Volume Title: NBER Macroeconomics Annual 1995, Volume 10

Volume Author/Editor: Ben S. Bernanke and Julio J. Rotemberg, eds.

Volume Publisher: MIT Press

Volume ISBN: 0-262-02394-6

Volume URL: http://www.nber.org/books/bern95-1

Conference Date: March 10-11, 1995

Publication Date: January 1995

Chapter Title: Inflation Indicators and Inflation Policy

Chapter Author: Stephen G. Cecchetti

Chapter URL: http://www.nber.org/chapters/c11019

Chapter pages in book: (p. 189 - 236)

*Stephen G. Cecchetti*
OHIO STATE UNIVERSITY

# *Inflation Indicators and Inflation Policy*

## 1. *Introduction*

Since late 1979, the goal of monetary policy in the United States has progressively shifted toward the reduction of the level and variation of inflation. Recent policy actions serve to highlight the desire of the Federal Reserve to keep inflation both low and stable, while downplaying the likely output and employment consequences. The purpose of this paper is to evaluate the practicality and desirability of inflation targeting.

Any attempt to control the path of the aggregate price level has several critical aspects. First, policy makers must be able to forecast inflation. To run a proactive policy that controls prices, it is crucial that one be able to assess the future path of prices in the absence of any immediate policy action. In other words, given some candidate inflation indicator, the relationship between future inflation and the indicator must be known with some degree of accuracy. This leads to the first major question: Are there any useful indicators of future inflation? If, for example, we see the level of capacity utilization or the price of gold rise, can this be translated into a forecast for the aggregate price level?

Once prices are forecast to rise above a target path, the next step is to figure out what to do about it. But the extent of a response requires an accurate estimate of the impact of the policy instrument on inflation. Is it possible to estimate the effect of policy actions on prices? What is the precision of the estimates?

Finally, given the estimated response of prices to both exogenous

Professor of Economics, Boston College; Professor of Economics, Ohio State University; and Research Associate, National Bureau of Economic Research. This paper was prepared for the NBER's Tenth Annual Conference on Macroeconomics, March 10–11, 1995. I would like to thank Ben Bernanke, Michael Bryan, John Campbell, Bruce Hansen, Anil Kashyap, Don Kohn, Julio Rotemberg, and Mark Watson for comments and suggestions. I am grateful to Michael Niemira and Roy Webb for providing their data, to Bruce Hansen for his programs, and to the National Science Foundation and the Federal Reserve Bank of Cleveland for financial and research support.

shocks and policy, and an objective for the policy makers, it is possible to formulate policy rules. With these in hand, a series of practical questions can be addressed. First, how quickly and by how much should policy react to perceived upward price pressures? What are the quantitative benefits of price-level targeting on the variance of inflation, and what are its costs in terms of increased variation of output? And finally, what are the consequences of shifting from price-level to nominal-income targeting?

The primary focus of this paper is on inflation. This can be justified by the fact that the reduced-form representation of a broad class of macroeconomic models depends only on inflation. A simple example can be constructed by starting with a standard staggered contract model in which nominal shocks have real effects that die out slowly.[1] Assume that the monetary authority minimizes a loss function that depends on current and future deviations of output ($y_t$) from its full-employment level ($y^*$), as well as inflation ($\pi$). This formulation ignores dynamic consistency problems, as it implicitly assumes that the loss is minimized at $y = y^*$.[2] Fixed nominal prices that change infrequently imply that output deviations are a distributed lag of unanticipated inflation, and so the loss function can be written in terms of inflation and expected inflation alone. It immediately follows that the optimal long-run policy objective should be zero inflation. The dependence of policy on expectations only affects optimal disinflationary paths.

The remainder of this paper is divided into five sections. Section 2 discusses the difficulties facing policy makers in their attempt to control high- and medium-frequency fluctuations in the aggregate price level. While evidence suggests that a $k$-percent rule may work well for controlling inflation in the long run (at horizons of five to ten years or more), it may result in substantial swings in prices over horizons of two or three years. This leads to an examination of a more sophisticated approach in which one first tries to forecast the path of inflation, and then formulates a reaction function to control it. Sections 3 and 4 examine different aspects of the forecasting problem, while Sections 5 and 6 study policy reaction functions. In Section 3 two sets of results are presented. The first examines commercial forecasts from various sources. This is followed by an examination of the reduced-form correlations between inflation and various candidate indicators. Section 4 examines structural changes in the relationship between inflation and the indicators. Section

1. Examples of such models are numerous. See for example the one in Ball and Cecchetti (1988).
2. Since price rigidity in staggered contract models usually results from monopolistic competition, social welfare will not in general be maximized at the general equilibrium level of output.

5 looks closely at the relationship between monetary policy and inflation, and Section 6 chooses a particular model to study optimal policy responses. Section 7 offers conclusions.

To anticipate the conclusions, I find that inflation is extremely difficult to forecast at horizons of even one quarter. One of the likely reasons for this is the fact that the relationship between inflation and various candidate indicators exhibits rather frequent structural breaks. Beyond the forecasting problem, I find that the relationship between inflation and policy is also difficult to estimate. Together, these lead to the final result: Shifting from price level to nominal-income targeting yields a substantial gain in real-income stabilization, while resulting in only a small loss in increased aggregate price variability.

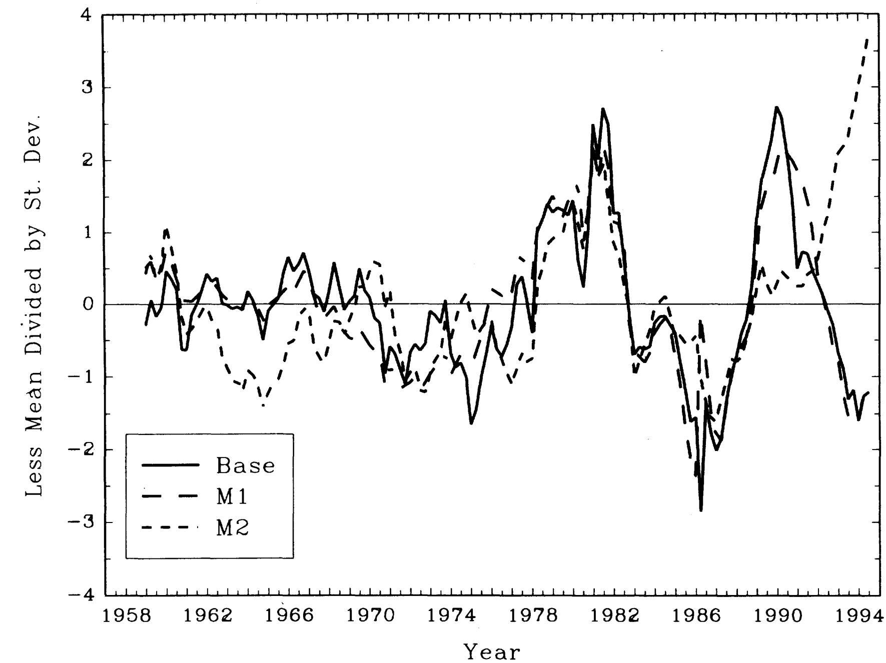
## 2. *The Nature of the Problem*

The apparent stability of long-run money demand functions implies that a Friedman-style *k*-percent rule for money growth could be used to reach either a long-run inflation or a long-run nominal GDP target.[3] But recent experience teaches that monetary policy aimed at controlling short-run movements in nominal variables faces formidable problems.

Figure 1 plots the nominal GDP velocity of the monetary base, M1, and M2, from the first quarter of 1959 through the third quarter of 1994. The data are standardized by removing a trend, subtracting the sample mean, and dividing by the sample standard deviation.[4] For the monetary base and M1, the trend has a break in 1986.[5] These data suggest two things. First, the short-run instability in velocity has increased substantially over the past decade. Second, over the past two years, the velocity of M2 has increased dramatically, while that of the monetary base and M1 has declined.[6] If one were to try to formulate a *k*-percent rule, the question would be: *k* percent of what?

3. See, for example, Lucas (1988), Hallman, Porter, and Small (1991), Stock and Watson (1993), and Feldstein and Stock (1994), among others.
4. The standardization makes the reported values similar to coefficients from a regression of the detrended log of nominal GDP on the detrended log of money.
5. The location of the break was determined using the Andrews (1993)–Quandt (1960) 'sup' test, robust to heteroscedasticity. For the monetary base, the break is estimated to occur in 1986:04, whereas for M1 the estimated break is in 1986:02. Obviously, real-time policy making would not have been able to take advantage of the shift that we can now estimate. This makes matters even worse than they appear here.
6. The use of M2 encounters the additional problem of the changes in definition over time. Recently, Duca (1992) has suggested including the stock of bond mutual funds in a revised measure. A very skeptical way to view these modifications is to observe that M2 seems to be constructed so that its implied velocity is stationary about a constant mean of 1.65.

Figure 1 GDP VELOCITY OF THE MONETARY BASE, M1 AND M2 STANDARDIZED, QUARTERLY, 1959Q1 TO 1994Q4




The implication of the recent instability in velocity is that it is hard to know how to run policy at high frequencies. To put it slightly differently, most economists would agree that if the monetary base grows at 10% rather than at 5%, then inflation will be higher in the long run. But this tells us very little about how the monetary authority should act on a month-to-month basis. A natural response to this is to seek a more sophisticated feedback rule that incorporates inflation, inflation indicators, and policy variables. That is the task addressed in the remainder of this paper.

## 3. *Forecasting Inflation*

The first step in formulating any policy aimed at reducing the level and variance of inflation is to forecast the evolution of aggregate prices both with and without policy interventions. The next two sections examine our ability to forecast inflation, and the following sections report evidence on the impact of policy changes on prices. Since economic theory implies that different indicators should forecast inflation at different horizons, I present results for forecasts of inflation over future periods of varying length.[7]

Section 3.1 examines the accuracy of contemporaneous inflation forecasts. Section 3.2 reports evidence on the simple correlation between indicators and inflation. Section 3.3 evaluates forecasts of inflation based on the indicators.

### 3.1 CONTEMPORANEOUS FORECASTS

There are several readily available sources for the history of commercial inflation forecasts. Table 1 reports the root-mean-square error of the quarterly and annual forecasts published by Data Resources Incorporated (DRI) and the consensus forecast from Blue Chip Economic Indicators. For comparison, the table also reports the results of using a simple random-walk model for inflation. This "naive" method takes current inflation as the forecast for all horizons.

Since forecasters report their expectations of the path for the price level into the future, it is possible to construct estimates of a term structure of expected future prices. Defining $E_t[\pi_{t+l,\,t+k}]$ as the expectation at $t$ of inflation from $t + l$ to $t + k$—the analog to a forward interest rate—we are able to compute the accuracy of inflation forecasts for various horizons.

The results in the table suggest several conclusions. First, while always better than the benchmark naive forecasts, the commercial forecasts are very poor, even at a one-quarter horizon. For example, the root-mean-square error of DRI's one-quarter-ahead forecast ($E_t[\pi_{t,\,t+1}]$) for 1982:01 to 1994:03 is 1.54, implying a 70% confidence interval of three percentage points and a 90% confidence interval in excess of five percentage points! The Blue Chip consensus forecast is only marginally more accurate (perhaps because it begins only in 1985).

Second, the accuracy of the DRI forecasts declines as the horizon increases. This is particularly true for the early part of the sample, which includes the large oil price shocks of the middle and late 1970s. But even over the past dozen years, the inaccuracy of the forecasts increases with the horizon, rising by one-third as the horizon increases from 1 to 10 quarters.

The table reports results for one-year forecasts, out up to three years. The results do show that the RMSE of the forecast one year ahead is quite a bit smaller than that of the forecast one quarter ahead. This suggests that forecasters might get the general trend in inflation roughly correct, while missing high-frequency movements that are subsequently reversed. Nevertheless, the forecasts still seem very inaccurate, implying 90% confidence intervals for one-year-ahead inflation of more than 3 percentage points.

7. Ball and Cecchetti (1990) make a similar point.

Table 1 ROOT-MEAN-SQUARE ERROR OF PUBLISHED FORECASTS

*Data Resources*

| *Horizon (Quarters):* | *1970 Q2 to 1994 Q3* Actual | Naive | *1970 Q2 to 1981 Q4* Actual | Naive | *1982 Q1 to 1994 Q3* Actual | Naive |
|---|---|---|---|---|---|---|
| $E_t[\pi_{t,t+1}]$: | 1.67 | 1.98 | 1.80 | 2.09 | 1.54 | 1.88 |
| $E_t[\pi_{t+1,t+2}]$: | 2.30 | 2.53 | 2.77 | 2.80 | 1.75 | 2.26 |
| $E_t[\pi_{t+2,t+3}]$: | 2.52 | 2.49 | 3.07 | 2.87 | 1.86 | 2.08 |
| $E_t[\pi_{t+3,t+4}]$: | 2.80 | 3.01 | 3.52 | 3.54 | 1.86 | 2.39 |
| $E_t[\pi_{t+4,t+5}]$: | 3.13 | 3.39 | 4.03 | 4.20 | 1.86 | 2.36 |
| $E_t[\pi_{t+5,t+6}]$: | 3.26 | 3.52 | 4.20 | 4.56 | 1.92 | 2.00 |
| $E_t[\pi_{t+6,t+7}]$: | 3.48 | 3.87 | 4.57 | 5.05 | 1.77 | 2.07 |
| $E_t[\pi_{t+7,t+8}]$: | 3.86 | 4.26 | 5.03 | 5.49 | 1.98 | 2.36 |
| $E_t[\pi_{t+8,t+9}]$: | 3.64 | 4.28 | 4.87 | 5.66 | 1.96 | 2.42 |
| $E_t[\pi_{t+9,t+10}]$: | 3.43 | 4.14 | 4.84 | 5.78 | 2.00 | 2.51 |
| $E_t[\pi_{t+10,t+11}]$: | 3.46 | 4.03 | 5.05 | 5.92 | 2.14 | 2.47 |
| $E_t[\pi_{t+11,t+12}]$: | 3.67 | 4.47 | 5.74 | 7.17 | 2.03 | 2.24 |
| Mean inflation: | 5.54 | | 7.66 | | 3.63 | |
| St. dev.: | 10.50 | | 10.76 | | 2.57 | |

| *Horizon (Years):* | *1970 Q2 to 1994 Q3* Actual | Naive | *1970 Q2 to 1981 Q4* Actual | Naive | *1982 Q1 to 1994 Q3* Actual | Naive |
|---|---|---|---|---|---|---|
| $E_t[\pi_{t,t+4}]$: | 1.79 | 2.13 | 2.32 | 2.43 | 1.06 | 1.79 |
| $E_t[\pi_{t+4,t+8}]$: | 3.09 | 3.54 | 4.09 | 4.57 | 1.46 | 1.95 |
| $E_t[\pi_{t+8,t+12}]$: | 3.48 | 4.13 | 5.65 | 6.61 | 1.61 | 2.06 |

*Blue Chip*

| *Horizon (Quarters):* | *1985 Q1 to 1994 Q3* Actual | Naive |
|---|---|---|
| $E_t[\pi_{t,t+1}]$: | 1.20 | 1.66 |
| $E_t[\pi_{t+1,t+2}]$: | 1.56 | 1.96 |
| $E_t[\pi_{t+2,t+3}]$: | 1.64 | 1.97 |
| $E_t[\pi_{t+3,t+4}]$: | 1.69 | 2.26 |
| $E_t[\pi_{t+4,t+5}]$: | 1.76 | 2.21 |
| Mean inflation: | 3.64 | |
| St. dev.: | 2.69 | |

| Horizon (Years): | 1985 Q1 to 1994 Q3 Actual | Naive |
|---|---|---|
| $E_t[\pi_{t,t+4}]$: | 1.02 | 1.65 |

Root-mean-square errors of "actual" forecasts are computed from published information. Root-mean-square errors of "naive" forecasts are computed assuming inflation is a random walk.

## 3.2 CORRELATION OF INFLATION WITH CANDIDATE INDICATORS

The next step is to examine how well different candidate indicators are correlated with inflation. To do this, I estimate the following simple regression:

$$\pi_{t+l,\, t+k}=a(L)\pi_{t-1} + b(L)x_{t-1} + \epsilon_t(l, k), \tag{1}$$

where $\pi_{t+l,\, t+k}$ is inflation from $t + l$ to $t + k$, $\pi_{t-1}$ is inflation from $t - 2$ to $t - 1$, $x$ is a candidate indicator, $a(L)$ and $b(L)$ are lag polynominals of order 6, and $\epsilon_t$ is a stationary moving average error of order $k - 1$ with i.n.i.d. innovations.[8]

Following the work of Niemira and Klein (1994), Webb and Rowe (1994), and others, the *x*s were chosen to include several commodity price indices, the price of gold, the price of oil, monetary aggregates, interest rates, interest-rate spreads, a wage index, the trade-weighted exchange rate, a weekly hours index, the employment population ratio, capacity utilization, and unemployment.[9] For the monthly sample beginning in January 1967, the set includes seventeen candidate variables.[10]

Table 2 reports results from estimating (1) for three horizons—1 year ahead ($k$ = 12, $l$ = 0), 1 to 2 years ahead ($k$ = 24, $l$ = 12), and 3 to 4 years ahead ($k$ = 48, $l$ = 36)—over two sample periods, 1967:01 to 1994:07 and 1982:01 to 1994:07. The numbers in the table are the *p*-values for the Wald form of the test that all of the elements of $b(L)$ are zero simultaneously, computed using a covariance matrix that is robust to heteroscedasticity and serial correlation.[11]

For the most part, both real variables, such as unemployment and capacity utilization, and material prices, such as the two spot-price indices and the prices of gold or oil, are correlated with inflation at horizons of 1 or 2 years, but *not* at horizons of 3 to 4 years. The same is not true for the National Association of Purchasing Managers diffusion

8. The order of the lag polynomial was chosen to enable estimation in the following section. Where they could be computed, results were shown to be equivalent to those using 12 lags.
9. It is worth noting that there is a vast literature on forecasting inflation turning points. This work employs techniques that are similar to that used in general business-cycle forecasting, and so is a bit removed from the work here. See Webb and Rowe (1994) and the citations therein.
10. Inflation is measured using the All Items CPI-U with rental equivalence. From 1967 to 1982, this is the experimental 'CPI-U X1'. All of the results are robust to using the weighted-median CPI described in Bryan and Cecchetti (1994).
11. The covariance matrix of the estimated coefficients is calculated using Newey and West (1987), with lags equal to 1.33$k$.

Table 2 CORRELATION OF INFLATION WITH VARIOUS INDICATORS

| Indicator | 3 mo | 1 yr | 1–2 yr | 3–4 yr |
|---|---|---|---|---|
| | *p-Value* | | | |
| | | *Sample 1967:01–1994:07* | | |
| J. of comm. indus. mater. | 0.01 | 0.00 | 0.01 | 0.99 |
| NAPM spot index | 0.00 | 0.00 | 0.00 | 0.86 |
| NAPM diffusion index | 0.00 | 0.00 | 0.02 | 0.02 |
| Price of gold, London fix | 0.34 | 0.00 | 0.00 | 0.97 |
| Price of oil, Brent North crude | 0.06 | 0.11 | 0.68 | 0.74 |
| Average hourly earnings | 0.03 | 0.19 | 0.18 | 0.02 |
| Exchange rate | 0.20 | 0.24 | 0.26 | 0.34 |
| Monetary base | 0.29 | 0.25 | 0.66 | 0.81 |
| M1 | 0.72 | 0.72 | 0.90 | 0.47 |
| M2 | 0.43 | 0.08 | 0.02 | 0.58 |
| Federal funds rate *rff* | 0.00 | 0.00 | 0.00 | 0.01 |
| 10-yr-bond–*rff* spread | 0.00 | 0.05 | 0.01 | 0.27 |
| Commercial paper–*rff* spread | 0.00 | 0.01 | 0.00 | 0.02 |
| Weekly hours index | 0.01 | 0.02 | 0.00 | 0.02 |
| Capacity utilization | 0.00 | 0.00 | 0.00 | 0.90 |
| Unemployment rate | 0.03 | 0.00 | 0.00 | 0.81 |
| Employment population ratio | 0.05 | 0.01 | 0.00 | 0.00 |
| | | *Sample 1982:01–1994:07* | | |
| J. of comm. indus. mater. | 0.25 | 0.04 | 0.01 | 0.00 |
| NAPM spot index | 0.07 | 0.11 | 0.06 | 0.00 |
| NAPM diffusion index | 0.09 | 0.00 | 0.02 | 0.00 |
| Price of gold, London fix | 0.33 | 0.61 | 0.75 | 0.01 |
| Price of oil, Brent North crude | 0.75 | 0.61 | 0.64 | 0.00 |
| Average hourly earnings | 0.85 | 0.64 | 0.02 | 0.02 |
| Exchange rate | 0.96 | 0.86 | 0.80 | 0.47 |
| Monetary base | 0.80 | 0.34 | 0.98 | 0.00 |
| M1 | 0.61 | 0.12 | 0.47 | 0.00 |
| M2 | 0.38 | 0.00 | 0.66 | 0.00 |
| Federal funds rate *rff* | 0.03 | 0.00 | 0.89 | 0.00 |
| 10-yr-bond–*rff* spread | 0.00 | 0.00 | 0.00 | 0.00 |
| Commercial paper–*rff* spread | 0.07 | 0.56 | 0.09 | 0.00 |
| Weekly hours index | 0.35 | 0.69 | 0.76 | 0.00 |
| Capacity utilization | 0.20 | 0.03 | 0.20 | 0.00 |
| Unemployment rate | 0.00 | 0.00 | 0.18 | 0.00 |
| Employment population ratio | 0.28 | 0.05 | 0.70 | 0.00 |

index, the average-hourly-earnings index, the weekly-hours index, and the employment-population ratio, which do have predictive power at the longer horizon. Increasing the horizon to 5 years does not change the results.

It is worthwhile examining the case of capacity utilization in more detail. In addition to computing the simple test for all of the *b*'s equalling zero simultaneously, it is possible to calculate the sum of the coefficients—$b(1)$.[12] Interestingly, an increase in capacity utilization is correlated with an increase in inflation at horizons of up to 3 years. For example, the $t$-ratio of $b(1)$ in the ($k = 24$, $l = 12$) case is + 6.97. But after that, at horizons of 3 and 4 years, $b(1)$ is negative—for ($k = 48$, $l = 36$) it is − 1.98.

The results for the different sample periods are dramatically different. For the more recent period (1982–1994), very few variables help forecast inflation at a 1-year horizon, but most seem to be useful at longer horizons. This is likely the result of the relatively small amount of independent information used in the longer-horizon estimates.

The main conclusion to be drawn from these simple correlations is that the different indicators provide information about inflation at different horizons, and that the information has changed over time.[13]

### 3.3 INFLATION INDICATORS AND INFLATION FORECASTING

The simple regressions of the previous section may not be representative of the actual ability of an indicator to forecast inflation. Within-sample statistics suffer from standard overfitting problems. These can be addressed by constructing out-of-sample forecasts. I do this with a series of rolling regressions, in which an equation is estimated over a sample of fixed length, a one-period-ahead forecast is computed, the next observation in the sample is added and the last one dropped, and the process is repeated. Table 3 reports the results from two such experiments. The first uses a 10-year sample beginning in 1967, while the second employs a 5-year window with data beginning in 1982. Again, the calculations are done for forecasts at various horizons. In addition to reporting the simple root-mean-square error of the forecasts, the table includes the rank correlation between the RMSE of a model and its Bayes Information Criterion (BIC) for the initial sample.[14]

12. There are a number of well-known pitfalls associated with the interpretation of b(1). Unless the sample includes periods in which capacity utilization movements are sustained for a number of months, then the sum of the coefficients is not a meaningful thing to compute. See Fisher and Seater (1993) for a discussion of related issues.
13. A similar result emerges from a more complex (and computationally intensive) exercise of estimating and ranking all of the possible models with all subsets of seventeen indicator variables. Using the Bayes information criterion (see footnote 14 below) as a ranking criterion, substantially different models are chosen for different horizons and sample periods. For example, using the entire sample period and a horizon of twelve months, the preferred model includes the NAPM spot index, the price of oil, M2, capacity utilization, and unemployment rate. But for a 36-month horizon and the full 1967:01–1994:07 sample the "best" model includes only M2 and the federal funds rate.
14. For a model with $p$ parameters estimated over a sample of length $T$, the BIC is defined as $[\ln \sigma_\epsilon^2 + {}^{p}\!/_{T} \ln T]$, where $\sigma_\epsilon^2$ is the error variance. These statistics, which are similar to an adjusted $R^2$, are only suggestive, as their relevance has been established only for the case in which the regression error process is not serially correlated.

Table 3 ROOT-MEAN-SQUARE ERRORS IN ONE-STEP AHEAD ROLLING FORECASTS

| | *Sample 1967:01–1994:07, 10-yr Window* | | | |
|---|---|---|---|---|
| *Indicator* | *3 mo* | *1 yr* | *1–2 yr* | *3–4 yr* |
| CPI only | 2.14 | 1.70 | 2.39 | 2.50 |
| J. of comm. indus. mater. | 2.15 | 1.64 | 2.38 | 2.63 |
| NAPM spot index | 2.07 | 1.63 | 2.36 | 2.60 |
| NAPM diffusion index | 2.09 | 1.48 | 2.19 | 2.54 |
| Price of gold, London fix | 2.24 | 1.68 | 2.35 | 2.73 |
| Price of oil, Brent North crude | 2.35 | 1.73 | 2.43 | 2.61 |
| Average hourly earnings | 2.09 | 1.71 | 2.43 | 2.44 |
| Exchange rate | 2.24 | 1.71 | 2.27 | 2.63 |
| Monetary base | 2.14 | 1.64 | 2.27 | 2.55 |
| M1 | 2.25 | 1.73 | 2.45 | 2.48 |
| M2 | 2.22 | 1.79 | 2.47 | 2.67 |
| Federal funds rate *rff* | 2.27 | 1.77 | 2.05 | 1.90 |
| 10-yr-bond–*rff* spread | 2.13 | 1.77 | 2.22 | 2.53 |
| Commercial–paper–*rff* spread | 2.27 | 1.87 | 2.25 | 2.19 |
| Weekly hours index | 2.28 | 1.81 | 2.34 | 2.08 |
| Capacity utilization | 2.14 | 1.52 | 2.24 | 2.83 |
| Unemployment rate | 2.20 | 1.65 | 2.33 | 2.89 |
| Employment population ratio | 2.27 | 1.75 | 2.52 | 2.17 |
| Rank correlation of initial sample | − 0.13 | 0.15 | 0.01 | 0.00 |
| BIC with RMSE | (0.25) | (0.25) | (0.25) | (0.25) |
| | *Sample 1982:01–1994:07, 5-yr Window* | | | |
| *Indicator* | | | | |
| *CPI only* | *1.54* | *1.28* | *1.43* | |
| *J. of comm. indus. mater.* | *1.60* | *1.45* | *1.30* | |
| *NAPM spot index* | *1.72* | *1.45* | *1.39* | |
| *NAPM diffusion index* | *1.70* | *1.42* | *1.49* | |
| *Price of gold, London fix* | *1.60* | *1.35* | *1.68* | |
| *Price of oil, Brent North crude* | *1.95* | *1.50* | *1.72* | |
| *Average hourly earnings* | *1.65* | *1.32* | *1.47* | |
| *Exchange rate* | *1.64* | *1.38* | *1.56* | |
| *Monetary base* | *1.68* | *1.14* | *1.61* | |
| *M1* | *1.54* | *1.08* | *1.58* | |
| *M2* | *1.55* | *1.16* | *1.32* | |
| *Federal funds rate rff* | *1.43* | *1.16* | *0.97* | |
| *10-yr bond–rff* spread | *1.43* | *0.90* | *1.30* | |
| *Commercial–paper–rff* spread | *1.61* | *1.35* | *1.48* | |
| *Weekly hours index* | *1.59* | *1.17* | *1.22* | |
| *Capacity utilization* | *1.56* | *1.06* | *1.03* | |
| *Unemployment rate* | *1.45* | *0.80* | *1.07* | |
| *Employment population ratio* | *1.28* | *0.77* | *1.12* | |
| *Rank correlation of initial sample* | *− 0.10* | *− 0.03* | *− 0.22* | |
| *BIC with RMSE* | *(0.25)* | *(0.25)* | *(0.24)* | |

The first line in each panel of the table reports the results of forecasting inflation using inflation alone. As is clear, additional variables can easily worsen the forecast. For example, when forecasting inflation from 3 to 4 years into the future, the addition of eleven of the seventeen indicators raises the root-mean-square-forecast error. Adding any of the price measures worsens the forecasts. Only the wage index, M1, the federal funds rate, the commercial paper spread, the weekly hours index, and the employment population ratio improve the forecast. At shorter horizons things are a bit better. Regardless of the sample period, only eight of the seventeen variables worsen forecasts over the next year.

A number of other interesting results emerge from this exercise. First, with the exception of the 3-month horizon, this forecasting method compares favorably with the commercial forecasts, though both are rather poor. Second, there is virtually no correlation between the ranking of the models by their RMSEs and their initial sample BICs.[15] Whether a model fits well in sample tells us virtually nothing about its out-of-sample forecasting ability.

## 4. *Structural Breaks in the Inflation Process*

A natural conclusion to draw from the simple examination of the previous section is that the inflation process is changing over time. If these changes were gradual, then the rolling-sample procedures used in Section 3.3 might take care of the problem. But there is a good reason to believe that this will not be the case, and that the correlation between inflation and candidate indicators, the $x$'s, will display structural breaks.[16]

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

$$\pi_{t+1} = \alpha(L)r_t + \beta(L)X_t + \omega_{t+1}, \tag{2}$$

where, $r_t$ is "policy," and $X_t$ is a vector of determinants.

Next write the policy reaction function as

$$r_t = \gamma(L)X_t + \nu_t. \tag{3}$$

15. Miyao (1994, Chapter 3) reports a similar result for the relationship between real and monetary variables.

16. A number of researchers have modeled changes in the inflation process. Caskey (1985) examines a linear model with Bayesian learning, and Evans and Wachtel (1993) investigate the possibility that inflation shifts between a stationary and unit root process.

The role of the policy maker is to choose $\gamma(L)$, the reaction of $r_t$, to observed $X$s. Since $\gamma(L)$ can contain zeros, a policy regime need not react to every element in $X$.

Now consider the reduced-form regression:

$$\pi_{t+1} = R(L)X_t + \eta_t. \tag{4}$$

Since

$$R(L) = \alpha(L)\gamma(L) + \beta(L), \tag{5}$$

changes in policy, which are changes in $\gamma(L)$, will change the correlation between $X$ and $\pi$. In effect, the reduced-form inflation regressions subsume the monetary policy reaction function (3), and so a change in the monetary authority's policy rule—which may be a change in the relative weight placed on various indicators—will cause changes in (4).[17]

Note also that if the policy objective is to minimize Var $(\pi_t)$, then the optimal policy reaction function sets $\gamma(L) = -\alpha(L)^{-1}\beta(L)$ and so inflation is uncorrelated with its determinants. This is one version of a point made by Kareken and Solow (1963) and Sims (1972), and more recently by Woodford (1994).

This suggests looking for structural breaks in the relationship between inflation and indicators. Shifts should occur near the times at which monetary policy procedures changed. To do this, I examine regressions of inflation 1 year ahead on candidate indicators—equation (1) in Section 3.2 with ($l = 0$, $k = 12$)—using a combination of the Andrews–Ploberger (1994) exponential Wald tests for structural stability, and Andrews (1993)–Quandt (1960) tests.

The Andrews–Quandt test is the maximum value of the statistic associated with the test that a break occurred at each point in the sample. It provides an estimate of the break date itself. The Andrews–Ploberger test is an exponentially weighted average of the statistics assuming a break at each date in the sample; it tests for structural stability generally. Calculation of the second of these requires the choice of a truncation parameter ($\tau$) denoting the proportion of the beginning and of the end of the sample not to be used in the computations. The results reported below choose $\tau$ to be the same proportion at the beginning and end of the sample, and equal to the number of right-hand-side variables plus 12 months, divided by the sample size.

17. This is yet another form of the Lucas (1976) critique.

As it is currently worked out, the econometric theory that forms the basis for these tests presumes that there is a single structural break in the relation under study. Since there is every reason to believe that the monetary policy process has changed more than once over the 1967–1994 sample, one would like a procedure that suggests more than one break date. To do this, I have employed these tests in a recursive manner. If the Andrews–Ploberger test rejects structural stability at the 5% level, then the sample is split at the date implied by the Andrews–Quandt test.[18] Assuming that enough data remain—in the results reported here, the sample must be a minimum of 3 years long—the tests are run again.

The results of this sequential procedure are reported in Table 4. The tests are robust to heteroscedasticity and serial correlation.[19] The full sample extends from January 1967 to July 1994. In addition to estimated break dates, the table reports the $p$-values for the Wald test that all of the elements of $b(L)$ are zero simultaneously for a sample beginning the month following the previous break (or the beginning of the full sample) until the next break (or the end of the full sample). Since the right-hand-side variable in the regressions is inflation over the preceding 12 months, these dates can be interpreted as suggesting a change in the inflation process sometime during the year following the reported date.

To understand how the table is constructed, take the example of the M2. The results suggest that the relationship between inflation and M2 changed four times over the sample, with estimated breaks in April 1972, October 1978, August 1983, and September 1989. This leaves five stable subsamples: the three between these four dates, the one from the beginning of the sample in January 1967 to the date of the first break in 1972, and the final one from September 1989 to the end of the sample in July 1994. The results from the Wald tests show that the coefficients on M2 in the inflation regression are significantly different from zero at standard levels of statistical significance in all but one of the subperiods. For the sample from 1983:08 to 1989:09, the $p$-value for the test of the coefficients on M2 is only 0.08. For the remainder of the subperiods, the estimated $p$-value is below 0.05.

The results show a number of interesting features. First, all of these relationships are highly unstable, with a minimum of three estimated

18. The $p$-values for these tests are constructed using the techniques described in Hansen (1995).

19. All test statistics are robust to both heteroscedasticity and serial correlation, using the Newey–West (1987) formulation with $m = 15$. The serial correlation correction is required, since the estimated regressions make use of overlapping data.

Table 4 ESTIMATED TIMING OF STRUCTURAL BREAKS IN THE INFLATION PROCESS, FULL SAMPLE 1967:01 to 1994:07: $\pi_{t,t+12} = a(L)\pi_{t-1} + b(L)x_{t-1} + e_t$

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CPI only | 0.58 | 72:04 | | | | | 0.93 | 79:06 | 0.00 | 82:12 | 0.25 | 86:08 | 0.33 | 90 .01 | 0.77 |
| J. of comm. indus. mater. | | | 0.00 | 73:12 | | | 0.00 | 80:12 | | | 0.00 | 85:09 | 0.21 | | |
| NAPM spot index | 0.00 | 72:09 | | | 0.00 | 76:10 | | | 0.00 | 82:11 | | | 0.23 | 89:08 | 0. 12 |
| NAPM diffusion index | 0.00 | 72:10 | | | | | 0.00 | 79:01 | 0.00 | 83:08 | | | 0.00 | 89: 07 | 0. 01 |
| Price of gold, London fix | 0.03 | 72:03 | | | 0.00 | 76:06 | | | 0.00 | 82:07 | | | 0.00 | 88:05 | 0.20 |
| Price of oil, Brent North crude | | | 0.90 | 73:01 | 0.00 | 77:02 | | | 0.00 | 81:05 | 0.71 | 86:05 | 0.04 | | |
| Exchange rate | 0.02 | 72:08 | | | 0.01 | 75:11 | | | 0.00 | 81:04 | | | 0.32 | 89:06 | 0. 58 |
| Average hourly earnings | 0.63 | 72:08 | | | 0.25 | 77:01 | | | 0.66 | 80:12 | 0.51 | | | | |
| Monetary base | | | 0.82 | 73:12 | | | 0.00 | 79:02 | 0.00 | 83:06 | | | 0.66 | 88:07 | 0. 00 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M1 | 0.07 | 72:03 | | | 0.54 | 78:04 | | | 0.03 | 82:06 | 0.87 | 86:10 | 0.00 | | |
| M2 | 0.00 | 72:04 | | | | | 0.00 | 78:10 | 0.04 | 83:08 | | | 0.08 | 89:09 | 0.00 |
| Federal funds rate $r^{ff}$ | 0.00 | 72:10 | | | 0.00 | 76:03 | 0.00 | 80:02 | | | 0.00 | 85:04 | 0.39 | 89:09 | 0.00 |
| 10-yr-bond–$r^{ff}$ spread | 0.00 | 72:09 | | | 0.00 | 76:03 | 0.00 | 80:11 | | | 0.00 | 85:02 | 0.00 | 89:02 | 0.00 |
| Commercial-Paper–$r^{ff}$ spread | 0.00 | 72:10 | | | 0.00 | | 0.00 | 78:12 | 0.57 | 83:04 | 0.00 | | | 89:07 | |
| Weekly hours index | | | 0.00 | 74:04 | 0.00 | | 0.00 | 79:12 | | | 0.00 | 84:09 | 0.00 | 89:04 | |
| Capacity utilization | 0.00 | 72:08 | | | 0.00 | 76:01 | | | 0.00 | 80:08 | 0.00 | 84:07 | 0.00 | 89:07 | 0.00 |
| Unemployment rate | 0.00 | 72:09 | | | 0.00 | 76:04 | | | 0.00 | 81:04 | 0.15 | 85:06 | 0.00 | 89:09 | 0.00 |
| Employment population ratio | 0.00 | 72:04 | | | 0.00 | 78:06 | | | 0.00 | 82:05 | 0.40 | 87 :10 | 0.00 | | |

breaks for a series.[20] In addition, the break dates show a distinct pattern. Nearly all of the series show shifts in early to mid-1972, between 1979 and 1982, and around 1989. A number of series also exhibit evidence of breaks in 1976. The four real variables show breaks in the mid-1980s as well.

Several of these shifts are as expected, since they are around the times of the three important changes in Federal Reserve operating procedures that occurred during the sample. These were: (1) the 1972 shift from free-reserve to federal-funds-rate targeting,[21] (2) the 1979 move to a nonborrowed-reserves operating procedure, and (3) the 1982 change back to federal-funds-rate targeting.[22]

The $p$-values for the test $H_0: \mathbf{b} = 0$, where $\mathbf{b}$ is the vector of coefficients in $b(L)$, also suggest some interesting conclusions. In only four cases—the NAPM diffusion index, the T-bond spread, capacity utilization, and weekly hours—are the $p$-values always near zero, implying that the indicators are always correlated with inflation over the next year, but that the correlation changes. For the remainder, the hypothesis that the correlation is zero is not rejected for at least one subperiod.[23] Policy makers are not just changing their emphasis on particular indicators; during some periods they appear to be ignoring some of them completely.

One criticism of the results in Table 4 is that they consider only bivariate relationships. There might be some multivariate inflation equation that is stable over the 27(½)-year sample period. But examination of multivariate models suggests that the problem persists. To establish this, I began by performing a model selection exercise in which I considered all 65,536 possible models comprising all of the subsets of the seventeen indicators variables, each entered with six lags. Ranking the models by their BIC, the "best" model for inflation one year ahead included the NAPM spot index, the price of oil, M2, capacity utilization, and the unemployment rate. The recursive structural break procedure in this specification reveals three break dates: 1974:12, 1981:01, and 1987:04. These clearly conform to the pattern in Table 4.[24]

Taken together, the results of the last two sections lead one to draw

20. These results are consistent with those of Stock and Watson (1994), who find widespread instability in bivariate relationships between macroeconomic time series.
21. This may also represent the imposition of the Nixon wage–price controls, which began in August 1971, and the end of the Bretton Woods system shortly thereafter.
22. See Strongin (1992) for a full description of these changes.
23. Both the signs and magnitude of the coefficients in these regressions change as well. For eleven of the seventeen indicators, there is a change in the sign of the sum of the response of inflation between subsamples. For the remaining six, coefficients often change by factors of 5 to 10.
24. The same result holds for the second and third specifications ranked by their full-sample BIC.

some fairly strong conclusions. First, inflation is extremely difficult to forecast with any accuracy. Second, different indicators provide information about future inflation over different horizons. And most importantly, shifts in the reduced form correlations occur frequently and at times that suggest they are the result of changes in Federal Reserve operating procedures.[25]

## 5. *The Relationship between Policy and Inflation*

The real-time conduct of monetary policy requires quantitative knowledge of the link between instruments and targets. For example, if the federal funds rate moves, then when and by how much does the price level change? An estimate of the impulse response of prices to policy innovations is one answer to this question.

There is now a vast literature on identification of monetary policy disturbances.[26] One currently popular technique is to examine a reduced-form vector autoregression (VAR) that includes measures of the log of output, the log of aggregate prices, the log of commodity prices, and a monetary policy indicator, such as federal funds rate or the log of nonborrowed reserves. The structural form of such a model can be written as

$$\begin{bmatrix} p_t \\ pf_t \\ y_t \\ r_t \end{bmatrix} = A(L) \begin{bmatrix} \epsilon_{pt} \\ \epsilon_{ct} \\ \epsilon_{yt} \\ u_t \end{bmatrix} \tag{6}$$

where $p_t$, $pf_t$, and $y_t$ are the logs of the aggregate price level, commodity prices, and output, respectively, $r_t$ is the policy indicator, $A(L)$ is a 4×4 matrix of lag polynomials in the lag operator $L$, the $\epsilon$s are the "exogenous" shocks, and $u$ is the policy innovation.

For the purposes of the exercise here, all that is needed is to identify the response of the four variables in the system to the policy innovation $u_t$. This limited identification is achieved by assuming that no other variables respond to monetary policy shocks contemporaneously. That is, the first three rows of the fourth column of $A(0)$ are each zero.

Given the results of the previous section, it seems foolish to presume that the response of the aggregate price level to innovations in the federal funds rate would be invariant over the past quarter century. In fact,

25. It is interesting to note that the same set of tests applied to the seventeen bivariate relationships of federal funds with the indicators (and inflation itself) show no evidence of structural breaks in any of the cases.

26. See Sims (1992) and Christiano, Eichenbaum, and Evans (1994a, 1994b) for discussions.

the estimates in Table 4 suggest that there have been five changes in the bivariate relationship between inflation and the federal funds rate—at 1976:03, 1980:02, 1985:04, and 1989:09. The average length of a regime is under 5 years.

In an attempt to take account of structural instability, I have computed the response of prices and output to policy innovations from a model estimated using a fixed 10-year rolling sample. I measure prices by the CPI, commodity prices by the *Journal of Commerce* index, output as industrial production, and monetary policy using the federal funds rate.[27]

Figure 2 plots the impulse response for a unit federal-funds-rate shock for horizons of 1, 2, and 3 years. The result of each regression is plotted on the final date of the sample used. For example, the responses plotted on December 1984 are computed from a VAR estimated over the ten years of data beginning with January 1975. The top panel plots results for the log of prices, and the bottom panel plots those for the log of industrial production. The results are roughly as expected. One year following a monetary tightening, prices are usually predicted to fall, but by only a small amount. By comparison, following a policy action, output is forecast to fall substantially. After three years, output will have recovered, while prices will have fallen substantially.

While the results in Figure 2 imply that the relationship between interest rates and inflation is unstable, they strongly suggest that monetary policy tightening leads to price declines. But these are only point estimates. The estimates of $A(L)$ have a sampling distribution that can be constructed using the delta method. Figure 3 reports the impulse responses for prices, with 2-standard-deviation bands. To aid in comparison, the vertical scale is the same in all three panels of the figure. The results are quite striking. In only a few isolated cases it is possible to reject (at conventional levels of statistical significance) the hypothesis that policy has no effect on prices.

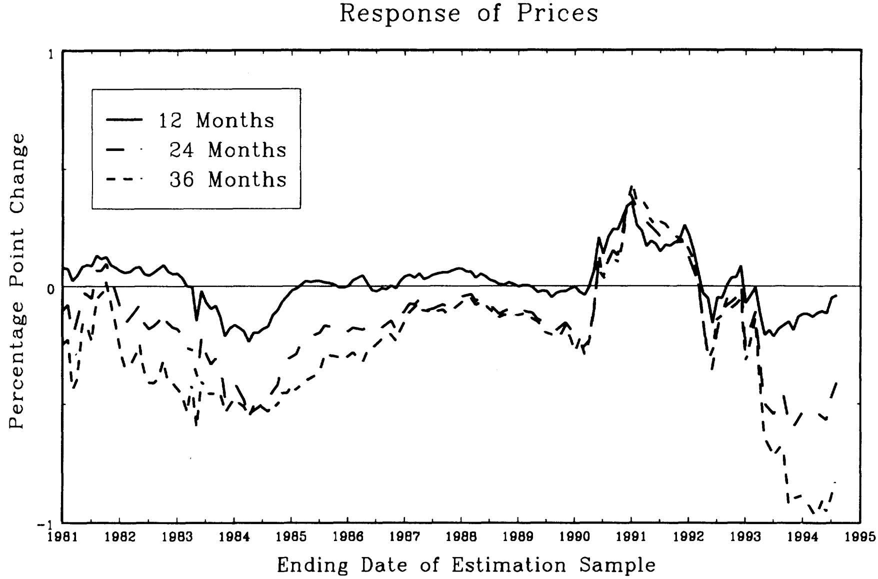
## 6. *Policy Rules*

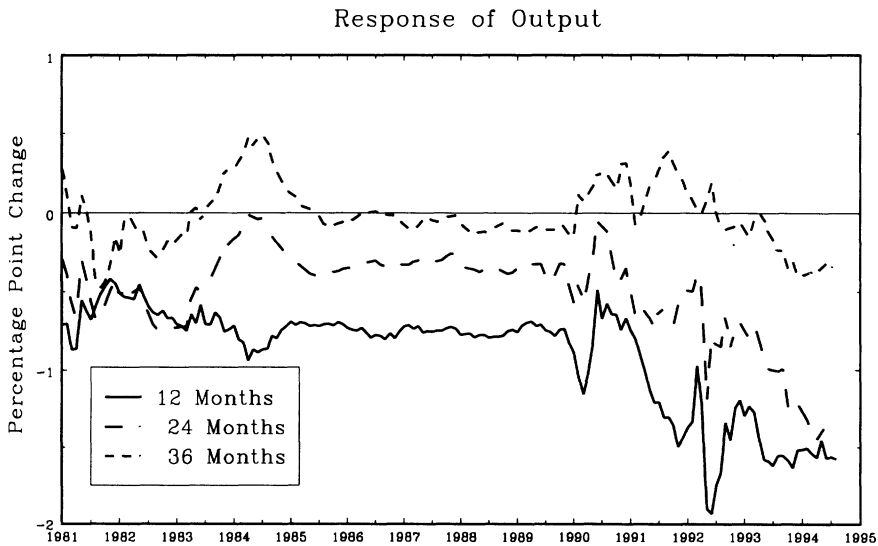
The results thus far highlight the importance of taking account of the dramatic imprecision both in the inflation forecasts and in the estimates of the effect of policy on prices in formulating policy rules.[28] The goal of this section is to compute several policy rules that take account of some of these sources of our ignorance. Once constructed, these rules provide

27. The general character of the results is robust to changes in the commodity price index, use of an interpolated estimate of monthly GDP, and the substitution of the log of nonborrowed reserves for the federal funds rate.

28. See Brainard (1967) for a discussion of this point.

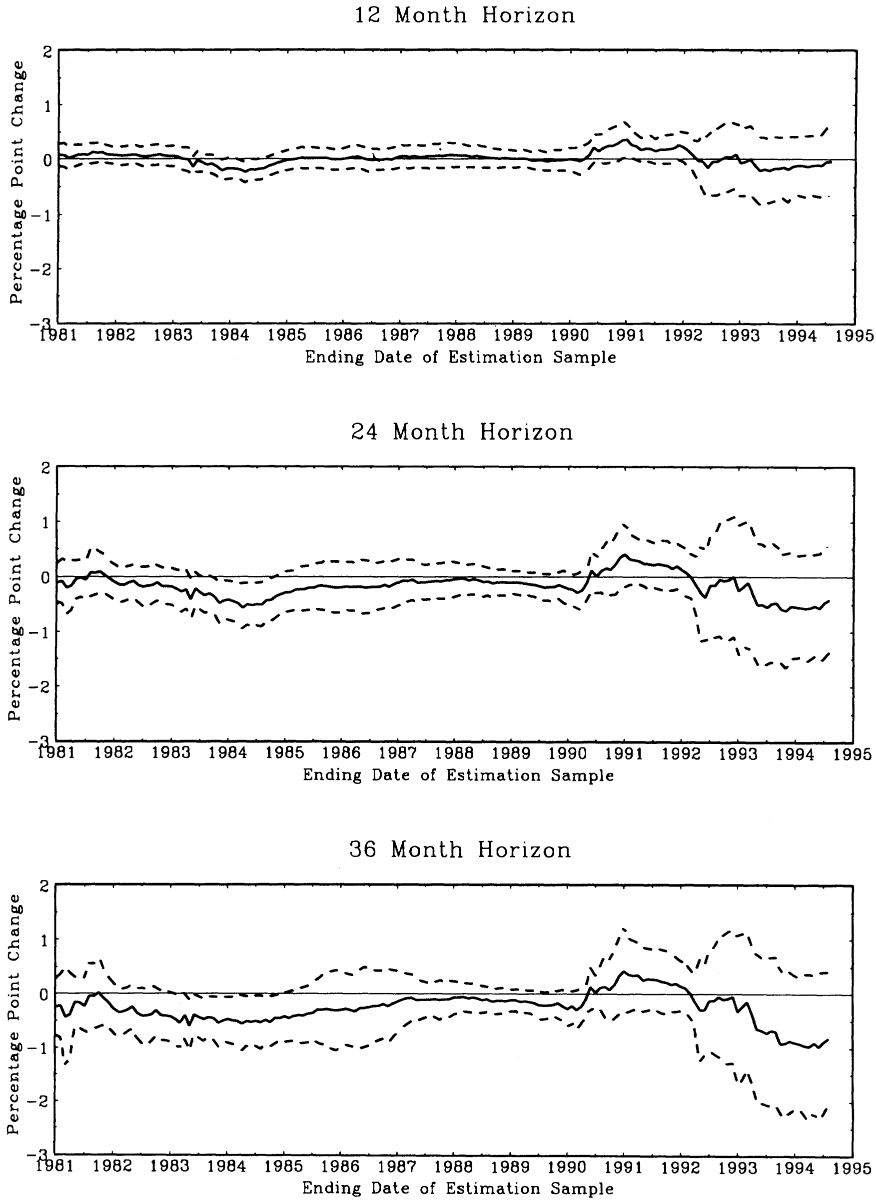
Figure 2 RESPONSE OF PRICES AND OUTPUT TO A UNIT FEDERAL FUNDS RATE SHOCK: 12-, 24- AND 36-MONTH HORIZONS; ROLLING REGRESSIONS USING 10 YEAR SAMPLES




Note: Response of prices to a one-percentage point federal funds rate shock at 12, 24, and 36 months, from a VAR with output, aggregate prices, commodity prices and the federal funds rate, estimated over a sample of ten years. The horizontal axis is the final month of the data sample used for estimation.

Figure 3 RESPONSE OF PRICES TO UNIT FEDERAL FUNDS RATE SHOCK: ROLLING REGRESSIONS USING 10-YEAR SAMPLES (WITH TWO STANDARD DEVIATION BANDS)



help in answering some immediately relevant questions. First, once policy makers sense that an exogenous shock has hit the economy, what is the time profile of the optimal response? Second, and more importantly, given the apparent practical limits, is price-level targeting really desirable, or might nominal income targeting be more sensible? And finally, has recent policy followed a path that is close to the one implied by one of the rules?

Formulation of a policy rule proceeds in several clear steps. First, an identified model must be specified. This allows estimation of the response of the variable of interest, i.e. prices and output, to both exogenous shocks and policy innovations. Next, it is necessary to assume a form for the loss function of the policy maker. Minimization of the loss function then yields appropriate responses to the shocks.

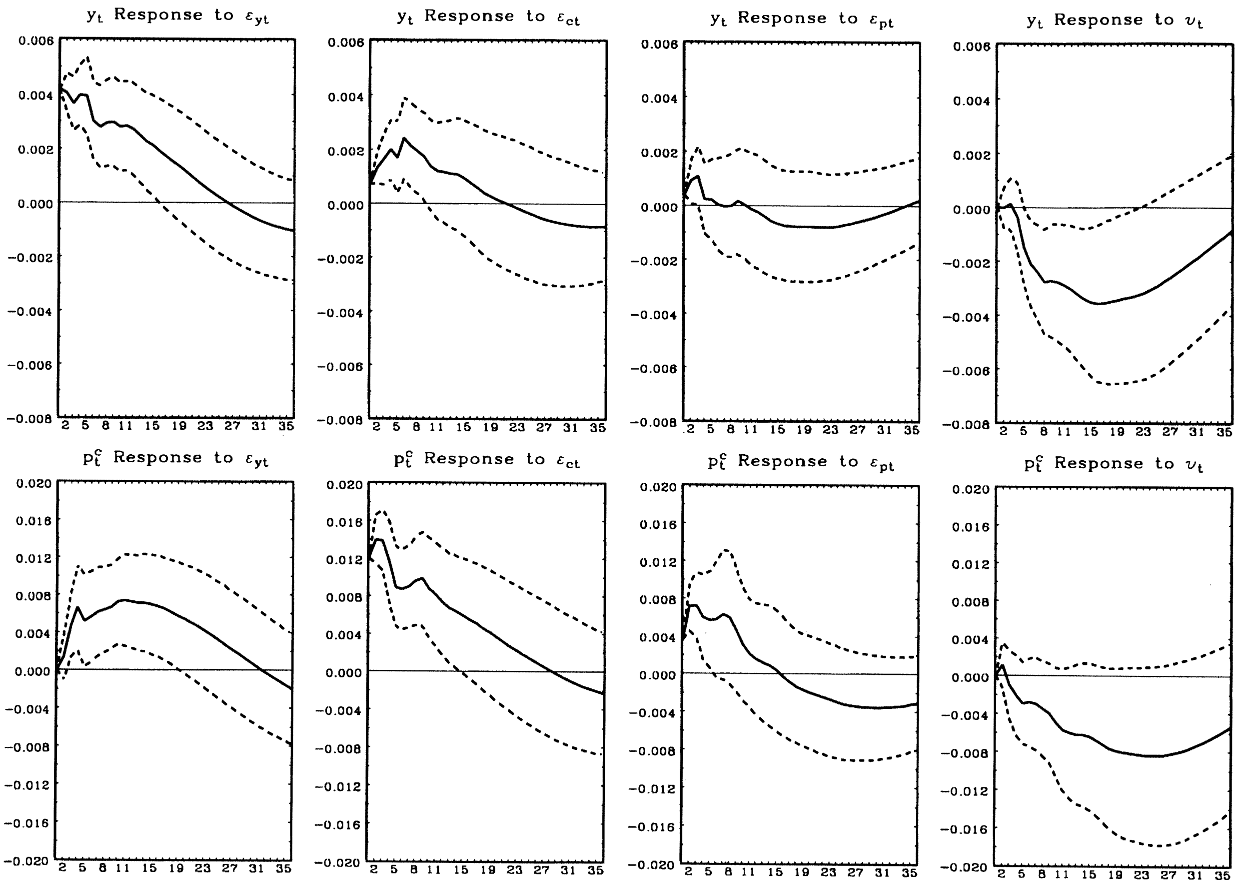
Application of this procedure begins with the VAR described in the previous section, the model (6), estimated over the sample from 1984:01 to 1994:07. Identification is achieved by assuming that VAR disturbances are related to the underlying economic shocks by a triangular (Choleski) decomposition of the covariance matrix of the reduced-form errors.[29] As described above, the monetary-policy shock is identified by assuming that no variable other than the federal funds rate responds to it contemporaneously. I label the portion of the reduced-form error in the output equation orthogonal to this as the "output" shock, the portion of the error in the commodity price equation orthogonal to these two as the "commodity price" shock, and the final part of the residual in the aggregate price equation that is orthogonal to all three of these as the "aggregate price" shock. Technically, the restriction is that $A(0)$ is lower triangular. This procedure, due originally to Sims (1980), has numerous shortcomings, but experimentation with an alternative suggested by Galí (1992) yields very similar results.

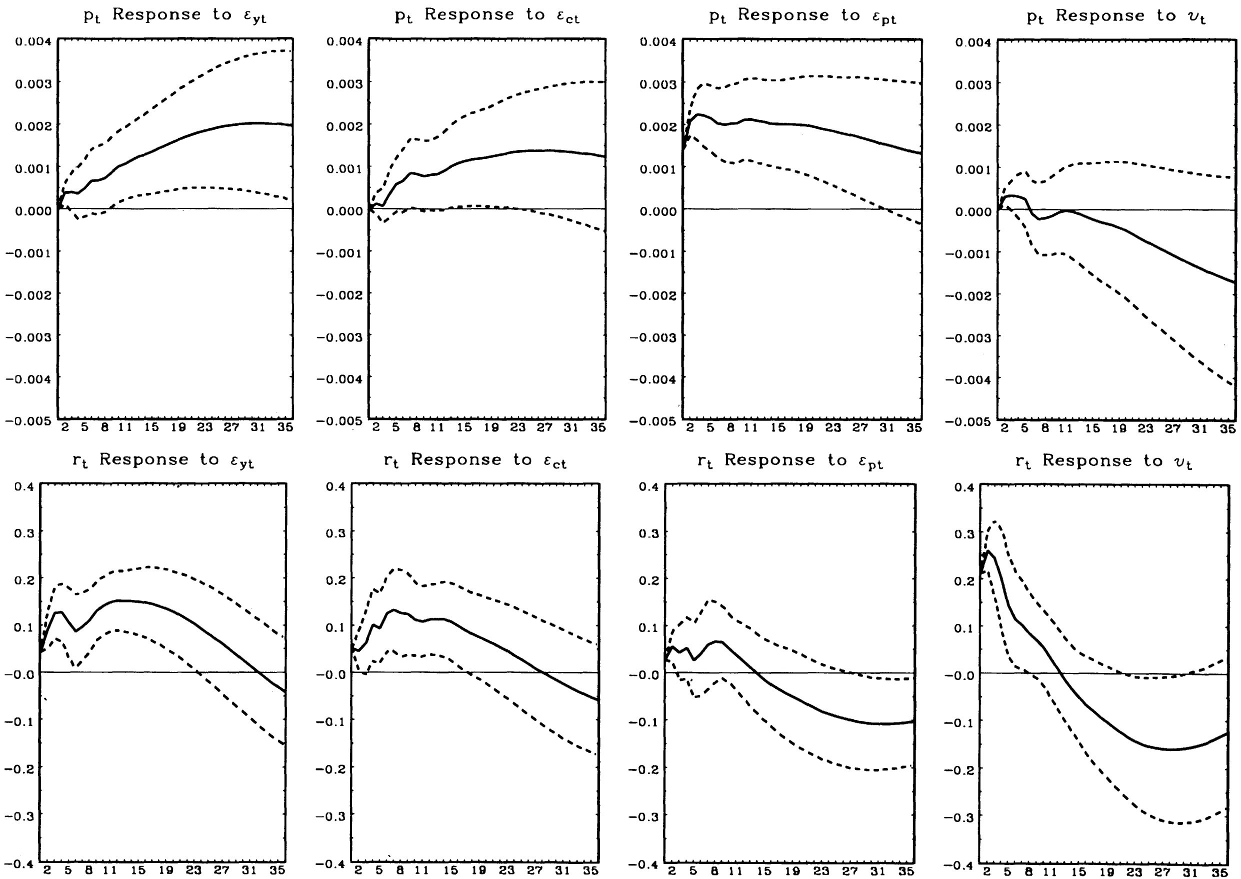
Figure 4 plots the impulse responses of output, commodity prices, aggregate prices, and the federal funds rate to "output," "commodity price," "aggregate price," and "funds rate" shocks.[30] Each response function is plotted with 2-standard-deviation bands constructed using the delta method. Two things are worth noting. First, in all cases commodity prices (plotted in the second row) respond more quickly than and in the same direction as aggregate prices (plotted in the third row). And second,

29. See Christiano, Eichenbaum, and Evans (1994b) for a discussion of this identification procedure.

30. These shocks have a very close correspondence to those that emerge from the Galí procedure, where they can be identified as a money-supply shock, a raw-material price shock, a general aggregate supply shock, and a general aggregate demand shock.

IMPULSE RESPONSE FUNCTIONS FOR FOUR-VARIABLE VAR (ESTIMATED RESPONSE, WITH TWO-STANDARD DEVIATION BANDS)






Note: Responses of output, $y_t$, commodity prices, $p_t^c$, aggregate prices, $p_t$, and the federal funds rate to one-standard deviations innovation in output, $\epsilon_{yt}$, commodity prices, $\epsilon_{ct}$, aggregate prices, $\epsilon_{pt}$ and the monetary policy, $u_t$.

for the three $\epsilon$-shocks, the response of output appears more precisely estimated than the response of aggregate prices.[31]

While they are somewhat interesting in and of themselves, the main use of these impulse responses is to construct policy rules given some objective function. To see how this is done, begin by noting that, for a given path of shocks, the first equation of the model (6) implies an estimated value for the aggregate price level of

$$\hat{p}_t = \hat{A}_{11}(L)\epsilon_{pt} + \hat{A}_{12}(L)\epsilon_{ct} + \hat{A}_{13}(L)\epsilon_{yt} + \hat{A}_{14}(L)u_t, \tag{7}$$

where the "$\hat{\ }$" denotes an estimated value. The $A_{1i}$'s are the impulse response functions plotted in the third row of Figure 4. In the context of the model (6), a policy rule is a sequence of $u_t$'s given the realization of the $\epsilon$'s, constructed to minimize a particular objective function.

One version of price-level targeting is to choose a sequence of $u$'s to minimize the average expected mean square error (MSE) of inflation over some future horizon:

$$\min_{\{u_i\}} \frac{1}{h} \sum_{i=1}^{h} E(\hat{p}_i - p_0)^2, \tag{8}$$

where $p_0$ is the log of the base-period price level, $h$ is the policy maker's horizon, and the expectation is over the sampling distribution of $\hat{p}$, which is related to the covariance matrix of the estimated coefficients in equation (7).[32] The $u$s are allowed to follow simple linear rules of the form

$$\bar{u}_i = \sum_{j=(p,c,y)} \sum_{k=0}^{\infty} \alpha_{jk}\epsilon_{jk}, \tag{9}$$

where the $\alpha_{jk}$'s are constants that constitute the rule, and $\bar{u}_i$ is the sequence of values that minimize the objective function (8). Taking account of the imprecision in the estimation of the impulse response functions goes only part of the way toward addressing the problems described in the previous sections. A full treatment of uncertainty would require accounting for the frequency and size of structural breaks as well.

I examine results based on three policy objectives. The first, which

31. This is consistent with Cochrane's (1993) observation that real output is forecastable with high $R^2$ at horizons of several years.

32. In the implementation of this technique, a heteroscedasticity-robust covariance matrix of the coefficient estimates in the reduced-form VAR is used.

might be termed passive, holds the federal funds rate fixed in the face of the shock.[33] The other two, which I will call active, minimize the average MSE of either price level or nominal output over a 36-month horizon ($h = 36$).[34] For each rule, I examine three experiments, one for each structural shock. In each of the nine resulting cases, $\epsilon_{j0} = 1$, and $\epsilon_{lk} = 0$ 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 estimates for the sequences of $\{\alpha_{jk}\}$ for each $j$ individually.

Figure 5 reports the implied path of the federal funds rate, aggregate prices, and industrial production for each policy objective in response to three structural shocks. The fixed federal funds rate policy results in consistently higher output and prices than either of the other two policies. The activist policies both have the same profile, regardless of the source of the shock. Output and prices both rise initially, and then fall, with output falling 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 Figure 4. The implication is that a policy maker wishing to stabilize prices must respond to exogenous shocks quickly in order to insure that future movements in prices are minimized. That is the argument for the Federal Reserve tightening immediately upon first sight of upward price pressure.

These calculations also 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 RMSE for inflation and nominal income for each policy. For the price-targeting case, these are the square root of the minimized objective function (8).

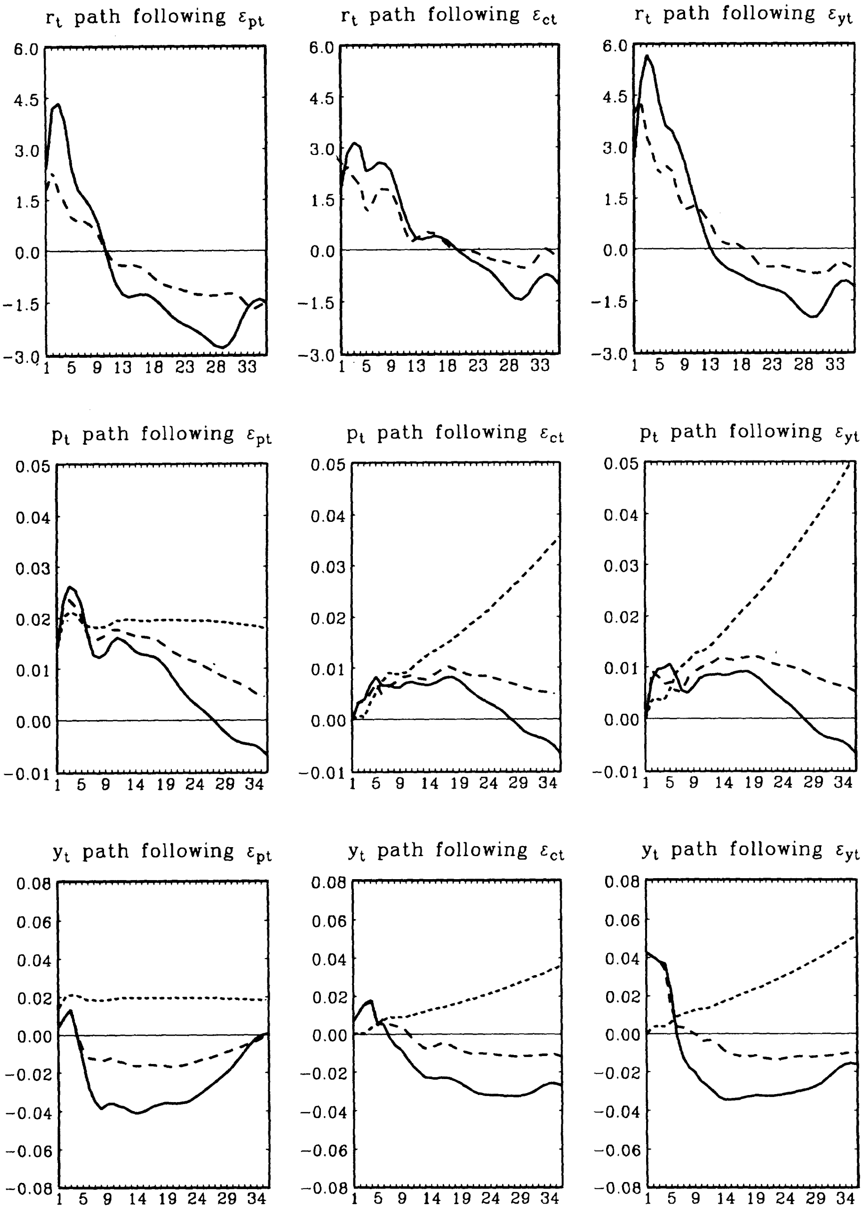
The results are reported in Table 5. As a baseline, I included the calculation of the RMSE for a case labeled "No shocks," in which all of the $u_i$s are set to zero. (This is a "policy" in which the authorities desist from introducing innovations into the Federal funds rate, and so it follows the path implied by the estimated reaction function.)

The computations suggest that nominal-income targeting has a type of robustness, as 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 0.90 to 1.09. But the move from price-level targeting to nominal-income targeting decreases the RMSE of nominal

33. As is clear from the model, this is not really a passive policy, as it involves shocks to overcome the estimated automatic reaction function.

34. Since the model is estimated in logs, the minimum MSE of nominal income policy minimizes the MSE of the sum of the log of industrial production and the log of the CPI.

Figure 5 INTEREST-RATE, OUTPUT, AND PRICE PATHS FOLLOWING SHOCKS AND POLICY RESPONSE




Note: Optimal policy responses to one-standard deviation shocks to output, $\epsilon_{yt}$, commodity prices, $\epsilon_{ct}$, and aggregate prices $\epsilon_{pt}$. The solid line is for a rule that minimizes the mean-square-error of inflation, the short-dashed line is for a rule that fixes the federal funds rate, and the long-dashed line is for a rule that minimizes the mean-square-error of nominal income.

Table 5 COMPARISON OF POLICY RESPONSES

| *Policy Rule* | Aggregate-Price shock | Commodity-Price Shock | Output Shock |
|---|---|---|---|
| | *Average RMSE of Inflation over 36-Month Horizon* | | |
| No shocks | 1.75 | 1.20 | 1.66 |
| Fixed interest rate | 2.01 | 2.13 | 2.78 |
| Min MSE ($\pi + \dot{y}$) | 1.51 | 0.93 | 1.09 |
| Min MSE ($\pi$) | 1.29 | 0.78 | 0.90 |
| | *Average RMSE of Nominal Income over 36-Month Horizon* | | |
| *Policy Rule* | Aggregate-Price Shock | Commodity-Price Shock | Output Shock |
| No Shocks | 1.92 | 2.16 | 3.25 |
| Fixed interest rate | 2.78 | 6.53 | 9.30 |
| Min MSE ($\pi + \dot{y}$) | 1.51 | 1.67 | 1.89 |
| Min MSE ($\pi$) | 2.64 | 2.53 | 3.01 |

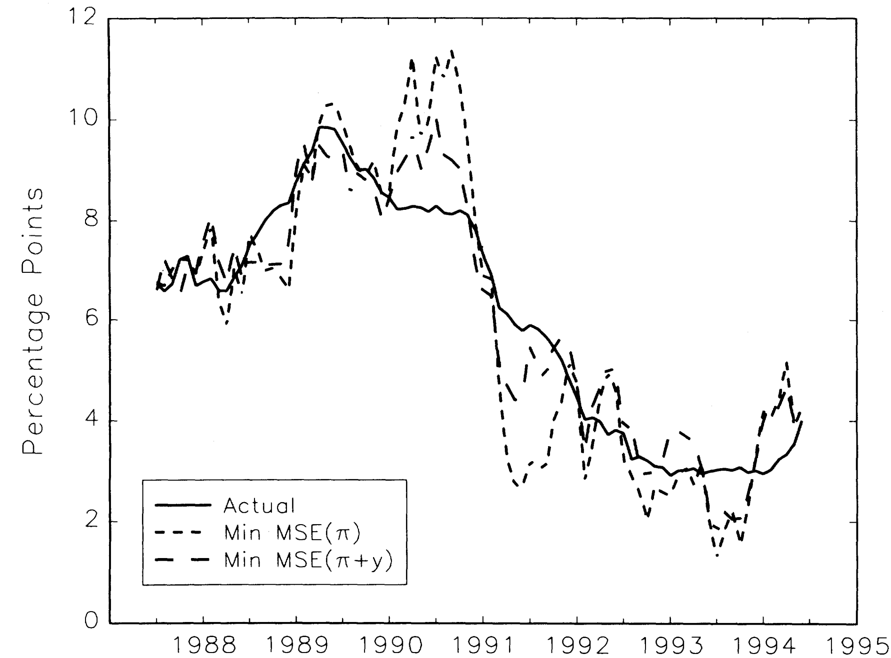
income substantially—from 3.01 to 1.89 when the output shock is the source of the instability. In other words, the inability to precisely estimate either the impact of shocks on prices or the response of prices to policy innovations provides a fairly strong argument for including real variables in the objective function.

Finally, one might ask how closely recent policy conforms to what would have been implied by either the price-level or the nominal-income targeting rules plotted in Figure 5. A simulated interest-rate path can be calculated by taking the estimated structural innovations, the $\hat{\epsilon}_{jt}$'s, to compute the optimal policy responses implied by equation (9), and then substituting the result into the equation for the federal funds rate, which is the equivalent of equation (7).[35]

Figure 6 compares the actual path of the federal funds rate with that implied by the estimated price-level and nominal-income targeting policies. Several findings emerge from examination of the figure. First, targeting the price level alone yields larger swings, as the funds rate reaches both higher and lower extremes. In addition, the actual funds rate is the least variable, looking like a smoothed version of the two

35. 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 a one unit shock in one $\epsilon_{jk}$ at a time will be robust to sequences of shocks in all the $\epsilon_{jk}$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.

Figure 6 COMPARISON OF OPTIMAL AND ACTUAL FEDERAL-FUNDS-RATE PATHS MONTHLY, 1987:06 TO 1994:07




Note: Optimal federal-funds-rate paths are computed using the policy rules in Figure 5, together with the realized output, commodity price, and aggregate price shocks from the VAR—the $\hat{\epsilon}_t$'s.

simulated paths. But the general character of the plot suggests that the optimal policy response simply involves faster and bigger movements than those exhibited by the actual path.[36]

## 7. *Conclusion*

Empirical analysis of the inflation process leads to a number of conclusions relevant for current policy formulation. First, inflation is extremely difficult to forecast. Even at horizons as short as 3 months, and for a sample including only the last 10 years, the root-mean-square error of inflation forecasts is always above 1 percentage point (at an annual rate).

Second, the relationship between potential inflation indicators and inflation is suspect as a basis for policy. This is true for two reasons.

36. As one would expect, these large policy innovations result in less stable real output, highlighting that the ultimate issue in policy making is still the relative weight on prices and output in the objective function.

First, whether an indicator fits well in sample says virtually nothing about its performance out of sample. Even more troubling is the fact that the relationship between inflation and inflation indicators is extremely unstable, shifting systematically when monetary policy operating procedures change.

The third major conclusion of the paper is that the relationship between monetary policy instruments, such as the federal funds rate, and inflation also varies substantially over time and cannot be estimated precisely. The point estimate of the response of prices to innovations in the federal funds rate changes sign over time, and is rarely significantly different from zero.

But policy makers must make decisions on a day-to-day basis, and so stopping here is not terribly useful. Instead, one can take these lessons to heart, and construct policy rules that incorporate the imprecision inherent in inflation forecasts and policy responses. This leads to advice of the following sort. Since prices take time to respond to all types of impulses, the federal funds rate should be raised immediately following a shock. One should not wait for prices to rise before acting. Furthermore, comparison of the results of price-level targeting with nominal-income targeting suggest that the difficulties inherent in forecasting and controlling the former provide an argument for focusing on the latter.

## *Data Appendix*

This appendix lists all of the data used in the paper. When the data are from CITIBASE, the mnemonics are in parentheses at the end of the descriptions. All data are available seasonally adjusted at least from 1967:01 to the present. The one exception is the average-hourly-earnings index, which is not seasonally adjusted, and was adjusted by using deterministic seasonal dummy variables.

1. Prices: CPI-X1 from 1967:01 to 1982:12, spliced to the All Items CPI-U (PUNEW).
2. Industrial production: Total index (IP).
3. Average hourly earnings: LEMXO average hourly earnings of production workers, excluding overtime (LEMXO).
4. *Journal of Commerce* spot price index: Industrial material price index, 18 commodities (FCJM).
5. NAPM spot index: National Association of Purchasing Managers spot price index, all commodities (PSCCOM).
6. Capacity utilization: Manufacturing, total (IPXMCA).
7. Unemployment rate: All workers, 16 years and over (LHUR).

8. Monetary base: Adjusted for reserve-requirement changes (FRB St. Louis) (FMBASE).
9. M1: Money stock: currency, travelers checks, demand deposits and other checkable deposits (FM1).
10. M2: Money stock: M1 plus overnight repurchase agreements, Eurodollars, etc. (FM2).
11. Federal funds rate (FYFF).
12. 10-year bond rate: U.S. Treasury constant maturities, 10-yr (FYGT10).
13. Weekly hours index: Total, private nonagricultural (LWHX).
14. Trade-weighted value of the dollar: Webb and Rowe (1994).
15. NAPM diffusion index: Niemira and Klein (1994).
16. Price of gold, London fix: Webb and Rowe (1994).
17. Price of oil, Brent North Sea crude: Webb and Rowe (1994).
18. Commercial-paper rate: Interest rate, commercial paper, 6-month (FYCP).
19. Nominal GDP: Gross domestic product, current dollars (GDP).
20. Nonborrowed reserves: Depository institutions' reserves, nonborrowed plus extended credit, adjusted for reserve-requirement changes (FMRNBC).

## REFERENCES


Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* 61:821–856.

———, and W. Ploberger. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62:1383–1414.

Ball, L., and S. G. Cecchetti (1990). Inflation and uncertainty at short and long horizons. *Brookings Papers on Economic Activity* 1990(1):215–245.

———, and ———. (1988). Imperfect information and staggered price setting. *American Economic Review* 78:999–1018.

Brainard, W. C. (1967). Uncertainty and the effectiveness of policy. *American Economic Review* 57: 411–425.

Bryan, M. F., and S. G. Cecchetti. (1994). Measuring core inflation. In *Monetary policy*, N. G. Mankiw (ed.). Chicago: University of Chicago Press for NBER, pp. 195–215.

Caskey, J. (1985). Modeling the formation of price expectations: A Bayesian approach. *American Economic Review* 75:768–777.

Christiano, L. J., M. Eichenbaum, and C. L. Evans. (1994a). The effects of monetary policy shocks: Evidence from the flow of funds, Federal Reserve Bank of Chicago Working Paper 94–2 (March).

———, ———, and ———. (1994b). Identification and the effects of monetary policy shocks, Federal Reserve Bank of Chicago Working Paper 94–2 (May).

Cochrane, J. H. (1994) Shocks. In *Carnegie–Rochester Conference Series on Public Policy* 41:275–364.

Duca, J. V. (1992). The case of the missing M2. *Federal Reserve Bank of Dallas Economic Review* (2nd Quarter):1–24.

Evans, M. D. D., and P. Wachtel. (1993). Inflation regimes and the sources of inflation uncertainity. *Journal of Money, Credit and Banking* 25 (August, Part 2).

Feldstein, M., and J. H. Stock. (1994). The use of monetary aggregates to target nominal GDP. In *Monetary Policy,* N. Gregory Mankiw (ed.). Chicago: University of Chicago Press for NBER, pp.7–70.

Fisher, M. E., and J. Seater. (1993). Long-run neutrality and superneutrality in an ARIMA framework. *American Economic Review* 83:402–415.

Galí, J. (1992). How well does the IS–LM model fit postwar U.S. data? *Quarterly Journal of Economics* 107:709–738.

Hallman, J. J., R. D. Porter, and D. H. Small. (1991). Is the price level tied to the M2 monetary aggregate in the long run? *American Economic Review* 81(4): 841–858.

Hansen, B. E. (1995). Approximate asymptotic *p*-values for structural change tests. Department of Economics, Boston College. Mimeo.

Kareken, J. H., and R. M. Solow. (1963). Lags in monetary policy. In *Stabilization Policies.* Commission on Money and Credit, pp.14–96.

Lucas, R. E., Jr. (1976). Econometric policy evaluation: A critique. In *Carnegie–Rochester Conference Series on Public Policy,* No. 2, pp. 19–46.

———. (1988). Money demand in the United States: A quantitative review. In *Carnegie–Rochester Conference Series on Public Policy,* No. 29, pp. 137–167.

Miyao, R. (1994). Essays on money and output. Department of Economics, Harvard University. Unpublished Ph.D. dissertation.

Newey, W. K., and K. D. West. (1987). A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–708.

Niemira, M. P., and P. A. Klein. (1994). *Forecasting financial and economic cycles.* New York: John Wiley and Sons.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55:324–330.

Sims, C. A. (1972). Money, income, and causality. *American Economic Review* 62:540–552.

———. (1980). Macroeconomics and reality. *Econometrica* 48:1–48.

———. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review* 36:975–1000.

Stock, J. H., and M. W. Watson. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61:783–820.

———, and ———. (1994). Evidence on structural instability in macroeconomic time series relations. Kennedy School of Government, Harvard University. Mimeo.

Strongin, S. (1992). The identification of monetary policy disturbances: Explaining the liquidity puzzle. Federal Reserve Bank of Chicago Working Paper 92–27 (November).

Webb, R. H., and T. S. Rowe. (1994). An index of leading indicators for inflation. Research Department, Federal Reserve Bank of Richmond. Mimeo.

Woodford, M. (1994). Nonstandard indicators for monetary policy: Can their usefulness be judged from forecasting regressions? In *Monetary Policy,* N. Gregory Mankiw (ed.). Chicago: University of Chicago Press for NBER, pp. 95–115.

## *Comment*

MARK W. WATSON
Northwestern University and Federal Reserve Bank of Chicago

In this paper, Stephen Cecchetti carefully outlines and evaluates the challenges of designing short- and medium-run monetary policy for controlling inflation. A convenient framework for discussing the set of issues that Cecchetti addresses is the control problem

$$\min_{m_t} E_t \sum_{i=0}^{\infty} \beta^i \pi_{t+i}^2, \tag{1}$$

where $\pi_t$ is the rate of price inflation (or deviation of inflation from some target value), $m_t$ is a monetary control variable (like the federal funds rate, the monetary base, or the quantity of nonborrowed reserves), and $\beta$ is a discount rate. The first-order conditions for the problem are

$$E_t \sum_{i=0}^{\infty} \beta^i \pi_{t+i} \frac{\partial \pi_{t+i}}{\partial m_t} = 0.$$

Evidently, to implement the optimal policy a policymaker needs to construct $E_t\pi_{t+i}$ and $E_t\ \partial\pi_{t+i}/\partial m_t$ in addition to the covariance between these two variables. Cecchetti's paper focuses on three aspects of this control problem: (1) accuracy of the inflation forcasts $E_t\pi_{t+i}$, (2) accuracy of estimates of the dynamic multipliers $E_t\ \partial\pi_{t+i}/\partial m_t$, and (3) robustness of the results to the choice of the objective function. The paper reaches four important conclusions:

1. *Forecasts of inflation are inaccurate.*
2. *Inflation forecasting rules and the dynamic multipliers $\partial\pi_{t+i}/\partial m_t$ are temporally unstable.*
3. *Even if dynamic multipliers are assumed stable, they are imprecisely estimated.*
4. *An objective function targeting nominal income is more robust than an objective function targeting inflation.*

In my comments I discuss each of these conclusions in turn.

### *1. Accuracy of Inflation Forecasts*

Cecchetti's paper reports large root-mean-square errors (RMSEs) for inflation forecasts. It shows that these large RMSEs obtain for commercial forecasts, naive time-series forecasts, and bivariate regression forecasts constructed using a large set of inflation leading indicators. The large RMSEs obtain for forecasts of different horizons and different sample periods.

The first row of my Table 1 provides results that are quite consistent with Cecchetti's systematic analysis. Letting $P_t$ denote the date $t$-value of the monthly consumer price index, this table shows the RMSE of annual inflation forecasts, $\pi_{t+l,t+l+12} = E_t \log (P_{t+l+12}/P_{t+l})$, for $l = 0$ and $l = 12$. The forecasts were constructed from a regression model using lags of inflation as well as building permits and the commercial-paper–Treasury-bill spread, two leading indicators that Stock and Watson (1995) found to be useful predictors for a number of macroeconomic time series. The forecast period was 1979:1–1993:12, and forecasts were constructed from models estimated recursively using data from 1959:1 through date $t$. The RMSEs are large over this sample period, roughly 2 percentage points at an annual rate, consistent with Cecchetti's findings.

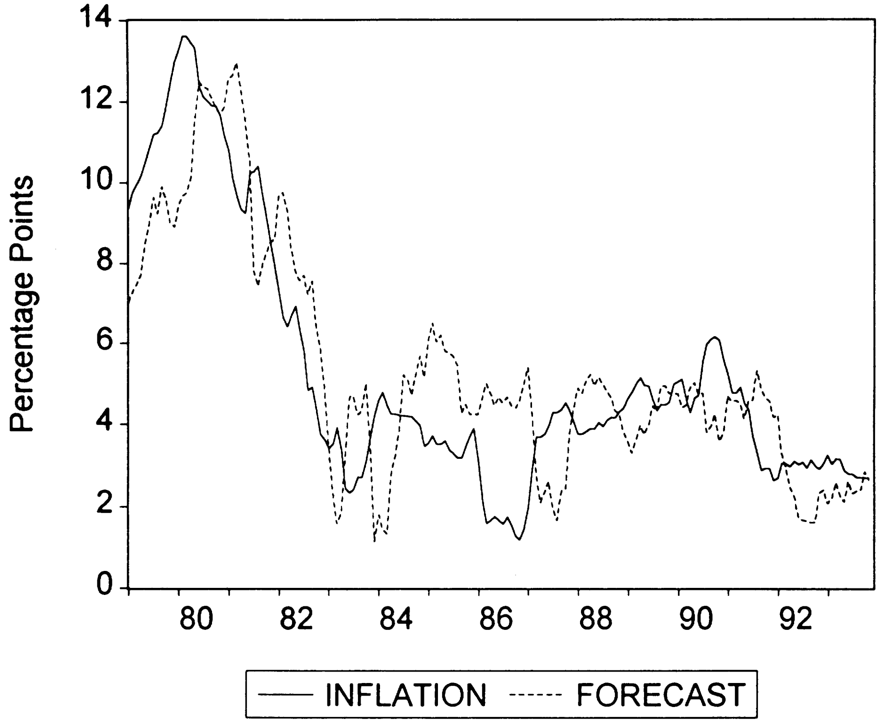
My Figure 1 plots the 1-year-ahead inflation forecasts together with the realized values of inflation. The forecasts are aligned with the actual data so that the vertical distance between the two series is the forecast error. Two things are evident from this figure. First, large absolute errors are apparent. For example, forecasts for 1986 called for inflation to exceed 4 percentage points, while the realized value of inflation was under 2 percentage points. Second, while there are large absolute errors, the forecasts do track the general trend in inflation. For example, they show inflation in the mid-1980s moderating from the very high level in the early 1980s, together with the further decline in trend inflation in the early 1990s. These results suggest that trend inflation forecasts paint an accurate picture of inflation for monetary policy actions.

Table 1 FORECAST ROOT MEAN SQUARE ERRORS, 1979:1–1993:12, IN FORECASTS OF log $(X_{t+l+2}/X_{t+l})$

| | RMSE | |
|---|---|---|
| *Variable* | $l = 0$ | $l = 12$ |
| Consumer price index | 1.82 | 2.30 |
| Industrial production | 4.07 | 4.42 |
| Real personal income | 2.51 | 2.83 |
| Nominal personal income | 2.35 | 2.61 |

*Notes:* These results show the RMSE from regression models computed recursively from 1959:1 through date t. Each regression included lags of the variable being forecast, a constant, and lags of building permits and the commercial-paper–Treasury-bill spread.

Figure 1 ONE-YEAR-AHEAD INFLATION FORECASTS




While not directly related to the focus of Cecchetti's paper, it is also important to point out that short-run inflation forecasts are more accurate than forecasts of other important macroeconomic aggregates. The final three rows of Table 1 show forecast RMSEs for two real output measures, industrial production and real personal income, together with nominal personal income. These forecasts were constructed over the same sample period using the same leading indicators as the inflation forecast and using the same recursive procedure. Forecasts for these measures of real and nominal activity are markedly less accurate than inflation forecasts. Forecasting macroeconomic aggregates is difficult, and inflation is no exception.

## 2. *Temporal Stability of Forecasting Rules*

Cecchetti reports dramatic evidence of temporal instability in forecasting rules. Using tests for time-varying regression coefficients developed in Andrews and Ploberger (1994), Cecchetti investigates stability in the

Table 2 TESTS FOR TIME-VARYING COEFFICIENTS IN THE INFLATION PROCESS

| | Test Statistics | | |
|---|---|---|---|
| | L | QLR | EW |
| *A. Univariate Tests* | | | |
| 10% critical value | 1.7 | 19.0 | 6.8 |
| Test statistic | 1.3 | 24.3 | 8.1 |
| *B. Bivariate Tests* | | | |
| Fraction of rejections | 16 | 81 | 61 |

*Notes:* These results are taken from Stock and Watson (1995). Panel A shows results for a univariate autoregression including a constant term and lags of inflation. The test was computed using data from 1959:1 to 1993:12. *L* denotes the Nyblom (1989) test, QLR is the Wald version of Quandt's (1960) test, and EW is the exponential test from Andrews and Ploberger (1994). Critical values were determined by simulation. Panel B shows summary results from 76 bivariate autoregressions used to forecast inflation. Variables included in these regressions are listed in Stock and Watson (1995).

univariate autoregression explaining inflation and in 17 bivariate forecasting models constructed from leading indicators of inflation. The null hypothesis of stability is rejected in all of the models.

Results shown in my Table 2 support these conclusions. This table presents results for a variety of time-varying coefficient tests calculated from six-lag univariate and bivariate autoregressions. The list of 17 leading indicators in Cecchetti's paper is expanded to the 76 indicators studied in Stock and Watson (1995), and three different tests for time-varying parameters are considered: Nyblom's (1989) *L*-test, the Wald version of the Quandt (1960) likelihood ratio (QLR) test for a single break at an unknown time, and the Andrews–Ploberger (1994) exponential Wald (EW) test considered by Cecchetti.[1] Panel A shows results for the univariate autoregression; the null hypothesis of stability is rejected using the QLR and EW tests, but is not rejected using the *L*-test. Panel B shows summary results for the 76 bivariate autoregressions. Stability is rejected in a large fraction of the bivariate models using the EW and QLR tests, but far fewer using the *L*-test. This difference in the rejection rates for the different tests says something about the form of instability in the models. The *L*-test is an optimal test for random-walk coefficient variation. The QLR and EW tests are formed from sequences of standard Chow tests for one-time breaks in the regressions. The results suggest sharp breaks in the forecasting models, for which the Chow tests have

1. Heteroscedastic versions of the statistics are used. The critical values for the test statistics were constructed using a simulation experiment described in Stock and Watson (1995). The asymptotic critical values for heteroscedastic robust versions of the QLR and EW statistics are considerably different from the critical values determined by simulation.

high power, rather than smoothly varying coefficients, for which the $L$-test has high power.

Another way to investigate time variation is to ask whether models with time-varying parameters produce more accurate forecasts than those with fixed parameters. Stock and Watson (1995) compare the forecasting performance of fixed-parameter models and a set of time-varying-parameter models for a large number of macroeconomic variables. Interestingly, in each of the 76 forecasting models that they construct for inflation, time-varying-parameter models produced more accurate forecasts than fixed-parameter models.

In summary then, both time-varying-parameter tests and forecasting experience suggest important instability in the inflation process.

## 3. *Accuracy of Estimates of* $\partial\pi_{t+i}/\partial m_t$

Cecchetti's paper shows wide confidence intervals for dynamic multipliers (impulse response functions) calculated from vector autoregressions. This implies large uncertainty in the sequence of values for $E_t\, \partial\pi_{t+i}/\partial m_t$. These results are consistent with the large body of empirical research in VARs that has accumulated over the past fifteen years. This literature suggests that these pessimistic results are robust to a wide range of changes in the specification of the VAR. Indeed, Cecchetti's results probably overstate the precision of the estimates, since they abstract from uncertainty in the identification of monetary shock. Evidently, the data do not tell us much about the values of these important multipliers.

## 4. *Robustness of Results to Choice of Objective Function*

One of the paper's most interesting conclusions is the robustness of the nominal-income targeting control rule and the nonrobustness of the inflation control rule. Specifically, the paper considers two distinct problems defining optimal monetary policy:

$$\min_{m_t} E_t \sum_{i=0}^{h} p_{t+i}^2 \tag{2}$$

and

$$\min_{m_t} E_t \sum_{i=0}^{h} (p_{t+i} + y_{t+i})^2 , \tag{3}$$

where $p_t$ and $y_t$ denote the logarithms of the price level and output relative to target values. Equation (2) is the inflation–price-level objective function that motivates the other analysis in Cecchetti's paper; equation (3) is an objective function that targets nominal income. Let $\{m_{2,t}\}$ denote the sequence of solutions to (2), and $\{m_{3,t}\}$ denote the sequence of solutions to (3). Cecchetti's calculations show that $\{m_{3,t}\}$ approximately solves (2), but that $\{m_{2,t}\}$ produces a value of the objective function in (3) that is far from the minimum. In this sense, the nominal-income targeting rule is robust (since it almost minimizes the inflation objective function as well), but the inflation-targeting rule is not robust (since it does a very poor job of minimizing the nominal-income objective function). Since there is no widespread agreement on the appropriate objective function for monetary policy, these robustness considerations suggest that the nominal-income rule dominates the inflation rule. That is, people who favor an inflation target lose little if a nominal income target is used, while people who favor a nominal income target (and thus put some weight on real activity) lose a lot if an inflation target is used.

Since this is arguably the most important result in Cecchetti's paper, I want to expand on some of the caveats that Cecchetti lists about his calculation. Cecchetti's calculation differs from the usual textbook control problem in two ways. First, he recognizes that optimal control must be carried out using estimated values of the parameters of the model, and allows uncertainty in these parameters to affect the control rule and resulting value of the objective function. Relaxing the assumption that the parameter values are known is useful, but it complicates the analysis and interpretation of the results for reasons that I discuss below. The second modification that Cecchetti introduces is a control rule specified in terms of the shocks in the model [see Cecchetti's equation (9)], rather than the directly observed variables ($y_t$, $p_t$, etc.). Such a change is inconsequential if the VAR parameters are known, because the VAR can be used to construct the observed variables from the shocks and vice versa. However, this is not the case when the VAR parameters are unknown; the exact values of the shocks cannot be recovered from the estimated VAR. Indeed, this introduces a logical tension in the analysis: if the monetary authority can observe the shocks for the purposes of control, it could also use these shocks together with the observed data to eliminate any estimation error in the parameters. Thus, the two modifications introduced by Cecchetti are in some sense mutually inconsistent.

A more important concern is the robustness of the calculations underlying Cecchetti's conclusions about the relative merits of the nominal-income and inflation-targeting rule. When the calculation is carried out using estimated parameters, the optimal control rule and the resulting value of the objective function can depend critically on the assumed realizations of the shocks. To see this, notice that in the control rule [Cecchetti's equation (9)] the shocks multiply the control coefficients $\alpha_{jk}$. Since the $\alpha_{jk}$'s are estimates, different linear combinations of these estimates (associated with different assumed values of the $\epsilon$'s) will be subject to different amounts of estimation error. Thus, as Cecchetti warns (see his footnote 35), his conclusions may not be robust to changes in the assumed values of the shocks $\epsilon$.

For computational reasons, Cecchetti's experiments involve a particularly simple realization of the shocks: $\epsilon_{it} = 1$ for one $i$ and one $t$, and all other realizations of the shocks are zero. While computationally convenient, this is much different from the historical realizations of the shocks or, more importantly, the realizations that will occur in the future. A more useful calculation would average over realizations drawn from the estimated probability distribution of the shocks. This would provide an estimate of the value of the objective that we could expect to obtain for the U.S. economy, and allow evaluation of policy rules for this economy. Cecchetti's calculations provide a first step in this process, and the power of his tentative conclusion provides a strong motivation for a more detailed analysis of the question.

## 5. *Summary*

Stephen Cecchetti presents a careful empirical analysis of important practical problems associated with controlling inflation. His paper presents several challenges for future researchers. First, RMSEs for short-run inflation forecasts are large, suggesting considerable room for improvement. Second, the inflation process (or at least the linear models of inflation considered here) are unstable, suggesting a need for time-varying or adaptive models for improving inflation forecasts. Third, this instability also suggests instability in the VARs that are often used to evaluate theoretical macroeconomic models, suggesting a need to modify these procedures. Finally, Cecchetti raises the possibility that some control rules may be more robust than others, in the sense that they are nearly optimal for many different objective functions, suggesting a useful criterion for choosing control rules.

REFERENCES

Andrews, D. W. K., and W. Ploberger. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62:1383–1414.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association* 84:223–230.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55:324–330.
Stock, J. H., and M. W. Watson. (1995). Evidence on structural instability in macroeconomic time series relations. Department of Economics, Northwestern University. Manuscript.

# Comment

DONALD L. KOHN[1]
Federal Reserve System

A little over a year ago the Federal Reserve began raising short-term interest rates. At the time, inflation was relatively low and not accelerating, and the unemployment rate was still above most estimates of the NAIRU. These actions were based on a projection that if the federal funds rate were held at its existing level, inflation would pick up. Given the lags in the effects of policy, acting in advance of rising inflation was seen as necessary to minimize variation in prices and output around desirable paths.

Anticipatory monetary policy was not universally popular or understood. A number of people asked how the Fed could be confident that inflation would have accelerated in the absence of tightening. Moreover, this was not the first time policy actions had been taken in advance of economic developments. For example, short-term rates were raised in 1984 while unemployment was still elevated, helping to prolong the expansion. More recently, the Federal Reserve eased policy in 1989 when economic activity softened even as inflation was relatively high and rising. More generally, Steve McNees's (1992a) work on reaction functions suggests that the Federal Reserve regularly uses staff forecasts in addition to recent data to adjust its policy stance. A recent study by Hall and Mankiw (1994) recommends that the Federal Reserve rely on private-sector forecasts in formulating monetary policy.

In this context, Cecchetti's paper asks exactly the right questions: How good are the inflation forecasts available to the Federal Reserve—from professional forecasters such as DRI or the Blue Chip contributors, or from statistical forecasting exercises using a few indicator variables in addition to the behavior of prices themselves? And, for a given deviation

1. Director, Division of Monetary Affairs, Board of Governors of the Federal Reserve System. The views expressed are those of the author and do not necessarily represent the views of the Board or its other staff. Thanks to Steve McNees and Flint Brayton for useful discussions and material.

of inflation from its desired level, what do the data tell us about how to change the federal-funds-rate operating target? He also uses a structural VAR model to look at responses to policy actions and to test alternative reaction functions. However, my comments will concentrate on his analysis relating to the two key questions.

Steve's answers to those questions are pretty discouraging. Forecast accuracy is poor, and relationships among variables seem to shift frequently over time, including the all-important relationship between inflation and the federal funds rate.

My comments start with a disclaimer: they are not based on a technical evaluation of the statistical tests and techniques used in the paper, but rather on the underlying nature of the tests and the conclusions drawn from them. In brief, I believe Steve's pessimism is not entirely justified, but derives in part from the "indicator" approach he follows and, in certain sections of his paper, from his concentration on the results of short-term forecasts. Nonetheless, Steve's work highlights important caveats for the policymaker using inflation forecasts. Their accuracy is not very good, and they can be difficult to interpret; consequently, policymakers need to be quite cautious in how much weight they put on them.

I am neither surprised nor convinced by the problems Steve encounters in the exercise in which he predicts future inflation by adding one indicator to past inflation. These tests suffer from a lack of consideration of structural issues. Indicator exercises by their nature submerge the fundamental behavioral relationships among economic and financial variables. In addition, even some of the indicator variables themselves seem not to have been specified with close regard for a sense of the underlying structure of the economy or theory that might link the indicator to inflation. For these reasons, the lack of predictive power of the equations and their many so-called structural shifts are difficult to evaluate, and the results could well be understating economists' abilities to predict and interpret key economic variables.

One example is the unemployment rate. A proper test of the unemployment rate and inflation would focus not on the level of the rate but on its deviation from the NAIRU. This is particularly a problem for subperiods when the realized unemployment rate may deviate substantially and persistently from its natural value, or when the NAIRU is shifting owing to changing demographic trends or other structural influences—as it did in the 1970s. The relationship of the unemployment gap to increases and decreases in inflation has been tested extensively by a number of different researchers over time, and generally has been found to be reasonably stable. Steve, in contrast, finds four shifts in twenty years in the relationship of inflation to unemployment, raising questions in my mind about the specification of the variable and the nature of the test.

M2 is another example. Unlike Steve, Hallman, Porter, and Small (1991) found a stable relationship between M2 and inflation (at least into the early 1990s), but in an accelerationist context that took account of the deviations of output and velocity from long-term equilibriums. These are simply examples of how some indicators might have performed better with different specifications.

For some of the indicator variables he looks at, one would not expect a stable bivariate relationship with inflation, even in a fully stable world. One such relationship involves the exchange rate. This price reflects a whole host of factors other than actual or prospective U.S. inflation, including macroeconomic policies and inflation abroad.

Another, more critical example of a relationship that should not be expected to be stable is that between inflation and the federal funds rate. Steve's downbeat conclusions about the ability of monetary policy to achieve its objectives derive importantly from the unstable relationship he finds between these two variables. But we know a nominal interest rate does not provide a nominal anchor. The rate of inflation associated with a particular nominal interest rate depends both on inflation expectations and on the level of the implied real rate relative to the value of that real rate consistent with production at its long-run potential. The latter can change in response to fiscal policy and other factors affecting spending, but even some of these changes embody predictable responses to evolving developments, rather than structural shifts. If a given nominal rate implies a real rate that pushes output beyond its potential, inflation will rise; the same nominal funds rate could imply falling inflation if inflation expectations were considerably lower. Similarly, a given change in the nominal federal funds rate will be associated with varying changes in inflation rates depending on the evolution of inflation expectations and equilibrium real rates.

In this context, Steve's findings of a shifting relationship between the funds rate and inflation is hardly surprising, and no bar to a sensible, forward-looking monetary policy. Policy makers understand that interactions between the funds rate and inflation are complex, and they take inflation expectations and possible changes in the relationship of real rates to aggregate demand into account in their decisions. To be sure, this is no simple task, and it is prone to error. Inflation expectations are hard to measure, and shifts in spending propensities that would affect equilibrium real rates are frequently difficult to forecast or even to detect as they are occurring. Nonetheless, simply by being aware of the underlying concepts and the pitfalls in implementing them, policy makers may be able to avoid repeated or cumulating policy errors.

Steve finds structural shifts in all his equations, and he tends to ascribe them to a changing Federal Reserve reaction function. Several of the dates he identifies, however, do not coincide, in my reading of Federal Reserve history, with changes in operating procedures or policy objectives. Instead, I suspect that many of the shifts are a consequence of changes in inflation expectations, of readily identified alterations in the structure of the economy and financial markets (such as the breakdown of Bretton Woods or the demise of Regulation Q), or simply of the particular specification. The point is that from a structural perspective the world is not as unstable as it would appear from the exercises—and that some sources of instabilities can be identified and taken into account as or even before they occur.

With some knowledge of the forces at work, forecasts can give policy makers a rough idea of the general outlook for inflation. Steve emphasizes the poor forecasting capability of forecasts for one quarter ahead. But as he also notes, quarter-to-quarter noise does cancel out to an extent, and four- or eight-quarter forecast errors are considerably smaller. These are the time periods relevant to policy making, since it is only broad trends in prices that are subject to influence from policy actions. At these longer horizons, forecast errors for some of Steve's formulations are on the order of 1 percentage point or less since the early 1980s. Interestingly, among the smallest errors are those for forecasts using the unemployment rate—an approximation of the Phillips curve.

These results are consistent with experience in "real time" forecasting. For example, the inflation forecasts of the members of the FOMC, which have been made since 1980 and published in our semiannual monetary policy reports, have RMSEs of 1 and 1¼ percentage points for one-year forecasts made at the beginning of the year and middle of the previous year, respectively.[2] Blue Chip forecasts are slightly less accurate, perhaps because the FOMC has "inside information" on its objectives and how it would respond to deviations from expectations. The RMSEs of Board staff forecasts for 1980 to 1989 (the former date is the first available currently in easily retrievable form; the latter date is the last publicly available) are somewhat smaller for the next four quarters (around 0.6); errors of staff forecasts for the next eight quarters (relative to average inflation at an annual rate) are the same as the one-year errors or some-

2. The data are courtesy of the Federal Reserve Bank of Boston and will be published in the Bank's *New England Economic Review* for July–August 1995.

what larger, depending on whether they are measured against current BEA estimates (that is, after benchmarking, etc.) or against data released closer to the time of the forecast.

Whatever the problems in predicting inflation, we know that today's policy actions will have their effect over a considerable period into the future. From the policymaker's perspective, the issue becomes whether it is worthwhile to use predictions of future conditions to gauge the need for policy actions and the consequences of alternative policy choices, and if so, how to make such predictions and how much weight to put on them. The alternative is to run policy simply by looking at past inflation, or at inflation and real output as in John Taylor's rule or nominal-income targeting, or at one or a few data series believed to be reliable intermediate targets or indicators. Even without an explicit forecasting exercise, however, policy makers need some sense of likely responses to policy actions in order to judge how much to change their instrument or proximate target variables when intermediate or final targets deviate from desired paths. Thus, policy makers cannot avoid looking into the future, at least implicitly.

Most policy makers prefer to look at a broad array of data and at explicit forecasts made in the context of a structural model—econometric or judgmental. This is what they get from the staff and from most outside forecasters. Policy makers are looking for guidance as to the broad trends in inflation and economic activity and to the major risks to the outlook. Projections provide benchmarks against which to judge incoming data to assess whether the economic and financial situation is evolving differently than they had expected.

In this regard, the structural, rather than reduced-form, nature of the forecasting exercise is important. Policy makers recognize the tendency of models with a few variables to break down or shift over time, as Steve has pointed out. Their experience setting targets for the monetary aggregates has only reinforced their concerns in this regard. "Telling stories" about why certain outcomes are expected in a structural sense is valuable because it enables policy makers to make their own judgments about the channels of policy influence and to feel more comfortable about the rationale for revising forecasts and instrument settings. This is reinforced by the need to explain policies to the public—an essential part of the accountability and credibility of an independent central bank. Convincing rationales for policy actions require explanations of how the actions are expected to affect national economic goals, and the likely outcome of failing to take the action. Black boxes do not provide these services.

As I noted, Steve's work, however, highlights a number of caveats about the use of forecasts—internal or external, reduced-form or structural—in policy.

For one, he emphasizes the key role of Federal Reserve reaction functions in economic forecasts. Private forecasters (and perhaps, to a lesser extent, the Federal Reserve's staff) need to make assumptions about the intermediate-term objectives of the Federal Reserve and how it might respond to various configurations of actual and prospective economic developments. One might ask whether the Federal Reserve couldn't be more explicit and precise about its objectives and likely reactions. My perception is that that would be quite difficult. Policy actions need to take account of a broad array of the characteristics of the current and prospective economic situation. In effect the policy reaction function is too complex to spell out adequately in advance as a mechanical, predetermined response to a handful of variables.

Nonetheless, when the Federal Reserve is looking at an economic forecast, unavoidably it is in part looking in a mirror—that is, at an estimate of its own reaction function. It needs to assess and separate carefully the information of most value to it—i.e. the forecaster's assessment of the strength of aggregate demand and associated price pressures and the likely response to any changes in policy. Again, this probably is easiest in a structural context, where the policy assumptions may be more readily ascertained and backed out.

In addition, as Steve notes, even the smaller RMSEs for longer-horizon forecasts are sizable. According to conventional inflation–unemployment relationships, it takes perhaps on the order of 2 percentage-point–years of deviations of unemployment from its natural rate to change the inflation rate by 1 percentage point. So misses in inflation projections can have notable effects on the policy stance and real activity if the FOMC places considerable weight on those projections. Moreover, the largest errors are generally made around cycle turning points, just when the FOMC may be most in need of accurate forecasts to damp unwanted variations in output and prices (McNees, 1992b).

Finally, the evaluation and use of some private forecasts is sometimes complicated by difficulties in discerning the structure of the economy and the behavioral relationships they are based on. For example, from 1991 to 1993 many forecasts were calling for a rise in inflation when the unemployment rate was above most estimates of the NAIRU and projected to stay there, since these same forecasts were calling for continued moderate economic growth. I suspect that such forecasts may be based on rules of thumb (perhaps in this case inflation on average tends to pick up in expansions) that are not firmly based in economic theory. To be sure, the rationales of many private forecasts are fully explained so that their underlying structure can be evaluated, but there is a risk in particular when using averages of forecasts, such as the Blue Chip or ASA–NBER extended by the Federal Reserve Bank of Philadelphia, where structure cannot be determined.

For all these reasons, inflation forecasts, while potentially helpful to policy makers, have to be used cautiously, carefully, and gingerly. Given the size of the errors, policy makers need to be flexible in revising forecasts and the policy stance in response to new information contradicting their previous predictions.

I conclude as I began. Steve has asked exactly the right questions. I particularly applaud policy-relevant academic research—an all too rare commodity these days, and one that has potential public as well as private benefits. As I noted, I'm not as pessimistic as Steve is about his results, though I recognize the considerable risks and problems in the policy process. Consequently, I would hope Steve and others would continue to think about the difficulties of forward-looking policy making and the role forecasts might play in improving the conduct of monetary policy. For example, along these lines, future research might test the "value added" of explicit forecasts in achieving policy objectives, as opposed to reacting only to incoming data, unprocessed through a forecasting exercise. In the context of Steve's work and further efforts that help define the limits of our knowledge, one might take a further look at the principles guiding policymaking in an uncertain world. Embedding this research in a structural context would be especially helpful to policy makers.

## REFERENCES

Hall, R. E., and N. G. Mankiw. (1994). Nominal income targeting. In *Monetary Policy*, N. G. Mankiw (ed.). Chicago: University of Chicago Press for NBER, pp. 71–92.

Hallman, J. J., R. D. Porter, and D. H. Small. (1991). Is the price level tied to the M2 monetary aggregate in the long run? *American Economic Review* 81(4):841–858.

McNees, S. K. (1992a). A forward-looking monetary policy reaction function: Continuity and change. *New England Economic Review*, November–December, pp. 3–13.

———. (1992b). How large are forecast errors? *New England Economic Review*, July–August, pp. 25–42.

# *Discussion*

Cecchetti responded to Kohn's criticism concerning the absence of more explicitly structural models from his paper. He argued that he intentionally avoided building a structural model because of the general lack of agreement about which model was the most appropriate. Further, he indicated that in this work he was primarily interested in high-frequency events and the appropriate policy responses to those events, and not so much in the longer-term trends in inflation and economic activity about which structural models have the most to say. Kohn countered by saying that policy advice based on reduced-form estimation was not generally convincing to policy makers; they need some kind of "story" about what is happening in the economy to help rationalize their decisions.

Greg Mankiw said he found the paper interesting and useful but expressed puzzlement that Cecchetti introduced commercial inflation forecasts early in the paper, only to drop them later on. Mankiw noted that, in joint work, he and Bob Hall had found that the commercial forecast consensus seemed to dominate other standard indicators. In a changing world, Mankiw argued, it is sensible that commercial forecasts—which reflect the forecasters' subjective judgment—should outperform mechanical rules; this observation is confirmed by Cecchetti's finding that reduced-form prediction equations lack robustness. Following up on this comment, Kohn suggested that the Fed's forecasts from the Green Book might be added to Cecchetti's Table 1, along with the commercial forecasts, and then used later in the estimation. Cecchetti agreed that making more extensive use of commercial forecasts would have value, but noted the problem that commercial forecasts are typically published quarterly, while his work in this paper emphasized monthly data.

Jim Stock pointed out that the RMSEs of regression forecasts of inflation presented by Watson appeared to be much larger than those of contemporaneous Federal Reserve forecasts discussed by Kohn. Possible explanations were that Watson's list of indicators is deficient, or that adding more structure would help. Another possibility suggested by Stock is that Kohn's forecasts refer to more recent periods, during which inflation has been easier to forecast. Kohn noted that one difference was that the Fed forecast inflation in the GDP deflator, while Watson's and Cecchetti's forecasts are for the CPI. Cecchetti suggested that, as Fed Green Book forecasts are available from 1969 on, one could compare the Fed's forecasts with those of Watson and himself for the earlier as well as the more recent period.

Carlos Végh felt that a more explicit theoretical framework was needed; in particular, there was some inconsistency between the standard theoretical analysis, which treats money as exogenous and interest rates as endogenous, and the empirical work, which treats the interest rate as the exogenous policy instrument. He noted that there exist theoretical analyses which treat the interest rate as the policy instrument, and which have found that the short-term response of inflation to a policy change may be in the opposite direction to that in the long run. Cecchetti replied that, according to recent research, using nonborrowed reserves rather than a short-term interest rate as a proxy for monetary policy does not change the qualitative results about the responses of the economy to policy shocks.

Michael Woodford questioned the paper's assumption that the Fed is concerned about stabilizing inflation only, without concern for the real side of the economy. Accounts of the Fed's actions suggest that its decisions to tighten or loosen depend in part on where unemployment is, relative to the natural rate. Subsequent discussion suggested that the Fed's attention to unemployment might either indicate that it is concerned about unemployment per se, or, alternatively, that it was using unemployment only as an indicator variable for predicting future inflation.

Martin Feldstein noted that although the forecasts were not very precise, Cecchetti's paper showed that monetary policy could be used constructively, e.g., it could reduce the variability of the nominal GDP. Perhaps Milton Friedman was wrong when he argued that activist monetary policy must be destabilizing. Feldstein also asked how the fact that the Fed gradually raised the federal funds rate in 1994 could be reconciled with the proposition in the paper that the optimal policy is a sharp tightening followed by gradual easing. Kohn replied that two factors might account for the Fed's behavior: First, new information came in over the year which led the Fed to raise rates more than had been expected at the beginning of the process. Second, the Fed was concerned about the implications for financial markets of rapid interest-rate changes in the initial stages of a tightening that followed a prolonged period of declining and low short-term interest rates. Cecchetti took issue with the latter justification for the Fed's behavior, arguing that the speed of the rate increase affects only the identities of the winners and losers in financial markets, and that these distributional effects should not be an important policy consideration.