Research Program on Forecasting

Are 'unbiased' forecasts really unbiased? Another look at the Fed forecasts[1]

Tara M. Sinclair
Department of Economics
George Washington University
Washington DC 20052
tsinc@gwu.edu

Fred Joutz
Department of Economics
George Washington University
Washington DC 20052
bmark@gwu.edu

H. O. Stekler
Department of Economics
George Washington University
Washington DC 20052
hstekler@gwu.edu

RPF Working Paper No. 2008-010
http://www.gwu.edu/~forcpgm/2008-010.pdf

August 19, 2008

RESEARCH PROGRAM ON FORECASTING
Center of Economic Research
Department of Economics
The George Washington University
Washington, DC 20052
http://www.gwu.edu/~forcpgm

[1] We would like to thank Peter Tinsley for helpful discussion and Lauren Taylor for excellent research assistance. M. Al Sadek's undergraduate thesis on this subject provided the stimulus for this paper.

Research Program on Forecasting (RPF) Working Papers represent preliminary work circulated for comment and discussion. Please contact the author(s) before citing this paper in any publications. The views expressed in RPF Working Papers are solely those of the author(s) and do not necessarily represent the views of RPF or George Washington University.

Are 'unbiased' forecasts really unbiased?
Another look at the Fed forecasts

Abstract

This paper reconciles contradictory findings obtained from forecast evaluations: the existence of systematic errors and the failure to reject rationality in the presence of such errors. Systematic errors in one economic state may offset the opposite types of errors in the other state such that the null of rationality is not rejected. A modified test applied to the Fed forecasts shows that the forecasts were ex post biased.

Are 'unbiased' forecasts really unbiased?
Another look at the Fed forecasts

Monetary policy decisions are based on forecasts of future economic activity. It is for this reason that economists have evaluated the forecasts made by the staff of the Board of Governors of the Federal Reserve System in preparation for the Open Market Committee meetings. The forecasts are quarterly predictions made several times each quarter with horizons of one to eight periods. The staff predicts GDP, its components, various price indices, unemployment, etc.

Previous evaluations have primarily focused on the predictions of real and nominal GDP growth rates, inflation rates and the unemployment rate. The studies of Karamouzis and Lombra (1989), Scotese (1994), and Jansen and Kishan (1996, 1999) yielded mixed results about the rationality of these forecasts. Romer and Romer (2000) and Sims (2002) found that the Fed's forecasts did not reject rationality and that they were as good as the predictions made by the private sector. On the other hand, while Joutz and Stekler (2000) found that the forecasts were unbiased, there was evidence of inefficiency, and, more importantly, the Fed staff made systematic errors.[1]

This finding that the forecasts displayed systematic errors, such as reported by Joutz and Stekler (2000), has also been observed in the results of other forecast evaluations (See Zarnowitz and Braun, 1992; Fildes and Stekler, 2002). Forecasters overestimated the rate of growth during slowdowns and recessions and underestimated it during recoveries and booms. Similarly, inflation was underpredicted when it was rising and overpredicted when it was declining.

[1] In addition, Clements et al. (2007) used a non-traditional approach to test for bias and found that the forecasts were biased, whereas the traditional approach had not rejected the null of no bias.

Particularly large errors occurred during the periods when prices were rising rapidly during the 1970s and early 1980s. Hanson and Whitehorn (2006) also observed these systematic errors but associated them with particular time periods rather than with stages of the business cycle.

Despite these systematic errors, the null of unbiasedness or rationality is frequently not rejected when the forecasts are tested. In particular, Joutz and Stekler (2000, pp. 30-31) note that, even though there were systematic errors, the null of unbiasedness was not rejected for any of the sets of forecasts of nominal or real GNP (GDP) or of the deflator made by the Federal Reserve. This result held for all forecast horizons. These two findings are mutually inconsistent and create a puzzle that we investigate.

The inconsistency is that forecasts that are considered unbiased should not display systematic errors. Since the systematic errors are observations that are clearly present and observable, the inconsistency must arise from the statistical methodology that is used to test for unbiasedness.

This paper will, therefore, reexamine the Fed forecasts to determine whether they are biased or unbiased and whether or not they display systematic errors. We will also attempt to reconcile any discrepancies in the results that we discover. In the next section, we present the data. This is followed by an analysis of the errors that are observed when applying the standard statistical methodology. Given the inconsistency in the results, we modify the statistical procedure and show that the Fed forecasts are ex post biased.

I. Data

We examine the Fed Greenbook forecasts that are made in preparation for the FOMC meetings. We analyze forecasts made between 1965IV and 2001IV for the current quarter and

one quarter ahead for three variables: the real output growth rate (GNP from 1965IV to 1991QIII and GDP from 1991IV on), the GNP/GDP deflator inflation rate, and the unemployment rate. During the entire sample period, there is not a continuous series of forecasts for the longer horizons. Whenever there are multiple forecasts in each quarter, we use the last one.[2] The actual data are the NIPA estimates that are released approximately 45 days after the quarter to which they refer.[3] All data, with the exception of the unemployment rate, are converted into annualized growth rates.

II. Characteristics of the Forecast Errors

Figures (1a-1d,.., 3a-3d) present data for each variable: the annualized growth rate of real GNP/GDP, the annualized inflation rate as measured by the GNP/GDP deflator, and the unemployment rate. The upper left hand graphs display the historical data used in the analysis. The other quadrants of these figures display, for each series: the current quarter and one quarter ahead forecast errors and the forecast revisions between the current and one quarter ahead forecasts.

There were six recessions in this period. The shaded areas of Figures 1 and 3 represent the dates of these recessions, as defined by the NBER. In Figure 2 the shaded areas represent periods when the inflation rate was increasing (moving from trough to peak). The dates for the periods of the inflation "cycle" were obtained from Dr. Anirvan Banerji of the Economic Cycle

[2] We use the last forecast for each quarter because those are the ones for which the Fed had the maximum amount of information on which to base their current and next quarter forecasts. If forecasts were made within the first 10 days of the next quarter, they are considered made in the previous quarter because there would be no new information for the subsequent quarter. We also analyzed forecasts made in the middle of the quarter. The results were similar.

[3] Use of the real time data avoids definitional and classification changes and is the most consistently available data set for our sample. The terminology for these data releases varied over the sample: Before 1974, the "final" data were released 45 days after the end of the quarter. Starting in 1974, releases at approximately the same time became known as "first revision" data (with a second revision about 75 days after the quarter). Since 1988, the "preliminary" data are released approximately two months after the quarter. We obtained similar results using the 90 day releases.

Research Institute (ECRI).

The forecast errors are defined as actual minus forecast with a positive error meaning the variable was underpredicted. In order to explore the role of the state of the economy, the errors for the entire sample were divided into periods of expansion and recession (based on the NBER business cycle dates) or periods when the rate of inflation was increasing or decreasing (based on the ECRI inflation cycle dates). The results (Table 1) show that the mean errors differed between the two states of the economy. Both the current and one quarter ahead forecasts underpredicted the real rate of growth in expansionary periods and overpredicted it during recessions. Consistent with those errors, the unemployment rate and the inflation rate were overestimated in periods of expansion and conversely during recessions.

For the current quarter there was no significant difference between the mean errors of the two periods. For the one quarter ahead forecasts, the null that there was no significant difference in these mean errors was rejected for two of the three variables based on the NBER dates. There was no significant difference in the mean inflation errors based on the NBER dates, but there was a significant difference when we used the ECRI inflation cycle dates (Table 1b), where inflation was underestimated when it was increasing and conversely when it was decreasing. These errors indicate that the Fed forecasts exhibit a systematic bias that is a function of the different states of the economy.

III. Bias Tests and Results

A. Methodology

The most frequently used test is the Mincer-Zarnowitz **(**1969**)** regression (1).[4]

$$A_t = \beta_0 + \beta_1 F_{t,t-i} + e_{t,t-i}; \quad i = 0,1 \tag{1}$$

where A_t and $F_{t,t-i}$ are the actual and predicted values for time *t*. The forecast is conditional on the information available at time *t-i*. When $i = 0$, it refers to a current quarter forecast. The null hypothesis is that $\beta_0 = 0 \; and \; \beta_1 = 1$. A rejection of this hypothesis indicates that the forecasts are biased and/or inefficient. The Wald test and the F distribution are used to test this null.

In applying (1) to the forecasts that have a one quarter lead, an econometric correction is made. In that case, the error term of the equation theoretically follows a first order moving average. The Newey-West procedure was used to estimate HAC consistent standard errors in (1) for the one quarter ahead forecasts.

C. Results

At the 5% significance level, the null for the current quarter forecasts is only rejected for inflation (Table 2a). When applied to the one quarter ahead forecasts, the null is rejected for unemployment but not for the other variables (Table 2b). In this case, we cannot reject the hypothesis that the forecasts of real GDP growth and of inflation are unbiased.

This poses the puzzle that we seek to reconcile: observed systematic biases based on Table 1, but statistical tests do not reject the null of rationality. There is an explanation for these conflicting findings. If, over time, the systematic under- and overestimates in the forecasts offset

[4] Similar results with respect to the role of the state of the economy are obtained using the bias test of Holden and Peel (1990).

each other, there might not be observable bias in a series which combined these periods of offsetting systematic errors. This idea is related to work by Swanson and van Dijk (2006) who show that the state of the economy affected the accuracy of the early releases of industrial production and producer price index data.

IV. Modified Procedure

The results suggest that the biases depended upon the state of the economy. We, therefore, divided our data sets accordingly, and used (2) to test whether or not the forecasts were ex post biased.

$$A_t = \beta_0 + \beta_1 F_{t,t-i} + \beta_2 D_t + e_{t,t-i}; \quad i = 0,1 \tag{2}$$

where in the real GDP growth and unemployment rate regressions D_t is a dummy that reflects the state of the economy where it takes on the value 1 if the economy was in a recession as dated by the NBER and is zero otherwise. In the inflation equation, the dummy takes on the value 1 from the trough to the peak of the inflation cycle as dated by Economic Cycle Research Institute (ECRI), zero otherwise. The other variables are defined as before. The joint null hypothesis is that $\beta_0 = 0$, $\beta_1 = 1$, *and* $\beta_2 = 0$. If the coefficients associated with the dummies are not zero, the dummies contain information that can explain the errors, indicating an ex post bias.

The results and probability values of the F statistic are reported in Tables 3a and 3b. The coefficients of the dummy variables are not significant in any of the current quarter equations (Table 3a). This result is consistent with our failure to find any significant systematic differences in the current quarter forecasts related to economic conditions (Table 1). The results are, however, different in the equations explaining the one quarter ahead predictions (Table 3b). The

dummy is significant in all three equations with the probability values all less than 5% (using the ECRI dummy for inflation).

The results in Table 3b are consistent with our expectations. In the real GDP growth equation, the NBER dummy coefficient is negative, suggesting that the Fed overestimated real GDP during recessions. In addition, the constant in that equation is positive and significant, showing that the Fed forecasts underestimated real GDP during expansions. In the inflation equation, the ECRI dummy coefficient is positive, indicating that the Fed underestimated inflation when it was increasing (moving from trough to peak in the inflation cycle). Although the constant in that equation is insignificant, it is negative. This suggests that the Fed overestimated inflation when it was decreasing (moving from peak to trough in the inflation cycle). We also obtain the expected results in the unemployment equation: the NBER dummy coefficient is positive.

We undertook this analysis to reconcile two findings in previous studies that evaluated forecasts. These studies had found that there were systematic errors but the conventional tests showed that the forecasts were unbiased. By including dummy variables that referred to alternative states of the economy in those tests, we found that there was information that (ex post) explained the forecast errors. This indicates there is no inconsistency in our forecast evaluation analyses: the forecasts are clearly ex post biased and exhibit systematic errors related to the state of the economy.

V. An Interpretation

We have observed that none of the current quarter predictions display ex post bias, but that all of the one quarter ahead predictions do. We, therefore, examine the revisions in the

forecasts for the current quarter to try to explain why the bias was eliminated in the current quarter estimates. The current quarter forecasts in our sample are the last ones that were made in each quarter. Consequently, at the end of the quarter the staff would have virtually complete knowledge about the state of the economy in period t-1 and considerable information about the current quarter.

We, therefore, determine how knowledge about the state of the economy affected the revisions of the forecasts. The revision in the forecast, $F_{t,t} - F_{t,t-1}$, is regressed on the state of the economy to determine whether this knowledge affected the revision and affected the predictions of real growth (inflation) in period t. We will consider both knowledge of the current quarter (D_t) and knowledge of the previous quarter (D_{t-1}).

$$F_{t,t} - F_{t,t-1} = \delta_0 + \delta_1 D_t + \upsilon_t \tag{3a}$$

$$F_{t,t} - F_{t,t-1} = \theta_0 + \theta_1 D_{t-1} + \eta_t \tag{3b}$$

Table 4 shows that the coefficients associated with the state of the economy are all significant and have the expected signs. Knowing that the economy was in a recession either in the previous or in the current quarter,[5] the GDP forecasts for the current quarter are revised downwards and the unemployment prediction is revised upwards. When inflation was increasing in the previous or current quarter, the forecast for the current quarter is revised upwards. The evidence indicates that the Fed forecasters had information about the state of the economy and used the information correctly because these revisions are in the right direction. Thus there is no evidence of ex ante bias.

[5] Although the NBER may not have yet officially called the recession, the forecasters have access to the same data as the members of the NBER business cycle dating committee have, so our results suggest that they know the state of the economy in the current quarter, but that they cannot predict it one quarter ahead.

VI. Implications and Conclusions

We have shown that it is possible to reconcile two contradictory findings: the existence of systematic errors and the failure to reject rationality in the presence of such errors. The data indicate that the systematic errors in one state of the economy offset the opposite types of errors in the other state. When a Mincer-Zarnowitz type equation was used, the null was not rejected. This equation was then modified by introducing a dummy variable that reflected the state of the economy. The application of this modified test to the forecasts of the Federal Reserve showed that conditional on the state of the economy the forecasts were ex post biased.

Our results indicate that in testing forecasts for bias and/or rationality, one can also check whether there are systematic offsetting errors that affect the results. Swanson and van Dijk (2006) already have explored this problem in the context of data revisions. We recommend that studies that investigate the rationality of forecasts always consider this possibility.

References

Fildes, R., Stekler, H., 2002. The state of macroeconomic forecasting. Journal of Macroeconomics 24, 435-468.

Holden K., Peel, D. A., 1990. On testing for unbiasedness and efficiency of forecasts. Manchester School, 48, 120–127.

Jansen, D., Kishan, R., 1996. An evaluation of Federal Reserve forecasting. Journal of Macroeconomics 18, 89-100.

Jansen, D., Kishan, R., 1999. An evaluation of Federal Reserve forecasting: reply. Journal of Macroeconomics 21, 189-203.

Joutz, F., Stekler, H., 2000. An evaluation of the predictions of the Federal Reserve. International Journal of Forecasting 16, 17-38.

Karamouzis, N., Lombra, N., 1989. Federal Reserve policymaking: an overview and analysis of the policy process. Carnegie Rochester Conference Series on Public Policy 30, 7-62.

Romer, C. D., Romer, D. H., 2000. Federal Reserve private information and the behavior of interest rates. American Economic Review, 90, 429-457.

Scotese, C., 1994. Forecast smoothing and the optimal underutilization of information at the Federal Reserve. Journal of Macroeconomics 16, 653-670.

Sims, C.A., 2002. The role of models and probabilities in the monetary policy process. Brookings Papers on Economic Activity: 2, 653-670.

Swanson, Norman R. and Dick van Dijk, 2006, "Are Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry", Journal of Business and Economic Statistics, 24, 24-42.

Zarnowitz,V., Braun P., 1992. Twenty two years of the NBER-ASA Quarterly Outlook Surveys: aspects and comparison of forecasting performance. NBER Working Paper 3965.

Figure 1.a 1-.d

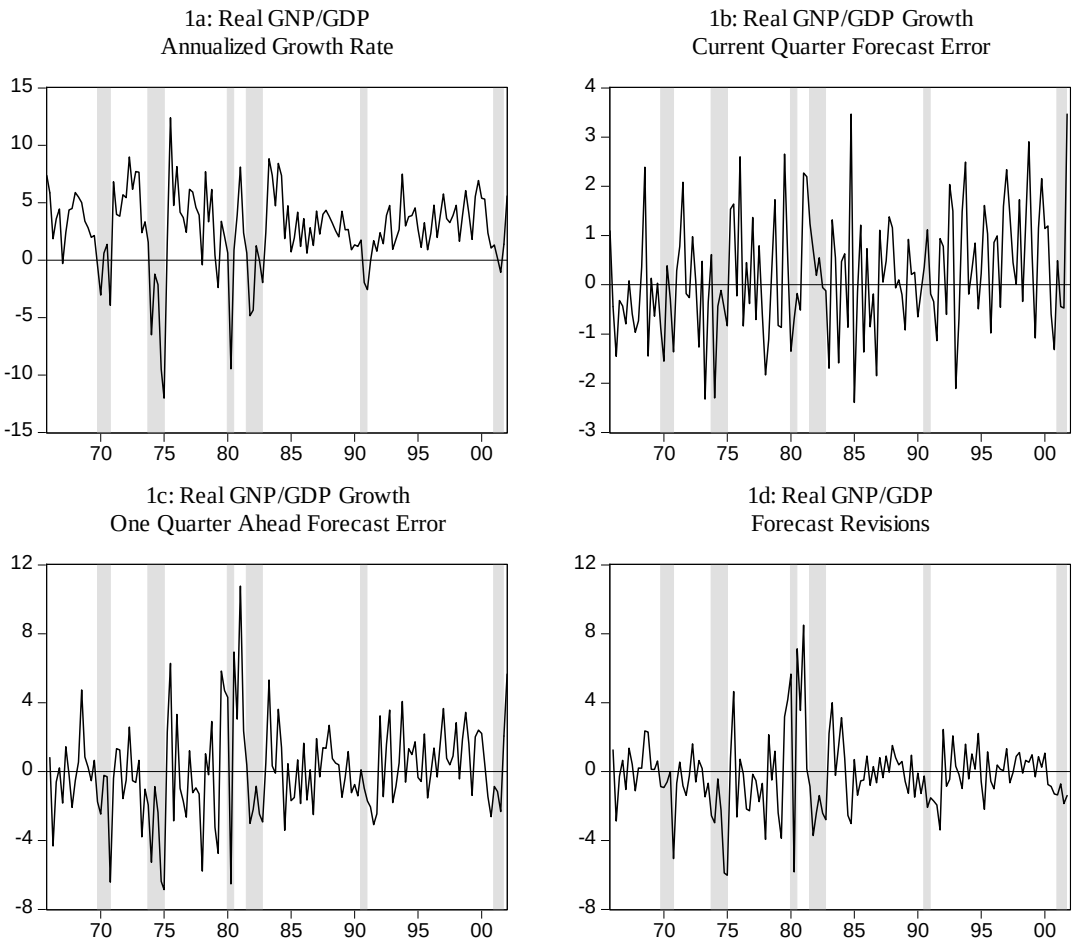
1a: Real GNP/GDP
Annualized Growth Rate
15
10
5
0
-5
-10
-15
70
75
80
85
90
95
00
1b: Real GNP/GDP Growth
Current Quarter Forecast Error
4
3
2
1
0
-1
-2
-3
70
75
80
85
90
95
00
1c: Real GNP/GDP Growth
One Quarter Ahead Forecast Error
12
8
4
0
-4
-8
70
75
80
85
90
95
00
1d: Real GNP/GDP
Forecast Revisions
12
8
4
0
-4
-8
70
75
80
85
90
95
00

Figure 2.a-2.d

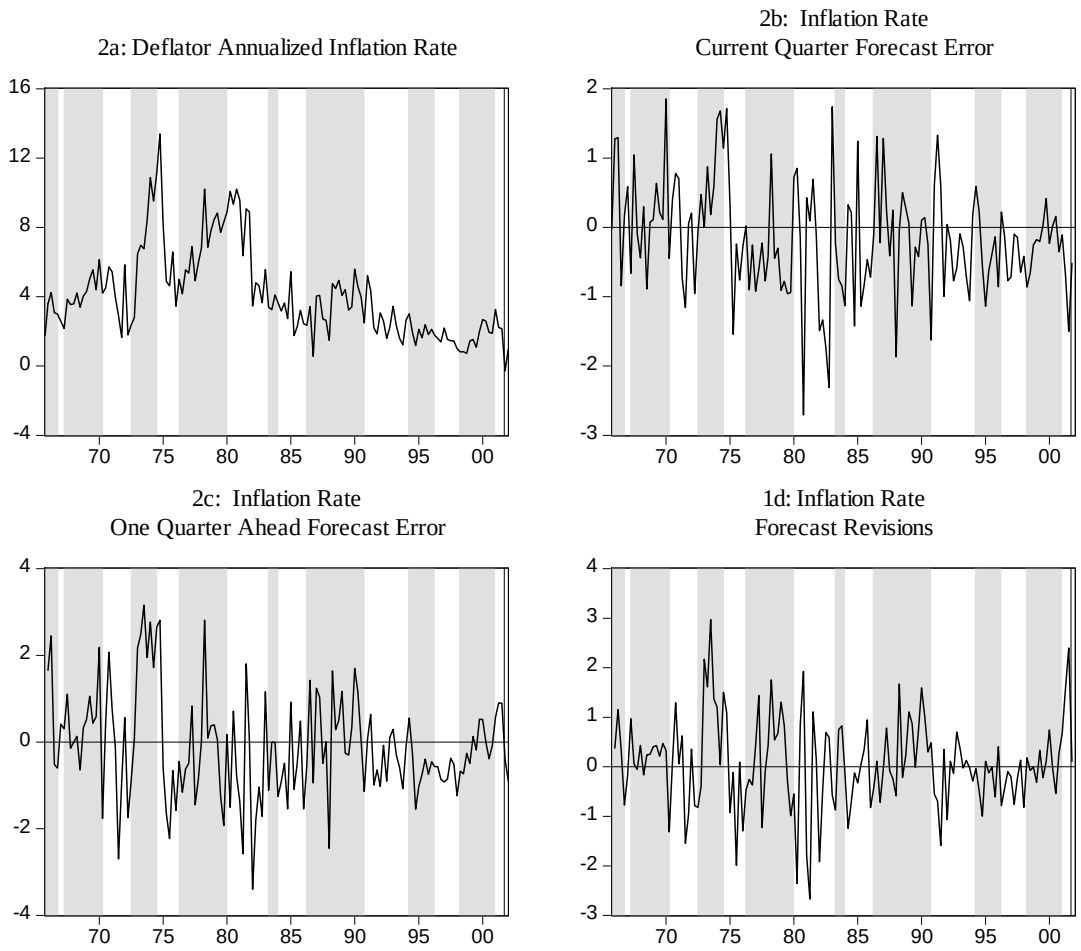
2a: Deflator Annualized Inflation Rate
2b: Inflation Rate
Current Quarter Forecast Error
2c: Inflation Rate
One Quarter Ahead Forecast Error
1d: Inflation Rate
Forecast Revisions
70
75
80
85
90
95
00

Figure 3.a-3.d

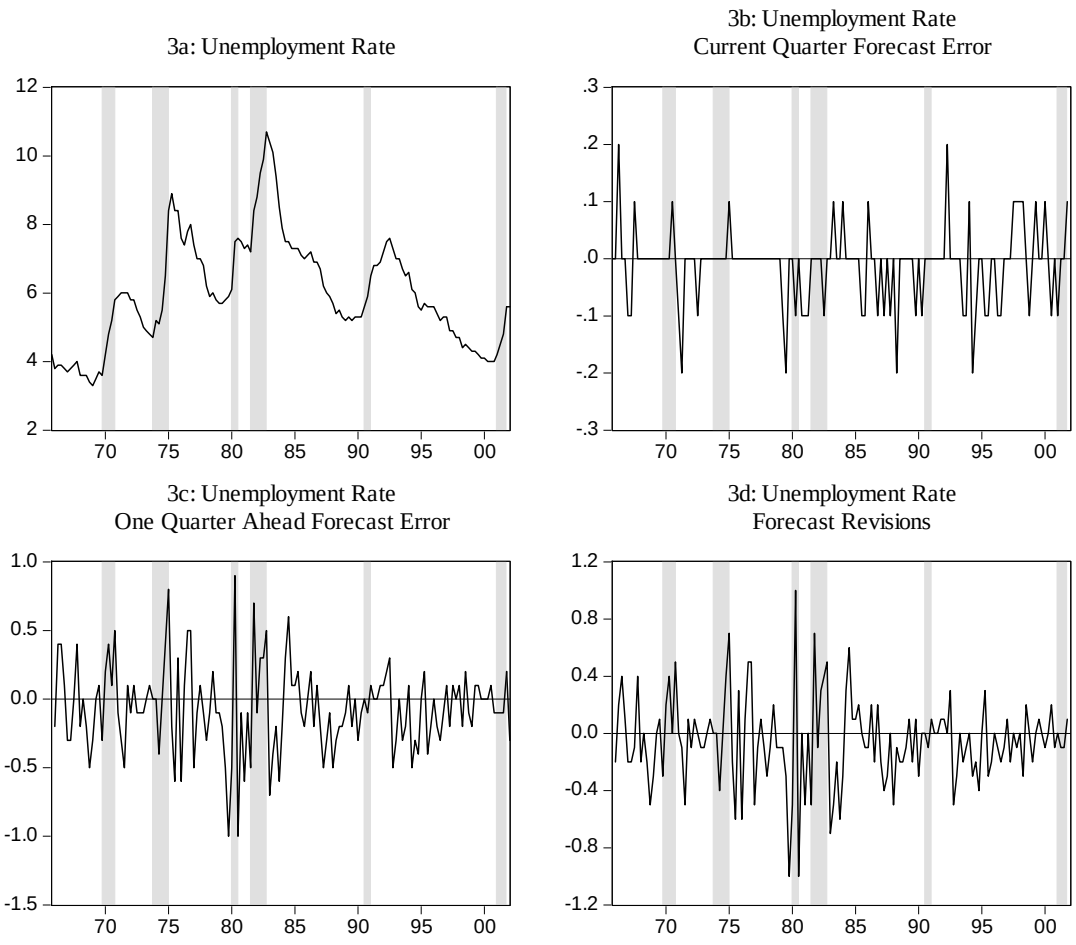
3a: Unemployment Rate
12
10
8
6
4
2
70
75
80
85
90
95
00
3b: Unemployment Rate
Current Quarter Forecast Error
.3
.2
.1
.0
-.1
-.2
-.3
70
75
80
85
90
95
00
3c: Unemployment Rate
One Quarter Ahead Forecast Error
1.0
0.5
0.0
-0.5
-1.0
-1.5
70
75
80
85
90
95
00
3d: Unemployment Rate
Forecast Revisions
1.2
0.8
0.4
0.0
-0.4
-0.8
-1.2
70
75
80
85
90
95
00

Table 1a
Greenbook Forecast Errors, 1965IV-2001IV

Variable	NBER State of the Economy: Expansion			Recession			Value of t-stat: Difference of means test
	Mean	Std Error	Obs	Mean	Std Error	Obs	
Current Quarter							
RealGrowth	0.260	1.209	118	-0.097	1.086	27	1.51
INF	-0.208	0.708	118	0.028	1.153	27	-1.02
UN	-0.014	0.070	118	0.000	0.048	27	-1.24
One Quarter Ahead							
RealGrowth	0.424	2.429	118	-1.706	3.112	27	3.33*
INF	-0.199	1.066	118	0.333	1.614	27	-1.63
UN	-0.102	0.267	118	0.081	0.423	27	-2.15*

Table 1b
Greenbook Forecast Errors, 1965IV-2001IV

Variable	ECRI Inflation Cycle: Peak to Trough			Trough to Peak			Value of t-stat: Difference of means test
	Mean	Std Error	Obs	Mean	Std Error	Obs	
Current Quarter							
INF	-0.294	0.915	57	-0.080	0.727	88	-1.49
One Quarter Ahead							
INF	-0.499	1.171	57	0.159	1.150	88	-3.33*

*Significant at the 5% level.

Table 2a Current Quarter Greenbook Forecast Mincer-Zarnowitz Tests 1965IV - 2001IV (Standard Errors in Parentheses)			
	Constant	**Slope**	**Wald Test Probability**
RealGrowth	0.135 (0.126)	1.023 (0.031)	0.12
INF	-0.189 (0.137)	1.006 (0.028)	0.05
UN	0.002 (0.021)	0.997 (0.003)	0.09

Table 2b: One Quarter Ahead Greenbook Forecast Mincer-Zarnowitz Tests 1965IV - 2001IV (Standard Errors in Parentheses)‡			
	Constant	**Slope**	**Wald Test Probability**
RealGrowth	0.407 (0.640)	0.858 (0.171)	0.64
INF	-0.003 (0.251)	0.977 (0.069)	0.77
UN	0.098 (0.106)	0.973 (0.019)	0.04

‡All one quarter ahead regressions have Newey-West standard errors. The years listed are when the forecasts were made.

Table 3a: Current Quarter Greenbook Forecast Ex Post Bias Tests
1965IV - 2001IV (Standard Errors in Parentheses)

		Constant	Slope	NBER Dummy	Inflation Dummy	Wald Test Probability
RealGrowth		0.295	0.990	-0.411		0.13
		(0.183)	(0.042)	(0.339)		
INF	(1)	-0.176	0.992	0.258		0.05
		(0.137)	(0.029)	(0.186)		
	(2)	-0.326	1.007		0.216	0.04
		(0.162)	(0.027)		(0.138)	
UN		0.001	0.997	0.016		0.11
		(0.021)	(0.003)	(0.014)		

Table 3b: One Quarter Ahead Greenbook Forecast Ex Post Bias Tests
1965IV - 2001IV (Standard Errors in Parentheses)‡

		Constant	Slope	NBER Dummy	Inflation Dummy	Wald Test Probability
RealGrowth		1.952	(0.543)	-3.781		<0.01
		(0.597)	(0.162)	(0.586)		
INF	(1)	0.029	0.941	0.652		0.20
		(0.229)	(0.057)	(0.424)		
	(2)	-0.474	0.995		0.655	<0.01
		(0.311)	(0.067)		(0.224)	
UN		0.082	0.969	0.192		<0.01
		(0.096)	(0.017)	(0.079)		

‡All one quarter ahead regressions have Newey-West standard errors. The years listed are when the forecasts were made.

Table 4a

Bias Tests for Forecast Revisions 1966I - 2001IV (Standard Errors in Parentheses)

		Constant	NBER Dummy	Inflation Dummy
RealGrowth		0.127	-1.735**	
		(0.185)	(0.428)	
INF	(1)	0.017	0.289	
		(0.083)	(0.192)	
	(2)	-0.205		0.457**
		(0.116)		(0.149)
UN		-0.085**	0.167**	
		(0.027)	(0.063)	

**Significant at the 1% level.

Table 4b

Bias Tests for Forecast Revisions 1966I - 2001IV (Standard Errors in Parentheses)

		Constant	Lagged NBER Dummy	Lagged Inflation Dummy
RealGrowth		0.088	-1.589**	
		(0.187)	(0.439)	
INF	(1)	0.041	0.172	
		(0.083)	(0.196)	
	(2)	-0.194		0.439**
		(0.116)		(0.150)
UN		-0.088**	0.188**	
		(0.027)	(0.063)	

**Significant at the 1% level.