federal Reserve Bank of Cleveland.

Introduction

The yield curve, which plots the yield of Treasury bonds against their maturity, is one of the most closely watched financial indicators.[1] Many market observers carefully track the yield curve's shape, which is typically upward sloping and somewhat convex. At times, however, it becomes flat or slopes downward ("inverts," in Wall Street parlance), configurations that many business economists, financial analysts, and other practitioners regard as harbingers of recession (see figure 1).

A recent article in *Fortune* labeled the yield curve "a near-perfect tool for economic forecasting" (see Clark [1996]). In fact, forecasting with the yield curve does have a number of advantages. Financial market participants truly value accurate forecasts, since they can mean the difference between a large profit and a large loss. Financial data are also available more frequently than other statistics (on a minute-to-minute basis if one has a computer terminal), and such a simple test as an inversion does not require a sophisticated analysis.

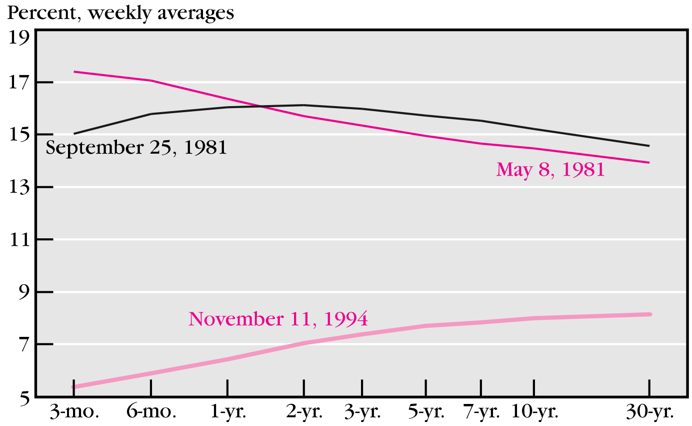
In this *Review,* we examine the yield curve's ability to predict recessions and, more generally, future economic activity. After comparing the curve's forecasts with the historical record, we judge its accuracy against other predictions, including naive forecasts, traditional leading indicators, and sophisticated professional projections. This article builds on a wide range of previous research, but, taking an eclectic approach, differs from the earlier work in a variety of ways. These differences show up mainly in the way we judge forecast performance.

Like the important early work of Harvey (1989, 1991, 1993) and Hu (1993), we use out-of-sample forecasts and compare yield curve forecasts with other predictions (including professional forecasts), but we extend our data set to the mid-1990s. In addition, we consider how adding the yield curve improves (or reduces) the accuracy of other forecasts. In this, we follow Estrella and Hardouvelis (1991), who do not, however, use out-of-sample forecasts. Finally, building on the recent work of Estrella and Mishkin (1995, 1996), we consider how well the yield curve predicts the *severity* of recessions, not just their probability, and compare the forecasts with a wider range of alternatives.

■ 1 Yield curve reports appear in the "Credit Markets" section of *The Wall Street Journal* and the "Business Day" section of *The New York Times.*

FIGURE 1

Yield Curves[a]



a. Three-month and six-month instruments are quoted from the secondary market on a yield basis; all other instruments are constant-maturity series.
SOURCE: Board of Governors of the Federal Reserve System.

The most distinguishing feature of this paper, however, is that it documents the decline in the yield curve's predictive ability over the past decade (1985–95) and discusses possible reasons for this phenomenon. By some measures, the yield curve should be an even better predictor now than it has been in the past.

Widespread use of the yield curve makes assessing its accuracy a worthwhile exercise for economists. But policymakers, too, need an accurate and timely predictor of future economic growth. The ready availability of term-structure data (as opposed to, say, quarterly GDP numbers) ensures a timely prediction, but accuracy is another question. Central bankers have an added incentive to understand the yield curve, since the federal funds rate and the discount rate are themselves interest rates. Uncovering the "stylized facts" about the curve can help the Federal Reserve to understand the market in which it operates.

With sophisticated macroeconometric models and highly paid professional forecasters, is there any place for a simple indicator like the yield curve? Aside from the knowledge gained about the curve itself, there are several reasons to answer that question affirmatively. Simple predictions may serve as a check on more complex models, perhaps highlighting when assumptions or relationships need rethinking. Agreement between predictions increases confidence in the results, while disagreement signals the need for a second look. A simple, popular indicator also provides some insight into market sentiment. Of course, it's always a good idea to check whether the expensive and complicated forecasts actually do perform better.

After first reviewing some basics about the yield curve and the reasons it might predict future growth, we look at the actual relationship and compare predictions from the yield curve to those generated by naive statistical models, traditional indicators, professional forecasters, and an econometric model.

I. Interest Rates and Real Economic Activity

While our main goal is a rather atheoretical assessment of the yield curve's predictive power, forecasts based on the yield curve are on a sounder economic footing than those based on hemlines or Superbowl victories. The best way to see this is to start with a simple theory of the term structure, called the *expectations hypothesis.*

Under this theory, long-term interest rates are the average of expected future short-term rates. If today's one-year rate is 5 percent and next year's one-year rate is expected to be 7 percent, the two-year rate should be 6 percent ($[7 + 5] \div 2 = 6$). More generally, the expectations hypothesis equates the yield (at time t) on an n-period bond, Y_{nt}, and a sequence of one-period bonds:[2]

$$Y_{nt} = E_t(Y_{1,\,t}\,Y_{1,\,t+1}\,Y_{1,\,t+2} \ldots Y_{1,\,t+n-1}).$$

If low interest rates are associated with recessions, then an inverted term structure—implying that upcoming rates will be lower—predicts a recession.[3]

One possible reason for expecting low interest rates in recessions might be termed the *policy anticipations hypothesis.* If policymakers act to reduce short-term interest rates in recessions, market participants who expect a recession would also expect low rates. The yield curve picks up the financial market's estimate of future policy.[4]

■ 2 See chapter 7 of Mishkin (1989) for a fuller treatment of this issue. Mishkin also points out the main flaw in the expectations hypothesis: The term structure normally slopes up, but interest rates do not trend up over time. Campbell (1995) offers a useful discussion of related points.

■ 3 For a classic documentation of this pattern, see Kessel (1965).

■ 4 Rudebusch (1995) takes this approach.

Another possibility is that current monetary policy may shift both the yield curve and future output. For example, tight monetary policy might raise short-term interest rates, flattening the yield curve and leading to slower future growth. Conversely, easy policy could reduce short-term interest rates, steepen the yield curve, and stimulate future growth. The yield curve predicts future output because each of these shifts follows from the same underlying cause: monetary policy. Taking this logic one step further, monetary policy may react to output, so that the yield curve picks up a complex intermingling of policy actions, reactions, and real effects.

In these explanations, the yield curve reflects future output indirectly, by predicting future interest rates or future monetary policy. It may also reflect future output directly, because the 10-year interest rate may depend on the market's guess of output in 10 years.

The expectations hypothesis certainly marks the beginning of wisdom about the yield curve, but only the beginning. The 30-year bond may have a high interest rate not because people expect interest rates to rise, but because such a bond must offer a high return to get people to hold it in the first place. (This is commonly called the risk premium, though for some theories that may be a misnomer.) Investors may dislike wide swings in prices as market expectations about the distant future change over time. Conversely, there may be reasons why some people would rather hold a 30-year bond than a one-year bond. For example, they may be saving for retirement and prefer the certain payoff on the longer-term note (this is sometimes called the *preferred habitat hypothesis*).

The risk premium provides another reason why the yield curve may be a useful predictor: The premium itself holds information. As a simple example, consider that recessions may make people uncertain about future income and employment, or even about future interest rates. The risk premium on a longer-term bond reflects this. In conjunction with changes working through the expectations hypothesis, the yield curve may take some very strange twists indeed, becoming inverted, humped, or even u-shaped.[5]

These explanations provide an additional motivation for investigating yield curve predictions. They also hint at the many important issues that transcend the yield curve's predictive power. It matters, for instance, if the curve reacts to future policy, to movements in output, or to some combination of the two. But these considerations fall by the wayside if the yield curve is not an accurate predictor of future economic activity. In this article, we concentrate on that more basic issue, leaving determination of the underlying causes for another day.

II. Data and Computation

There are many ways of using the yield curve to predict future real activity. One common method uses inversions (when short rates exceed long rates) as recession indicators. Is it possible, however, to predict the magnitude as well as the direction of future growth? Does a large inversion predict a severe recession? Does a steep yield curve predict a strong recovery?[6] Operationally, this means relating a particular measure of yield curve "steepness" to future real growth. In taking this route, we follow and build on the related work of Estrella and Hardouvelis (1991).

Obtaining predictions from the yield curve requires much preliminary work. Three principles guided us through the many decisions that were required: Keep the process simple, preserve comparability with previous work, and avoid data snooping. Thus, we avoided both complicated nonlinear specifications and a detailed search for the "best" predictor.

To begin with, there is no unambiguous measure of yield curve steepness. A yield curve may be flat at the short end and steep at the long end. The standard solution uses a spread, or the difference between two rates (in effect, a simple linear approximation of the nonlinear yield curve).[7] This means choosing a particular spread, in itself no trivial matter. Among the 10 most commonly watched interest rates (the federal funds rate and the three-month, six-month, and one-, two-, three-, five-, seven-, 10-, and 30-year Treasury rates), 45 possible spreads exist.[8]

An additional problem is that there are several types of yield curves or term structures. In fact, it sometimes helps to draw a distinction between the yield curve and the term structure. The yield curve is the relation between the

■ **5** Stambaugh (1988) makes this point. For a less technical description, see Haubrich (1991).

■ **6** Other approaches also exist. For example, Harvey (1988) examines whether the term structure predicts changes in consumption.

■ **7** Frankel and Lown (1994) is one of the few papers that considers nonlinear measures of steepness.

■ **8** If there are n rates, there are $n/2\,(n-1)$ spreads. This is the classic formula for combinations. See Niven (1965), chapter 2.

yield on Treasury securities and their maturity. The term structure is a particular yield curve—that for zero-coupon Treasury securities. The term structure is theoretically more interesting. It answers the question, "How much would I pay for one dollar delivered 10 years from today?" The problem is that a zero-coupon Treasury security rarely matures in exactly 10 years. What we actually observe in the market are prices (and thus yields) on existing Treasury securities. These may not mature in precisely 10 years (or whatever maturity you choose), and they often have coupon payments. That is, a 10-year Treasury note pays interest semiannually at a specified coupon rate, so its yield is not the yield called for in the term structure.

Finding the desired interest rate almost always involves estimation of some kind. Calculating the theoretically pure term structure is often quite difficult, as it must be estimated from coupon bonds of the wrong maturity, all subject to taxation. (Using zero-coupon bonds may help, but this approach introduces problems of its own, as the market is thinner and the tax treatment of coupons and principal differs.) This means that the pure term structure is not available in real time, when the Federal Reserve must attempt to discern the course of the economy. To avoid stale data, we must turn to the more "rough-and-ready" yield curve. Even here, the problem of matching maturities arises. Fortunately, the Treasury Department publishes a "constant-maturity" series, where market data are used to estimate today's yield on a 10-year Treasury note, even though no such note exists.[9]

For our study, we use data from the Federal Reserve's weekly H.15 statistical release ("Selected Interest Rates"), which compiles interest rates from various sources. For the spread, we chose the 10-year CMT rate minus the secondary-market three-month Treasury bill rate. In addition to allowing a comparison with the work of Estrella and Hardouvelis (1991), Harvey (1989, 1993), and Estrella and Mishkin (1995, 1996), choosing only one spread enables us to minimize the problems of data snooping (Lo and MacKinlay [1990]) and the associated spuriously good results.[10] That is, trying every single spread would produce something that looked like a good predictor, but it very likely would be a statistical fluke akin to Superbowl victories and hemlines. We then convert the bill rate, which is published on a discount rate basis, to a coupon-equivalent yield so that it is on the same basis as the 10-year rate.[11]

Also following Estrella and Hardouvelis, we use quarterly averages for the spread. This smoothes the anomalous rates that appear at the turn of each month.[12] A priori there is no presumption that GDP should correlate better with a particular date's spread than with the quarterly average.[13]

As our measure of real growth, we use the four-quarter percent change in real (fixed-weight) GDP. GDP is, of course, the standard measure of aggregate economic activity, and the four-quarter forecast horizon answers the "what-happens-next-year" type of question without embroiling us in data snooping issues regarding the optimal horizon choice.

Our sample period runs from 1961:IQ through 1995:IIIQ. This covers various inflationary experiences, episodes of monetary policy tightening and easing, and several business cycles and recessions. Included are five recessions, inflation rates from 1 percent to more than 13 percent, and a federal funds rate ranging from under 3 percent to over 19 percent.

Our basic model, then, is designed to predict real GDP growth four quarters into the future based on the current yield spread. Operationally, we accomplish this by running a series of regressions (detailed below) using real GDP growth and the interest rate spread lagged four quarters (for example, the interest rate spread used for 1961:IQ is actually from 1960:IQ).

The next step involves comparing the yield curve forecasts with a sequence of increasingly sophisticated predictions using other techniques. We start with a naive (but surprisingly effective) technique which assumes that GDP growth over the next four quarters will be the same as it was over the last four. (That is, the growth rate is a random walk.) We then regress real GDP growth against the index of leading economic indicators (lagged four quarters). This enables us to make a comparison with another simple and popular forecasting technique.

■ **9** See Smithson (1995) for a good description of constant-maturity Treasuries (CMTs).

■ **10** Work by Knez, Litterman, and Scheinkman (1994) suggests that using more or different rates would not capture much additional information.

■ **11** The three-month CMT rate was not published before May 1995. To keep the data consistent, we use the secondary-market three-month Treasury bill rate throughout. For such a short rate, the differences are minimal.

■ **12** Park and Reinganum (1986) document this calendar effect.

■ **13** We also reworked the results using data for the last week of each quarter for the 1963–95 period. The findings were comparable, although the predictive power of the spread decreased somewhat.

BOX 1

Forecasting Equations

Yield spread: in-sample

$$\frac{RGDP_{t+4} - RGDP_t}{RGDP_t} = \alpha + \beta\ spread_t$$

Yield spread: out-of-sample

$$\frac{RGDP_{t+4} - RGDP_t}{RGDP_t} = \alpha + \beta\ spread_t$$

Naive

$$\frac{RGDP_{t+4} - RGDP_t}{RGDP_t} = \frac{RGDP_t - RGDP_{t-4}}{RGDP_t}$$

Leading indicators

$$\frac{RGDP_{t+4} - RGDP_t}{RGDP_t} = \alpha + \beta\ index_t$$

Lagged GDP

$$\frac{RGDP_{t+4} - RGDP_t}{RGDP_t} = \alpha + \beta \frac{RGDP_t - RGDP_{t-4}}{RGDP_t}$$

Lagged GDP plus yield spread

$$\frac{RGDP_{t+4} - RGDP_t}{RGDP_t} = \alpha + \beta \frac{RGDP_t - RGDP_{t-4}}{RGDP_t} + \gamma\ spread_t$$

We next look at two additional forecasts generated by statistical procedures. We regress real GDP growth against its own lag (again, four quarters) and against its own lag and the 10-year, three-month spread.

The final, and most sophisticated, alternative forecasts we consider come from the Blue Chip organization and DRI/McGraw–Hill (hereafter referred to simply as DRI). We first compare the results of our model with forecasts from *Blue Chip Economic Indicators,* beginning with the July 1984 issue.[14] We use the one-year-ahead Blue Chip consensus forecasts for real GDP (or real GNP when GDP forecasts are unavailable), labeled "percent change from same quarter in prior year." These forecasts are taken from the Blue Chip newsletters corresponding to the first month of each quarter.

We next compare our results with predictions from DRI, reported in various issues of its *Review of the U.S. Economy*. DRI generates these forecasts from an econometric model. Although we tried to collect forecasts for the same period as our Blue Chip forecasts (that is, from issues corresponding to the first month of each quarter), our DRI data set is missing two points: 1985:IIIQ and 1987:IIQ. We use forecasts for real GDP (or real GNP when GDP forecasts are unavailable) one year ahead.

Box 1 summarizes the regressions used to forecast future real GDP growth.

III. Forecast Results

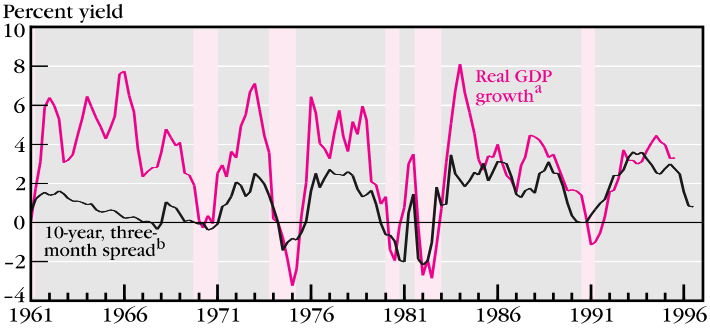
Does the yield curve accurately predict future GDP? First, look directly at the data. Figure 2 shows the growth of real GDP and the lagged spread between the 10-year and three-month Treasury yields. A decline in the growth of real GDP is usually preceded by a decrease in the yield spread, and a narrowing yield spread often signals a decrease in real GDP growth. A negative yield spread (inverted yield curve) usually precedes recessions, but not always. For example, the yield spread turned negative in the third and fourth quarters of 1966, but no recession occurred for the next three years. (The recession that began in late 1969 was preceded by two quarters of a negative yield spread.) The latest recession, which occurred in 1990–91, was preceded by a yield curve more accurately described as flat than inverted.

Figure 3 plots the same data in a different form. It shows a scatterplot, with each point representing a particular combination of real

■ **14** *Blue Chip Economic Indicators* is a monthly collection of economic forecasts by a panel of economists from some of the top firms in the United States (the so-called Blue Chip companies). The Blue Chip consensus forecast for real GDP is the average of about 50 individual forecasts. See Zarnowitz and Lambros (1987) for evidence that the consensus forecast predicts much better than individual forecasts, and Lamont (1994) for a possible explanation.

FIGURE 2

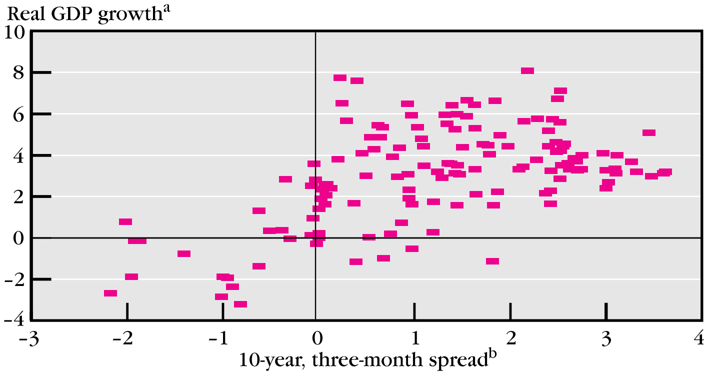
Real GDP Growth and Lagged Yield Spread



a. Four-quarter percent change.
b. Lagged four quarters.
NOTE: Shaded areas indicate recessions.
SOURCES: Board of Governors of the Federal Reserve System; and U.S. Department of Commerce, Bureau of Economic Analysis.

FIGURE 3

Scatterplot of Real GDP Growth and Lagged Yield Spread



a. Four-quarter percent change.
b. Lagged four quarters.
SOURCES: Board of Governors of the Federal Reserve System; and U.S. Department of Commerce, Bureau of Economic Analysis.

GDP growth and the lagged yield spread. Even a casual look at the results reveals that the relationship between the two variables is usually positive; that is, positive real GDP growth is associated with a positive lagged yield spread, and vice versa.

Plotting the data gives a strong, albeit qualitative, impression that the yield spread predicts future real activity. We desire a more quantitative prediction, one that says more than "The yield curve is steep; looks like good times." To generate the GDP predictions, we ran an in-sample regression, using the entire sample to generate each predicted data point. This is the sort of comparison Estrella and Hardouvelis (1991) make, and our results, presented below, confirm their assessment that the 10-year, three-month spread has significant predictive power for real GDP growth:[15]

$$\text{Real GDP growth} = \underset{(3.89)}{1.8399} + \underset{(4.50)}{0.9791}\,spread$$

$$R^2 = 0.291, \text{ D–W} = 0.352.$$

The yield spread emerges as statistically and economically significant, translating almost one for one into expected future growth. Thus, a spread of 100 basis points (1 percent) implies future growth of 2.8 percent (we derive this as 1.8 + 0.98 x 1). The R^2 indicates that much variation remains to be explained. Figure 4 plots our in-sample real GDP predictions versus actual real GDP growth.

The in-sample results are somewhat misleading, as the coefficients depend on information not available early in the sample. Figure 5 plots another series of predicted real GDP growth, this time generated from an out-of-sample regression. Each data point in this chart is based on a regression using only the data (yield spreads) before the predicted data point. That is, the predicted GDP growth rate for, say, 1980:IQ is based on the data sample from 1961:IQ through 1979:IVQ. Hence, this regression generates a true forecast because it uses available data to predict future (out-of-sample) real GDP growth.[16]

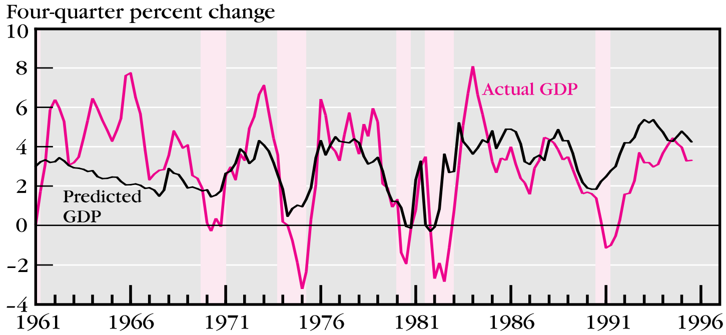
The predicted GDP series from the in-sample and out-of-sample regressions are broadly similar and generally follow the actual GDP data. The root mean square error (RMSE) of the predictions is 2.04 for the in-sample and 2.10 for the out-of-sample forecasts. It is not surprising that the in-sample regression performs slightly better. If we calculate the RMSE over the last 10 years of our data set (1985:IIIQ to 1995:IIIQ), the in-sample regression (RMSE 1.07) again does better than the out-of-sample regression (RMSE 2.09).

■ **15** The *t* statistics are Newey–West corrected with five lags. This offsets the bias created by overlapping prediction intervals (a serious problem in this case, as indicated by the Durbin–Watson statistic).

■ **16** We also corrected the regression for a more subtle problem. Because first-quarter GDP numbers become available only after the first quarter ends, we should not use those numbers in a regression. (That is, as of the first quarter, we still don't know the four-quarter growth rate over last year's first quarter.)

FIGURE 4

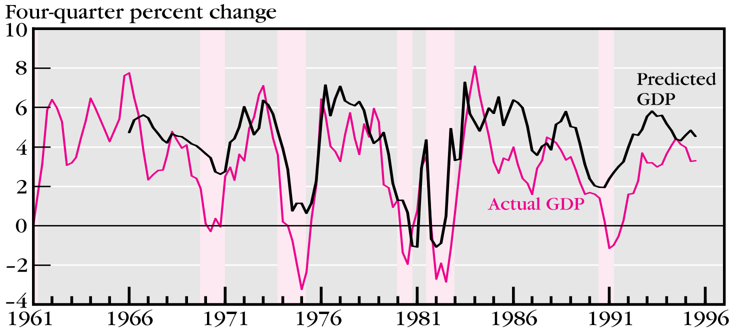
Real GDP Predictions: In-Sample



NOTE: Shaded areas indicate recessions.
SOURCES: U.S. Department of Commerce, Bureau of Economic Analysis; and authors' calculations.

FIGURE 5

Real GDP Predictions: Out-of-Sample



NOTE: Shaded areas indicate recessions.
SOURCES: U.S. Department of Commerce, Bureau of Economic Analysis; and authors' calculations.

TABLE 1

RMSE of Forecasts

	1965:IVQ–1995:IIIQ	1985:IIIQ–1995:IIIQ
Yield spread: in-sample	2.04	1.07
Yield spread: out-of-sample	2.10	2.09
Naive	3.15	1.66
Leading indicators	2.37	1.36
Lagged GDP	2.49	1.46
Lagged GDP plus yield spread	2.19	2.06
DRI[a]	n.a.	0.63
Blue Chip	n.a.	1.13

a. DRI forecasts are missing two quarters.
SOURCE: Authors' calculations.

We use this RMSE criterion to compare the yield spread forecasts with those derived from other techniques. The results are reported in table 1.

For the entire sample, the yield curve emerges as the most accurate predictor of real economic growth. Furthermore, adding the yield spread to a lagged GDP regression improves the forecast, while adding lagged GDP to the yield spread worsens the forecast. For in-sample regressions, adding variables never hurts, but it quite commonly reduces the performance of the out-of-sample regressions.

Curiously, the 1985–95 subsample completely reverses the results. The yield spread becomes the least accurate forecast, and adding it to lagged GDP actually worsens the fit. The leading indicators emerge as the best of the "low-cost" forecasts, and the two professional services do markedly better than the rest. In part, this may reflect the simple specifications used in the regression forecasts: With more lags, a simple regression forecast is often better than a sophisticated model (see Chatfield [1984], chapter 5). But the change is even more significant than that. Using unpublished data, Harvey (1989) finds that over the 1976–85 period, the yield spread performs as well as or better than seven professional forecasting services (including DRI but not Blue Chip).

The dramatic drop in forecasting ability may result from several factors. It certainly reflects a changing relationship between the yield curve and the economy. The coefficients in the term-spread regression demonstrate this. At the beginning of the sample, using only 20 data points (five years of quarterly data), the coefficient on the term spread is –0.14 (statistically insignificant). Midway through the sample, after 70 data points, the coefficient is 1.48, and for the whole sample, 0.98. While one advantage of an out-of-sample procedure is that it allows the coefficients to change, the influence of the first 20 years may force the "wrong" coefficient on the last 10. It is also quite reasonable that the relationship between the yield curve and the real economy might have changed over 30 years. Advances in technology, new production processes, changes in market organization or in the way the market reacts to new information, or even shifts in Federal Reserve policy (the famous "Lucas critique") might have altered the relationship between the yield curve and real activity.

Evidence suggests that both the timing and the size of the relationship between the yield curve and real activity has in fact changed. If

TABLE 2

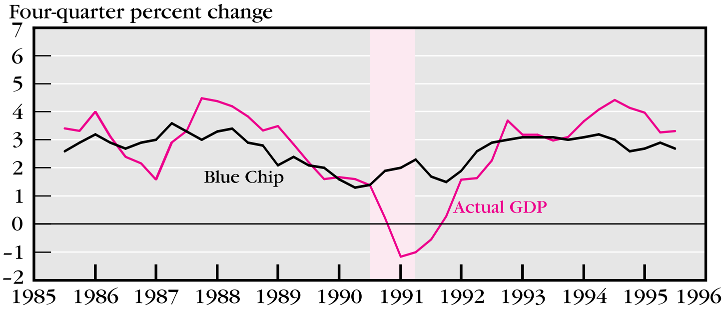
Correlations between Lagged Yield Spread and Real GDP Growth

Lag	1960:IQ–1995:IIIQ	1960:IQ–1985:IIQ	1985:IIIQ–1995:IIIQ
0	0.041	0.075	0.143
1	0.206	0.252	0.358
2	0.366	0.439	0.538
3	0.481	0.585	0.641
4	0.540	0.666	0.684
5	0.505	0.624	0.708
6	0.423	0.523	0.736
7	0.331	0.398	0.721
8	0.218	0.251	0.660

SOURCE: Authors' calculations.

FIGURE 6

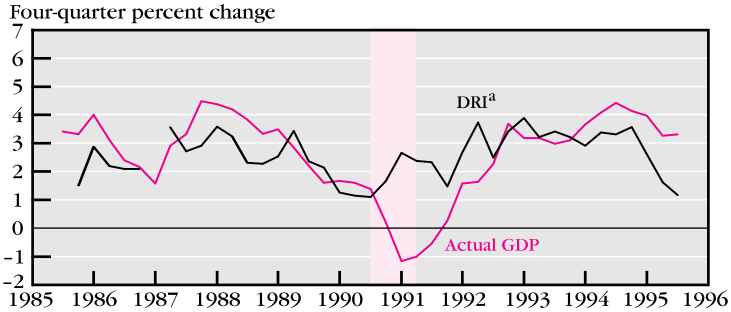
Real GDP Predictions: Blue Chip



NOTE: Shaded area indicates recession.
SOURCES: U.S. Department of Commerce, Bureau of Economic Analysis; and *Blue Chip Economic Indicators,* various issues.

FIGURE 7

Real GDP Predictions: DRI



a. 1985:IIIQ and 1987:IIQ data are missing.
NOTE: Shaded area indicates recession.
SOURCES: U.S. Department of Commerce, Bureau of Economic Analysis; and DRI/McGraw–Hill.

we look at the correlations between the yield spread and real GDP growth at different lags (table 2), we see that the recent period has higher correlations between lags of yield spreads and real GDP growth. We also see that the largest correlation for the early (and total) period, 0.666, occurs at a lag of four quarters, exactly the lag used in our regressions. In the recent period, however, despite a higher correlation at four lags, the highest correlation (0.736) is reached at six quarters. The correlations drop off more slowly in the latter period as well. This accounts for our somewhat paradoxical conclusion: Despite better correlations between the yield curve and real GDP—with an in-sample RMSE that beats even the Blue Chip forecast—regressions using past data are less reliable predictors.

It is somewhat instructive to make a more detailed comparison with the sophisticated forecasts. The Blue Chip predictions can be considered out-of-sample because each one is based on data available at the time of the prediction. Figure 6 plots the Blue Chip consensus forecasts against actual real GDP growth. The Blue Chip forecasts appear much smoother than GDP, as they consistently underpredict real GDP when economic growth is high and overpredict real GDP when economic growth is negative.

The DRI forecasts are plotted in figure 7. These appear broadly similar to the Blue Chip forecasts, although the DRI series is more volatile. Like the Blue Chip forecasts, the DRI forecasts generally underpredict GDP when economic growth is high and overpredict GDP when economic growth is low. Our out-of-sample forecasts based on the yield spread overpredict GDP growth for all but one quarter during the 1985:IIIQ–1995:IIIQ period (corresponding to the Blue Chip and DRI data sets).

IV. Conclusion

Does the yield curve accurately predict real economic growth? Answering this seemingly simple question requires a surprising amount of preliminary work. Much of this paper is devoted to refining the initial question to confront the realities of the financial marketplace.

Fortunately, the answer is less complex, if somewhat nuanced. The 10-year, three-month spread has substantial predictive power, and in this we confirm a variety of earlier studies. Over the past 30 years, it provides one of the best (in our sample, *the* best) forecasts of real growth four quarters into the future. Over the past decade, it has been less successful: Indeed, the yield curve was the *worst* forecast we examined. This shift seemingly results from a change in the relationship between the yield curve and real economic activity—one that has become closer, but nonetheless has made regressions based on past data less useful.

An interesting topic for future research would be to examine whether simple fixes, such as a rolling regression model or more lags, could improve the recent performance of the yield curve. Certainly the simple yield curve growth forecast should not serve as a replacement for the consensus predictions of the Blue Chip panel or the DRI econometric model. It does, however, provide enough information to serve as a useful check on the more sophisticated forecasts and to encourage future research into the reasons behind the yield curve's worsening performance.

References

Campbell, J.Y. "Some Lessons from the Yield Curve," *Journal of Economic Perspectives,* vol. 9, no. 3 (Summer 1995), pp. 129–52.

Chatfield, C. *The Analysis of Time Series: An Introduction,* 3d. ed. New York: Chapman and Hall, 1984.

Clark, K. "A Near-Perfect Tool for Economic Forecasting," *Fortune,* July 22, 1996, pp. 24–26.

Estrella, A., and G.A. Hardouvelis. "The Term Structure as a Predictor of Real Economic Activity," *Journal of Finance,* vol. 46, no. 2 (June 1991), pp. 555–76.

Estrella, A., and F. S. Mishkin. "Predicting U.S. Recessions: Financial Variables as Leading Indicators," National Bureau of Economic Research, Working Paper No. 5379, 1995.

___________, **and** ___________. "The Yield Curve as a Predictor of U.S. Recessions," Federal Reserve Bank of New York, *Current Issues in Economics and Finance,* vol. 2, no. 7, June 1996.

Frankel, J.A., and C.S. Lown. "An Indicator of Future Inflation Extracted from the Steepness of the Interest Rate Yield Curve along Its Entire Length," *Quarterly Journal of Economics,* vol. 109, no. 2 (May 1994), pp. 517–30.

Harvey, C.R. "The Real Term Structure and Consumption Growth," *Journal of Financial Economics,* vol. 22, no. 2 (December 1988), pp. 305–33.

___________. "Forecasts of Economic Growth from the Bond and Stock Markets," *Financial Analysts Journal,* September/October 1989, pp. 38–45.

___________. "The Term Structure and World Economic Growth," *Journal of Fixed Income,* June 1991, pp. 7–19.

___________. "Term Structure Forecasts Economic Growth," *Financial Analysts Journal,* vol. 49, no. 3 (May/June 1993), pp. 6–8.

Haubrich, J.G. "Wholesale Money Market," in *The New Palgrave Dictionary of Money and Finance.* New York: Stockton Press, 1992, pp. 798–800.

Hu, Z. "The Yield Curve and Real Activity," *International Monetary Fund Staff Papers,* vol. 40, no. 4 (December 1993), pp. 781–806.

Kessel, R.A. "The Cyclical Behavior of the Term Structure of Interest Rates," National Bureau of Economic Research, Occasional Paper No. 91, 1965.

Knez, P.J., R. Litterman, and J. Scheinkman. "Explorations into Factors Explaining Money Market Returns," *Journal of Finance,* vol. 49, no. 5 (December 1994), pp. 1861–82.

Lamont, O. "Macroeconomic Forecasts and Microeconomic Forecasters," Princeton University, unpublished manuscript, 1994.

Lo, A.W., and A.C. MacKinlay. "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies,* vol. 3, no. 3 (1990), pp. 431–67.

Mishkin, F. S. *The Economics of Money, Banking, and Financial Markets,* 2d. ed. Glenview, Ill.: Scott, Foresman, and Company, 1989.

Niven, I. *Mathematics of Choice: How to Count without Counting.* Washington, D.C.: The Mathematical Association of America, 1965.

Park, S.Y., and M.R. Reinganum. "The Puzzling Behavior of Treasury Bills that Mature at the Turn of Calendar Months," *Journal of Financial Economics,* vol. 16, no. 2 (June 1986), pp. 267–83.

Rudebusch, G.D. "Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure," *Journal of Monetary Economics,* vol. 35, no. 2 (April 1995), pp. 245–74.

Smithson, C. "ABC of CMT," *Risk,* vol. 8, no. 9 (September 1995), pp. 30–31.

Stambaugh, R.F. "The Information in Forward Rates: Implications for Models of the Term Structure," *Journal of Financial Economics,* vol. 21, no. 1 (May 1988), pp. 41–70.

Zarnowitz, V., and L.A. Lambros. "Consensus and Uncertainty in Economic Prediction," *Journal of Political Economy,* vol. 95, no. 3 (June 1987), pp. 591–621.