View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

FEDERAL RESERVE BANK of ATLANTA

Transparency, Expectations, and Forecasts

Andrew Bauer, Robert A. Eisenbeis,
Daniel F. Waggoner, and Tao Zha

Working Paper 2006-3
April 2006

WORKING PAPER SERIES

FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Transparency, Expectations, and Forecasts

Andrew Bauer, Robert A. Eisenbeis, Daniel F. Waggoner, and Tao Zha

Working Paper 2006-3
April 2006

Abstract: In 1994, the Federal Open Market Committee (FOMC) began to release statements after each meeting. This paper investigates whether the public's views about the current path of the economy and of future policy have been affected by changes in the Federal Reserve's communications policy as reflected in private sector's forecasts of future economic conditions and policy moves. In particular, has the ability of private agents to predict where the economy is going improved since 1994? If so, on which dimensions has the ability to forecast improved? We find evidence that the individuals' forecasts have been more synchronized since 1994, implying the possible effects of the FOMC's transparency. On the other hand, we find little evidence that the common forecast errors, which are the driving force of overall forecast errors, have become smaller since 1994.

JEL classification: E59, C33

Key words: transparency, common errors, idiosyncratic errors

This article was presented at the fourth European Central Bank Workshop on Forecasting Techniques, held on December 14–15, 2005, in Frankfurt, Germany. It is forthcoming in the Federal Reserve Bank of Atlanta *Economic Review*. The authors thank John Robertson and Ellis Tallman for critical comments and Cindy Soo and Eric Wang for excellent research assistance. A similar version of this research is also published with the same title in the first quarter issue of the Atlanta Fed's *Economic Review*. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Andy Bauer, Senior Economic Analyst, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-7906, andy.bauer@atl.frb.org; Robert Eisenbeis (corresponding author), Executive Vice President and Director of Research, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8824, robert.a.eisenbeis@atl.frb.org; Daniel Waggoner, Assistant Policy Adviser, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8278, daniel.f.waggoner@atl.frb.org; or Tao Zha, Policy Adviser, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8353, tzha@earthlink.net.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's Web site at www.frbatlanta.org. Click "Publications" and then "Working Papers." Use the WebScriber Service (at www.frbatlanta.org) to receive e-mail notifications about new papers.

Transparency, Expectations, and Forecasts

I. Introduction

Many macroeconomists have argued that a central bank should be transparent about its objectives, its views about the economic outlook, and the reasoning behind its policy changes (see Faust and Leeper 2005). In 1994 the Federal Open Market Committee (FOMC) began to release statements accompanying changes in the federal funds rate target. Since then, the degree of specificity of the statements and the guidance provided on the likely course of future policy have evolved significantly.[1] In a recent paper, Woodford (2005) discusses two kinds of central-bank communications: current policy decisions and the central bank's view of likely future policy. He articulates four categories of information—the central bank's view of current economic conditions, current operating targets, strategies guiding policy decision making, and the outlook for future policy—that a central bank might seek to communicate to the public. Woodford argues that these open communications are "beneficial, not only from the point of view of reducing the uncertainty with which traders and other economic decision makers must contend, but also from that of enhancing the accuracy with which the FOMC is able to achieve the effects on the economy that it desires, by keeping the expectations of market participants more closely synchronized with its own."

[1] Kohn and Sack (2003) characterize several distinct periods of increasing transparency in FOMC statements: statements on changes in the discount rate (1989–93), statements on changes in the federal funds rate (1994–98), statements including policy tilt (1998–99), and statements including assessment of the balance of risks (2000–04). In May 2003 a further refinement was added to separately state the committee's views on the risks to inflation and growth. And, finally, in August 2003 the committee provided explicit guidance on the likelihood that policy would remain accommodative.

This article investigates whether the public's views about the economy's current path and about future policy have been affected by changes in the Federal Reserve's communications policy as reflected in private-sector forecasts of future economic conditions and policy moves. In particular, has private agents' ability to predict the direction of the economy improved since 1994, when the FOMC began to publicly state its views of the economic outlook? If so, on which dimensions has the ability to forecast improved? The analysis focuses on both the short-term and longer-term economic forecasts of key macroeconomic variables—such as inflation, gross domestic product (GDP) growth, and unemployment—and of policy variables such as short-term interest rates. Private agents' current-year and next-year forecasts are used as proxies for the public's short-term and longer-term expectations, and empirical evidence is presented regarding whether such forecasts have performed better in predicting future economic and policy conditions since 1994.

The private-agent forecasts used in this article are those of individual participants as well as the consensus (average) forecasts contained in the monthly Blue Chip Economic Indicators surveys from 1986 to 2004, which include both the pre-FOMC-statement subperiod (1986:01–1993:12) and the post-FOMC-statement subperiod (1994:01–2004:12). We employ the econometric methodology of Eisenbeis, Waggoner, and Zha (2002), which permits us to evaluate the accuracy of forecasts both in cross section and across time and to examine the errors in forecasting key economic variables on both a univariate and a multivariate basis. The latter is important because agents are not simply forecasting one economic variable but rather a set of variables that presumably are interrelated and jointly capture important dimensions of economic

performance. Good forecasts on one dimension but poor overall performance may provide some indication of the internal consistency of the forecaster's approach.

This cross-sectional data set enables us to decompose forecast accuracy into two components: the common error that affects all individual participants and the idiosyncratic error that reflects discrepant views across individuals about future economic and policy conditions. According to Woodford (2005), one should expect the idiosyncratic error to become smaller as FOMC open communications become more transparent. But the common error may not change much because it is likely to be affected by factors other than changes in policy transparency, such as unforeseen business cycles.

To preview the main result, we find that since 1994 the idiosyncratic errors for key macroeconomic variables have steadily declined and the expectations of market participants are more closely synchronized to one another. We find no evidence, however, that the common error has become smaller since 1994, especially for the longer-term forecasts.

II. Methodology

Let μ_t be a $n \times 1$ vector of economic variables at time t, y_t be the realized value of these economic variables, and y_t^i is the *i*th individual's forecast value of the variables. Assume that y_t is normally distributed with mean μ_t and economy-wise (common) covariance matrix Ω_t^R and that y_t^i is normally distributed with mean μ_t and forecast-wise covariance matrix Ω_t^F. The super-scripts *R* stands for "realized" and *F* for

"forecast." The covariance matrix Ω_t^R reflects the aggregate shocks that affect the realized value of μ_t; the covariance matrix Ω_t^F captures the discrepancy in forecasts across individual participants. The assumption that the mean forecast among individual participants is μ_t is reasonable because previous work has suggested that the Blue-Chip consensus forecast as a proxy to the mean forecast is close to being an unbiased estimate of μ_t (Bauer, Eisenbeis, Waggoner, and Zha, 2003). Denote the forecast error for the *i*th forecaster by $x_t^i = y_t^i - y_t$. It follows that the individual forecast error x_t^i has mean zero and variance matrix

$$\Omega_t = \Omega_t^R + \Omega_t^F,$$

which indicates that x_t^i is subject to both idiosyncratic and common shocks.[2] The standard statistical theory implies that

$$\chi_t^i \equiv x_t^{i\prime} \Omega_t^{-1} x_t^i \sim \text{chi}^2(n),$$

where $\text{chi}^2(n)$ denotes the chi-square distribution with *n* degrees of freedom and χ_t^i is a square error weighted by Ω_t. The above expression says that the weighted square error χ_t^i follows the chi-square distribution with *n* degrees of freedom. To measure the forecast accuracy for each individual participant, we compute a score value (*p*-value) associated with this chi-square distribution and call it an "accuracy score." The score for individual forecaster *i* at forecast time *t* is a function of χ_t^i and *n*:

$$p\left(\chi_t^i, n\right) = 1 - \text{chi2cdf}\left(\chi_t^i, n\right),$$

[2] In future research, we intend to relax the assumptions that the Consensus forecast is equal to μ_t and idiosyncratic shocks are independent of common shocks.

where chi2cdf $\left(\chi_t^i, n\right)$ is the probability that a random observation from the chi-square distribution with *n* degrees of freedom falls in the interval [0 χ_t^i].[3]

As Eisenbeis, Waggoner, and Zha (2002) pointed out, the summary measure $p\left(\chi_t^i, n\right)$ is a probability that is invariant to the underlying scales of error variances. It can be interpreted that the i th participant's forecast is closer to the realized value than do 100 $p\left(\chi_t^i, n\right)$ percent of all possible forecasters. Moreover, the score $p\left(\chi_t^i, n\right)$ can be compared across forecasters, within a forecast period, and across periods.

Bauer, Eisenbeis, Waggoner, and Zha (2003) show how to estimate the covariance matrices Ω_t^R and Ω_t^F . The matrix Ω_t^R can be estimated as the sample covariance matrix of the Blue-Chip Consensus forecast errors across time under the assumption that Ω_t^R is the same across years for each month but varies across months within a year. Thus, the variances on the diagonal of Ω_t^R become smaller as t gets closer to the end of the year, for more information becomes available to forecast economic conditions for the current year. The covariance matrix Ω_t^F can be estimated as the sample covariance matrix of forecast errors across individual forecasters; this covariance varies both across months and across years.[4] The estimate of Ω_t , denoted by $\hat{\Omega}_t$, is the sum of the estimates of Ω_t^R and Ω_t^F . Given this estimate, the weighted square error can be calculated as

$$\hat{\chi}_t^i = x_t^{i\prime} \hat{\Omega}_t^{-1} x_t^i .$$

At each time *t*, the average accuracy score is

[3] If the assumptions used are valid, the distribution of accuracy scores from 1986 to 2004 should be uniform. We have verified that such a distribution is more or less uniform, taking into account small-sample uncertainty.

$$\hat{p}_t(n) = \frac{1}{N_t}\sum_{i=1}^{N_t} p\left(\hat{\chi}_t^i, n\right),$$

where N_t is the number of individual forecasters at time *t*. One can also calculate the cross-sectional distribution of accuracy scores, which is described in detail in Box I.

III. Vintage Data and Forecast Errors

The monthly Blue-Chip Economic Indicators report the forecasts of key macroeconomic variables for the current and next years. We study the annual average forecasts of five key variables: the 3-month treasury bill rate, the consumer price index (CPI) inflation rate, real gross national product (GNP) for 1986 to 1995 or real gross domestic product (GDP) from 1996 to 2004, the unemployment rate, and the long-term bond yield (the corporate bond yield from 1986 to 1995 or the ten-year treasury note yield from 1996 to 2004. The three-month T-bill rate, the CPI inflation rate, the unemployment rate, and the long-term bond yield are monthly variables while real GNP/GDP is a quarterly variable. This frequency difference is important to note when evaluating forecasts.

As the year gets close to the end, more information is available about the actual current-year data and therefore the forecast errors for both the current and next years get smaller. For example, the forecasters participating in the December Blue Chip survey will have data monthly data on the three-month T-bill rate and the long-term bond yield through November. They will have data on the unemployment rate through October or November. They will have data on the CPI inflation rate through October. However, since GNP/GDP is released quarterly, forecasters will only have information regarding

[4] Other estimates can also be constructed using model-based methods.

GNP/GDP through the third quarter of the year. The weighted square error $\hat{\chi}_t^i$ is designed to avoid the influence of different amounts of available data so that the errors are comparable across time.

To gauge forecast errors, the realized values of each variable at a given time must be used. The values of some of variables are revised over time by the agencies responsible for reporting those variables. In particular, real GNP/GDP is reported quarterly and revised twice. Every year additional benchmark revisions may be made in July to the past data of GDP. Hence, what is reported are the continuously changing estimates of the final values of many key economic variables. Finally, sometimes the definition of GDP is changed and the series is completely revised. With such revisions taking place, the question arises as to what vintage data should one use to evaluate forecast errors? From a macro policy perspective, we would argue that the focus should be on the "best" estimate of the final value of the variable of interest. However, often that value is not known for several years, and sometimes the difference between even a preliminary estimate and its nearest neighbor estimates can be very large. For example, the advanced estimate for real GDP for Q1 2005 was 3.1% which was revised up by the Bureau of Economic Analysis from 3.4% and finally to 3.8% as more data on the performance of the economy became available. The difference between the first and most recent estimate could cause policy makers to infer that the economy was growing below trend according to the first number, but above trend based on the final estimate. Such differences could have significantly different implications for policy. For this reason, we would argue that the focus should be on forecast methods that best approximate the final number rather than the initial estimate. Moreover, with such a

focus, a priori knowledge of the expected performance of a model or forecasting method, can help inform the policy maker as to which evidence to give greater weight to, when there are significant differences between the initial releases of data and forecasts of those data.

For the purposes of this study of the current-year forecasts, we use the vintage data available at the end of January following the current year; and for the next-year forecasts, we use the data available at the end January following the next year. We use vintage data so that the results here will be comparable with previous studies. We also provide a comparison between the average Blue Chip Consensus score using vintage and final data, using January 2005 data as our final data.

IV. Accuracy Scores

In this section we look at the distribution of scores at each month and examine whether the distribution has changed over time, especially from the pre-statement sub-period to the post-statement period. The technical details of how to characterize the cross-sectional distribution of scores are provided in Box I.

Chart 1A shows the time-series paths of average scores and standard deviations of scores for the current year and Chart 1B shows the time-series paths of skewness and kurtosis for the current year. Charts 2A and 2B show the similar paths for the next year. The measure of standard deviation is often used to approximate the volatility of the public's expectations or forecasts at each point in time. As it can be seen from Chart 1A, the average score fluctuates over time and so does the standard deviation of scores. There are no noticeable differences in the degree of fluctuation before and after 1994, and

nor are there differences for any sub-periods after 1994. There is no trend in which the average score has increased or the standard deviation of scores had decreased since 1994. There are clearly periods when forecasters made big errors, such as missing the onset of the recessions in 1990 and 2001. In addition, while the average scores have increased in the 2004, so have the standard deviations of those scores. Similarly, the average scores dropped significantly in 1995, which is mainly caused by the definition change of the GDP series. In January of 1996 the Bureau of Economic Analysis changed the measurement of GDP to a chain-weighted system, but the forecasts made before January 1996 might be based on the non-chain-weighted series. Interestingly, this change seems having relatively less effect on the longer-term forecast errors (Chart 2B).

The average score for the next year (Chart 2A) shows no improvement since 1994 and in fact appears to have drifted lower since 1996. There has been a steady upward drift in the standard deviation of the scores since 2001. The pattern of the drift in the standard deviation is similar to that which occurred just prior to and coming out of the 90-91 recession. As will be discussed further in the next section, these lower scores after 1996 are most likely associated with the nature of the business cycle and unexpected growth in productivity that surged in the late 1990s.

We now look at the skewness and kurtosis of accuracy scores (Charts 1B and 2B). Skewness measures how asymmetric the score distribution is. The more negative this measure is, the more scores spread out toward 0%. Conversely, the more positive this measure is, the more scores spread out toward 100%. Kurtosis measures how likely the score distribution has extreme outliers that may affect the average score. The bigger the value of this measure is, the more likely we have outliers in the score distribution. For

the current-year forecasts, the skewness and kurtosis have remained stable except for a few periods (Chart 1B). The spike that occurred in 1995 is due to the redefinition of GDP and the small spikes around 2001 are associated with the recent recession. For the next-year forecasts, again, there is no clear pattern or trend in which skewness and kurtosis have changed since 1994 (Chart 2B). There were a couple of spikes in skewness and kurtosis, whose periods correspond to the Asian financial crisis and the recent recession.

To provide further information about distributional changes of accuracy scores, we display in Chart 3 the time-series paths of accuracy scores of Blue-Chip consensus forecast and the average of the top and bottom 5 forecasts for each month. The current-year results are reported in Chart 3A and the next-year results are in Chart 3B. The consensus forecast is of particular interest because its score is *on average* the highest (see Appendix II for details) and because it performs better than any single individual forecaster over the sample. Again, it can be seen from Chart 3 that there is no tendency that these scores have improved over time since 1994. In fact, the scores of consensus forecasts appear to be slightly lower after 1996 than before, especially for the next-year forecast. Moreover, the drop in the consensus scores around the recent recession and again following 9/11 in 2001, suggests that events and exogenous shocks affected forecast performance much more than FOMC statements. The drop in the scores towards the end of 1995 is due to the redefinition of GDP. We also show the average scores for the top five forecasters in each period as well as the average score for the 5 poorest performers. The evidence suggests that data have fat tails, with most of the forecasts being clustered at the high end with a few really poor performers on the bottom.

All these findings suggest that the individual participant's forecast performance *relative to* other participants has not improved between the pre-statement and post-statement periods. Although the accuracy score is a powerful summary measure of forecasting performance, it is a nonlinear function of the square forecast errors weighted by the overall covariance matrix Ω_t. It would be informative to separate Ω_t and forecast errors for further analysis. In the next section, we examine whether the covariance matrix Ω_t^F has changed over time and study the sources of forecast errors that do not depend on Ω_t^F and Ω_t^R.[5]

V. Transparency and Sources of Forecast Errors

Kohn and Sack (2003) and Woodford (2005) argue that the contents in FOMC statements have become more transparent since 1994. It is therefore important to see whether the expectations of market participants via the forecasts of key economic variables have become more synchronized in the post-statement period than in the pre-statement sub-period. If there is useful information content in the statement, then one might expect that there may be an overall improvement in forecast accuracy, ceteris paribus, or at least more agreement among forecasters (ie. a tighter distribution of idiosyncratic errors.) A positive answer may provide evidence about the effects of the FOMC statements on the private sector's agreement on the direction of the future economy.

We also examine the sources of forecast errors by directly decomposing the mean square error (MSE) into the idiosyncratic component reflecting the discrepancy in

[5] The reader may recall that by assumption Ω_t^R does not change from one year to another. We intend to

individual participants from the Surveys and the common component that is associated with unanticipated aggregate shocks and affects all participants. The technical details of this decomposition are provided in Box II.

The MSE is the average of square errors across individual forecasters. Arguably, both the idiosyncratic and common errors may show a decreasing trend if there is useful information in the statement and forecasters gain better understanding of the economy over time, especially since 1994. To the extent that the common error is affected by exogenous aggregate shocks, and the distribution of the shocks is not constant, the there may be no clear inference about the size of the common error. However, we hypothesize that the more important impact is likely to be seen for the idiosyncratic component, in that the idiosyncratic errors should be tighter – that is there should be greater agreement among the forecasters. The empirical results presented below confirm this hypothesis.

We first study how synchronized the expectations of market participants are. The degree of synchronization is measured by the cross-sectional standard deviations of all the variables, which are equal to square roots of the diagonal elements of Ω_t^F. Charts 4-8 report the cross-sectional standard deviation of each of the five macroeconomic variables considered in this paper. Charts 4A-8A display the standard deviations for both the current-year and next-year forecasts; Charts 4B-8B display the 12-month moving averages of the standard deviations to show the trend more clearly. It is clear from these charts that for not only the interest rates but also the other variables, the trend has been downward and the standard deviations after 1994 tend to be smaller than before 1994. These findings suggest that individual participants' forecasts have indeed been more

---

relax this assumption in future research.

synchronized since 1994, both in terms of their overall view of the economy and of the interest rate variable most closely tied to policy.

We now study the decompositions of forecast errors for each of the five key macroeconomic variables. Charts 9-14 show the time-series paths of decompositions for individual variables as well as all the variables jointly. As evident in Panel A of each chart of Charts 9-14, one uniform result is that the time path of idiosyncratic errors shows a pattern of steady decline as well as very seasonal pattern for the current-year forecasts. Within the current year, the individual participant's forecast error becomes much smaller as the time gets close to December. The seasonal pattern is much less obvious for the next-year forecasts (Panel B of each chart of Charts 9-14), partly because the uncertainty about the economy next year is still large even if one tries to forecast as of December last year. For both the current-year and next-year forecasts we see a clear pattern of smaller idiosyncratic errors after 1994. Again, these results are consistent with the hypothesis that individual forecasts have been more synchronized since 1994.

Patterns of common errors are distinctively different from those of idiosyncratic ones and the difference seems to be associated with business cycles unrelated to the FOMC statements. One can see from Charts 9-14 that the common errors in the current year forecast are large relative to the idiosyncratic errors, whereas the common errors are dominant in the next year forecasts. But there is no apparent pattern that the common errors are smaller after 1994 than before.

According to Chart 9A, the unusually large common errors for the current-year forecasts of the short-term interest rate occur in 2001. These common errors are associated with the unexpected sharp decline of the federal fund rate. The large common

errors of longer-term (next-year) forecasts seem to be associated with missing the turning point of the federal funds rate in the early 2000s and failing to predict the unchanged rate in 2002 and 2003 (Chart 9B).

For CPI inflation, except for a couple of unusually large common errors before 1994, the common errors of the current-year forecasts have the similar patterns before and after 1994 (Chart 10A). The common errors for the next-year forecasts tend to be larger in the period after 1996 than before (Chart 10B), and there shows no tendency that these errors have become smaller than before 1994.

Typically as the time gets closer to the end of the year, both idiosyncratic and common errors become smaller for the current year forecasts. But for 1995, there are unusually large common errors of the current-year forecasts of real GNP/GDP towards the end of 1995, caused mainly by the definition change of the GDP series. These errors are amplified when divided by the diminishing variances of forecast errors, which explains the steep drop of accuracy scores toward the end of 1995 in Chart 3A. In Chart 11A, the errors are not divided by the variances of forecast errors and thus are not as visually dramatic as in Chart 3A. The substantial, persistent common errors of the next-year forecasts in the late 1990s are consistent with the sustained increase in productivity growth largely unexpected by the public, while the federal funds rate did not change much in the late 1990s.

The common errors in forecasting the unemployment rate for the current year appear to be somewhat smaller after 1994 than before, but those errors for the next year have similar patterns before and after 1994 (Charts 12A and 12B). The large common

errors for the next-year forecasts have much to do with business cycles and with the errors in predicting output growth (Chart 12B).

There are no clear patterns in which the common forecast errors of the long-term bond yield have become smaller since 1994 (Charts 13A and 13B). In particular, the errors around the recent recession are relatively large in magnitude. It is interesting, however, that there was a noticeable drop in the idiosyncratic errors in both the current year and next year forecast after 1987 when Chairman Greenspan became chairman and the effects of the stock market problems dissipated.

Charts 14A and 14B summarize the decomposition of the MSE for all the five variables combined. For the current-year forecasts, the seasonal pattern is evident, as explained early in this article. For the next-year forecasts, the large common errors occurred in the periods around the last two recessions. The persistent and volatile common errors since 1994 are mainly due to the correlation effect among forecast errors across variables, because the forecast errors for individual variables other than GNP/GDP do not share these features. Overall there is no evidence that the public's forecasts of key macroeconomic variables have improved since 1994, following the FOMC's efforts to increase transparency.

Table 1 reports the average of percentages of the MSE that are attributed to the idiosyncratic component and the common component. Two methods are used to compute the average percent contributions. The first method is to calculate the percent contributions of idiosyncratic and common errors for each period and then average over all the periods. This method helps eliminate outliers of extremely large errors, so the

results may not conform to the impression by looking at the charts. The top panel of Table 1 reports these results.

The second method is to accumulate the forecast errors of both types throughout the entire sample and then calculate the percent contributions of idiosyncratic and common errors. The results generated by this method are reported in the bottom panel of Table 1. This method is likely to be influenced by outliers but will be consistent with the impression given by the charts.

Let us look at the top panel of Table 1 first. For the current-year forecasts, except for GNP/GDP the idiosyncratic errors contribute much more to the total errors than the common errors, despite the fact that the common errors are much larger at times. But when one examines all the variables jointly, the common errors become more important. This result implies that while it may be relatively easy to predict a single variable, the problem of predicting a set of economic variables may be more difficult.[6] As for the longer-term (next-year) forecasts, the picture is completely different: the common errors are clearly a driving force for almost all variables (except for CPI), individually and jointly.

Compared to the top panel of Table 1, the results from the bottom panel give more dominate role to the common errors, partly because the common errors are much larger than the idiosyncratic errors in some periods. All in all, it is clear that the common errors play a dominant role in overall forecast errors.

This finding suggests that unexpected shocks, which of course are also not anticipated in the FOMC statements, are dominant factors in affecting forecast

[6] One might also infer that there are different models being used and they perform better on some variables than others, but in aggregate there are significant differences among the forecasts.

performance, and improvements in policy transparency would be unlikely to make the forecast errors smaller except on the margins.[7] Another possibility is that clearer patterns may show up as more observations become available; the FOMC only began in August 2003 to provide explicit guidance on the likely path of future policy and state-contingent economic conditions in the future. Given the data available today, however, we find no empirical evidence of significant improvement in the common forecast errors over the period in which the FOMC attempted to clarify its views of the economy or the likely course for future policy. This finding does not necessarily suggest that the movement toward transparency has been a failure. It may simply indicate that no new information was provided in the statements that had not already been inferred by market participants. Given the unpredictable nature of business cycles, moreover, the common error may be mostly affected by factors other than monetary policy transparency.

VI. Vintage Data Versus Final Data

One could argue that whenever forecast errors are evaluated, final data available at that time should be used. The reason is obvious. From a policy perspective, being able to predict well initially released data that subsequently get revised, may lead to policy errors, especially when close to turning points or where the revisions may substantially change one's view of the economy. However, when policy formulation relies heavily upon model forecasts, it is important that those forecasts capture, as best as possible the true underlying paths for key economic variables. If they do not, then the risk of series policy errors may be increased. Furthermore, how to choose the vintage data at various

[7] This interpretation is consistent with the results of Stock and Watson (2003) and Sims and Zha (forthcoming).

points in time is completely arbitrary, and there is not statistical or economical foundation for such an arbitrary decision. The public knows that data such as GDP are often revised and sometimes thoroughly revised. They take such unpredictable outcomes into account and make their forecasts as accurately as possible *on average*.

In this section, we use the revised and most currently available data at the beginning of 2005 to re-compute the forecast errors. Chart 15 displays the Blue Chip Consensus accuracy scores with the vintage data and the final data for both the current-year and next-year forecasts. Interestingly, the scores using final data do not fall considerably. The average current-year score using vintage data is 70.9 while the average current-year score using final data is 67.0, just 3.9 points lower. For the next-year forecast scores there is very little difference between the scores using vintage and final data. The average score using vintage data is 57.4 while the average score using final data is 56.4. There are several periods, in 1992, 1995-1996, and in 1998 where the next-year forecast score was lower using final data, but at the same time there are several periods (1994, 1999 and 2002) where it is higher. These results indicate that future data revisions are random enough such that they do not introduce a bias that does not significantly negatively affect forecast scores on average. More important, it also suggests that the data revisions do not pose significant risks for policy makers.

One would expect, perhaps, a greater disparity between the two scores given that additional revision errors are unpredictable. However, an important advantage of using the final data is that one can avoid the distorted forecast errors of GDP caused by the data revision in 1995. Comparing Chart 6A with Chart 11A, one can see that the distortion is completely eliminated when the final data are used to measure the forecast accuracy.

Still, we find that when the 1995 period is excluded the difference between the current-year scores using vintage and final data increases from 3.9 to 7.7. Looking more closely at the source of this difference, we find that it can be mostly attributed to the forecast error for GNP/GDP.

Chart 16 displays the decompositions of forecast errors for GNP/GDP using the final data as realized values. Comparing this chart with Chart 11, one can see that there are some notable differences in the breakdown in the composition for both the current-year and next-year forecasts. In Chart 11A, we see larger overall errors in 1992 and in the 1996 to 2004 period. These larger errors are due to increases in error attributable to the common component of the forecast error. Consequently, a greater proportion of the error each period is due to the common component. The average contribution of the common component to the overall error rises to 73.9 percent from 56.7 percent. In addition, the overall error in 1995 using vintage data (which resulted from the changing to chain-weighted GDP) is no longer present. For the next-year forecasts in Chart 11B, we again see that the overall error has increased but to a considerably more modest degree. The overall forecast error prior to the 1990-91 recession is less using final data but is greater (on aggregate) for the 1996 to 2000 period. But once again, this increase in overall error is attributable to the common component of the forecast error. The average contribution of the common component of the overall error rises to 61.8 percent from 59.0 percent.

Our findings suggest there may make little difference to use final data or vintage data when evaluating forecasts. We have shown that the average Blue Chip Consensus score was modestly affected for current-year forecasts and was almost unchanged for

next-year forecasts. In addition, we have shown that the decrease in score for current-year and next-year forecasts was as a result of an increase in the common component of the forecast error and did not affect the idiosyncratic component. Therefore, the effect of a switch to final data for evaluating individual forecasts scores should be roughly equal across forecasts. And lastly, the use of final data eliminates the need for arbitrarily choosing among different vintages.

VII. Conclusion

In 1994 the FOMC began to release statements after each meeting. The amount of policy information released in the statements has increased and changed over time. The findings from Kohn and Sack (2003) and Ehrmann and Fratzscher (2004) suggest that financial markets are sensitive to the information revealed in these statements. While knowing whether the statements have affected markets is important, understanding whether the statements are providing strong signals concerning the FOMC's views about the future path of the economy or economic policy is also important. That is, has the public's ability to forecast future economic and financial conditions improved since 1994? This question is important because one hopes that transparency, if appropriately communicated, enhances market participants' ability to forecast (Woodford 2005).

This article analyzes the forecast errors across a large section of forecasters and for a set of five key macroeconomic variables. The analysis finds evidence that the individuals' forecasts have been more synchronized since 1994, implying the possible effects of the FOMC's transparency. On the other hand, we find little evidence that the common forecast errors, which are the driving force of overall forecast errors, have become smaller since 1994. In fact, common forecast errors have increased and have

become more volatile on several dimensions. These common errors seem to be associated with business cycles and other economic shocks. Transparent monetary policy may not necessarily enhance the public's predictability of business cycles.

On the other hand, it is possible that we do not have a long-enough sample to observe the effects of transparency because the FOMC just began in August 2003 to provide more explicit guidance on the likely path of future policy and its contingency on future economic conditions. We hope that our findings will generate more research on this important topic.

Box I Characterizing the Distribution of Accuracy Scores

The distribution of accuracy scores can be summarized by the first four moments. We show how to calculate the mean or average score $\hat{p}_t(n)$ in the text. In this appendix we show how to calculate the other three moments: standard deviation, skewnesss, and kurtosis. These three informative measures can be calculated as follows:

$$\hat{\sigma}_t(n) = \left[\frac{1}{N_t} \sum_{i=1}^{N_t} \left(p\left(\hat{\chi}_t^i, n\right) - \hat{p}_t(n) \right)^2 \right]^{\frac{1}{2}},$$

$$\hat{s}_t(n) = \frac{\frac{1}{N_t} \sum_{i=1}^{N_t} \left(p\left(\hat{\chi}_t^i, n\right) - \hat{p}_t(n) \right)^3}{\hat{\sigma}_t(n)^3},$$

$$\hat{u}_t(n) = \frac{\frac{1}{N_t} \sum_{i=1}^{N_t} \left(p\left(\hat{\chi}_t^i, n\right) - \hat{p}_t(n) \right)^4}{\hat{\sigma}_t(n)^4},$$

where σ stands for the standard deviation, s the skewness, and u the kurtosis.

Box II Decomposition of Mean Square Error

Let the estimate of μ_t be

$$\hat{\mu}_t = \frac{1}{N_t}\sum_{i=1}^{N_t} y_t^i .$$

Note that $\hat{\mu}_t$ is also the Blue-Chip consensus forecast. The weighted mean square error at time *t* can be decomposed as

$$\begin{aligned}\frac{1}{N_t}\sum_{i=1}^{N_t} x_t^{i\prime} x_t^i &= \frac{1}{N_t}\sum_{i=1}^{N_t}\left[\left(y_t^i - \hat{\mu}_t\right) - \left(y_t - \hat{\mu}_t\right)\right]'\left[\left(y_t^i - \hat{\mu}_t\right) - \left(y_t - \hat{\mu}_t\right)\right] \\ &= \frac{1}{N_t}\sum_{i=1}^{N_t}\left(y_t^i - \hat{\mu}_t\right)'\left(y_t^i - \hat{\mu}_t\right) + \frac{1}{N_t}\sum_{i=1}^{N_t}\left(y_t - \hat{\mu}_t\right)'\left(y_t - \hat{\mu}_t\right)\end{aligned},$$

where the first term at the right hand side is the MSE attributed to the idiosyncratic component and the second term is the MSE attributed to the common component. The cross term is zero because

$$\frac{1}{N_t}\sum_{i=1}^{N_t}\left(y_t^i - \hat{\mu}_t\right)'\left(y_t - \hat{\mu}_t\right) = \left(\hat{\mu}_t - \hat{\mu}_t\right)'\left(y_t - \hat{\mu}_t\right) = 0 .$$

Appendix I: Data Description

Three-month Treasury bill rate: 1986-2004. Secondary market, monthly average. Source: Board of Governors of the Federal Reserve System.

Consumer price index: 1986-2004. CPI-U (all urban consumers). Source: Bureau of Labor Statistics, U.S. Department of Labor.

Gross national/domestic product: 1986-1995, not chained; 1996-2004, chained. Source: Bureau of Economic Analysis, U.S. Department of Commerce.

Unemployment rate: 1986-2004. All 16 years or older workers. Source: Bureau of Labor Statistics, U.S. Department of Labor.

Corporate bond yield: 1986-1995. Aaa, monthly average. Source: Moody's Investors Service, Inc.

Ten-year Treasury note yield: 1996-2004. Constant maturity, monthly average. Source: Board of Governors of the Federal Reserve System.

Appendix II: Scores and Ranks for Individual Forecasters

In this appendix we report in Table 2 the average scores for all the individual forecasters who have continued to participate in the surveys in recent years. We include also the consensus forecast and the Bayesian vector autoregressive (BVAR) model. The BVAR model is often used in the empirical literature as a benchmark for model comparison (Robertson and Tallman 1999, 2001), and reporting the real-time forecasting performance of this model is of particular interest to academic researchers. We also report other forecasters' scores toward the end of the table for completeness. The years in which each forecaster participated in the Survey are also reported in the table.

References

Andy Bauer, Eisenbeis, Robert A. Daniel F. Waggoner, and Tao Zha, 2003. "Forecast evaluation with cross-sectional data: The Blue Chip Surveys," Federal Reserve Bank of Atlanta *Economic Review* (Q2), 17-31.

Ehrmann, Michael, and Marcel Fratzscher, 2004. "Central Bank Communication: Different Strategies, Same Effectiveness?" *Unpublished Manuscript* (November), the European Central Bank.

Eisenbeis, Robert A. Daniel F. Waggoner, and Tao Zha, 2002. "Evaluating Wall Street Journal Survey Forecasters: A Multivariate Approach," *Business Economics* 37(3), 11-21.

Faust, John and Eric M. Leeper, "Forecasts and Inflation Reports: An Evaluation," manuscript prepared for the Sveriges Riskbank conference *Inflation Targeting: Implementation, Communication and Effectiveness,* June 11-12, 2005.

Kohn, Donald L., and Brian P. Sack, 2003. "Central Bank Talk: Does It Matter and Why?" *Finance and Economics Discussion Series* 2003-55 (November), Board of Governors of the Federal Reserve System.

Robertson, John C. and Ellis W. Tallman, 1999. "Vector Autoregressions: Forecasting and Reality," Federal Reserve Bank of Atlanta *Economic Review* (Q1), 4-18.

Robertson, John C. and Ellis W. Tallman, 2001. "Improving Federal-Funds Rate Forecasts in VAR Models Used for Policy Analysis," *Journal of Business and Economic Statistics* 19(3), July, 324-330.

Sims, Christopher A. and Tao Zha, 2005. "Were There Regime Switches in US Monetary Policy?" *American Economic Review,* forthcoming.

Stock, James H. and Mark W. Watson, 2003. "Has the Business Cycles Changed? Evidence and Explanations," *Monetary Policy and Uncertainty: Adapting to a Changing Economy,* Federal Reserve Bank of Kansas City Symposium, Jackson Hole, Wyoming, August 28-30.

Michael Woodford, 2005. "Central-Bank Communication and Policy Effectiveness," manuscript prepared for the Federal Reserve Bank of Kansas City Conference on *The Greenspan Era: Lessons for the Future,* Jackson Hole, Wyoming, August 25-27, 2005.

Chart 1A

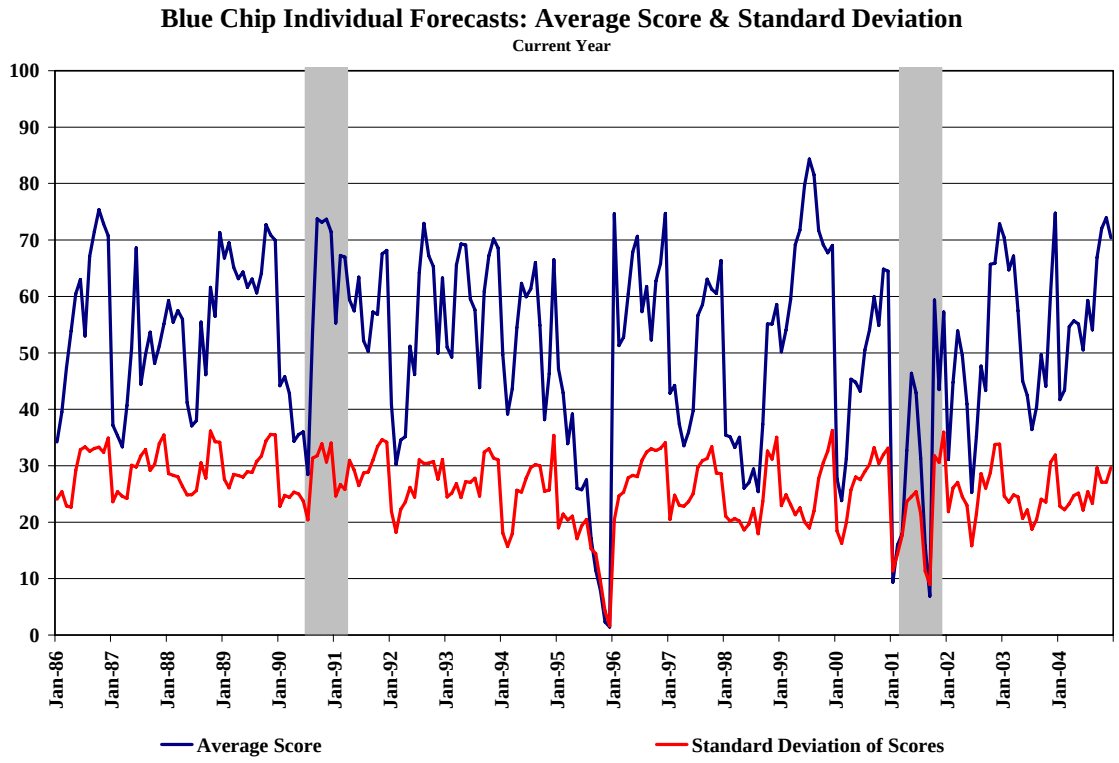

Chart 1B

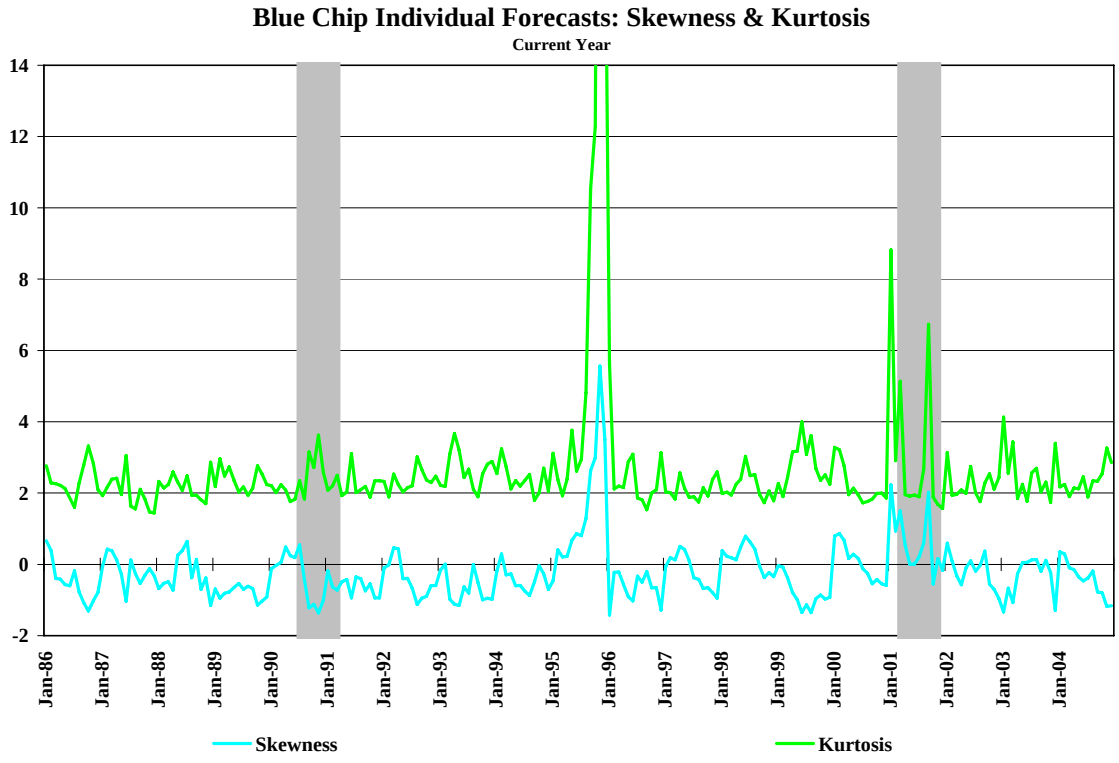

Chart 2A

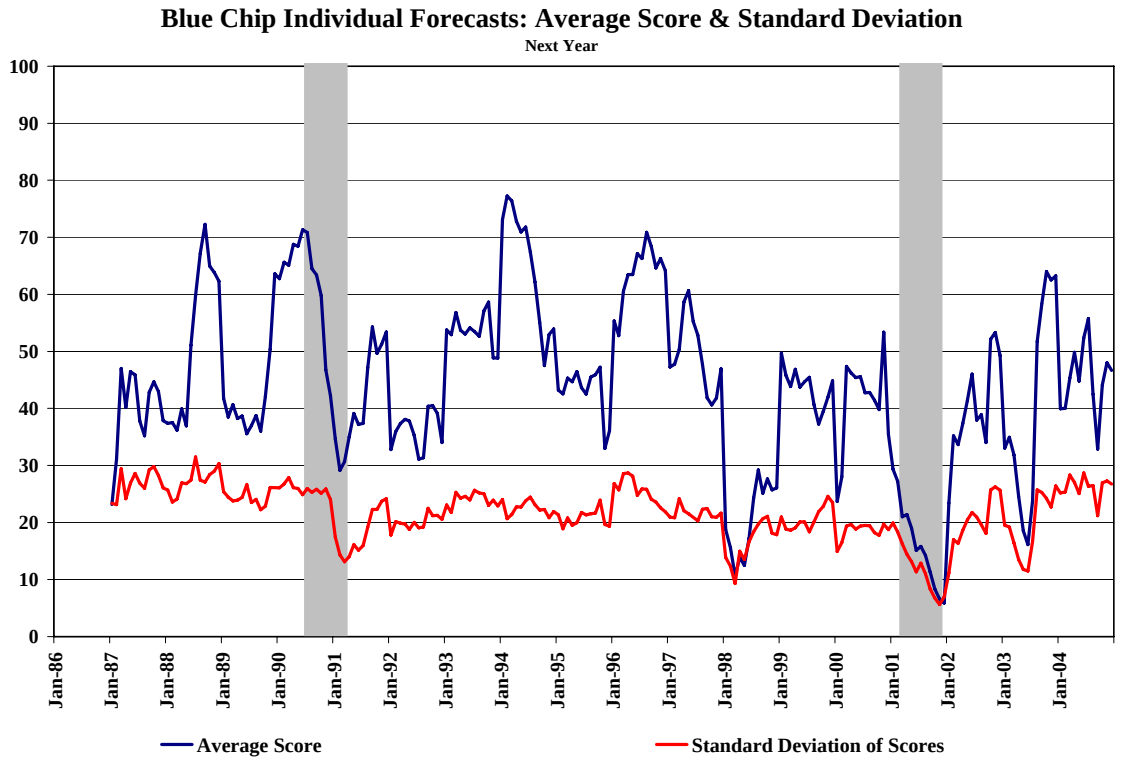

Chart 2B

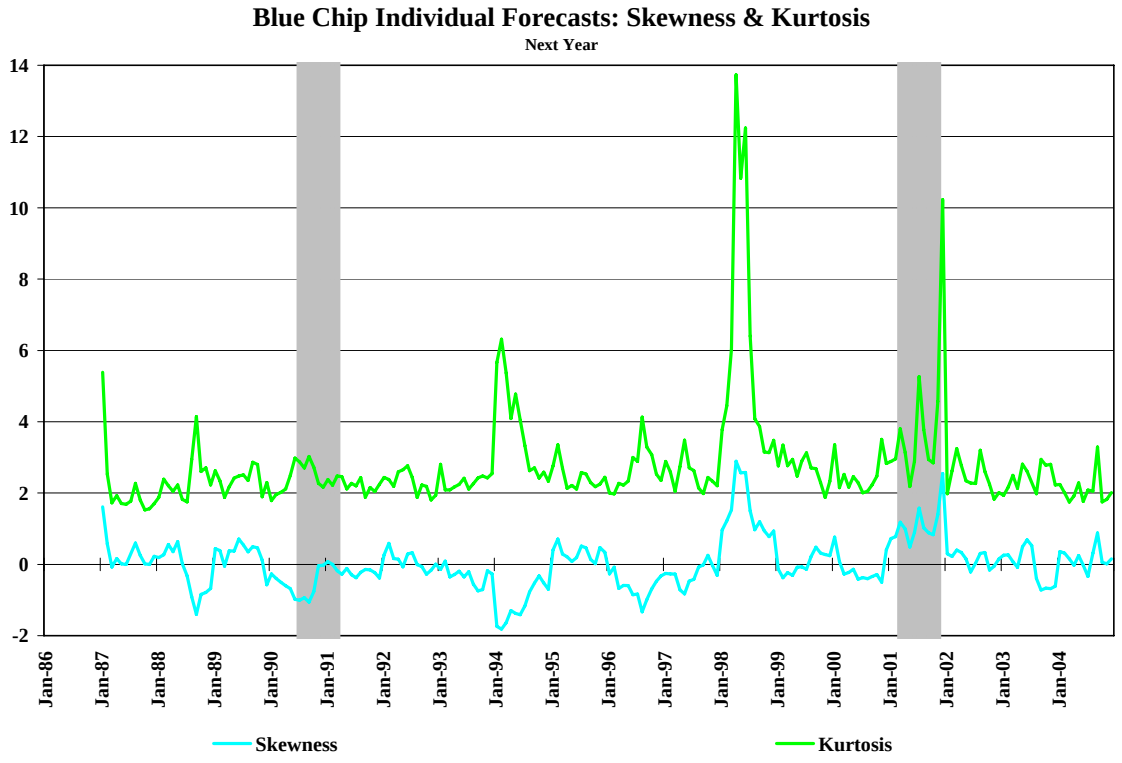

Chart 3A

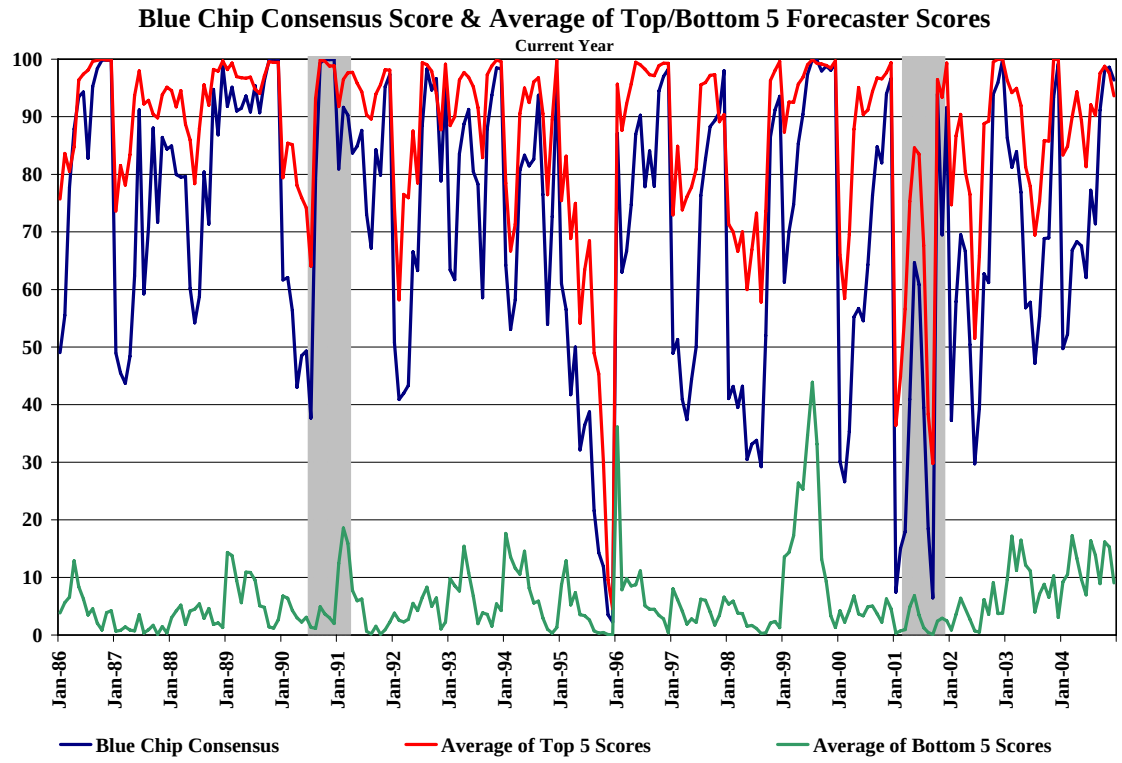

Chart 3B

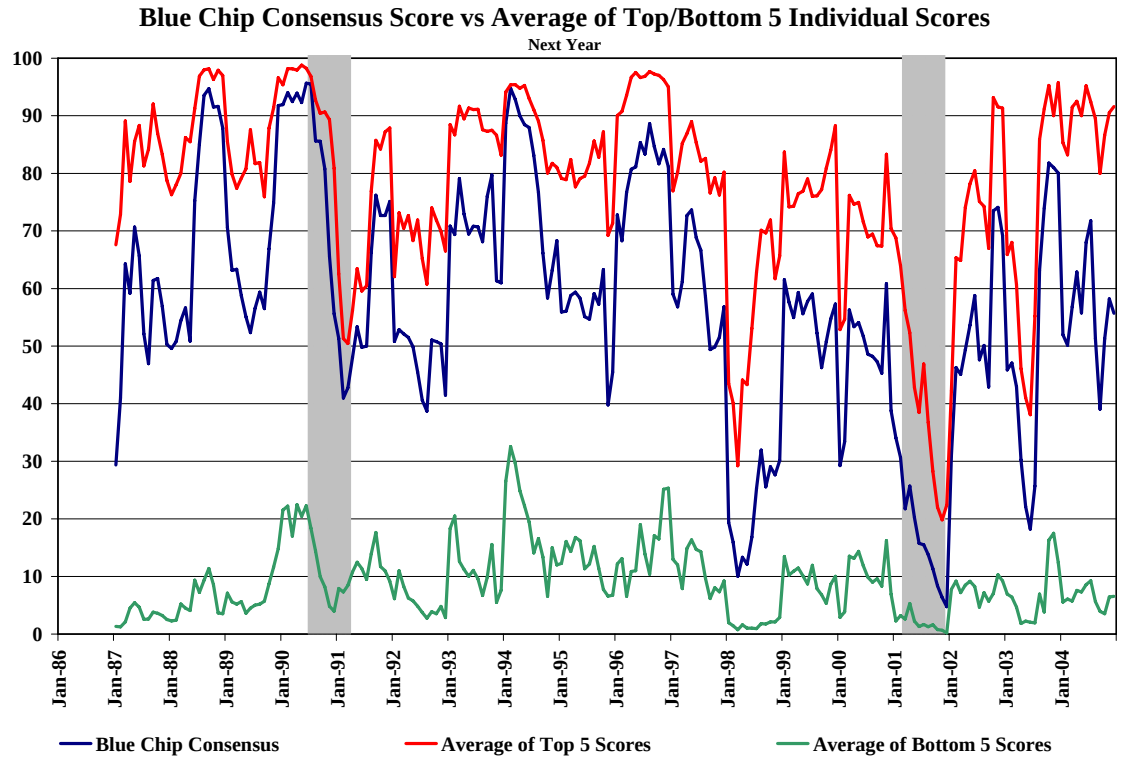

Chart 4A

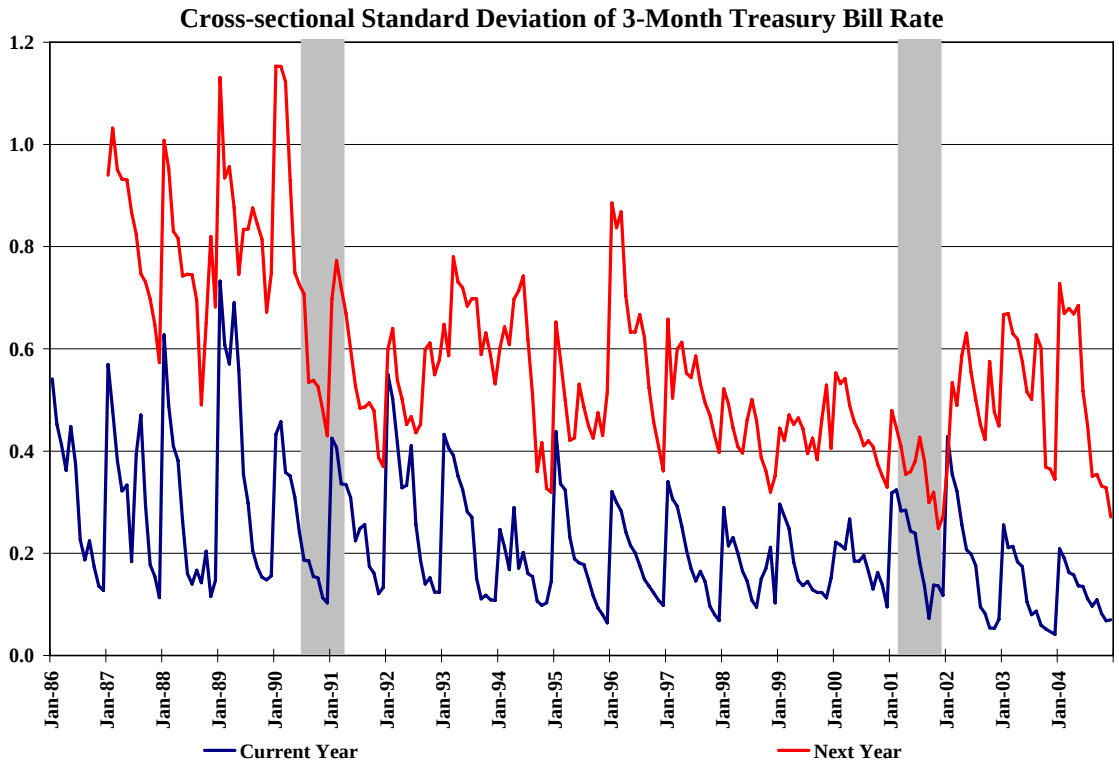

Chart 4B

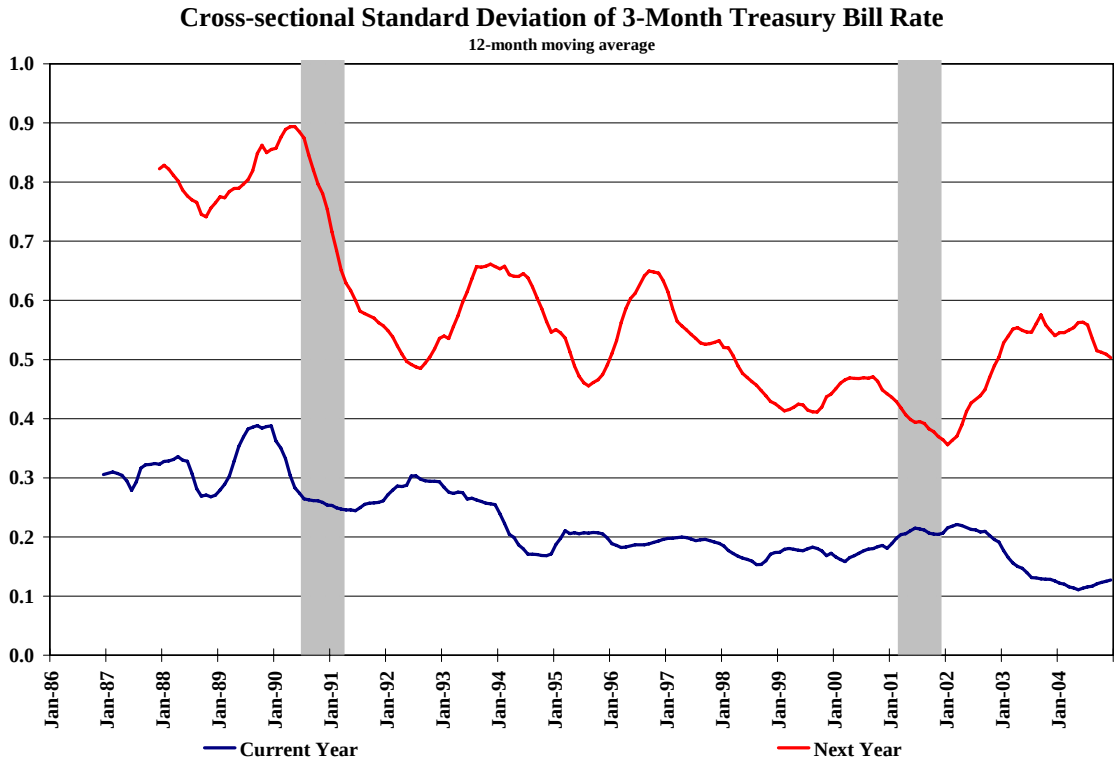

Chart 5A

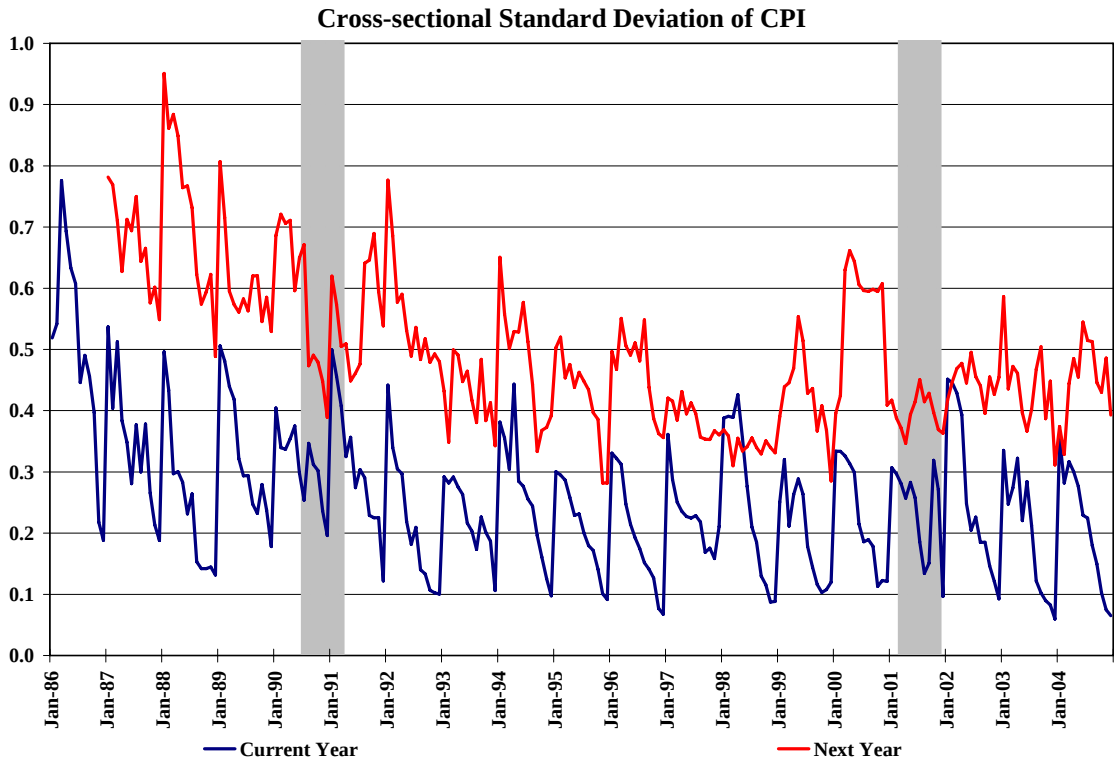

Chart 5B

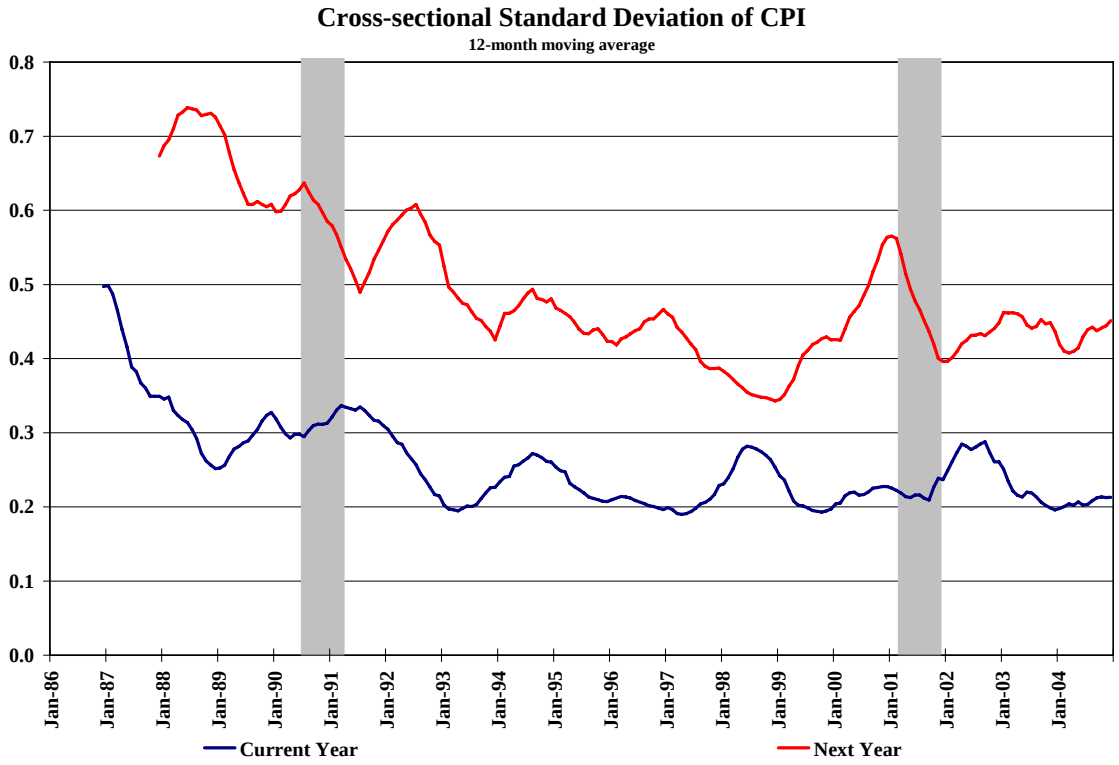

Chart 6A

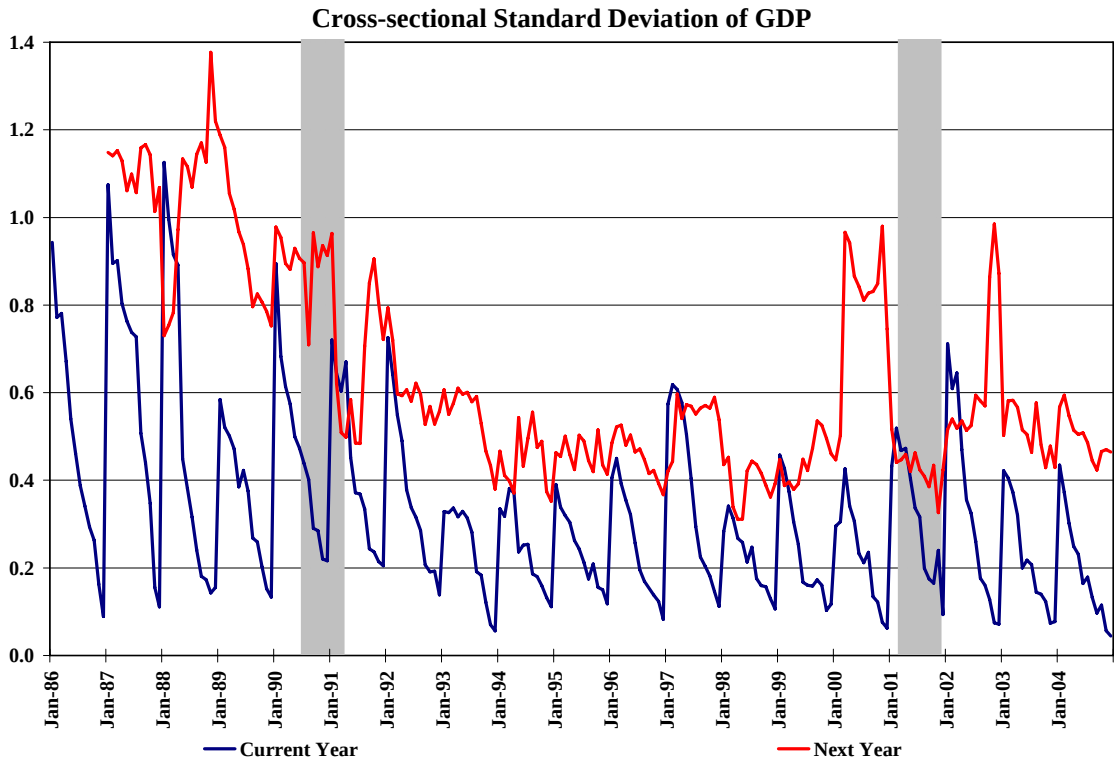

Chart 6B

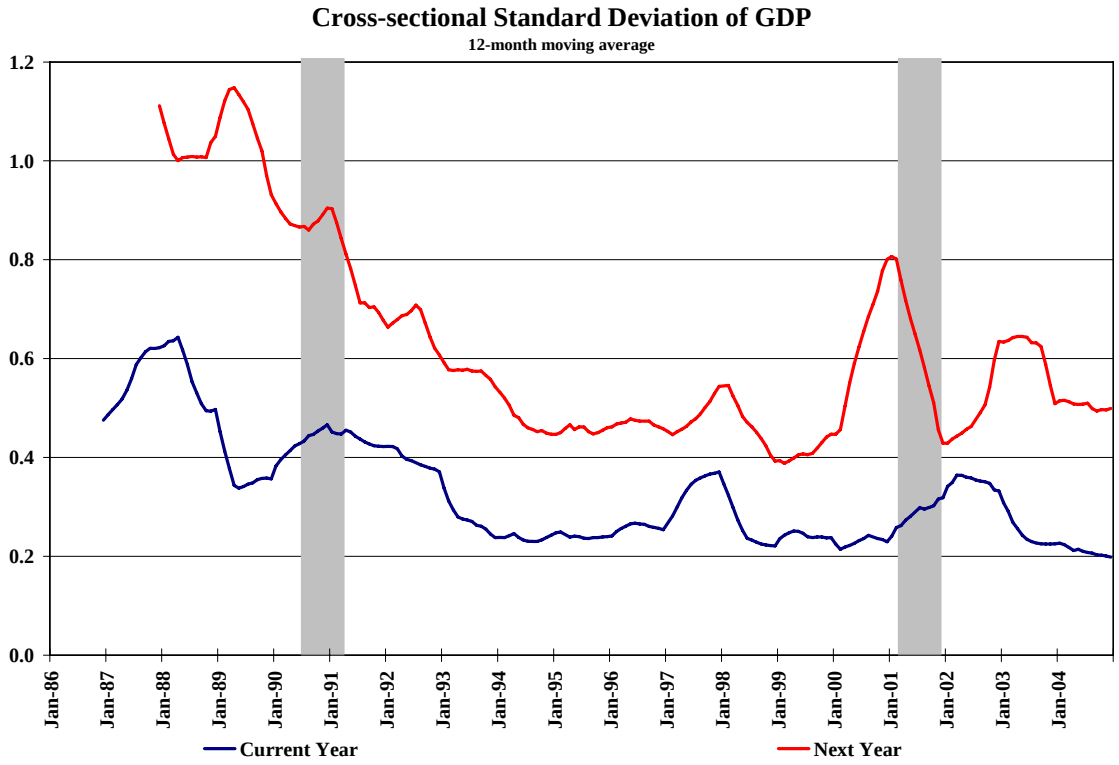

Chart 7A

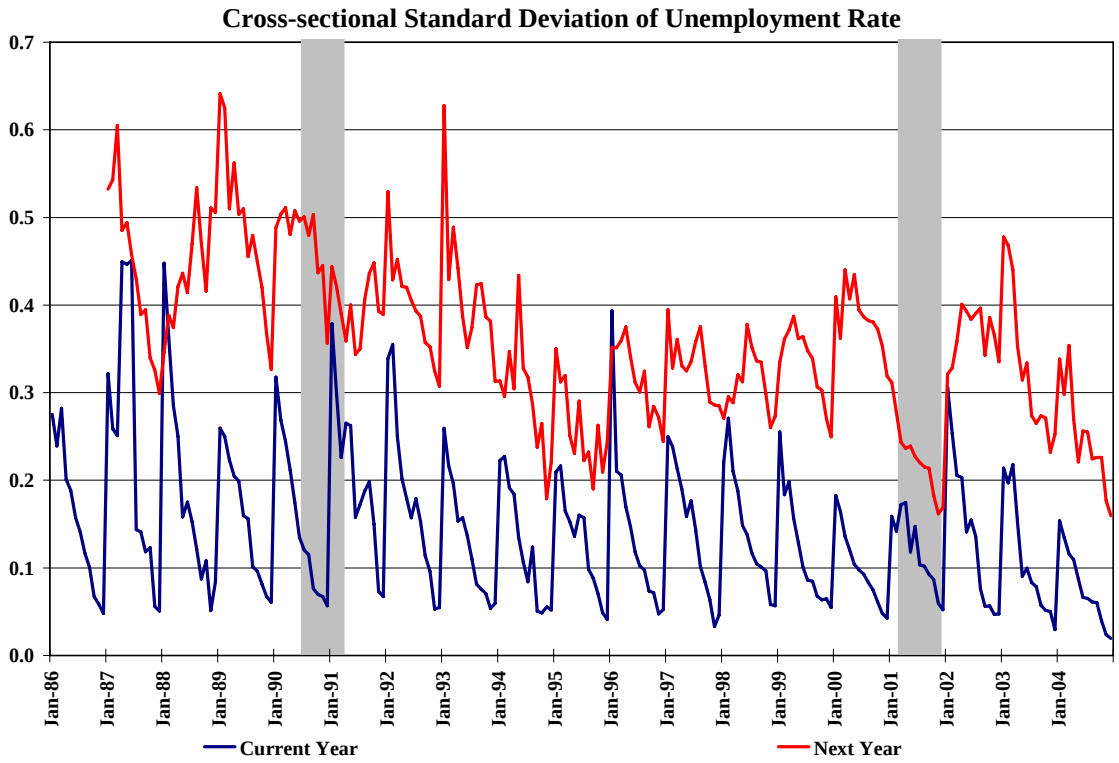

Chart 7B

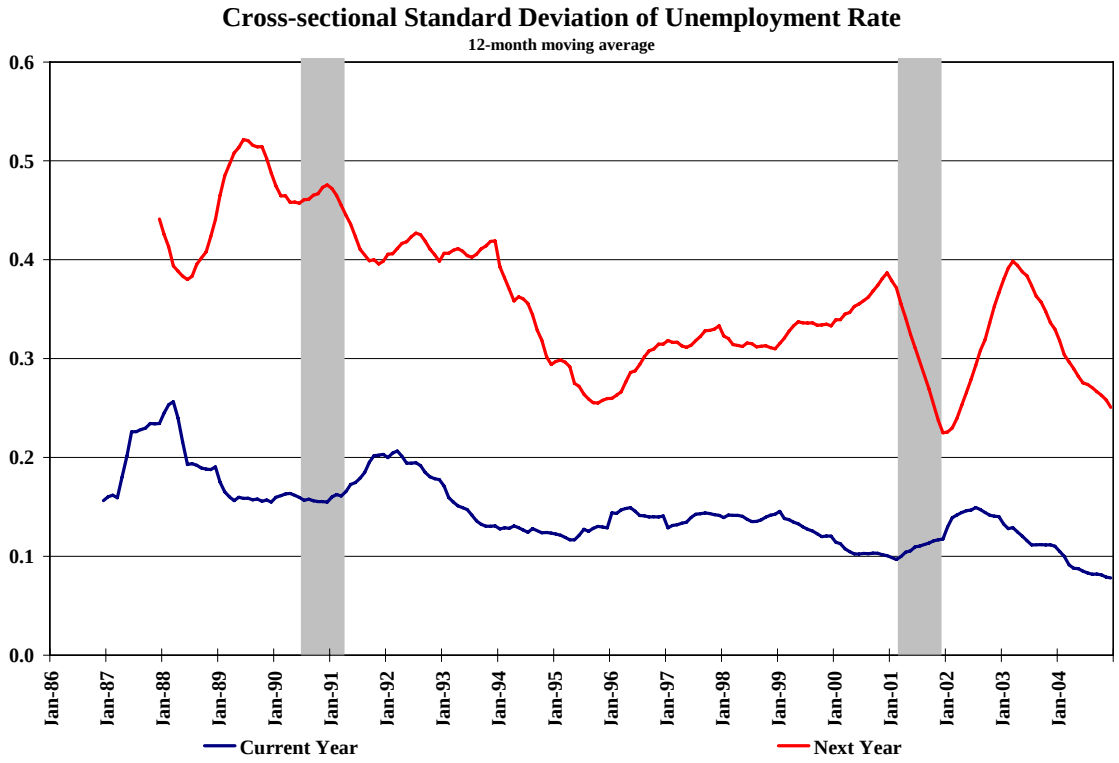

Chart 8A

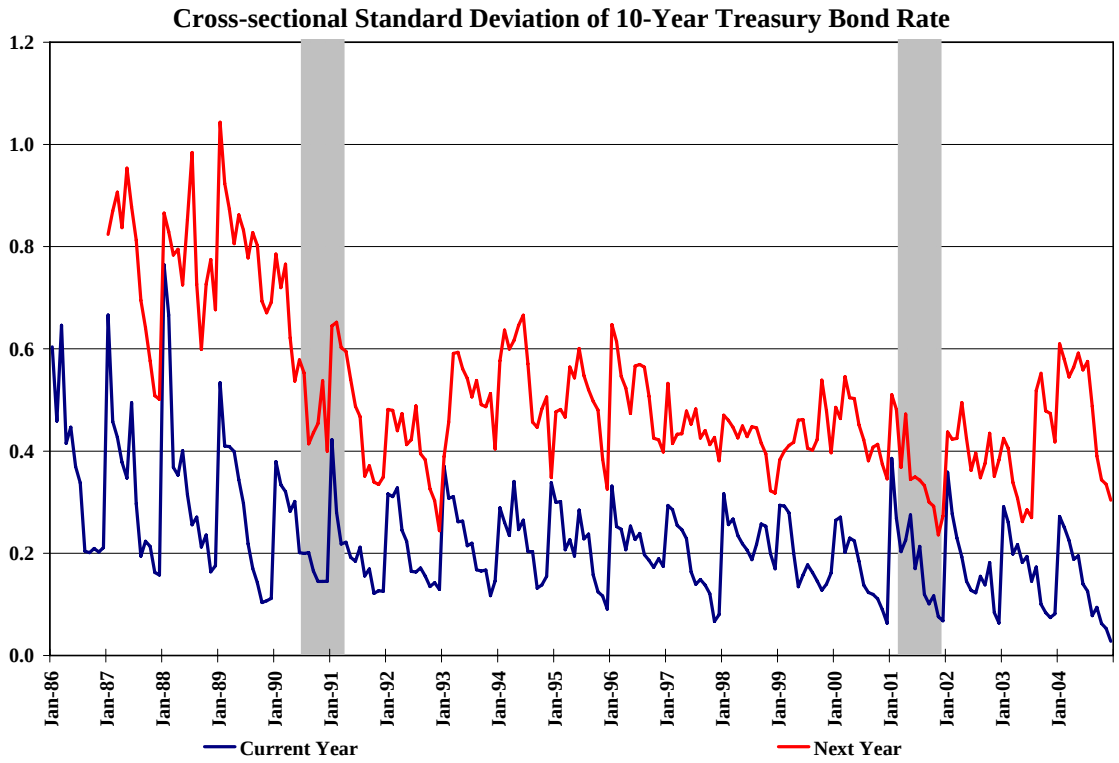

Chart 8B

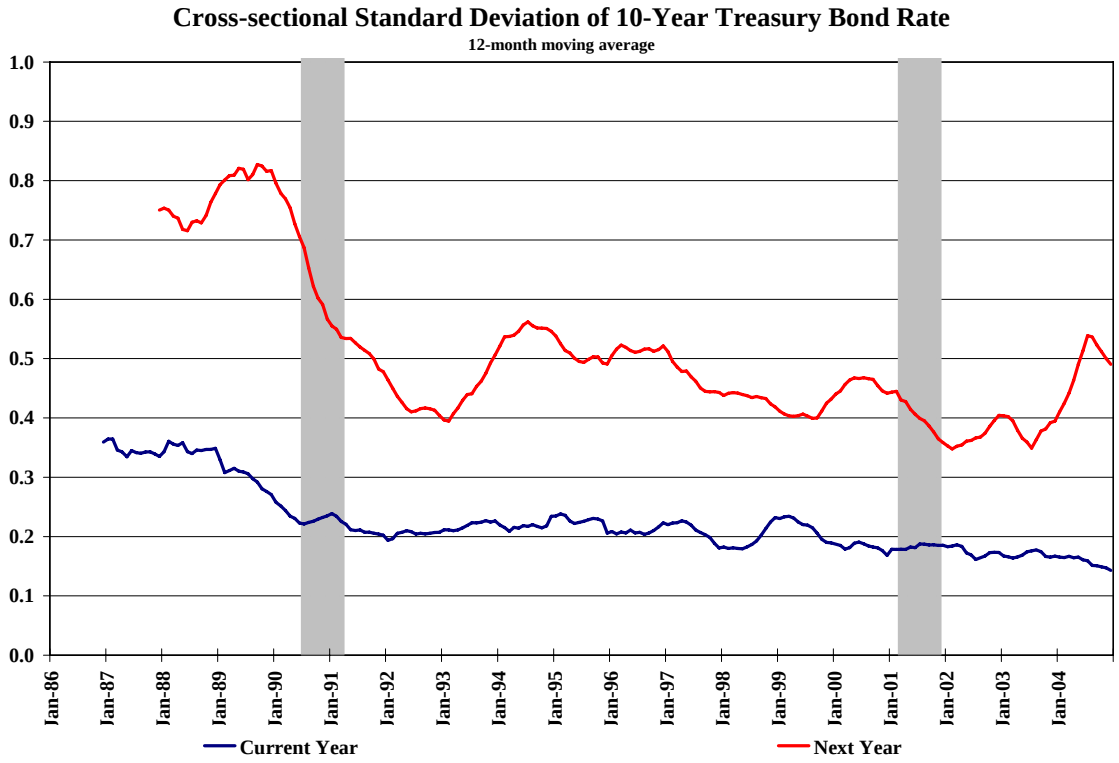

Chart 9A

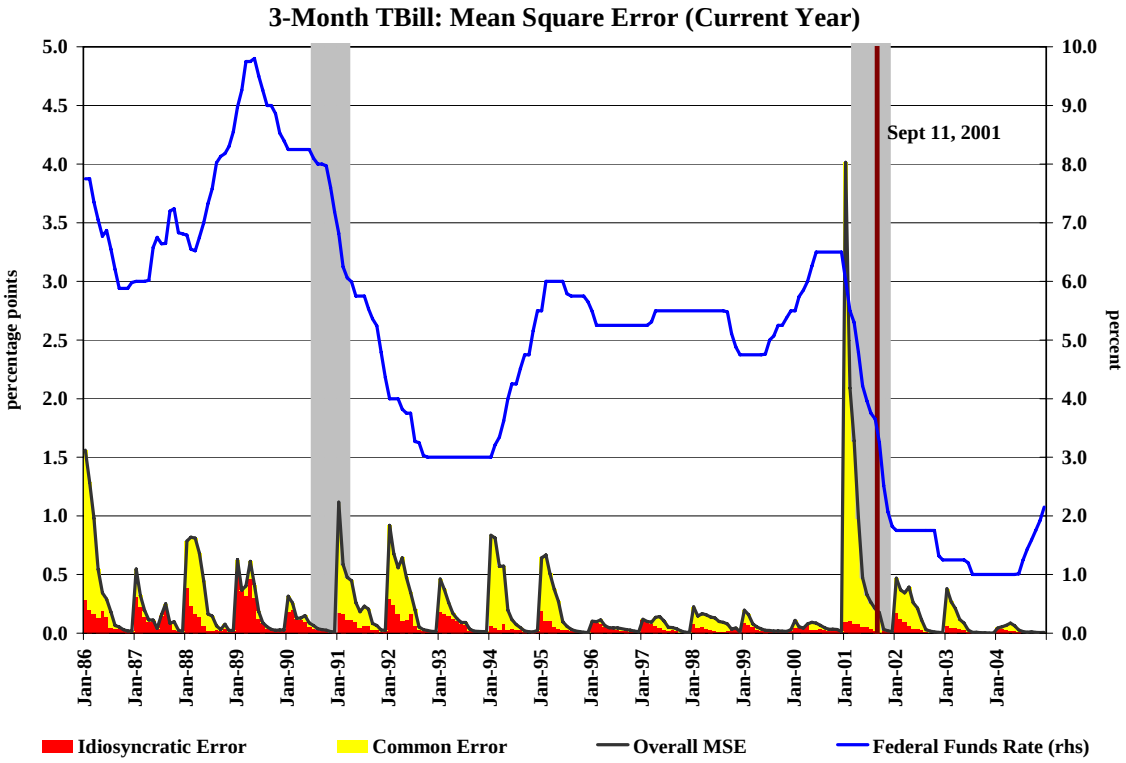

Chart 9B

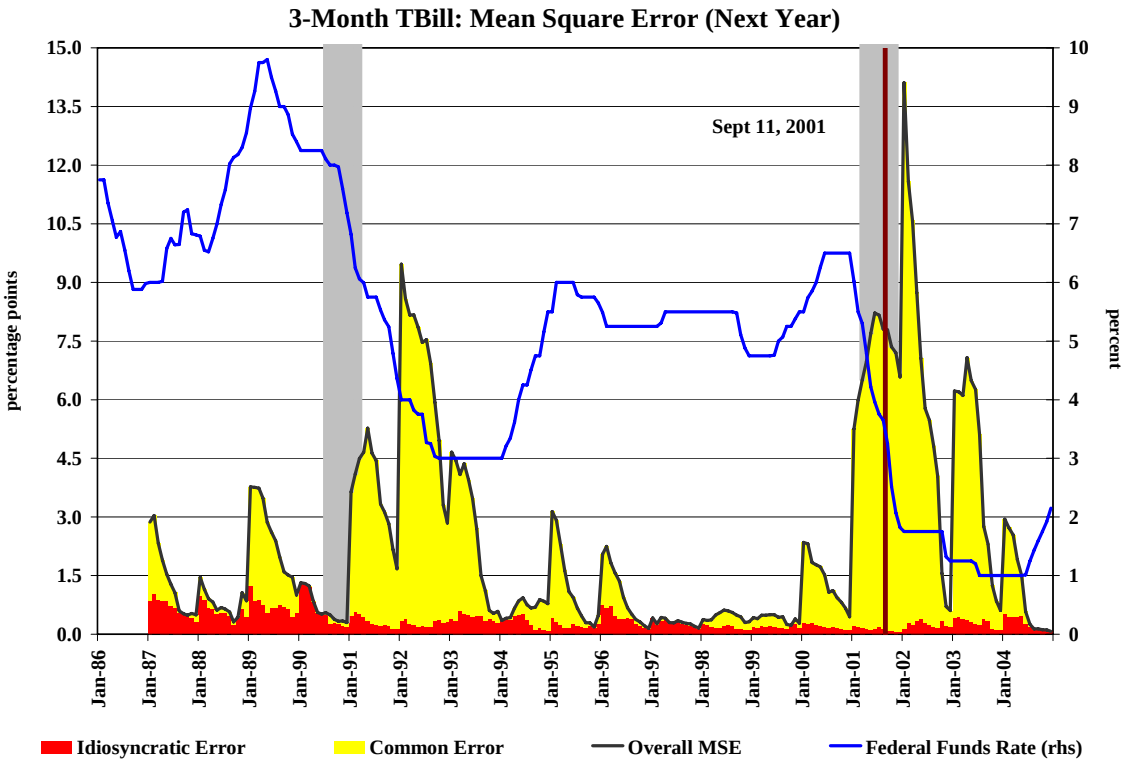

Chart 10A

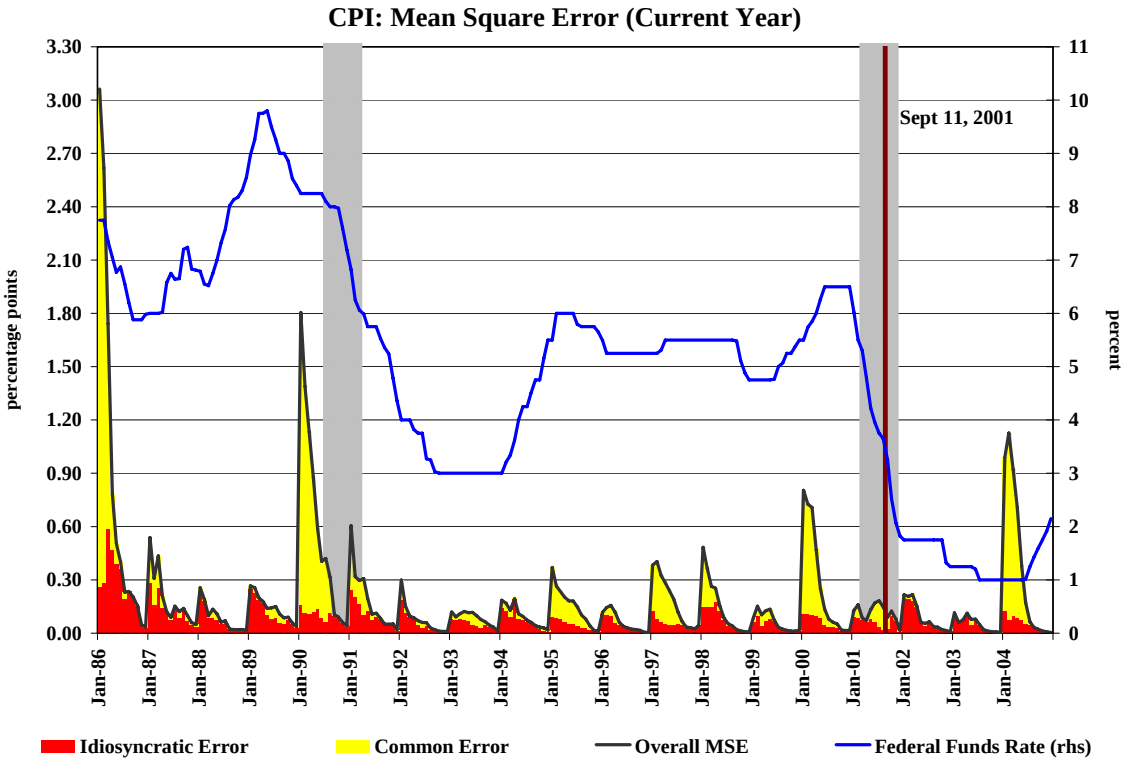

Chart 10B

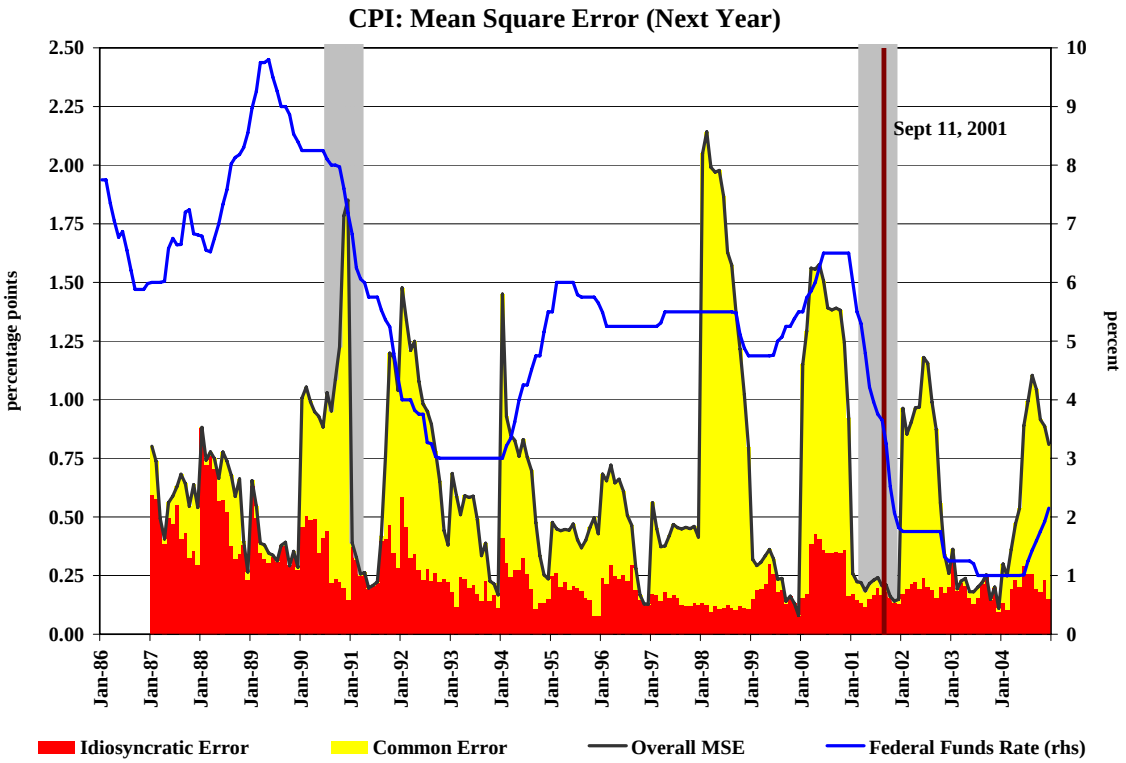

Chart 11A

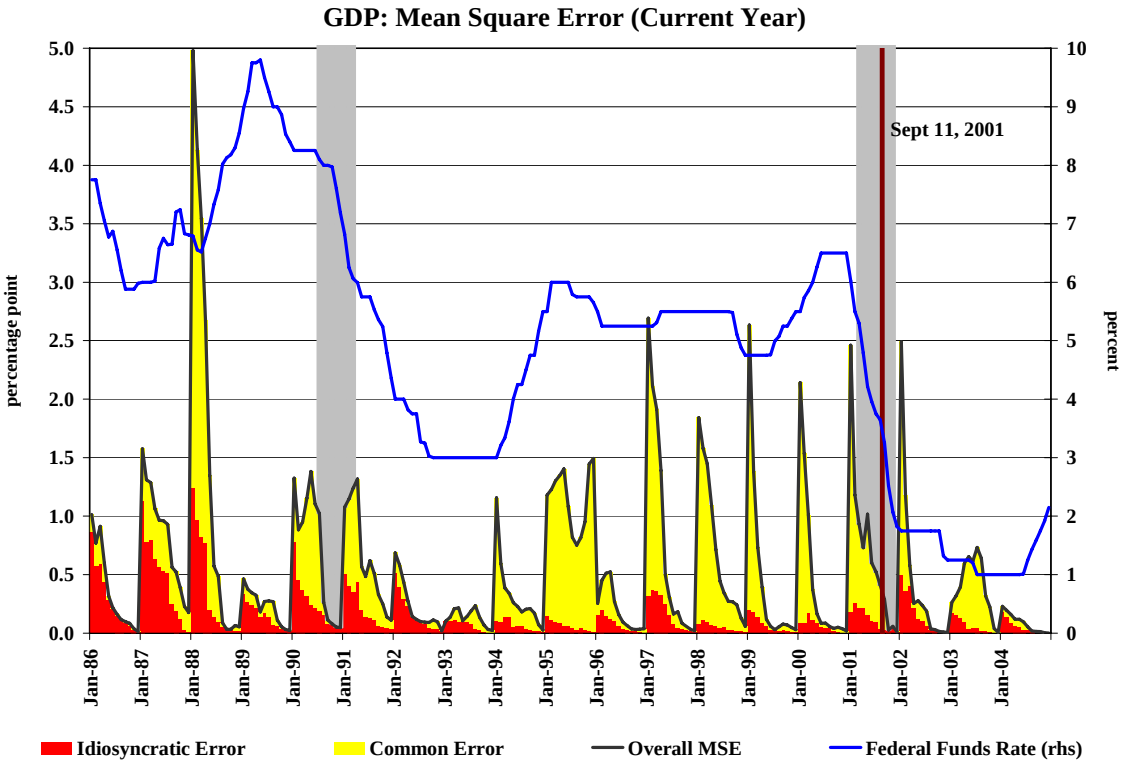

Chart 11B

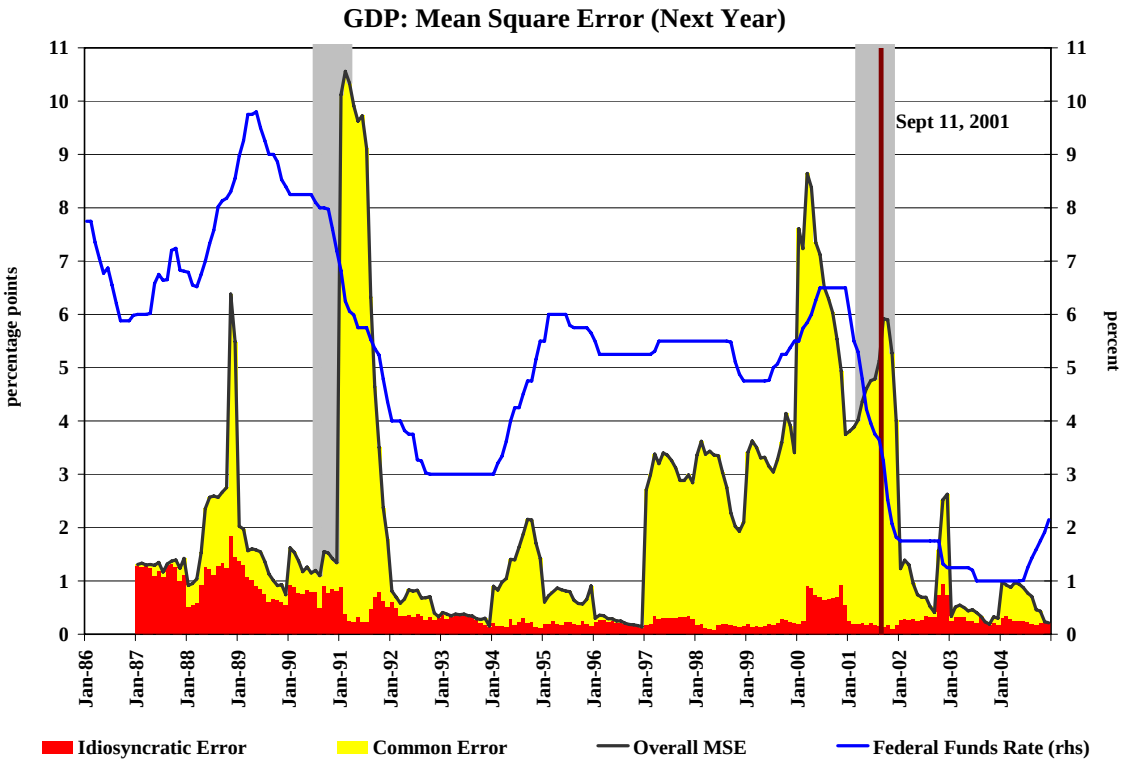

Chart 12A

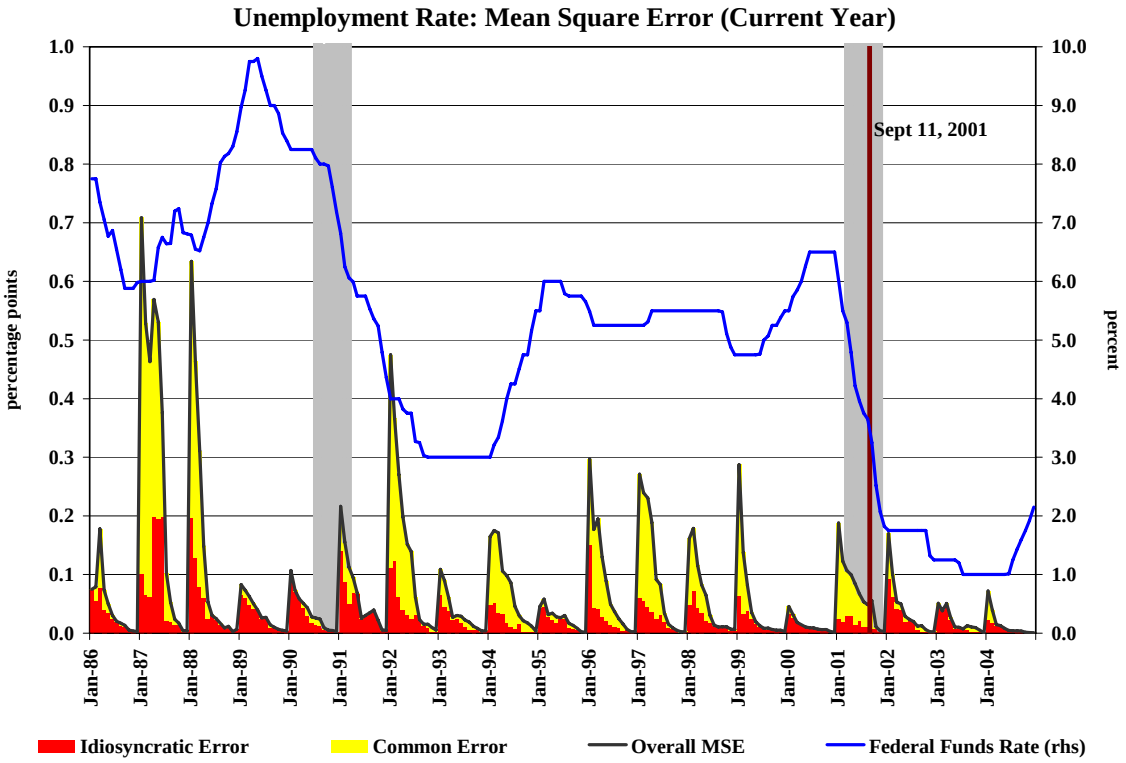

Chart 12B

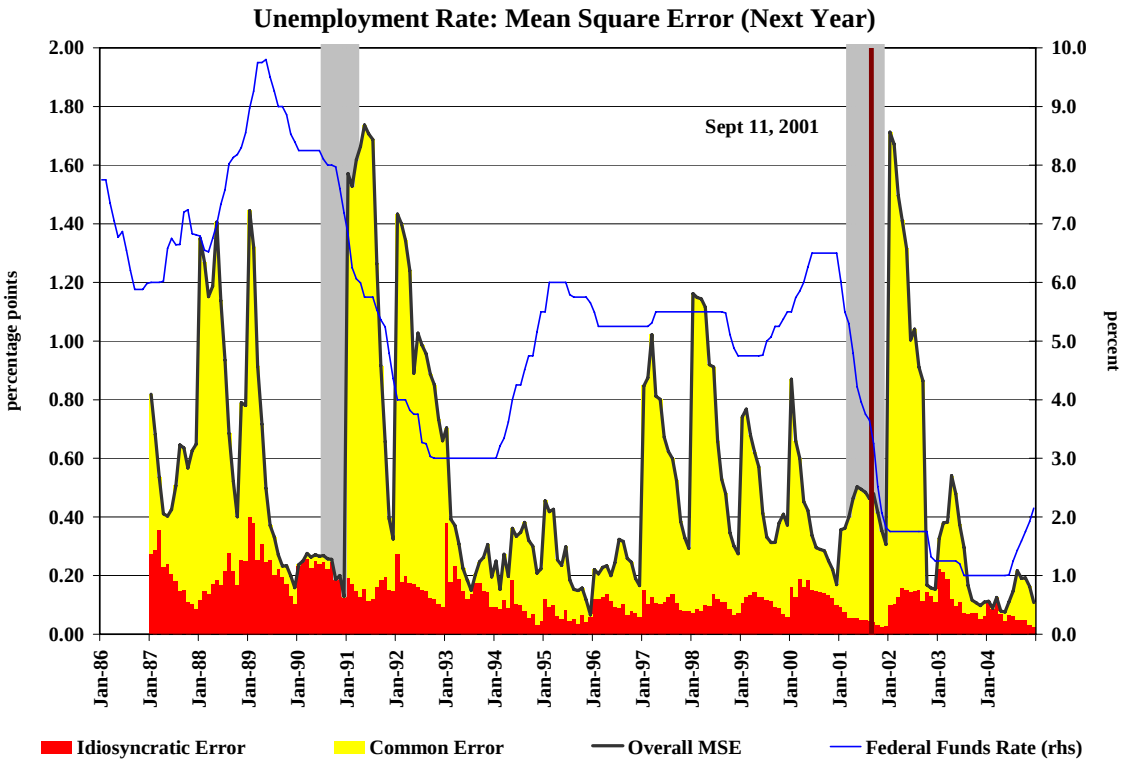

Chart 13A

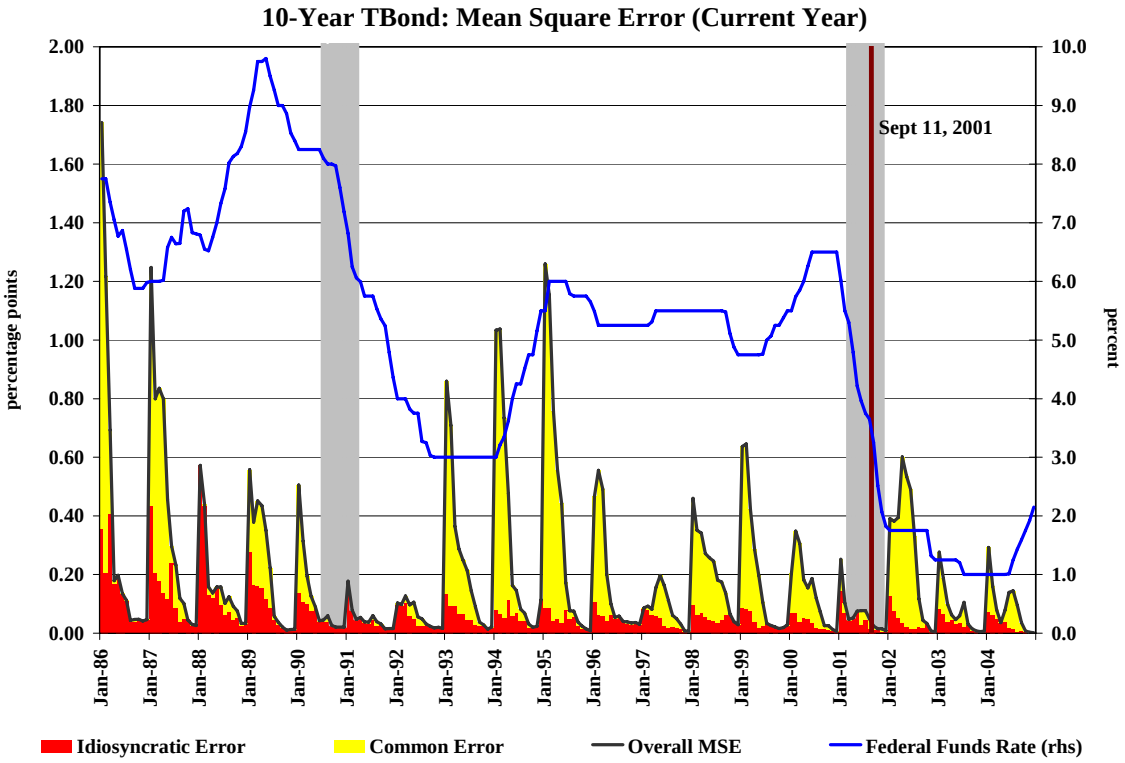

Chart 13B

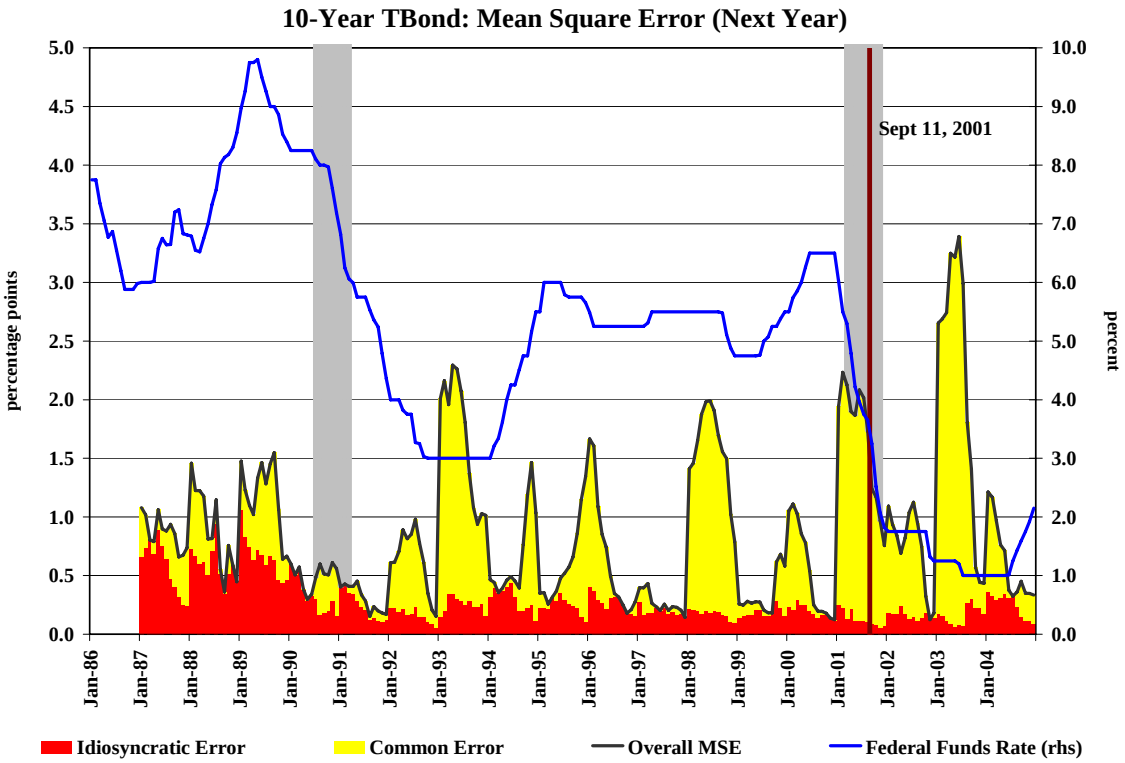

Chart 14A

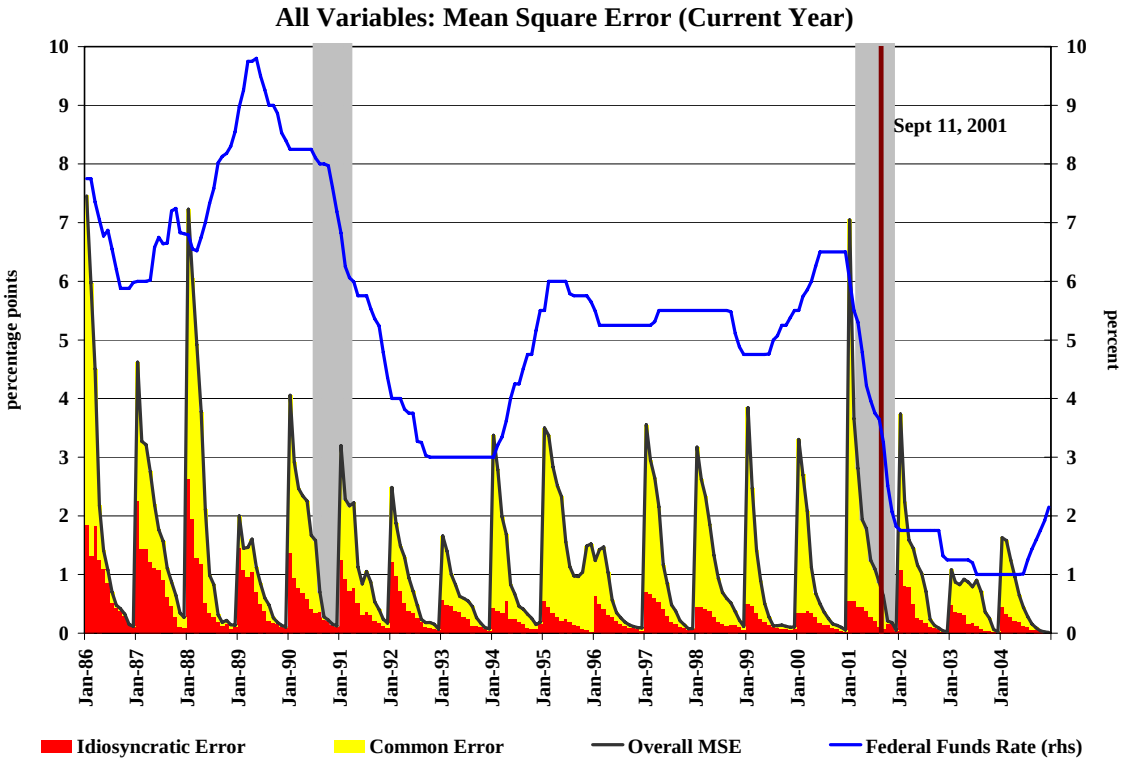

Chart 14B

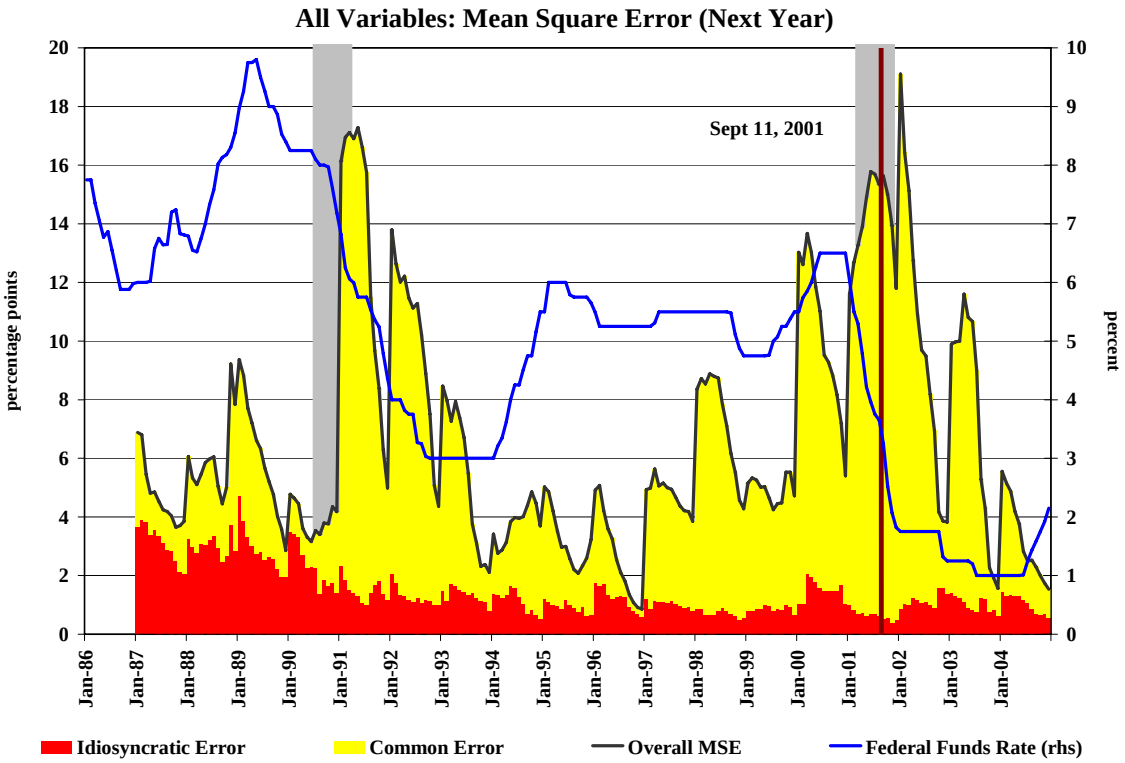

Table 1

Decomposition of Mean Square Error

By Average Percent Contribution to Error in Each Period

Current Year Forecasts (1986 - 2004)	All Variables	3MTB	CPI	GDP	UR	10YTB
Idiosyncratic Component	44.5	57.0	69.7	43.3	64.0	58.7
Common Component	55.5	43.0	30.3	56.7	36.0	41.3

Next Year Forecasts (1986 - 2003)	All Variables	3MTB	CPI	GDP	UR	10YTB
Idiosyncratic Component	30.0	40.0	52.7	41.0	36.6	48.5
Common Component	70.0	60.0	47.3	59.0	63.4	51.5

By Percent Contribution of Total Error Across Sample

Current Year Forecasts (1986 - 2004)	All Variables	3MTB	CPI	GDP	UR	10YTB
Idiosyncratic Component	31.9	30.9	40.6	28.0	39.6	32.0
Common Component	68.1	69.1	59.4	72.0	60.4	68.0

Next Year Forecasts (1986 - 2003)	All Variables	3MTB	CPI	GDP	UR	10YTB
Idiosyncratic Component	22.1	15.1	38.6	20.1	24.7	32.1
Common Component	77.9	84.9	61.4	79.9	75.3	67.9

Chart 15A

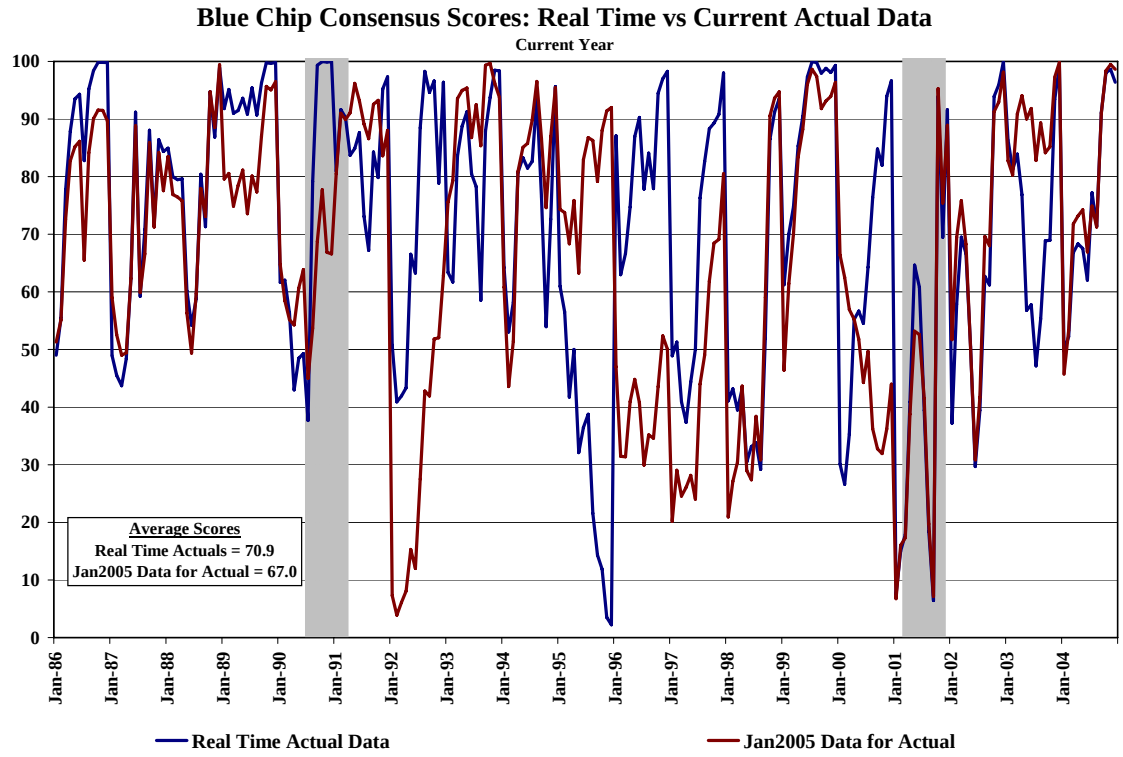

Chart 15B

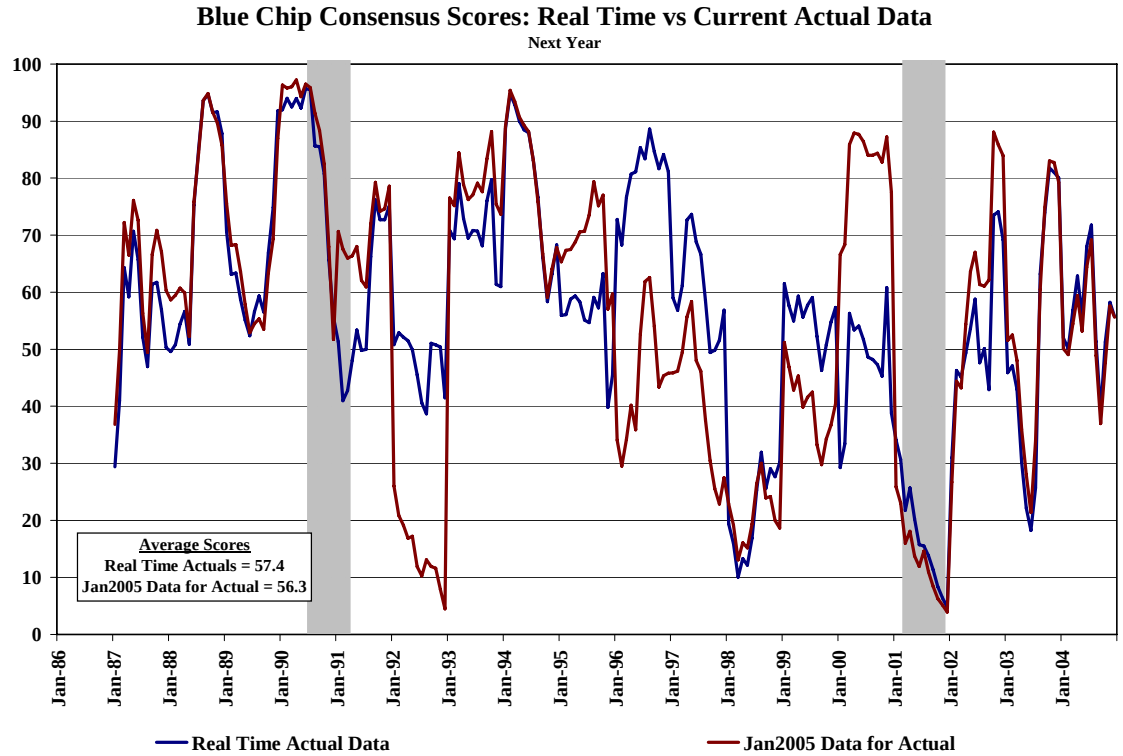

Chart 16A

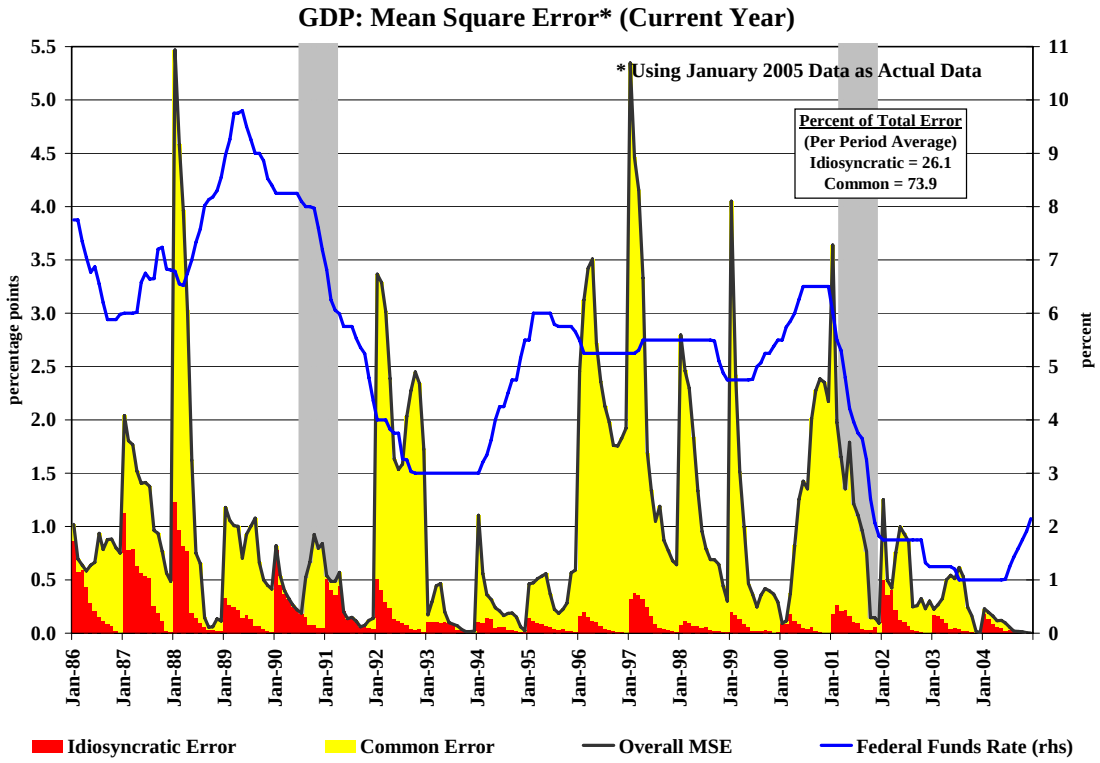

Chart 16B

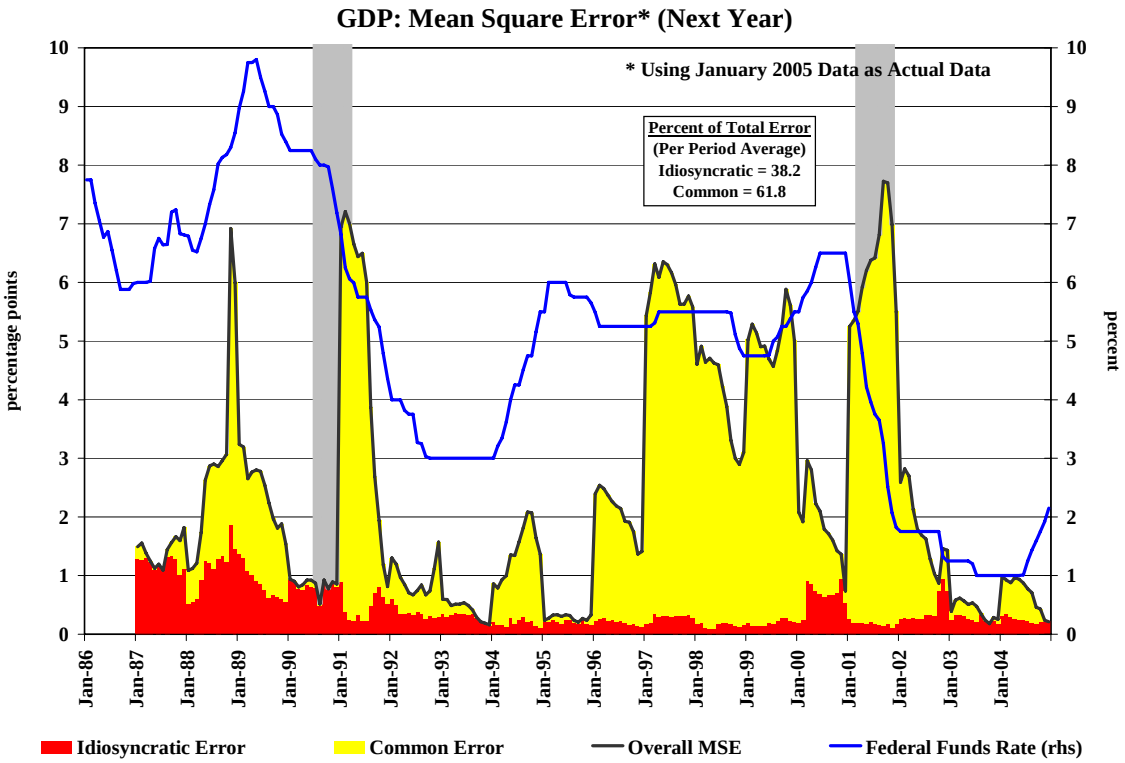

Table 2

Overall Performance: Score								
	Overall		Current Year		Next Year		Participation	
Forecaster Name	Average Score	Standard Deviation	Average Score	Standard Deviation	Average Score	Standard Deviation	CY	NY
BC- Average of Top 10	82.24	16.86	86.45	15.99	77.81	16.66	228	216
BC Consensus	64.36	23.49	70.92	24.07	57.43	20.77	228	216
Macroeconomic Advisers, LLC	62.58	27.71	71.57	26.25	53.10	26.06	227	215
Schwab Washington Research Grou	62.04	28.26	69.97	27.11	53.64	27.07	197	186
Atlanta BVAR	59.69	31.19	69.21	29.54	49.64	29.75	228	216
U.S. Trust Co.	59.25	27.15	64.61	26.25	49.96	26.25	227	131
ClearView Economics	59.23	28.94	66.69	27.72	50.10	27.99	66	54
Banc of America Corp.	59.22	27.10	63.28	27.82	54.87	25.68	204	190
Northern Trust Company	58.75	28.01	63.34	27.27	53.17	27.95	222	183
Wayne Hummer & Co.	55.89	27.27	58.05	27.61	53.58	26.78	228	214
Moody's Investors Service	55.04	28.03	65.77	28.63	42.35	21.34	78	66
Perna Associates	54.61	26.31	60.90	28.35	47.82	22.08	167	155
Merrill Lynch	54.50	27.41	58.36	28.97	50.32	25.02	206	190
Wells Capital Management	53.58	28.71	59.83	28.19	46.83	27.81	161	149
National Assn. of Home Builders	53.56	26.06	58.77	26.69	47.93	24.21	176	163
Nomura Securities	52.55	28.87	55.77	29.82	48.57	27.41	63	51
Nat. City Bank of Cleveland	52.01	26.08	56.75	26.56	46.93	24.61	224	209
DuPont	51.68	25.60	57.06	28.14	46.00	21.23	228	216
Georgia State	51.67	27.39	51.72	28.64	51.62	26.07	223	211
Fannie Mae	51.43	28.00	59.67	29.13	41.81	23.35	84	72
DaimlerChrysler AG	51.34	29.13	58.94	29.24	43.35	26.85	226	215
Standard & Poors	51.25	30.43	58.86	30.09	42.78	28.63	120	108
Eggert Economic Enterprises	50.79	25.90	50.12	27.56	51.48	24.09	225	215
Siff, Oakley, Marks, Inc.	50.66	28.19	56.56	27.41	44.77	27.78	197	197
Evans, Carrol and Associates	50.43	29.77	58.01	30.35	42.86	27.21	202	202
Bank One	49.82	31.39	56.87	31.75	42.21	29.22	205	190
Bear Steams & Co., Inc.	49.67	29.96	53.11	30.39	43.95	28.59	98	59
BC- Ave of Individual Scores	48.13	16.08	51.84	16.22	44.21	15.00	228	216
La Salle National Bank	47.47	29.73	54.13	32.16	40.22	24.97	158	145
Prudential Securities	47.07	31.41	47.40	33.06	46.57	28.88	175	117
Prudential Financial	47.01	26.68	50.54	28.97	43.31	23.55	201	192
Goldman Sachs & Co.	46.28	27.19	59.47	25.49	30.49	19.85	79	66
National Assn. of Realtors	46.10	29.24	51.08	29.08	40.10	28.56	64	53
Conference Board	45.08	29.38	52.22	31.03	37.46	25.46	224	210
Chamber of Commerce, USA	44.97	27.68	48.35	28.34	41.20	26.50	214	192
General Motors Corporation	44.30	28.03	46.05	29.42	42.42	26.40	162	150
Econoclast	43.29	27.13	42.32	30.94	44.32	22.44	227	215
Eaton Corporation	43.04	28.51	40.92	30.07	45.37	26.62	127	115
Turning Points (Micrometrics)	43.04	27.86	41.15	29.28	45.04	26.19	185	174
Comerica	42.41	25.44	43.88	29.34	40.84	20.41	178	166
UCLA Business Forecast	42.12	30.19	45.32	32.23	38.75	27.55	227	215
Motorola, Inc.	42.02	28.76	50.83	31.73	31.91	20.92	102	89
J P Morgan Chase	40.92	27.12	47.57	29.12	33.40	22.55	104	92
Kellner Economic Advisers	40.79	23.00	41.86	24.65	39.55	21.02	91	79
Genetski.com	40.46	32.50	50.61	32.88	29.53	28.37	154	143
Wachovia Securities	40.39	27.19	44.60	31.05	35.69	21.33	98	88
Federal Express Corp.	39.92	26.15	41.80	28.76	37.63	22.60	65	53
DRI-WEFA	39.02	27.09	48.32	28.48	27.99	20.65	77	65
Morgan Stanley & Co.	35.95	29.39	38.27	31.92	32.30	24.75	85	54
Inforum-U. of Md.	35.72	26.46	33.15	27.15	38.46	25.47	222	208
Deutsche Banc Alex Brown	30.71	28.22	31.86	26.76	29.08	30.33	91	64
Naroff Economic Advisors	29.96	28.67	33.36	32.65	25.86	22.59	70	58
Ford Motor Company	25.80	25.12	27.32	26.09	23.69	23.73	103	74
BC- Average of Bottom 10	7.50	6.35	6.12	6.32	8.96	6.05	228	216