Has the Economy Become More Predictable? Changes in Greenbook Forecast Accuracy

PETER TULIP

# Has the Economy Become More Predictable? Changes in Greenbook Forecast Accuracy

Several researchers have recently documented large reductions in economic volatility. But a more important question may be whether the economy has become more predictable. Using forecasts from the Federal Reserve Greenbooks, I find that inflation and output have become more predictable, though the results for output are somewhat mixed. The reductions in unpredictability (if any) are significantly smaller than reductions in volatility. Associated with this, the predictable component of fluctuations in output and inflation has virtually disappeared.

*JEL* code: E37
Keywords: predictability, volatility, forecast errors, Greenbook.

THE VOLATILITY OF the U.S. economy has declined dramatically. The standard deviation of annualized changes in quarterly real seasonally adjusted gross domestic product (GDP) declined from 1.1 percentage points in 1965–84 to 0.5 percentage points in 1984–2004. The stabilization of inflation has been similar. This "Great Moderation" has been described as one of the most striking changes in the business cycle in recent decades (Stock and Watson 2003, Bernanke 2004). It is the subject of a large and growing literature, of which McConnell and Perez-Quiros (2000), Kim and Nelson (1999), and Blanchard and Simon (2001) are prominent examples.

However, what matters to most people is not *volatility* but *uncertainty*. If a change in inflation or activity is expected, then people adjust to it. That may involve a change in prices or a shift in production from one period to another, but interviews with firms (Blinder et al. 1998) and everyday observation suggest these adjustment costs are

I would like to acknowledge helpful comments from Dave Reifschneider, David Wilcox, Neil Ericsson, Charles Goodhart, Michael Kiley, Mike McCracken, seminar participants at the Federal Reserve Board, the referees, and, especially, John Roberts. The views presented in this paper are solely those of the author and do not necessarily represent those of the Federal Reserve Board or its staff.

PETER TULIP *is from the Division of Research and Statistics, Federal Reserve Board (E-mail:* ptulip@frb.gov).

Received August 7, 2006; and accepted in revised form December 29, 2008.

*Journal of Money, Credit and Banking*, Vol. 41, No. 6 (September 2009)

often small. In contrast, when people do not know the future, investments occur when they should not, goods and assets are mispriced, precautions are taken against events that do not occur, and so on.

Perhaps the clearest evidence of the importance of uncertainty relative to volatility is the lack of interest in seasonal economic variations. Seasonal variations are huge, accounting for about 85% of the quarterly variability of output (Beaulieu and Miron 1992, Table 1). But because they are predictable, almost no one pays attention to them at a macro-economic level. Even the studies of so-called "volatility" use seasonally adjusted data. They do not measure the total variation in the data, only the variation not accounted for by one specific influence. But there is no obvious reason for singling out seasonality. Just as predictable seasonal variations are appropriately removed from the data, so should other predictable influences.

If one is interested in unpredictability, one can measure it directly, as the difference between actual outcomes and what people were expecting. There are many available measures of expectations. I use the forecasts of the staff of the Federal Reserve Board of Governors, as published in a document called the Greenbook. Differences between these forecasts and actual outcomes are the Greenbook errors.

The Greenbook errors provide an interesting measure of uncertainty for several reasons. Previous researchers have found that the Greenbook forecasts have been at least as accurate as other forecasts (Romer and Romer 2000, 2008, Sims 2002, Bernanke and Boivin 2003, Reifschneider and Tulip 2007). So they can be viewed as a good summary of the available information and as the near the frontier of best performance in forecasting. Furthermore, the data on Greenbook forecasts are richer than for many private sector forecasts. The forecast horizon is longer and the data extend further back in time.[1]

One limitation of Greenbook forecasts is that they have sometimes been based on unusual assumptions, such as a constant funds rate.[2] However, the competitive forecasting performance of the Greenbook suggests that any such handicap was unimportant: it still provided at least as reliable a guide as alternatives. A stronger limitation is that the forecasts are confidential. So although they are interpretable as reflecting what people could or should have expected, they do not directly measure what private sector decisionmakers *did* expect. But then, it is not clear that any point estimate serves this purpose.

The Greenbook errors are especially interesting for analysis of monetary policy, given the document's authorship and intended readership. From a historical perspective, changes in the quality of the forecasts might help explain changes in policy performance. From a normative perspective, the accuracy of forecasts and its stability

1. For example, whereas the Greenbook forecasts for real GDP began in 1965, the Survey of Professional Forecasters began in 1968, DRI forecasts began in 1970, Blue Chip forecasts began in 1977, and the *Wall Street Journal* survey began in 1986.

2. The importance of the funds rate assumption is often exaggerated. The funds rate was only assumed to remain constant for a few forecasts in the late 1990s. Greenbook four-quarter-ahead forecasts of 3-month bills and 10-year bonds, from 1986 to 2006, have the same RMSE as those of the Survey of Professional Forecasters (Reifschneider and Tulip 2008).

help determine the extent to which monetary policy should be "forward looking." Last, if the forecast errors are stable over time then the distribution of outcomes about previous forecasts would provide a reliable guide to the distribution of possible outcomes about the current forecast. In this context, the Federal Reserve's Federal Open Market Committee (FOMC) now regularly publishes a *Summary of Economic Projections*, in which it compares its projections with historical forecast errors averaged over the last two decades, an approach partly based on research reported in this paper (FOMC 2007, Reifschneider and Tulip 2007).

Although this paper is partly motivated by these monetary policy issues, its primary focus is whether uncertainty has declined. I find that there has been a large reduction in uncertainty regarding inflation across horizons and output at short horizons. However, I do not find a reduction in uncertainty about output at longer horizons. I find that the reduction in uncertainty is significantly less than the reduction in volatility at both short and long horizons, for both inflation and output. The differing changes in uncertainty and volatility arise because the predictive power of forecasts has virtually disappeared. Whereas the Fed predicted a large share of the fluctuations in output and inflation in the 1970s and 1980s, more recent fluctuations have been surprises.

## 1. RELATED LITERATURE

The view that unpredictability is more interesting than volatility is not new. As noted above, almost all of the studies of volatility remove predictable seasonal influences from the data. Many others remove the predictions of a vector autoregression. Several papers in this literature—for example, Stock and Watson (2003, 2007)—focus on unpredictability.

However, when measures of uncertainty are presented, they are typically derived from the errors of an econometric model. That is, they examine what agents *could* have expected rather than what they actually did expect. After-the-event regression residuals are easier to compile than real-time forecast errors, they facilitate decomposition and analysis, and they provide an indication of expectations when direct measures may be missing. But otherwise, they provide an unsatisfactory measure of the uncertainty facing decisionmakers in real time.

One problem with using models to measure changes in uncertainty over time is the limited role of learning. Coefficients may be updated, but the specification and underlying theory do not change. For measuring uncertainty at a point in time, models suffer biases in both directions. Models understate real-time uncertainty because they are estimated after the event and so benefit from hindsight, particularly when they are complicated. On the other hand, they tend to overstate uncertainty because they are simple. Even the largest models incorporate much less information than the Greenbook forecast, which reflects the pooling of many variables, models, and statistical methods by a large team of economists. These biases are unlikely to exactly cancel. Indeed, previous forecast comparisons (cited above) have found the Greenbook and private sector forecasts to be more accurate than econometric models, which suggest the models have overstated uncertainty. The results below suggest that this may no longer be true. Although that may allay some concerns about bias, the instability makes comparisons over time even harder.

Many earlier papers have analyzed real-time forecast errors, but these have not focused on changes over time. Schuh (2001), Reifschneider and Tulip (2007), and Faust and Wright (2007) provide overviews and discussion. Campbell's (2007) work, circulated as the first draft of this paper was being completed, overlaps to a greater extent. Campbell observes that the forecast errors for output growth of the Survey of Professional Forecasters (SPF) approximately halve after 1984. Output volatility declined even more, and the predictive content of the SPF disappeared. Among the more important differences between this paper and Campbell's are that I consider inflation, I examine whether the reduction in uncertainty is statistically significant, I allow for the timing of changes to be uncertain, and I examine forecasts over longer horizons.

## 2. DATA

Before scheduled meetings of the FOMC, the staff of the Board of Governors publishes a detailed forecast in a document universally, though unofficially, called the Greenbook. The Greenbook plays a central role in the policy deliberations of the FOMC (see Meyer 1998, Woodford 2008, or FOMC transcripts). The FOMC ultimately reports its own forecasts, though Romer and Romer (2008) find these add little, if any, information to the Greenbook.

The Greenbook forecasts are available at the website of the Federal Reserve Bank of Philadelphia, except for those from the last 5 years, which are confidential. The first current-quarter forecasts for real gross national product (GNP) and the GNP deflator were published in November 1965. The forecast horizon has been gradually extended since then. The horizon typically rolls forward to cover a new calendar year approximately every 12 months, so the data are discontinuous at longer horizons. I use forecasts through October 2002, which have a horizon extending to 2004q4.

Previous researchers have often focused on forecasts of quarterly changes, an approach that places considerable weight on transient errors. But an error that is reversed the following quarter is less important than one that is sustained. One-quarter blips rarely cause noticeable interest rate responses. Accordingly, I also examine forecasts of cumulative changes and I examine longer horizons (up to nine quarters) than is customary. Cumulative errors can be much larger, and hence are more closely related to recessions, deflations and other major concerns of policy. The disadvantage of longer-horizon forecasts is that they are available for a shorter period of time and their overlap creates more serial correlation. So statistical inference is harder, though as shown below, still possible.

I use one Greenbook per quarter (that closest to the middle of the quarter), although the actual frequency of publication is higher. I assume that the potential loss of information is outweighed by the convenience of measuring forecasts and outcomes at the same frequency. I focus on the forecasts for real output, defined as GNP prior to 1991, then GDP, and the forecast for inflation, defined as the deflator for these series. The GDP deflator does not now receive much attention in policy circles, but it is the price measure available over the longest time span. I assume the forecast commences with the current quarter, so a "four-quarter-forecast," for example, refers to the forecast for the current quarter and three subsequent quarters.

To calculate forecast errors, I compare these predictions with real-time data. Specifically, I use estimates from about 20 weeks after the relevant quarter, published in the Federal Reserve Bank of Philadelphia's real-time data set. Hence, "truth" for, say, the change in the four quarters to 2000q1 is the estimate as of mid-August 2000. Typically, these estimates represent the "first final" estimate (also called the "second revision") of the Bureau of Economic Analysis. These data reflect a more comprehensive analysis of source data than earlier estimates while usually adhering to the same data definitions as at the time of the forecast.

In contrast, other researchers (e.g., Sims 2002, p. 7, and a referee) prefer to use latest available estimates. Using recent estimates usually incorporates more information and it is easier. However, it involves treating changes in data definitions as forecast errors. There are at least two important problems with this approach, for my purposes. First, use of recent data would bias results toward showing that predictability has increased over time because recent forecasts would use data definitions that were closer to the "truth" than earlier forecasts. Second, using later data definitions would make forecast errors correlated, lowering the information content of individual errors. Other reasons for preferring real-time to current data are noted in Romer and Romer (2000), Robertson and Tallman (1998), and several references cited by Schuh (2001, n.14).

Figures 1 and 2 show some illustrative data. The line in Figure 1 is the four-quarter percentage change in output. The line in Figure 2 is the four-quarter percentage change in the GDP/GNP deflator, less the nonoverlapping preceding four-quarter change. Both series use real-time data. The dots show corresponding forecasts from the beginning of the four-quarter period. Throughout the paper I date errors by time

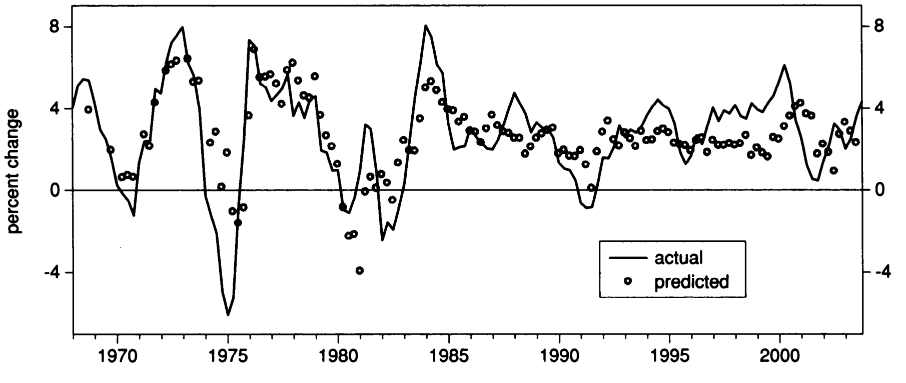


FIG. 1. Four-Quarter Output Growth.

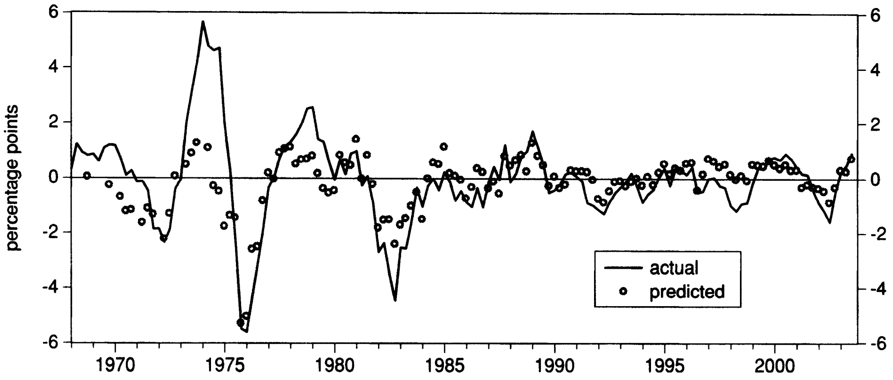


FIG. 2. Change in Four-Quarter Inflation (Four-Quarter Change in GDP Deflator, Less Preceding Four-Quarter Change).

of the event, not the time of the forecast. Forecast errors are simply the difference between the forecast and outcomes. A dot close to the line represents an accurate forecast.

As the figures show, there were large swings in activity and inflation in the 1970s and 1980s. Importantly, the Fed staff expected a large share of these. That is, the pre-1984 volatility documented by the literature on the Great Moderation was, to a substantial extent, anticipated. But successful predictions are harder to see over the most recent two decades.

## 3. KEY RESULTS

Some of the main features of the data can be seen in the following figures. I begin with inflation, where the story is simpler. The solid gray line in Figure 3 is the mean squared error (MSE) of forecasts of four-quarter inflation. The dashed black line is the variance of changes in inflation, that is, of four-quarter changes less the preceding nonoverlapping four-quarter change, measured using the same real-time observations as used in calculating forecast errors. For ease of later interpretation, the variance is uncentered (the mean is set at zero) and equals the MSE of a random walk forecast, though a conventional variance looks very similar. Both series are measured (in the figures, not the subsequent statistical analysis) using 5-year rolling windows, following Blanchard and Simon (2001), with the figure beginning 5 years from the first forecast.

Figures 4–6 show corresponding figures for real output. In contrast to inflation, where a four-quarter horizon is both standard and representative, the story for output varies somewhat with the horizon. I show errors and changes over one, four, and eight quarters, commencing from 1969, 1972, and 1984, respectively. The (centered or conventional) variance is of the deviation of percentage changes in output from the 5-year moving average.

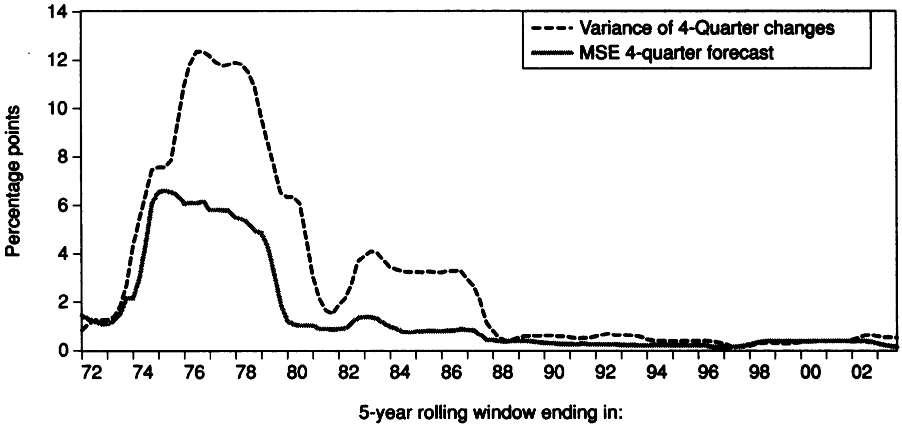


FIG. 3. Variance and Unpredictability of Changes in Inflation.

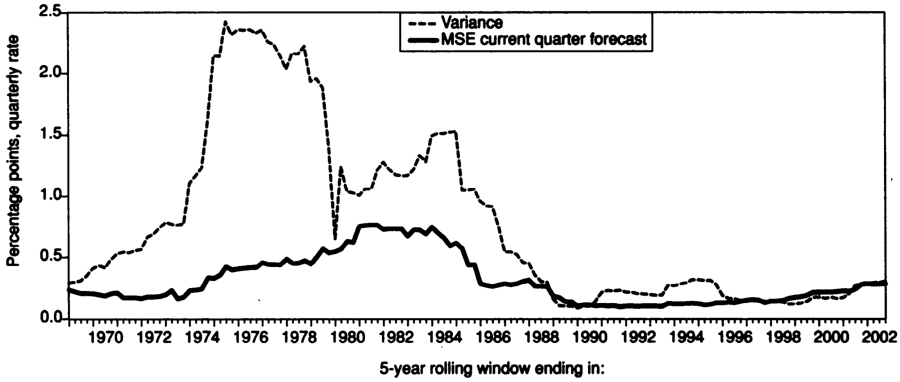


FIG. 4. Variance and Unpredictability of Quarterly GDP Growth.

The variances and MSEs shown in the figures are algebraically related. For illustration, let $y_t$ represent actual output growth in quarter $t$ and $f_t$ its forecast. The forecast error is then $e_t = y_t - f_t$. Rearranging, subtracting the mean $\bar{y}$ from each side, squaring and averaging over $n$ quarters ($n = 20$ for a 5-year window), gives

$$\frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})^2 = \frac{1}{n}\sum_{t=1}^{n}e_t^2 + \frac{1}{n}\sum_{t=1}^{n}(f_t - \bar{y})^2 + \frac{2}{n}\sum_{t=1}^{n}(f_t - \bar{y})e_t,$$

$$\text{Variance} = \text{MSE} + \text{predicted variation} + 2 \times \text{covariance}.$$

An analogous decomposition holds for changes in inflation, substituting the preceding change in prices for the mean. The distance between the two lines in each

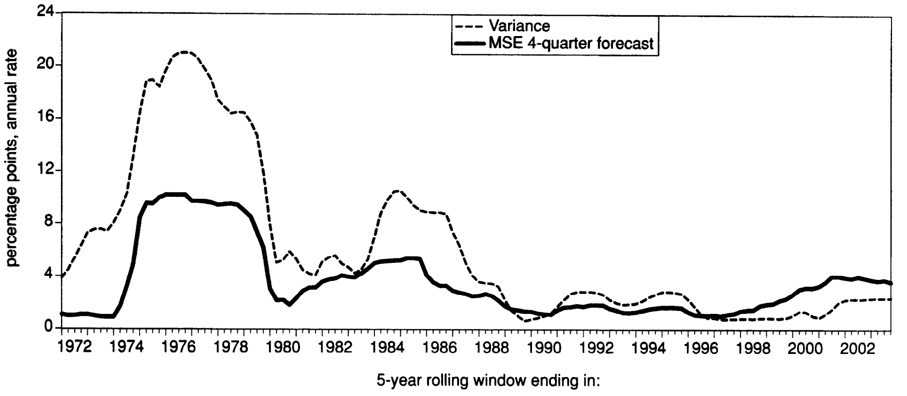


FIG. 5. Variance and Unpredictability of Four-Quarter GDP Growth.

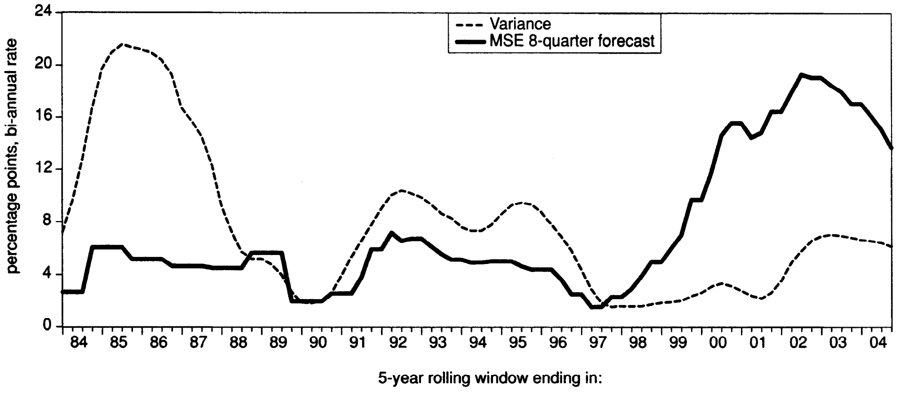


FIG. 6. Variance and Unpredictability of Eight-Quarter GDP Growth.

figure equals the sum of the last two terms in the equation. This can be called the predictable component of fluctuations in inflation or output.[3]

Four key results are illustrated in the figures. The statistical basis of these results, and those for other horizons, is established formally and precisely in Section 4.

(i) As the literature on the Great Moderation has documented, the variance of changes in output and inflation declines substantially, in the sense that it has been much smaller in the last two decades than it was in the previous two decades.

3. This terminology has some risk of ambiguity, given that it differs slightly from measures based on after-the-event econometric analysis, where the event $y$ is known before its prediction $f$. Then error minimization means the covariance of predictions and errors is zero (otherwise, errors could be reduced by changing the prediction). But when $f$ is determined before $y$, as in forecasting, the forecast does not minimize errors *ex post* (though it presumably minimizes them *ex ante*) and the *ex post* covariance need not equal zero. In forecast analysis, the "predictable component" may be negative.

(ii) Most of the decline in volatility reflects a decline in its predictable component, which seems to have virtually disappeared. That is, the decline in uncertainty (if any) is much smaller and less clear than that for volatility.

(iii) Nevertheless, uncertainty also declines for inflation and (at least for the current-quarter) output.

(iv) However, on some measures, no decline in uncertainty is evident. For example, eight-quarter output errors have increased over time.

Three secondary results also can be seen in the figures, though I do not pursue these. First, several of the time series are dominated by the outliers of 1974. Second, there is no clear trend in MSEs over the last two decades, with the possible exception of eight-quarter output errors, discussed below. Third, as can be seen in Figure 4, the Greenbook predicted most variations in current-quarter output growth in the 1970s and 1980s. Sims (2002), Faust and Wright (2007), and D'Agostino, Giannone, and Surico (2006) have emphasized this finding, suggesting it reflects a comparative advantage of the Greenbook. However, the result seems no longer valid. Figure 4 also shows (allowing for the 5-year averaging) that the Greenbook has not successfully explained current-quarter GDP since 1992. Although not shown, the same is true for inflation: the Greenbook successfully predicted changes in current-quarter inflation up till the early 1990s but has not done so since.

## 4. STATISTICAL TESTS

Do the estimates shown in the previous section represent changes in the distribution of forecast errors, or do they reflect a few lucky forecasts?

Table 1 shows tests of a change in forecast accuracy at unknown breakpoints. These tests involve regressing squared forecast errors on a constant and a postbreak

TABLE 1

CHANGES IN FORECAST ACCURACY

| (1) | (2) | (3) | (4) Inflation | (5) | (6) | (7) Output | (8) |
|---|---|---|---|---|---|---|---|
| Horizon | First forecast | $p$-value (%) | Most likely break | % change in RMSE | $p$-value (%) | Most likely break | % change in RMSE |
| 1 (current) quarter | 1965q4 | <0.1 | 1988q4 | −51 | 0.9 | 1985q2 | −38 |
| 2 quarters | 1966q1 | 0.8 | 1983q2 | −59 | 6.5 | 1985q3 | −39 |
| 3 quarters | 1968q1 | 5.0 | 1983q3 | −65 | 17.3 | 1984q4 | −41 |
| 4 quarters | 1968q1 | 9.6 | 1983q3 | −68 | 33.1 | 1984q3 | −38 |
| 5 quarters | 1969q4 | 10.8 | 1987q2 | −68 | 34.3 | 1984q3 | −37 |
| 6 quarters | 1972q3 | 11.1 | 1987q3 | −65 | 47.8 | 1984q4 | −25 |
| 7 quarters | 1979q1 | 2.1 | 1986q3 | −54 | 13.1 | 1997q3 | 61 |
| 8 quarters | 1979q1 | 3.9 | 1988q4 | −58 | 3.7 | 1997q4 | 89 |
| 9 quarters | 1989q4 | 60.5 | 2001q4 | −33 | 1.7 | 1997q4 | 94 |

NOTE: The $p$-values are of Andrews–Ploberger tests of instability, as discussed in the text. Most likely breaks are dated by the time of the event, not the time of the forecast. The sample extends from the first forecast, as given in column 2, through the forecast of October 2002, which had a horizon extending to 2004q4.

dummy. The coefficient on the constant equals the prebreak MSE, while the coefficient on the dummy is the change in the MSE. Regressions are run for a large range of possible breakdates, and the Exp $F$-statistic of Andrews and Ploberger (1994) is calculated (this can be thought of as a weighted average of the $t$-statistics on each of the dummies). Standard errors are calculated according to Newey and West (1987), with lag length of 1.5 times the horizon, rounded up, following the approach of Clark and McCracken (2005).[4] These standard errors are robust to ARCH effects and to the moving average nature of the errors that arises due to the overlap of multiple-quarter forecasts. Columns 3 and 6 report Hansen's (1997) $p$-value of the Exp $F$-statistic. This is the approximate probability of observing the sample, were there no change in forecast accuracy. Columns 4 and 7 report the breakdate with the highest $t$-statistic, this is, the most likely date for a structural change in the MSE. Columns 5 and 8 report the percentage change in root mean square error (RMSE) at this breakpoint. This is a different scaling to the units of measurement in the regressions but seems more intuitive.

The test statistics reported in Table 1 are only valid asymptotically. In small samples, Monte Carlo experiments using Newey–West standard errors tend to overreject (Newey and West 1994). Furthermore, squared errors are skewed, and hence the individual $t$-statistics used to form the Exp $F$-statistic have a nonstandard distribution, albeit one that resembles a normal moderately quickly.[5] Neither of these issues is critical for analysis of current-quarter errors, where I have 149 observations, weak autocorrelation, and strong rejections of the relevant null hypotheses. But as the horizon lengthens, one should treat the results more cautiously. Furthermore, breaktests near the ends of the sample are less reliable. To reduce this latter problem, I conduct the Andrews–Ploberger tests over 60% of possible breakpoints, whereas standard practice is to use 70%.[6]

Taken at face value, the clearest results in Table 1 are for inflation, where there is evidence of a structural break across forecast horizons. For example, the hypothesis of stability of current-quarter errors can be rejected at a significance level of 0.1%. At longer horizons, the overlap of forecasts and shortening of the sample reduce the effective number of observations, but most $p$-values are at least marginally significant. Generalizing across horizons and sample starting points, the most likely date of a reduction in forecast errors is around the mid 1980s with the reduction in RMSE at this point being between one-half and two-thirds.

4. Least squares regressions imply very similar breakdates and, unsurprisingly, smaller standard errors. Alternative Newey–West lag truncation rules, such as those based on the number of observations or Andrews (1991), can set the bandwidth narrower than the horizon, which seems improbable. Nevertheless, they give similar results for Table 1. An alternative approach would be a moving blocks bootstrap, but that is complicated by the discontinuous nature of the data.

5. Another possible concern is persistence. However, limited Monte Carlo evidence suggests this might be less important. O'Reilly and Whelan (2005) report that Andrews tests perform well except when their experimental data have higher autoregressive coefficients than do most Greenbook error series.

6. This matters. Tests of stability of output errors with a horizon of three or four quarters, using the conventional 70% trim, imply a breakpoint in 1973, after which errors *increase* with a $p$-value of 10%.

Evidence of a change in output uncertainty is mixed. Up to horizons of six quarters, the estimated timing and magnitude of breaks are similar: they occur in 1984 or 1985, after which RMSEs were about 40% smaller. That reduction is slightly smaller than the 50% reduction in RMSE found by Campbell (2007, Table 1). There is relatively strong evidence that these reductions represent a change in the distribution. The hypothesis that current-quarter output errors have been stable can be rejected at the 1% significance level. For next-quarter errors, stability can be rejected at the 7% level. At longer horizons, $p$-values are higher. But this is to be expected, given that there are fewer independent observations. If one assumes that the persistence of errors did not change (the similarity of the point estimates of the size of the change is consistent with this assumption), then a reduction in uncertainty could also be inferred at these horizons from the break in short-horizon errors. That is, the same structural change may have occurred as the current-quarter and next-quarter errors, but tests of longer-horizon errors lack enough independent data to clearly detect this.

In contrast, at horizons beyond six quarters, point estimates suggest that output uncertainty may have *increased*. For example, the RMSE of eight-quarter errors was 89% higher after 1997q4, with a $p$-value of 4%. The different trends in long- and short-horizon errors could reflect that the forecast horizon did not extend beyond six quarters until 1979, or (contrary to the assumption above) an increase in the persistence of errors over time. Or it may be that the increase in long-horizon errors is just a fluke, with the relatively high reported significance levels being discounted due to the small-sample problems discussed above. Although this last explanation deserves some weight, it does not mean that the long-horizon output data should be ignored. They provide much stronger results below, particularly in Table 3.

The results in Table 1 differ from those in Tulip (2005) where I reported stronger evidence of a reduction in uncertainty in output. Although the approaches differ in several respects, the most important is that my earlier paper took a breakpoint in 1984 as given. However, this date was chosen (e.g., by McConnell and Perez-Quiros) essentially because it generates the most significant break in related data. Clearly this biases the results toward finding evidence of instability.

For comparison, Table 2 presents the same tests of instability for volatility. For inflation, the dependent variable is the squared difference in the $h$-quarter percentage change in the GNP/GDP deflator from the preceding $h$-quarter change. For output, the dependent variable is the squared deviation of $h$-quarter changes from the 1965 to 2004 mean. Only those changes in output and inflation for which there is a Greenbook forecast are used, so the tests cover the same sample and number of observations as the tests for uncertainty. The $p$-values, shown in columns 2 and 5, are calculated in the same manner as Table 1.

Table 2 replicates earlier findings of a Great Moderation (see references in the opening paragraph). The contribution is the contrast with the weaker results of Table 1. As an example, consider output at a four-quarter horizon. Whereas Table 1 reported a 38% reduction in RMSE, Table 2 shows a larger 60% fall in the standard deviation. Whereas the corresponding $p$-value for a break in uncertainty was an insignificant 33%, that for volatility is 2%. Such differences occur throughout.

TABLE 2

CHANGES IN VOLATILITY

| (1) | (2) | (3) Inflation | (4) | (5) | (6) Output | (7) |
|---|---|---|---|---|---|---|
| Frequency of change | $p$-value (%) | Most likely break | % change in standard deviation | $p$-value (%) | Most likely break | % change in standard deviation |
| 1 quarter | <0.1 | 1992q2 | −56 | 0.1 | 1984q3 | −57 |
| 2 quarters | 1.0 | 1995q2 | −55 | 1.0 | 1984q3 | −59 |
| 3 quarters | 1.6 | 1983q3 | −64 | 1.1 | 1984q4 | −60 |
| 4 quarters | 1.1 | 1984q2 | −72 | 1.7 | 1992q3 | −60 |
| 5 quarters | 0.1 | 1977q3 | −64 | 1.2 | 1992q4 | −57 |
| 6 quarters | 0.8 | 1985q2 | −74 | 0.2 | 1985q1 | −51 |
| 7 quarters | 13.5 | 1985q3 | −71 | < 0.1 | 1985q3 | −51 |
| 8 quarters | 11.9 | 1985q3 | −75 | 1.3 | 1985q4 | −45 |
| 9 quarters | 0.7 | 1995q4 | −52 | 42.9 | 1994q3 | −50 |

NOTE: The $p$-values are of Andrews–Ploberger tests of instability in squared deviations from previous changes (for inflation) or the mean (for output), as discussed in the text. Dates of most likely breaks are for changes to the reported quarter. Samples range from changes beginning at the time of the first forecast for each horizon, as shown in Table 1, through to those beginning in 2002q4.

TABLE 3

CHANGES IN PREDICTABLE COMPONENT OF OUTPUT AND INFLATION

| (1) | (2) Inflation | (3) | (4) Output | (5) |
|---|---|---|---|---|
| Horizon | $p$-value (%) | Most likely break | $p$-value (%) | Most likely break |
| 1 quarter | 0.8 | 1993q3 | 1.3 | 1992q1 |
| 2 quarters | 34.9 | 1995q1 | 1.2 | 1992q4 |
| 3 quarters | 25.6 | 1992q3 | < 0.1 | 1992q3 |
| 4 quarters | 6.8 | 1984q2 | < 0.1 | 1996q3 |
| 5 quarters | 2.0 | 1984q4 | < 0.1 | 1997q1 |
| 6 quarters | 2.8 | 1985q2 | < 0.1 | 1993q2 |
| 7 quarters | 17.2 | 1985q2 | < 0.1 | 1985q2 |
| 8 quarters | 19.2 | 1985q3 | < 0.1 | 1993q3 |
| 9 quarters | 7.5 | 1995q4 | 1.4 | 1998q3 |

NOTE: The $p$-values are of Andrews–Ploberger tests of instability in the difference between squared forecast errors and deviations from previous changes (for inflation) or the mean (for output), as discussed in the text. Samples are the same as in Tables 1 and 2.

In general, the reductions in volatility are clearer, larger, and more uniform than the reductions in uncertainty. The difference is mild for inflation and dramatic for output.

Table 3 tests whether the difference between the change in uncertainty and that in volatility is statistically significant. It repeats the tests of Tables 1 and 2 using the difference between squared errors and squared deviations (from the mean for output and from the previous change for inflation) as the dependent variable.[7] This difference was described above as the predictable component of economic fluctuations (though this interpretation does not affect these results). As shown by the low $p$-values, there

7. Campbell (2007, Table 2) reports a Wald test like this for output, taking a breakdate of 1984 as given.

is relatively strong evidence of a reduction in the predictable component of inflation (column 2) and overwhelming evidence of a reduction in the predictable component of output (column 4), including, notably, at long horizons. Put another way, the reduction in volatility is clearly larger than the reduction in uncertainty.

The estimated size of the reduction in the predictable component (not shown) is around 100% at most frequencies. In other words, the tests find breakpoints after which the predictable component is approximately zero. After these breakpoints, the mean of output growth, or the preceding rate of inflation, would forecast about as well as the Greenbook. This lack of predictive power can be seen in Figures 3 through 6. It has also been found for other U.S. forecasters (Campbell, 2007, Reifschneider and Tulip, 2007, Stock and Watson, 2007), a wider range of variables and econometric models (Atkeson and Ohanian 2001, D'Agostino, Giannone, and Surico 2006), and the other large industrialized countries (Vogel 2007).

## 5. CONCLUSION

Table 1 reported statistically and economically significant reductions in uncertainty about inflation and (at least for short horizons) output. Table 3 showed that these reductions in uncertainty were significantly smaller than the reductions in volatility across variables and horizons. Although some implications of these results were mentioned in the introduction, two pointers for further research might be worth noting.

Perhaps the most immediate implications are for measures of uncertainty. Whereas Table 1 found breaks in inflation and short-horizon output errors in the mid-1980s, Figures 3– 5 suggest little change since then. (A possible exception is long-horizon output errors that have widened). On balance, these results provide some support for the FOMC's approach of only averaging errors from the last two decades in assessing "typical" uncertainty. Future research will hopefully uncover a richer set of variables, beyond a post-1980s dummy, for estimating conditional uncertainty.

A more fundamental set of issues surrounds the apparent disappearance of the predictable component of economic fluctuations. Future research may examine why this occurred and how decision makers and forecasters should react to it.

## LITERATURE CITED

Andrews, Donald. (1991) "Heteroskedasticity and Autocorrelation Consistent Covariance Estimation." *Econometrica*, 59, 817–58.

Andrews, Donald, and Werner Ploberger. (1994) "Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative." *Econometrica*, 62, 1383–414.

Atkeson, Andrew, and Lee E. Ohanian. (2001) "Are Phillips Curves Useful for Forecasting Inflation?" *FRB Minneapolis Quarterly Review*, Winter, 2–11.

Beaulieu, J. Joseph, and Jeffrey A. Miron. (1992) "A Cross Country Comparison of Seasonal Cycles and Business Cycles." *Economic Journal*, 102, 772–88.

Bernanke, Ben S. (2004) "The Great Moderation." Speech to the Eastern Economic Association, Washington, DC, February 20.

Bernanke, Ben S., and Jean Boivin. (2003) "Monetary Policy in a Data-Rich Environment." *Journal of Monetary Economics*, 50, 525–46.

Blanchard, Olivier, and John Simon. (2001) "The Long and Large Decline in U.S. Output Volatility." *Brookings Papers on Economic Activity*, 1, 135–64.

Blinder, Alan S., Elie R. D. Canetti, David E. Lebow, and Jeremy B. Rudd. (1998) *Asking about Prices: A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.

Campbell, Sean D. (2007) "Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters." *Journal of Business and Economic Statistics*, 25, 191–200.

Clark, Todd E., and Michael W. McCracken. (2005) "Evaluating Direct Multi-Step Forecasts." *Econometric Reviews*, 24, 369–404.

D'Agostino, Antonello, Domenico Giannone, and Paulo Surico. (2006) "Unpredictability and Macroeconomic Stability." European Central Bank, Working Paper No. 605.

Faust, Jon, and Jonathan H. Wright. (2007) "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset." NBER Working Paper Series No. 13397.

Federal Open Market Committee. (2007) *Summary of Economic Projections for the Meeting of October 30-31, 2007*. Washington, DC: Federal Reserve Board of Governors.

Hansen, Bruce, E. (1997) "Approximate Asymptotic *P*-Values for Structural Change Tests." *Journal of Business and Economic Statistics*, 15, 60–67.

Kim, Chang-Jin, and Charles Nelson. (1999) "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." *Review of Economics and Statistics*, 81, 608–16.

McConnell, Margaret, and Gabriel Perez-Quiros. (2000) "Output Fluctuations in the United States: What Has Changed Since the Early 1980s?" *American Economic Review*, 90, 1464–76.

Meyer, Laurence H. (1998) "Come with Me to the FOMC." Speech at Willamette University, Salem Oregon, April 2.

Newey, Whitney K., and Kenneth D. West. (1987) "A Simple, Positive, Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55, 703–8.

Newey, Whitney K., and Kenneth D. West. (1994) "Automatic Lag Selection in Covariance Matrix Estimation." *Review of Economic Studies*, 61, 631–53.

O'Reilly, Gerard, and Karl Whelan. (2005) "Has Euro-Area Inflation Persistence Change Over Time?" *Review of Economics and Statistics*, 87, 709–20.

Reifschneider, David, and Peter Tulip. (2007) "Gauging the Uncertainty of the Economic Outlook from Historical Forecasting Errors." In Finance and Economics Discussion Series 2007-60, Federal Reserve Board.

Reifschneider, David, and Peter Tulip. (2008) "Gauging the Uncertainty of the Economic Outlook from Historical Forecasting Errors." Mimeo.

Robertson, John C., and Ellis W. Tallman. (1998) "Data Vintages and Measuring Forecast Model Performance." *Economic Review, Federal Reserve Bank of Atlanta, Fourth Quarter*, 83, 4–20.

Romer, Christina D., and David H. Romer. (2000) "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review,* 90, 429–57.

Romer, Christina D., and David H. Romer. (2008) "The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?" Presented at the American Economic Association Annual Meeting, January 5.

Schuh, Scott. (2001) "An Evaluation of Recent Macroeconomic Forecast Errors." *New England Economic Review*, Federal Reserve Bank of Boston, January-February, 35–56.

Sims, Christopher. (2002) "The Role of Models and Probabilities in the Monetary Policy Process." *Brookings Papers on Economic Activity,* 2, 1–62.

Stock, James H., and Mark W. Watson. (2003) "Has the Business Cycle Changed? Evidence and Explanations." In *Monetary Policy and Uncertainty: Adapting to a Changing Economy.* Jackson Hole, WY: Federal Reserve Bank of Kansas City Symposium.

Stock, James H., and Mark W. Watson. (2007) "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking,* 39, 3–33.

Tulip, Peter. (2005) "Has Output Become More Predictable? Changes in Greenbook Forecast Accuracy." In Finance and Economics Discussion Series 2005-31, Federal Reserve Board.

Vogel, Lukas. (2007) "How Do OECD Growth Projections for the G7 Economies Perform?" OECD Economics Department Working Paper No. 573.

Woodford, Michael. (2008) "The Fed's Enhanced Communication Strategy." VoxEU January 8, 2008. Available at http://www.voxeu.org/index.php?q=node/860.