Does the Federal Reserve Have Information Advantage?: A New Evidence at the Individual Level

JOB MARKET PAPER

Somchai Amornthum*
University of Hawai'i at Mānoa

November 12, 2006

Abstract

In this paper, we test the claim that the Federal Reserve (Fed) possesses information advantage over the public. We compare accuracy of forecasts made by the Fed to those made by the public at the individual level. Past evidences show that the Fed forecasts are more accurate than "consensus" forecasts. However, our results suggest that the Fed is more accurate than only half, but not all, of private-sector forecasters. We further examine the cause of the Fed's relative forecast accuracy. Inside information about future policy does not seem to be the only source of the advantage. The more important explanation is the inefficient use of publicly available information by the private-sector forecasters.

Keywords: Asymmetric Information, Inflation Forecasts, Monetary Policy, Rational Expectations

JEL classification: D82, E37, E52

*I am grateful to Carl Bonham for his valuable suggestions and the amount of time that he spends on this paper. I also thank Byron Gangnes, Ilan Noy, Xiaojun Wang and Tom Ramsey for many helpful suggestions and comments. The usual disclaimer applies. Address: Department of Economics, University of Hawaii at Manoa, 2424 Maile Way, Saunders Hall 542, Honolulu, HI 96822. Email: amornthu@hawaii.edu. Homepage: http://www2.hawaii.edu/~amornthu.

1 Introduction

The effectiveness of discretionary monetary policy is often tied to the existence of asymmetric information. The Federal Reserve (Fed) is frequently assumed to have information about the state of the economy that is not available to market participants.[1] However, empirical evidence to support the assumption is scarce. Perhaps, the reason is because it is impossible to measure the amount of information held by the Fed and the public. Romer and Romer (2000) overcome the problem by evaluating forecasts made by the Fed and the public instead. The "information the Federal Reserve has about the economy that is not known to market participants is likely to be reflected in the Federal Reserve's internal forecasts" (Romer and Romer, 2000, p. 429). Given that all forecasts are rational (i.e. they are unbiased and informationally efficient), the Fed forecasts can be more accurate than the public forecasts if the Fed has an information advantage over the public. Romer and Romer show evidences supporting the claim that the Fed has the information advantage. Other recent studies that look into this topic include Joutz and Stekler (2000), Sims (2002) and Carmona (2005).

We note that all of the studies mentioned above compare the Fed forecasts to the "consensus" forecasts.[2] We argue in this paper that the consensus forecasts should *not* be used in this type of comparison. The fact that the Fed forecasts are more accurate than the consensus forecasts does not guarantee that the Fed has more information than all (or even half) of the private-sector forecasters.[3] Comparing the Fed forecasts to the consensus forecasts can conceal individual "acceptance" of indifferent accuracy hypothesis. The main question is how many of the forecasters can be as accurate as the Fed and how many cannot. If the Fed is not more

[1] See for example Sargent and Wallace (1975); Barro (1976); Fischer (1977); Barro and Gordon (1983); Canzoneri (1985).

[2] The "consensus" forecast is referred to the central tendency (mean, median or others) of survey forecasts. The term may have been popularized by Joseph Livingston (see Croushore, 1997 for the overview of the Livingston survey), and it has stuck despite there being little real consensus among professional forecasters (see Zarnowitz and Braun, 1993).

[3] The consensus forecasts in different periods may come from different forecasters. As a result, it is not possible to pinpoint to which forecaster or how many of them the Fed dominates when it dominates the consensus.

accurate than most of the forecasters, the asymmetric information argument may have little support from the evidence, and so does the use of discretionary monetary policy. Figlewski and Wachtel (1981), Keane and Runkle (1990) and Bonham and Cohen (2001) argue in the case of rationality test that the consensus forecasts should not be used. Similarly, we argue here that comparing forecast accuracy, especially between the Fed forecasts and the public forecasts, should be done only at the individual level.

The usual practice in the literature is to first find which set of forecasts is more accurate. Then, it is concluded that the one producing relatively accurate forecasts (usually the Fed) has more information. However, we argue that the information advantage is not the only factor contributing to the relative accuracy. There are other factors causing it as well. In this paper, we go one step further by proving what is the source of Fed relative forecast accuracy (if it exists). In fact, many explanations have been raised in the literature, but a few have been tested. For instance, the Fed staff can be relatively accurate because of the inside information about future policy (Romer and Romer, 2000). By directly involved in the policy making process, the staff of the Board of Governors may have knowledge about the future path of policy. Since the FOMC meeting is closed to the public, this inside information may give the Fed staff the advantage in forecasting future economic variables. Another possible explanation raised by Romer and Romer (2000) is the amount of resources spent by the Fed on forecasting. Such amount is far greater than what is paid by any forecaster. The Fed is therefore better at extracting from equally available information. We also raise another possibility that some forecasters may simply be irrational and do not use information efficiently. The assumption that the public forecast is rational may not hold in practice.

The objective of this paper is twofold. First, we compare the accuracy of the Fed forecasts to that of the private-sector forecasts at the individual level. Second, we explain what factor gives the Fed the relative accuracy (if it exists). This paper focuses primarily on the inflation forecasts because the inflation is often viewed as the intermediate target of monetary policy. Should the Fed have the inside information about future policy, such information would affect the accuracy of inflation forecasts more than the accuracy of output forecasts.

This paper is organized as follows. Section 2 discusses the data used in this paper including the data on the Fed inflation forecasts, the inflation forecasts from the Survey of Professional Forecasters, and the actual inflation. Section 3 provides a brief background on the rational expectation hypothesis. Section 4 evaluates the Fed and commercial forecasts based on the loss associated with forecast errors. The evidence from Diebold and Mariano (1995) test with Harvey, Leybourne and Newbold (1997)'s modification suggests that the Fed's mean square forecast error (MSFE) is statistically smaller than only half of those from commercial forecasters. In section 5, we evaluate the accuracy using the forecast encompassing test proposed by Harvey, Leybourne and Newbold (1998). Here too, we do not find evidence that the Fed is more accurate than all forecasters. The Fed encompasses only half of the commercial forecasters. Section 6 then turns to the source of the information advantage. Mainly, we want to find a piece of information that is used by the Fed but not by most of the private-sector forecasters. We focus specifically on the inside information about the future policy and the inefficient use of information by the public. We argue that the better explanation seems to be the mixture of these two explanations. Conclusions and suggestions in section 7 then complete this paper.

2 The Data

We use forecasts from the Survey of Professional Forecasters (SPF), which begain in 1968Q4, to represent the public forecasts.[4] Survey participants include forecasters from financial firms, banks, consulting firms, university research centers, and other private firms. Since these forecasters tend to receive monetary rewards from producing accurate forecasts, they

[4]The survey was formerly conducted by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER) and was known as the ASA/NBER Survey. The Federal Reserve Bank of Philadelphia took over the survey from NBER in 1990Q2. See Croushore (1993) for detailed introduction of the survey.

should have incentives to minimize forecast errors. Keane and Runkle (1990) also argue that these forecasters should have incentive to report the true forecasts as well. Therefore, the mis-measurement problem is minimized with the SPF.

In the second month of each quarter, the participants are asked submit their forecasts of many economic variables. Along with the forecasts, they need to submit realizations of variables in the quarter preceding the survey quarter as well. We therefore calculate the inflation forecasts as the annualized quarterly growth rates of the GNP/GDP price deflator forecasts.[5] In this way, we use only data from the SPF to calculate the forecasts of inflation. Forecast horizons have been fixed over time at 0 to 4 quarters ahead. The number of respondents in each survey ranges from 131 forecasters in 1970Q4 to 9 forecasters in 1990Q2. The usual practice is to drop those who respond less than 20 times from the sample (Keane and Runkle, 1990; Bonham and Cohen, 2001). Also, following Bonham and Cohen (2001), we exclude the 1990Q2 survey because the questionnaires were sent out late in 1990Q3. A few observations with input errors are dropped. In the end, we have 94 forecasters in our sample with an average of 35.9 current-quarter forecasts per forecaster.

For the Fed inflation forecasts, we collect data from different issues of the Greenbooks. Before each FOMC meeting, the staff of the Board of Governors needs to prepare forecasts of many variables to assist the committee in decision making process. These forecasts are presented in a document, which becomes widely but unofficially known as the “Greenbook”. Since the releases of the Greenbooks are tied to the FOMC meetings, the forecasts are not available in standard frequency. The committee held 12 or more meetings a year during the 1960s and 1970s, and have met 8 times a year since 1982. There are more than one Greenbook in most of the quarters.[6] Since our goal is to compare accuracy of the forecasts, the Fed and

[5]We focus only on the forecasts of price deflator inflation and not the forecasts of CPI inflation mainly because of the data length. The survey of CPI forecasts did not start until 1981Q3, and the Fed did not start forecasting CPI until October 1979. The price deflator forecasts on the other hand have been available from the beginning of our sample period.

[6]Different researchers use different approaches to obtain quarterly Fed forecasts. For instance, Karamouzis and Lombra (1989) use the Greenbook prepared for the first meeting of each quarter, Romer and Romer (2000)

the SPF forecasters should be approximately on an equal stance. Neither of them should be allowed to have an information advantage because one consistently makes forecasts after the other. Thus, we use forecasts from the Greenbook, which is dated closest to the 15th day of the middle month of each quarter, to represent the Fed forecasts for the quarter.[7]

The Greenbooks are made available to the public only after a 5-year lag. The last Greenbook that we have was made in December 1999. As a result, our sample ends in 1999Q4. Further, the forecast horizons are not fixed over time. The maximum horizon in each Greenbook can be as short as the current quarter or as long as 9 quarters ahead. For consistency with the forecasts from the SPF, we focus only on the forecasts made for the current and the next 4 quarters. We calculate the forecasts of inflation as the annualized quarterly growth rate of the GDP/GNP price deflator forecasts as well.

The last piece of information is the actual inflation. In this study we evaluate the forecasts against the real-time data (as opposed to the current-vintage data). Croushore and Stark (2000) show that data from different vintages can produce different outcomes. For example, growth of real output in 1977Q1 could be between 5% and 10% depending on what the vintage of data is being observed. As argued by Keane and Runkle (1990), the forecasters should not be responsible for the revisions. For simplicity, we use the data available from the Real-Time Data Set for Macroeconomists compiled by the Federal Reserve Bank of Philadelphia.[8] The data set includes data as they exist in the middle of quarter. We use the data in the quarter in which they first come out to compare with the forecasts. Given the current schedule of data release, we effectively compare the forecasts with the first release of the GNP/GDP data.[9] The

select only the forecasts made in the second month, while Jansen and Kishan (1996) use the data from Greenbook prepared for the last meeting in a quarter.

[7]We use the date on the Greenbook, usually a few days before the meeting date, as the date when the forecasts are made.

[8]see Croushore and Stark (2001) for detailed description of the data set.

[9]In recent years, the U.S. Department of Commerce releases initial estimate of quarterly GDP at the end of the first month in the subsequent quarter. This release is called "advance" estimate. Then, the data are revised at the end of the second month (called "preliminary" estimate) and again at the end of the third month (called "final" estimate).

actual inflation is calculated in the same fashion as the forecasts of inflation.

3 Forecast Rationality and Unbiasedness Test

The link between relative forecast accuracy and information advantage depends heavily on rationality of the forecasts. We present an overview of the rational expectation hypothesis in section 3.1. A discussion about unbiasedness test is also presented in this section. Then, we show the results of the unbiasedness tests in section 3.2. We use the unbiasedness results in the following sections where we also divide results between biased and unbiased forecasters.

3.1 Rational Expectation Hypothesis and Unbiasedness Test

The essence of rational expectations hypothesis (REH), according to Muth (1961), is that the unobservable subjective expectations of economic agents are the same as the objective expectations implied by the relevant economic theory; that is, $\pi^e_{i,t,h} = E(\pi_{t+h}|\Omega_{i,t})$ where π_{t+h} is the actual value of inflation in period $t+h$, $\pi^e_{i,t,h}$ is the forecast of π_{t+h} made by forecaster i at time t, and $\Omega_{i,t}$ is the information set held by forecaster i at time t.[10] Let $\eta_{i,t,h} = \pi_{t+h} - \pi^e_{i,t,h}$ be the forecast error. Pesaran (1987) shows that the REH requires the expected forecast errors conditioned on the available information set to have zero mean; that is, $E(\eta_{i,t,h}|\Omega_{i,t}) = 0$.[11] He calls the condition as the *orthogonality* property.

The orthogonality property implies two further properties—*unbiasedness* and *efficiency*. The unbiasedness simply requires the forecast errors to have zero unconditional mean or $E(\eta_{i,t,h}) = 0$. The efficiency property requires forecasters to use all available information when forming the forecasts. As a result, the forecast errors are uncorrelated to any variable known at the time of forecast. Assuming an information set $S_{i,t}$ such that $S_{i,t} \subset \Omega_{i,t}$, the efficiency property suggests that the mean of forecast errors conditioned on $S_{i,t}$ must be equal

[10]Note that we use t as the time when forecast is made. Readers should not be confused with notations used in other studies where t sometimes represents the target period (Davies and Lahiri, 1995).

[11]See the proof of this property and other properties implied by REH in Pesaran (1987, ch.2)

to zero; that is, $E(\eta_{i,t,h}|S_{i,t}) = 0$. The efficiency property can be further divided into a weak and a strong form (see Lovell, 1986; Nordhaus, 1987; Bonham and Dacy, 1991). Under the weak form of efficiency, $S_{i,t}$ contains only the past forecast errors. Instead, if $S_{i,t}$ contains any variable known at the time of forecast, the condition is known as the strong form of efficiency. We use the concept of forecast efficiency to examine the source of forecast accuracy in section 6. At this point, we discuss only the unbiasedness test.

The usual test for unbiasedness is conducted by regressing realizations on the forecasts as $\pi_{t+h} = \alpha_i + \beta_i \pi^e_{i,t,h} + \epsilon_{i,t,h}$.[12] The null hypothesis of unbiasedness is H_0: $\alpha_i = 0$ and $\beta_i = 1$. However, Holden and Peel (1990) argue that the test above provides only a sufficient, but not a necessary, condition. Unbiased forecasts may fail the test above under a certain circumstance. They instead suggest, and we follow, that the unbiasedness can be tested by running the regression

$$\eta_{i,t,h} = \alpha_{i,h} + e_{i,t,h}. \tag{1}$$

Our null hypothesis is simply H_0: $\alpha_{i,h} = 0$. We can test the hypothesis using the standard t-test. A difficulty arises in estimating (1) because the residuals may not be white noise. Hansen and Hodrick (1980) demonstrate that forecast error $\eta_{i,t,h}$ follows an MA(h) process where h is the forecast horizon.[13] We correct the serial correlation in $e_{i,t,h}$ by using Newey and West (1987)'s Heteroscedasticity and Autocorrelation Consistent (HAC) variance in (1).

Another important problem is with the missing observations and how to deal with them in our sample. The problem is especially severe in the SPF because the survey respondents are asked, but not required, to return the questionnaires. A forecaster may respond to the survey in a quarter but skip many following quarters before another response. Some researchers may

[12] See for instance Mincer and Zarnowitz (1969); Figlewski and Wachtel (1981); Zarnowitz (1985); Keane and Runkle (1990); Davies and Lahiri (1995); Bonham and Cohen (2001)

[13] Other studies may indicate that forecast errors are MA($h-1$), and neither one of us is wrong. The difference is whether the current-period realization is known at the time of forecast or not. Theoretical studies usually assume that the current-period realization is known at the time of forecast; thus, the forecast error follows MA($h-1$) process. In our case, the forecast is made in the middle of the quarter when the inflation rate for the same period is not yet known. Thus, our forecast error follows MA(h) process.

ignore the problem entirely and simply collapse the data (i.e. treat periods with no observation as if they do not exist.) Such method, however, can lead to incorrect calculation of the HAC variances. For instance, the 2nd-order autocorrelation can be falsely interpreted as the 1st-order autocorrelation if the forecasts are available every other period. We follow the literature in replacing missing observations of $e_{i,t,h}$ with zeros, its expected value under the classical linear regression assumption, when we compute the variances.

3.2 Results of Unbiasedness Tests

Results of the unbiasedness tests on the Fed forecasts are presented in Table 1. In general, the means of forecast errors or $\alpha_{f,h}$ are very close to zero and not statistically significant for all horizons. The null hypothesis of unbiasedness is never rejected in any case. Further, the standard errors of the coefficients or $\sigma_{\alpha_{f,h}}$ tends to be larger for longer horizons. This indicates uncertainty involved with forecasting inflation at long horizons. The evidence is also well documented in Karamouzis and Lombra (1989) and Joutz and Stekler (2000).

TABLE 1: UNBIASEDNESS TESTS ON FED FORECASTS

Eq. (1): $\eta_{f,t,h} = \alpha_{f,h} + e_{f,t,h}$; H_0: $\alpha_f = 0$

Horizon (h)	$\alpha_{f,h}$	$\sigma_{\alpha_{f,h}}$	p-value	$T^*_{f,h}$
0	-0.12	0.09	0.205	124
1	-0.07	0.14	0.621	123
2	0.00	0.18	0.985	122
3	0.01	0.21	0.979	116
4	-0.07	0.23	0.748	108

Notes: $\eta_{f,t,h}$ is the forecast error of the Fed for an h-quarter-ahead inflation forecast made at time t; $\sigma_{\alpha_{f,h}}$ is the HAC standard error of $\alpha_{f,h}$; and $T^*_{f,h}$ is the sample size.

Next, we conduct the unbiasedness tests on the forecasts from each SPF forecaster. We report only the averages and standard deviations of the individual test results in Table 2. Out of 94 forecasters, 74% of them (70 forecasters) are able to produce unbiased forecasts of current-quarter inflation. The proportion of the unbiased forecasters declines as the horizon increases.

It becomes harder for them to make consistent forecasts at longer horizons. Still, even with the longest horizon in this study, more than half of the SPF forecasters can form unbiased forecasts. Averages of $\alpha_{i,h}$ among biased and unbiased are close to zero. Figure 1 shows that the averages of $\alpha_{i,h}$ among the biased forecasters are close to zero because the numbers of negatively and positively biased forecasters are approximately the same. As a result, the standard deviations of $\alpha_{i,h}$ in parentheses are relatively large among the biased forecasters.

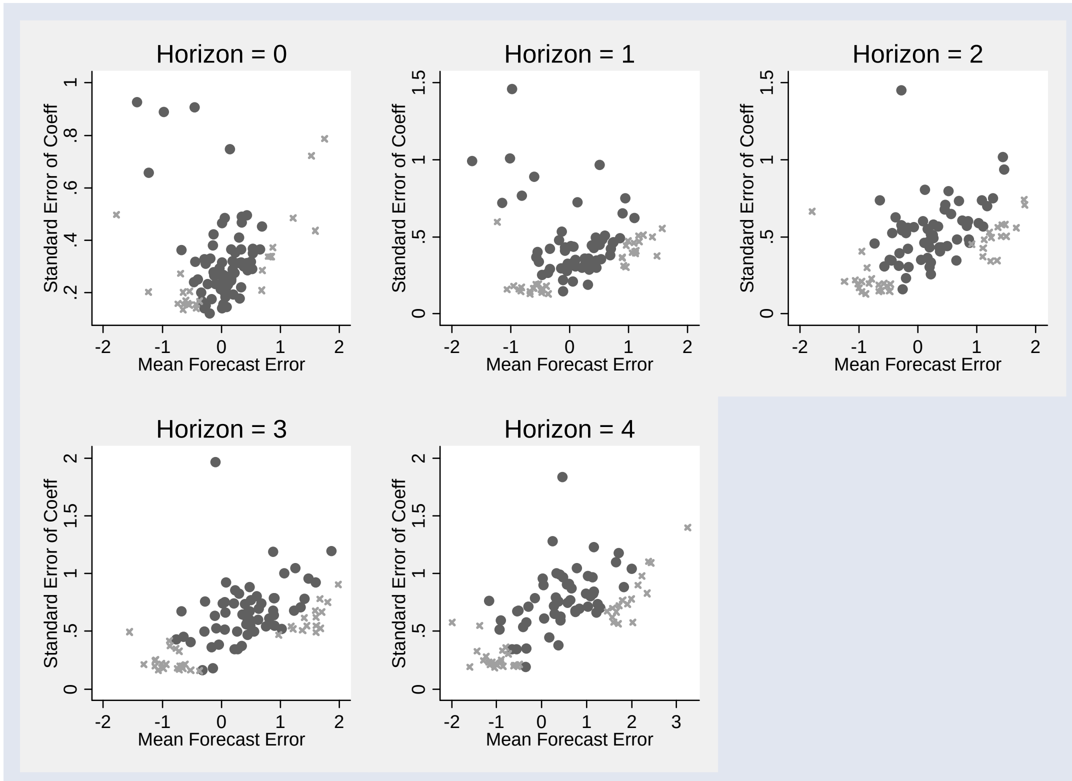
TABLE 2: UNBIASEDNESS TESTS ON SPF FORECASTS

Eq. (1): $\eta_{i,t,h} = \alpha_{i,h} + e_{i,t,h}$; H_0: $\alpha_{i,h} = 0$

Horizon (h)	$\bar{\alpha}_{i,h}$	$\bar{\sigma}_{\alpha_{i,h}}$	$\bar{T}^*_{i,h}$	N
	Unbiased SPF Forecasters (H_0 Not Rejected)			
0	0.03 (0.39)	0.33 (0.17)	35.97 (15.65)	70
1	0.10 (0.55)	0.46 (0.23)	35.13 (15.39)	60
2	0.26 (0.54)	0.54 (0.21)	36.58 (15.93)	59
3	0.44 (0.58)	0.67 (0.28)	35.71 (15.37)	58
4	0.44 (0.73)	0.78 (0.27)	33.13 (14.74)	54
	Biased SPF Forecasters (H_0 Rejected)			
0	0.04 (0.97)	0.30 (0.18)	35.04 (9.56)	24
1	0.22 (0.94)	0.31 (0.15)	37.26 (13.09)	34
2	0.15 (1.16)	0.36 (0.19)	34.71 (12.03)	35
3	0.12 (1.23)	0.40 (0.21)	35.86 (13.20)	36
4	0.26 (1.55)	0.49 (0.32)	33.05 (11.23)	40

Notes: $\eta_{i,t,h}$ is the forecast error of forecaster i for an h-quarter-ahead inflation forecast made at time t; $\sigma_{\alpha_{i,h}}$ is the HAC standard error of $\alpha_{i,h}$; $T^*_{i,h}$ is the sample size; and N is the number of forecasters. In column 2 to 4, numbers not in parentheses are averages of individual results, and number in parentheses are standard deviations of the individual results.

Figure 1 also shows a positive relationship between $\widehat{\alpha}_{i,h}$ and $\widehat{\sigma}_{\alpha_{i,h}}$. The pattern is more obvious for the longer horizons. The forecasters who underpredict the inflation tend to have high variances of forecast errors. From our detailed results not shown here, we find a tendency that the forecasters who on average underpredict are those in the early surveys, especially in the 1970s. During the time, the actual inflation was rising, and there were many episodes of economic shocks. Their forecasts could be incorrect by a great extent causing the variances of



Note: Dark circles represent the mean and standard error pairs from unbiased forecasters, while light "X" markers are the pairs from biased forecasters.

FIGURE 1: MEAN FORECAST ERRORS AND STANDARD ERRORS OF COEFFICIENTS FROM UNBIASEDNESS TESTS ON EACH SPF FORECASTER

their forecast errors to be high. In contrast, the forecasters who overpredict are mainly those in the latter surveys, mostly in the 1990s when the inflation was low and stable. As a result, their forecast errors tend to have low variances. Mankiw, Reis and Wolfers (2003) present a similar evidence based on consensus forecasts. They show that the consensus forecasts tend to underpredict when the actual inflation is rising and overpredict when the actual rate is falling. Mankiw et al.'s and our findings may raise a question on the credibility of the unbiasedness test results. A forecaster could be *cyclically biased*. That is, he is biased in the opposite directions during periods of rising or falling target. However, his forecasts may prove to be unbiased because the data cover a full cycle of rising and falling target. We are not aware of any study on this issue. Thus, we comply with the conventional method as in (1). The conclusion is that the Fed is unbiased, and some of the SPF forecasters can produce unbiased forecasts.

4 MSFE Dominance

One of the methods for comparing accuracy between two forecasts is to evaluate the loss associated with forecast errors. The forecast with lower loss are said to dominate the other forecast. The most frequently used loss functions are the quadratic and absolute functions. These functions imply that forecasters try to minimize mean square forecast error (MSFE) and mean absolute forecast error (MAFE) when making the forecasts. The former is especially appealing because of its similarity to the well known least square principle. We thus assume the quadratic loss function in this paper.

Comparison of forecast accuracy of the Fed and public forecasts under this method can date back to Lombra and Moran (1980) and Karamouzis and Lombra (1989). They respectively use averages of forecasts from various sources collected by McNees (1974) and McNees (1985) as the public forecasts. Lombra and Moran show that the MAFE of Fed inflation forecasts is lower than that of the consensus forecasts over 1970-1973 period. However, an update by Karamouzis and Lombra (1989) show a conflicting result that the consensus forecasts have lower MAFE over 1974-1983 period. The problem with these studies is that the MAFE is only a point estimate. It is possible that the ranking may switch due to random sampling. Diebold and Mariano (1995, DM hereafter) propose a way to test the null of no difference in the accuracy of two competing forecasts. Joutz and Stekler (2000) use the DM test to compare accuracy of the Fed and the consensus forecasts. They conclude that the Fed does not statistically dominate the consensus under the MSFE criterion. In this paper, we also use the DM test but, as mentioned earlier, apply it with the individual forecasts.[14] We describe the DM test in section 4.1, and the

[14]There are other methods of statistically evaluating the losses from two forecasts such as the F-test, the Morgan-Granger-Newbold test (Granger and Newbold, 1977), and Meese and Rogoff (1988) test. However, Monte Carlo simulations from DM prove that the DM tests maintain approximately correct size under the presence of contemporaneous and serial correlations, while other tests can be seriously mis-sized. Since both types of correlations are likely to exist in our study, we employ only the DM test here. See Mariano (2002) for a recent survey of the methods.

results are shown in section 4.2.

4.1 Diebold-Mariano Test

According to Diebold and Mariano, we can evaluate accuracy of two competing forecasts by comparing loss differentials. Under the null of no difference in the accuracy, the expected loss differential should be equal to zero, or $H_{0,i,h}$: $E(d_{i,t,h}) = 0$ where $d_{i,t,h} = \eta^2_{i,t,h} - \eta^2_{f,t,h}$ is the loss differential, $\eta^2_{i,t,h}$ is the loss made by forecaster i for $i = 1, \ldots, N$, and $\eta^2_{f,t,h}$ is the loss made by the Fed. The null hypothesis can be tested by the statistic[15]

$$S_{i,h} = \frac{\bar{d}_{i,h}}{\sqrt{\widehat{V}(\bar{d}_{i,h})}} \tag{2}$$

where $\bar{d}_{i,h} = \frac{1}{T^*_{i,h}} \sum_{t=1}^{T} d_{i,t,h}$ is the sample mean of the loss differentials, T is the number quarters in our sample, and $T^*_{i,h}$ is the number of effective sample of $d_{i,t,h}$.[16] The term $\widehat{V}(\bar{d}_{i,h})$ is the variance of $\bar{d}_{i,h}$, which is given by

$$\widehat{V}(\bar{d}_{i,h}) = T^{*-1}_{i,h} \sum_{\tau=-h}^{h} w \widehat{\gamma}_{i,h}(\tau)$$

where $\widehat{\gamma}_{i,h}(\tau)$ is the τ^{th} order autocovariance of $d_{i,t,h}$, and w is the lag window. Although Diebold and Mariano originally suggest the use of uniform lag window, we instead choose the Bartlett lag window because the uniform window sometimes yields negative value of $\widehat{V}(\bar{d}_{i,h})$. Our w is therefore $1 - \frac{|\tau|}{(h+1)}$.

Diebold and Mariano however point out that $S_{i,h}$ statistic can be seriously oversized in small samples.[17] To improve the test's behavior, Harvey, Leybourne and Newbold (1997) propose a modification of (2) that has a better size than the original test in the small samples.

[15] Diebold and Mariano call this statistic as S_1, but many studies refer to it as the DM statistic.

[16] We replace the missing observations with zeros. Given that we divide the sum of $d_{i,t,h}$ with $T^*_{i,h}$ instead of T, our $\bar{d}_{i,h}$ is the same as not using this observation when calculating the mean.

[17] The test size is the probability of the test statistic rejecting the null hypothesis when the null is correct. The test is said to be missize if the test reject the true null hypothesis more often (oversize) or less often (undersize) than the preset significance level (nominal size).

We note that their modification is based on the unrealistic assumption that the realization of the current period is known at the time of forecast. Obviously, the current-quarter inflation is not known when the forecasts are made. Thus, the forecast errors follow an MA(h) process instead of MA($h-1$) suggested by Harvey et al. The appropriate form of the modified DM test for our study is therefore

$$S^*_{i,h} = \left[\frac{T^*_{i,h} - 1 - 2h + T^{*-1}_{i,h} h(h+1)}{T^*_{i,h}}\right]^{1/2} S_{i,h}. \quad (3)$$

Under the null hypothesis, the $S^*_{i,h}$ statistic has the Student's t distribution with $(n-1)$ degrees of freedom.[18]

4.2 Results of Diebold-Mariano Test

The results from the DM tests are presented in Table 3. Instead of showing all individual tests, we present the results only in term of the proportion of forecasters under different categories. Although these proportions change with the significance levels, the overall pattern does not change much. Thus, in the analysis below we set the significance level at the conventional level of 5%. The results show that the Fed has statistically lower MSFE than 60-75% of the commercial forecasters. The proportion is highest at 73.85% for the forecasts of inflation in 4 quarters ahead. Clearly, no SPF forecaster can dominate the Fed at any horizon under the MSFE criterion. A few forecasters may have lower MSFE than the Fed's MSFE, but none of them has statistically lower MSFE.

Although the proportions presented in Table 3 are high, the evidence is also clear that under the MSFE criterion the Fed does not dominate *all* of the SPF forecasters. Bonham and Cohen (2001) show that the microhomogeneity does not hold among the SPF forecasters; that is, these forecasters are not drawn from the same random distribution of equally accurate forecasters. Thus, we can *not* interpret the results from Table 3 as we find enough evidence to reject the null hypothesis for the entire population. Rather, we should interpret the results as

[18] We modified the Stata program *dmariano* written by Christopher F. Baum to take into account the missing observations and the Harvey et al.'s modification.

TABLE 3: DIEBOLD-MARIANO TESTS ON FED AND SPF FORECASTS

$\bar{d}_{i,h} = T^{*-1}_{i,h} \sum_{t=1}^{T} (\eta^2_{i,t,h} - \eta^2_{f,t,h})$; H_0: $E(\bar{d}_{i,h}) = 0$

	Horizon (h)				
	0	1	2	3	4
All Forecasters					
Fed has lower MSFE	97.87	98.91	98.91	97.53	100.00
Fed has statistically lower MSFE	65.96	59.78	60.87	64.20	73.85
SPF has lower MSFE	2.13	1.09	1.09	2.47	0.00
SPF has statistically lower MSFE	0.00	0.00	0.00	0.00	0.00
Number of Forecasters	94	92	92	81	65
Unbiased Forecasters					
Fed has lower MSFE	97.14	98.31	100.00	95.83	100.00
Fed has statistically lower MSFE	60.00	50.85	51.72	52.08	70.59
SPF has lower MSFE	2.86	1.69	0.00	4.17	0.00
SPF has statistically lower MSFE	0.00	0.00	0.00	0.00	0.00
Number of Forecasters	70	59	58	48	34
Biased Forecasters					
Fed has lower MSFE	100.00	100.00	97.06	100.00	100.00
Fed has statistically lower MSFE	83.33	75.76	76.47	81.82	77.42
SPF has lower MSFE	0.00	0.00	2.94	0.00	0.00
SPF has statistically lower MSFE	0.00	0.00	0.00	0.00	0.00
Number of Forecasters	24	33	34	33	31

Notes: Fed has lower MSFE if $\bar{d}_{i,h} > 0$; Fed has statistically lower MSFE if $\bar{d}_{i,h} >$ and H_0 is rejected; SPF has lower MSFE if $\bar{d}_{i,h} < 0$; and SPF has statistically lower MSFE if $\bar{d}_{i,h} < 0$ and H_0 is rejected. The null hypothesis is rejected if $S^*_{i,h}$ in (3) is significant at 5% level. Except for the number of forecasters, the numbers in table are the proportion of SPF forecasters in the category.

the Fed dominates approximately 60% of the sample of commercial forecasters, and there are 40% of the forecasters who can be as accurate as the Fed.

We further divide the results between biased and unbiased forecasters. The Fed dominates larger proportion of the biased forecasters than that of the unbiased ones. For instance, the Fed has statistically lower MSFE of current quarter forecasts than 83.33% of the biased forecasters but lower than only 60% of the unbiased forecasters. In fact, the proportion is always higher among the biased forecasters than the unbiased.[19] Next, we turn to the forecast

[19]One can easily prove that MSFE $= E(\eta_{i,t,h})^2 + \text{Var}(\eta_{i,t,h})$. Thus, by being biased, the forecaster already gives an advantage to the Fed under the MSFE-based tests. However, a biased forecaster can have approximately the same MSFE as the Fed forecast if (1) his degree of biasedness, $E(\eta_{i,t,h})^2$, is not very large or (2) the forecast-

encompassing, another method of comparing the forecast accuracy.

5 Forecast Encompassing

The use of MSFE as the criterion in the comparison of forecast accuracy is criticized by Clements and Hendry (1993). The main argument is on its lack of invariance to isomorphic transformation. The Fed may be more accurate than a forecaster in predicting inflation, but the forecaster may be better at predicting the price index. The lack of invariance may pose a problem because we do not know what is being predicted. Instead, Clements and Hendry support the use of forecast encompassing, which is invariant to scale-preserving linear transformation. In section 5.1, we first introduce the forecast combination, the predecessor of forecast encompassing, and then the forecast encompassing itself. The results of the forecast encompassing tests are shown in section 5.2.

5.1 Forecast Combination and Forecast Encompassing Test

The concept of forecast encompassing was born out of the forecast combination literature.[20] Bates and Granger (1969) propose that a combination of two unbiased forecasts can be more accurate than either one alone. The combination can take a simple form of a weighted average of the two forecasts. The weights can be chosen by the following regression

$$\pi_{t+h} = \omega\pi_{1,t,h} + (1-\omega)\pi_{2,t,h} + u_{t,h} \tag{4}$$

where $\pi_{1,t,h}$ and $\pi_{2,t,h}$ are the forecasts of π_{t+h} made in period t by forecaster 1 and 2, respectively. Romer and Romer (2000) follow this approach to test asymmetric information between the Fed and the public. They find that the weights on the Fed forecasts are highly significant and positive, while the weights of the consensus forecasts are mostly negative and insignificant.

error variance, Var($\eta_{i,t,h}$), is smaller than the Fed's. Lamont (1995) argue that some forecasters can be biased but still produce small variance of forecast errors because they want to distinguish themselves from the herd.

[20] See Newbold and Harvey (2002) for a recent survey of literature on forecast combination and encompassing.

"The estimates indicate *overwhelmingly* that the Federal Reserve possesses valuable information not contained in the commercial forecasts" (Romer and Romer, 2000, p. 435, emphasis added).

We however argue that the evidence found by Romer and Romer does not support their conclusion. First, the usual t-test does not indicate that the two coefficients are statistically different from each other. It only tells whether each coefficient is different from zero, and nothing more. Second, the forecast with a negative weight in the combine forecast does not necessarily have less information. In fact, Granger and Newbold (1977) show that the negative weight is possible if errors from two forecasts are highly correlated. The negative weight on the forecast with high variance (but not necessarily statistically higher) can "pull down" the average when both forecasts overpredict the inflation.

In recent years, many researchers also argue that the regression as in (4) should not be used to combine forecasts, but rather it should be used to improve the preferred forecast. "[T]wo models which individually fail to capture the salient features of the data are unlikely to combine on a systematic basis to produce good forecasts" Clements and Hendry (1998, p. 233). The forecast encompassing test should be used to improve the preferred forecasts instead. There are many forms of the forecast encompassing test, but perhaps the most well-known one is from Clements and Hendry (1998). In our context, the test can be done by regressing forecast errors from forecaster i on the difference between Fed and commercial forecasts; that is, $\eta_{i,t,h} = \omega_{i,f}(\pi_{f,t,h} - \pi_{i,t,h}) + \eta_{i,f,t,h}$. By adding and subtracting the actual inflation on the right hand side, we obtain a more familiar form as

$$\eta_{i,t,h} \quad = \quad \omega_{i,f}(\eta_{i,t,h} - \eta_{f,t,h}) + \eta_{i,f,t,h}. \tag{5}$$

If the Fed can explain forecast errors made by the forecaster i, we should be able to reject the null of $H_{0,i,f}$: $\omega_{i,f} = 0$. We also need to test if the forecaster i can explain the errors made by the Fed. This can be done by estimating a similar equation of the form

$$\eta_{f,t,h} = \omega_{f,i}(\eta_{f,t,h} - \eta_{i,t,h}) + \eta_{f,i,t,h}. \tag{6}$$

The null hypothesis is $H_{0,f,i}$: $\omega_{f,i} = 0$. Then, the Fed encompasses forecaster i if $H_{0,i,f}$ is rejected but $H_{0,f,i}$ is not rejected. On the contrary, forecaster i encompasses the Fed if the first hypothesis is not rejected but the second test is rejected. The test results become inconclusive when neither or both hypotheses are rejected.

This form of the test is less prone to the problem of integration. Both sides of (5) and (6) tend to be balanced.[21] The remaining problem is that the residuals in (5) and (6) may not be white noise. Both heteroscedasticity and autocorrelation problems can exist. An "obvious" approach is to estimate the equation with HAC variance. However, Harvey, Leybourne and Newbold (1998) show that this "obvious" approach can give seriously oversized results in small samples. They propose another form of forecast encompassing test that has a better size in small samples.

In this paper, we follow the forecast encompassing test proposed by Harvey, Leybourne and Newbold (1998). They note that testing $\omega_{i,f}$ in (5) is similar to testing the correlation between $\eta_{i,t,h}$ and $(\eta_{i,t,h} - \eta_{f,t,h})$, while testing $\omega_{f,i}$ in (6) is the same as testing the correlation between $\eta_{f,t,h}$ and $(\eta_{i,t,h} - \eta_{f,t,h})$. Thus, one can test for forecast encompassing by defining

$$d_{i,f,t,h} = \eta_{i,t,h}(\eta_{i,t,h} - \eta_{f,t,h}) \quad \text{and} \tag{7}$$

$$d_{f,i,t,h} = \eta_{f,t,h}(\eta_{i,t,h} - \eta_{f,t,h}), \tag{8}$$

and test the null hypotheses of $H_{0,i,f}$: $E(d_{i,f,t,h}) = 0$ and $H_{0,f,i}$: $E(d_{f,i,t,h}) = 0$. To test these hypotheses, we can use the DM test with Harvey et al. (1997)'s modification as described in section 4.1. The Fed encompasses forecaster i if we can reject $H_{0,i,f}$ but not $H_{0,f,i}$.

[21] A balanced regression is the regression in which regressand is integrated of the same order as the regressors or any linear combination of the regressors (see Banerjee, Dolado, Galbraith and Hendry, 1993).

5.2 Results of Forecast Encompassing Test

The results of the forecast encompassing tests are shown in Table 4. Again, it is clear that the Fed does not encompass all of the SPF forecasters. Proportions of SPF forecasters who are encompassed by the Fed are approximately between 50% and 90%. It is highest at 87.23% for the current-quarter forecasts. However, the number quickly drops to 64.13% for the one-quarter-ahead forecasts, and it decreases further as the horizon increases. At the longest horizon, the Fed encompasses 49.23% of the SPF forecasters. We do not find a strong evidence that the SPF forecasters can encompass the Fed. Only 2.13% of them can encompass the Fed for the current-quarter forecasts, and the number reduces to zero for other horizons.

TABLE 4: FORECAST ENCOMPASSING TESTS ON FED AND SPF FORECASTS

Eq. (7): $d_{i,f,t,h} = \eta_{i,t,h}(\eta_{i,t,h} - \eta_{f,t,h})$; $H_{0,i,f}$: $E(d_{i,f,t,h}) = 0$
Eq. (8): $d_{f,i,t,h} = \eta_{f,t,h}(\eta_{i,t,h} - \eta_{f,t,h})$; $H_{0,f,i}$: $E(d_{f,i,t,h}) = 0$

	Horizon (h)				
	0	1	2	3	4
All forecasters					
Fed encompasses SPF	87.23	64.13	60.87	53.09	49.23
SPF encompasses Fed	2.13	0.00	1.09	0.00	0.00
Inconclusive Result	10.64	35.87	38.04	46.91	50.77
Number of Forecasters	94	92	92	81	65
Unbiased forecasters					
Fed encompasses SPF	82.86	61.02	60.34	43.75	47.06
SPF encompasses Fed	2.86	0.00	1.72	0.00	0.00
Inconclusive Result	14.28	38.98	37.94	56.25	52.94
Number of Forecasters	70	59	58	48	34
Biased forecasters					
Fed encompasses SPF	100.00	69.70	61.76	66.67	51.61
SPF encompasses Fed	0.00	0.00	0.00	0.00	0.00
Inconclusive Result	0.00	30.30	38.24	33.33	48.39
Number of Forecasters	24	33	34	33	31

Notes: Fed encompasses SPF if $H_{0,i,f}$ is rejected and $H_{0,f,i}$ is not rejected; SPF encompasses Fed if $H_{0,i,f}$ is not rejected and $H_{0,f,i}$ is rejected; inconclusive result when neither or both hypothesis is rejected. $H_{0,i,f}$ is rejected if $S^*_{i,h}$ in (3) based on $d_{i,f,t,h}$ in (7) is significant at 5% level, and $H_{0,f,i}$ is rejected if $S^*_{i,h}$ in (3) based on $d_{f,i,t,h}$ in (8) is significant at 5% level. Except for the number of forecasters, the numbers in table are the proportion of SPF forecasters in the category.

Next, we divide the test results between biased and unbiased forecasters. The results are also presented in Table 4. Two noteworthy conclusions stand out from the table. First, for all horizons, the proportions of SPF forecasters encompassed by the Fed are higher among the biased forecasters than the unbiased ones. As the horizon increases, the proportions decrease in both groups, but it is always higher in the group of biased forecasters. Second, the Fed encompasses *all* of the biased forecasters when it comes to making forecasts of current-quarter inflation. No biased forecaster can encompass the Fed at any horizon.

Lastly, we combine the results from the DM tests and the forecast encompassing tests and show the results in Table 5. Approximately, 45-65% of the SPF forecasters are dominated by the Fed under both criteria. The proportion is the highest for the current-quarter forecast at about 65%. However, the number quickly drops to about 50% for the one-quarter-ahead forecast and remains approximately at this level for other horizons. Our results here also contradict to the results from Romer and Romer (2000) who find that the weight given to the Fed forecast in the combined forecast is highest at the longest horizon.

TABLE 5: COMBINED RESULTS FROM DIEBOLD-MARIANO AND FORECAST ENCOMPASSING TESTS (UNIT=%)

	Horizon (h)				
	0	1	2	3	4
Fed dominates under both criteria	65.96	51.09	45.65	48.15	49.23
Fed dominates under one criterion	21.28	21.74	30.44	20.99	24.62
Fed dominates under neither criterion	12.77	27.17	23.91	30.86	26.15
Number of forecasters	94	92	92	81	65

Notes: Fed dominates under both criteria if Fed has statistically lower MSFE in Table 3 *and* Fed encompasses SPF in Table 4; Fed dominates under one criterion if either one is true; and Fed dominates under neither criterion if neither is true. Except for the number of forecasters, the numbers in table are the proportion of SPF forecasters in the category.

6 Source of Forecast Accuracy

In the previous sections, we show that the Fed can be more accurate than some SPF forecasters. Our next task is to explain why that is possible. What is the source of Fed relative forecast accuracy? Clearly, it is appropriate to reduce our attention to forecasters who are dominated by the Fed under both MSFE and forecast encompassing criteria. We call this group of forecasters as the "less able" forecasters.[22] The strategy here is to find information used by the Fed but not by most of the less able forecasters. If there exists such information, and it is known only to the Fed, our results would then support the existence of asymmetric information. However, we argue that the Fed's relative accuracy may rise from inefficient use of information by the less able forecasters as well. We use the orthogonality test as our main methodology in this section. The test is discussed in section 6.1. Then, we test if inside information about future policy can be the source of asymmetric information in section 6.2. The inefficient use of publicly available information is the focus of section 6.3. Lastly, we summarize the results of the orthogonality tests and discuss other possible sources of forecast accuracy in section 6.4.

6.1 Orthogonality Test

The orthogonality property follows from the REH as discussed in section 3.1. If forecasts are rational, the forecast errors should be orthogonal to the information known at the time of forecast. The orthogonality test can be implemented by regressing forecast error on a variable of interest as (see for example Batchelor and Dua, 1991).

$$\eta_{i,t,h} = \alpha_{i,h} + \theta_{i,h} x_t + u_{i,t,h} \tag{9}$$

where x_t is the variable that we want to test. The equation is estimated using the OLS with HAC variance. If x_t is known at the time of forecast, then x_t should *not* be correlated to $\eta_{i,t,h}$.

[22]We include both biased and unbiased forecasters in our sample because we want to compare our results to the entire population not just a group of unbiased forecasters. However, we explicitly take into account the forecast biasedness that may exist in the tests below.

The null hypothesis is therefore H_0: $\theta_{i,h} = 0$.[23]

One potential problem with (9) is that the equation may not be balanced if x_t is non-stationary. The unbalanced equation will lead to a false "acceptance" of the forecast efficiency hypothesis (Bonham and Cohen, 1995). We therefore perform necessary transformations to reduce the order of integration of the interested variable to zero. By using different variables as x_t, we can test different claims about the source of Fed relative forecast accuracy.

6.2 Inside Information about Future Path of Policy

One source of Fed's information advantage may be the inside information about the future policy (Romer and Romer, 2000). By directly involved in the policy making process, the Fed staff may know what the committee will agree upon in the next meetings. Such information can definitely improve their projections of the future interest rates as well as the forecasts of inflation. Since data on the interest rate projections are not available, we test this claim indirectly by using the future actual interest rates. Specifically, we use future changes of effective fed funds rates ($\Delta rfed_{t+k}$) and future changes of indicated target rates ($\Delta targ_{t+k}$) for k = 0, 1, 4 and 8.[24] The results are shown in Table 6.

The inside information does not seem to be the main source of Fed's relative forecast accuracy. We do not observe in Table 6 that the Fed passes most of the tests, and most of the less able forecasters fail them.[25] Instead, the evidence shows that the Fed itself does not pass some of the tests. For instance, the Fed's 4-quarter-ahead forecast errors are not orthogonal to

[23] Usually, the null hypothesis for the rational forecast is $\alpha_{i,h} = 0$ and $\theta_{i,h} = 0$. However, we are interested only if x_t is used in the forecast making process or not. Therefore, our null hypothesis is only H_0: $\theta_{i,h} = 0$. Further, since our less able forecasters include both biased and unbiased forecasters, the constant term is added to capture the degree of biasedness. The constant may not be zero when biased forecasts are being tested.

[24] Data on the monthly effective fed funds rate are collected from the Federal Reserve Bank of St. Louis. We average the monthly rates to obtain the quarterly rates. The change in the intended target rate is the total change from all meetings in the quarter. Data between 1968Q4 to 1996Q4 are drawn from Romer and Romer (2004), and we update the data through 1999Q4. We note that the Fed changed its operating procedure to target monetary aggregates from October 1979 to December 1981. However, Romer and Romer (2004) among others argue that the fed funds rate is still by and large a good measure of policy during the time.

[25] We tend to use the word "pass" the test to mean that we can not rejecting the null hypothesis. Also, "failing" the test means that we can reject the null hypothesis.

the current change of the interest rates. For the tests that the Fed passes, only a small portion of forecasters (about 10%) fail them. For the tests that the Fed fails, more than three quarters of the less able forecasters manage to pass them. For example, 75% of the less able forecasters have their 4-quarter-ahead forecast errors orthogonal to the current policy change.

TABLE 6: ORTHOGONALITY TESTS USING FUTURE INTEREST RATES

Eq. (9): $\eta_{i,t,h} = \alpha_{i,h} + \theta_{i,h} x_t + u_{i,t,h};\ H_0 : \theta_{i,h} = 0$

	Horizon (h)				
x_t	0	1	2	3	4
			Federal Reserve, p-values		
Δrfed$_t$	0.61	0.18	0.51	0.07	0.00
Δrfed$_{t+1}$	0.98	0.48	0.14	0.51	0.15
Δrfed$_{t+4}$	0.00	0.50	0.10	0.75	0.97
Δrfed$_{t+8}$	0.30	0.95	0.77	0.86	0.25
Δtarg	0.32	0.63	0.46	0.05	0.03
Δtarg$_{t+1}$	0.46	0.25	0.63	0.31	0.16
Δtarg$_{t+4}$	0.03	0.90	0.30	0.64	0.75
Δtarg$_{t+8}$	0.33	0.69	0.51	0.76	0.30
			Less Able SPF Forecasters, % Not Rejecting H_0		
Δrfed$_t$	93.55	87.23	85.71	79.49	75.00
Δrfed$_{t+1}$	93.55	97.87	73.81	87.18	84.38
Δrfed$_{t+4}$	83.87	100.00	95.24	89.74	84.38
Δrfed$_{t+8}$	85.48	72.34	85.71	82.05	84.38
Δtarg	93.55	93.62	83.33	94.87	81.25
Δtarg$_{t+1}$	93.55	97.87	90.48	84.62	84.38
Δtarg$_{t+4}$	91.94	97.87	95.24	97.44	84.38
Δtarg$_{t+8}$	87.10	93.62	95.24	92.31	84.38

Notes: $\eta_{i,t,h}$ is the forecast error of forecaster i for an h-quarter-ahead inflation forecast made at time t. Forecaster i include the Fed and the "less able" forecasters. Top panel presents marginal significance level of $\theta_{f,h}$ when Fed forecast error is used. Bottom panel shows proportion of less able forecasters who do not reject H_0 at 5% significance level.

Next, we calculate the proportion of the less able forecasters who manage to pass the tests that the Fed passes at each horizon. For instance, out of 62 forecasters who are dominated by the Fed under both criteria, we count how many of them have and do not have their forecast errors orthogonal to all future changes of interest rates (used in this study) except for $\Delta rfed_{t+4}$ and $\Delta targ_{t+4}$. The results are shown in Table 7. We use the proportion of the less able forecasters who fail at least one of the tests that the Fed passes as the explanatory power of this claim. In the case of current-quarter forecasts, only 35.5% of the less able forecasters fail at least one of the tests that the Fed passes. In other words, the claim of inside information can explain only 35.5% of the cases where the Fed dominates under both criteria. Although the explanatory power increases with the forecast horizon, the best it can do is to explain 53.8% of the cases. We support Romer and Romer (2000)'s conclusion that the inside information is not the main source of Fed relative forecast accuracy.

TABLE 7: PROPORTION OF SPF FORECASTERS PASSING OR FAILING ORTHOGONALITY TESTS USING FUTURE INTEREST RATES WHEN THE FED PASSES

	Horizon (h)				
	0	1	2	3	4
Pass all tests that Fed passes	64.52	59.57	47.62	46.15	46.88
Fail at least 1 test that Fed passes	35.48	40.43	52.28	53.85	53.12
Fail 1 test	19.35	23.40	23.81	20.51	25.00
Fail 2 tests	14.52	14.89	16.67	28.21	21.88
Fail 3 tests or more	1.61	2.13	11.90	5.13	6.25
Number of Forecasters	62	47	42	39	32

Notes: A forecaster "passes all tests that Fed passes" if his forecast errors are orthogonal to variables, to which the Fed forecasts errors are also orthogonal, in Table 6. Numbers are proportions of less able forecasters under each category.

6.3 Inefficient Use of Information

In this paper, we raise another explanation that the Fed is relatively accurate because some forecasters are irrational. They may not use the publicly available information efficiently. In contrast, the Fed utilizes all of the available information, and as a results it can produce superior forecasts. We test for both weak and strong efficiency conditions in this section. We note that this argument is observationally equivalent to the argument on the amount of resources spent on forecasting [26] The Fed has more information because it "commits far more

[26]To test this claim directly, one need to look at the correlation between forecast accuracy and cost of forecasting (e.g. man-hour used in forecasting). If the correlation exists, the evidence supports the resource argument; otherwise it supports the simply inefficient use of information by the forecasters. Lacking such data, we leave this task for future work.

resources to forecasting than even the largest commercial forecasters" (Romer and Romer, 2000, p. 437). By having large troops trying to extract information from equally available data, the Fed can produce forecasts relatively accurate. Also, with large resource, the Fed may be able to monitor more data series than any SPF forecaster. Stock and Watson (1999) and Bernanke and Boivin (2003) show that the use of large number of data series can greatly improve forecasts of inflation.

We now test the weak efficiency condition. The error of zero-quarter-ahead forecast made in the preceding quarter ($\eta_{i,t-1,0}$) is used as x_t in (9). The test results are shown in Table 8. The Fed seems to be using information from its own forecast errors efficiently, although it comes close to not being weakly efficient for the current-quarter forecasters with the p-value of 0.08. For the less able forecasters, only small proportion of them are not weakly efficient. Approximately 87% of them learn from their own mistakes when making new current-quarter forecasts. The proportion of the less able forecasters "accepting" the weak efficiency condition decreases slightly to approximately 65-70% for the forecasts of 3- or 4-quarter-ahead inflation.

TABLE 8: ORTHOGONALITY TESTS USING OWN FORECAST ERRORS

Eq. (9): $\eta_{i,t,h} = \alpha_{i,h} + \theta_{i,h} x_t + u_{i,t,h}$; $H_0 : \theta_{i,h} = 0$

	Horizon (h)				
x_t	0	1	2	3	4
Federal Reserve, p-values					
$\eta_{f,t-1,0}$	0.08	0.72	0.29	0.29	0.61
Less Able SPF Forecasters, % Not Rejecting H_0					
$\eta_{i,t-1,0}$	87.50	86.84	87.50	65.63	69.23

Notes: $\eta_{i,t,h}$ is the forecast error of forecaster i for an h-quarter-ahead inflation forecast made at time t. Forecaster i include the Fed and the "less able" forecasters. Top panel presents marginal significance level of $\theta_{f,h}$ when Fed forecast error is used. Bottom panel shows proportion of less able forecasters who do not reject H_0 at 5% significance level.

Next, we test the strong efficiency condition. We employ 20 variables from the real-time data set. They include many variables important to inflation forecasting such as M1, M2, output, nominal profits, import price, unemployment rate.[27] We also add the past fed funds rate, the target rate, the oil price, and the past error of the consensus forecast into the list. The list of all variables used to test the strong efficiency is presented in Table 9. The results of strong efficiency tests on the Fed forecasts and the forecasts from less able forecasters are presented in Table 10 and 11, respectively.

TABLE 9: VARIABLE DESCRIPTIONS

Variables	Transformation	Descriptions
$m1_t$	$\Delta^2 ln$	M1
$m2_t$	$\Delta^2 ln$	M2
$ncprofat_t$	Δln	Nominal corporate profits after tax
$noutput_t$	Δln	Nominal output
$oilp_t$	$\Delta^2 ln$	Spot oil price: West Texas intermediate
p_t	$\Delta^2 ln$	Output-price index
$pimp_t$	$\Delta^2 ln$	Price index for imports
$rcon_t$	Δln	Real personal consumption expenditure
$rcond_t$	Δln	Real personal consumption expenditure on durables
$rconnd_t$	Δln	Real personal consumption expenditure on nondurables
$rcons_t$	Δln	Real personal consumption expenditure on services
rex_t	Δln	Real export of goods and services
rg_t	Δln	Real government purchases of goods and services
$rimp_t$	Δln	Real imports of goods and services
$rinvbf_t$	Δln	Real business fixed investment
$rinvchi_t$	none	De-trended real change in inventories
$rinvresid_t$	Δln	Real residential investment
$routput_t$	Δln	Real output
ruc_t	Δ	Unemployment rate
$rcprofat_t$	Δln	Corporate profits after tax deflated by output price
$rfed_t$	Δ	Fed funds rate
$targ_t$	Δ	Target fed funds rate
$\eta_{c,t,h}$	none	Forecast errors of h-quarter-ahead consensus forecast

Notes: Δ means first difference; Δln means first difference of logarithm multiplied by 400; and $\Delta^2 ln$ means second difference of logarithm multiplied by 400.

[27] All variables are the realizations in the preceding quarter, except the nominal and real profits which are the realizations in the previous 2 quarters due to the lag in data release. The real profits is calculated using the nominal and output price index, both available from the data set. Most of the variables are transformed into growth rates to create stationary variables. The exceptions are the price variables, which are treated as I(2) and enter (9) in a difference of growth rate, and the interest rate variables, which enter (9) in a difference form.

TABLE 10: ORTHOGONALITY TESTS ON FED MQF USING PUBLIC INFORMATION

Eq. (9): $\eta_{f,t,h} = \alpha_{f,h} + \theta_{f,h} x_t + u_{f,t,h}$; $H_0 : \theta_{f,h} = 0$

	Horizon (h)				
x_t	0	1	2	3	4
Fed MQF, p-values					
$m1_{t-1}$	0.11	0.72	0.68	0.05	0.40
$m2_{t-1}$	0.96	0.14	0.75	0.00	0.74
$ncprofat_{t-2}$	0.31	0.07	0.03	0.65	0.12
$noutput_{t-1}$	0.44	0.47	0.02	0.04	0.49
$oilp_{t-1}$	0.30	0.30	0.01	0.13	0.75
p_{t-1}	0.24	0.73	0.37	0.16	0.97
$pimp_{t-1}$	0.10	0.18	0.02	0.49	0.53
$rcon_{t-1}$	0.09	0.81	0.90	0.93	0.91
$rcond_{t-1}$	0.05	0.46	0.69	0.91	0.50
$rconnd_{t-1}$	0.48	0.54	0.62	1.00	0.84
$rcons_{t-1}$	0.44	0.89	0.06	0.03	0.04
rex_{t-1}	0.29	0.19	0.11	0.35	0.53
rg_{t-1}	0.05	0.81	0.15	0.78	0.06
$rimp_{t-1}$	0.16	0.67	0.83	0.13	0.15
$rinvbf_{t-1}$	0.42	0.70	0.04	0.07	0.31
$rinvchi_{t-1}$	0.86	0.77	0.77	0.61	0.48
$rinvresid_{t-1}$	0.10	0.82	0.82	0.81	0.81
$routput_{t-1}$	0.26	0.99	0.29	0.07	0.55
ruc_{t-1}	0.84	0.77	0.71	0.08	0.50
$rcprofat_{t-2}$	0.30	0.09	0.03	0.69	0.11
$rfed_{t-1}$	0.69	0.55	0.04	0.01	0.85
$targ_{t-1}$	0.32	0.39	0.03	0.12	0.70
$\eta_{c,t-1,0}$	0.48	0.48	0.36	0.36	0.62

Notes: $\eta_{f,t,h}$ is the forecast error of the Fed for an h-quarter-ahead inflation forecast made at time t. Numbers are marginal significance levels of $\theta_{f,h}$ when Fed forecast error is used.

In general, our evidence does not point to any single piece of publicly available information as the main source of the relative accuracy. The Fed is strongly efficient with respect to most of the variables, but it also fails some of the tests. For example, the past changes in M1 and M2 are not fully used in the three-quarter-ahead forecasts of inflation. The two-quarter-ahead forecasts could have been improved if the information on nominal corporate profit, nominal output, oil price, import price, real investment, real corporate profit, or target

TABLE 11: ORTHOGONALITY TESTS ON LESS ABLE SPF FORECASTERS USING PUBLIC INFORMATION

Eq. (9): $\eta_{i,t,h} = \alpha_{i,h} + \theta_{i,h} x_t + u_{i,t,h};\ H_0 : \theta_{i,h} = 0$

	Horizon (h)				
x_t	0	1	2	3	4
$m1_{t-1}$	93.55	93.62	100.00	89.74	90.63
$m2_{t-1}$	92.31	88.37	100.00	91.89	100.00
$ncprofat_{t-2}$	95.16	95.75	92.86	97.44	87.50
$noutput_{t-1}$	88.71	97.87	95.24	69.23	93.75
$oilp_{t-1}$	74.19	76.60	57.14	69.23	84.38
p_{t-1}	91.94	89.36	95.24	46.15	81.25
$pimp_{t-1}$	94.83	93.62	73.17	81.58	93.75
$rcon_{t-1}$	98.28	93.62	97.56	97.37	96.88
$rcond_{t-1}$	98.28	95.75	100.00	94.74	96.88
$rconnd_{t-1}$	98.28	97.87	95.12	94.74	96.88
$rcons_{t-1}$	86.21	100.00	85.37	60.53	53.13
rex_{t-1}	86.21	89.36	85.37	44.74	84.38
rg_{t-1}	82.76	97.87	97.56	86.84	93.75
$rimp_{t-1}$	79.31	100.00	97.56	97.37	90.63
$rinvbf_{t-1}$	87.93	91.49	80.49	63.16	78.13
$rinvchi_{t-1}$	87.93	74.47	65.85	60.53	65.63
$rinvresid_{t-1}$	91.38	95.75	95.12	92.11	90.63
$routput_{t-1}$	88.71	97.87	97.62	79.49	96.88
ruc_{t-1}	90.32	91.49	95.24	69.23	90.63
$rcprofat_{t-2}$	93.55	95.75	92.86	97.44	84.38
$rfed_{t-1}$	85.48	78.72	85.71	74.36	84.38
$targ_{t-1}$	88.71	87.23	83.33	87.18	100.00
$\eta_{c,t-1,0}$	96.55	80.00	92.68	87.18	68.75

Notes: $\eta_{i,t,h}$ is the forecast error of forecaster i for an h-quarter-ahead inflation forecast made at time t. Numbers are proportions of less able forecasters who do not reject H_0 at 5% significance level.

rate had been used efficiently. A similar conclusion holds for the less able forecasters. Most of them are strongly efficient with respect to most variables. For each variable, there may be some forecasters failing the test. Yet, there is no single variable that could explain the forecast errors made by most of the less able forecasters. The variables that come close are the oil price, the real export and the real consumption on services for forecasts of inflation in 2-, 3-, and 4-quarter ahead, respectively. About 40-60% of the less able forecasters fail these variables at these horizons.

Table 12 shows the proportion of the less able forecasters who produce inferior forecasts but still manage to pass all strong efficiency tests that the Fed passes. For the current quarter forecasts, only 25.8% of less able forecasters pass all of the tests that the Fed passes. It means that 74.2% of them are dominated by the Fed because they do not use information efficiently. For other horizons, we can explain 64.3-92.3% of the sample using the same argument. The explanatory power of this argument is much higher than the inside information argument.

TABLE 12: PROPORTION OF SPF FORECASTERS PASSING OR FAILING ORTHOGONALITY TESTS USING PUBLIC INFORMATION WHEN THE FED PASSES (UNIT=%)

	Horizon (h)				
	0	1	2	3	4
Pass all tests that Fed passes	25.81	19.15	35.71	7.69	18.75
Fail at least 1 test that Fed passes	74.19	80.85	64.29	92.31	81.25
Fail 1 test	20.97	29.79	42.86	12.82	15.63
Fail 2 tests	20.97	21.28	19.05	17.95	28.13
Fail 3 tests or more	32.26	29.79	2.38	61.54	37.50
Number of Respondents	62	47	42	39	32

Notes: A forecaster "passes all tests that Fed passes" if his forecast errors are orthogonal to variables, to which the Fed forecasts errors are also orthogonal, in Table 10. Numbers are proportions of less able forecasters under each category.

6.4 Other Explanations

The results from section 6.2 and 6.3 suggest that neither the inside information nor the inefficient use of information is the main cause of Fed relative forecast accuracy. In this section, we expand the possibility. Perhaps, it is not any single one of these two arguments, but it is both of them that matter. From all 32 tests conducted in section 6.2 to 6.3 (8 tests on the inside information, 1 test on the weak efficiency, and 23 tests on the strong efficiency), we again calculate the proportions of forecasters who pass or fail the tests that the Fed passes. The results are shown in Table 13.

The results from Table 13 support our argument above. For most of the horizons, less

TABLE 13: PROPORTION OF SPF FORECASTERS PASSING OR FAILING ALL ORTHOGONALITY TESTS WHEN THE FED PASSES (UNIT=%)

	Horizon (h)				
	0	1	2	3	4
Pass all tests that Fed passes	9.68	8.51	16.67	2.56	6.25
Fail at least 1 test that Fed passes	90.32	91.49	83.33	97.44	93.75
Fail 1 test	20.97	23.40	19.05	5.13	12.50
Fail 2 tests	25.81	21.28	33.33	7.69	12.50
Fail 3 tests	14.52	12.77	11.90	15.38	18.75
Fail 4 tests	12.90	17.02	16.67	12.82	6.25
Fail 5 tests	9.68	8.51	0.00	33.33	25.00
Fail 6 tests or more	6.45	8.51	2.38	23.08	18.75
Number of Respondents	62	47	42	39	32

Notes: A forecaster "passes all tests that Fed passes" if his forecast errors are orthogonal to variables, to which the Fed forecasts errors are also orthogonal, in Table 6, 8 and 10. Numbers are proportions of less able forecasters under each category.

than 10% of the less able forecasters pass all of the tests that the Fed passes. We can explain why more than 90% of them produce less accurate forecast using the inside information and the inefficient use of information arguments. The number of tests that the SPF forecasters fail but the Fed passes increases as the forecast horizon increases. For the current quarter forecast, about 45% of the forecasters fail a test or two. For the four-quarter-ahead forecasts, more than 40% of forecasters fail 5 tests or more. It should be noted that the tests that these forecaster fails do not need to be the same test. This explains why we could not find any single argument as the main reason for Fed accuracy. We conclude that the Fed is better than the private forecasters because (1) some of them do not have inside information about the future policy; and (2) some of them do not use information efficiently (both weakly and strongly efficiently). These reasons together causes the Fed to make better forecasts than some of the SPF forecasters.

Another source of Fed's information advantage can come from unofficial or confidential information that the Fed receives from business leaders and bankers. Peek, Rosengren and Tootell (1999, 2003) specifically point to the supervisory knowledge about troubled, non-publicly traded institutions. Because the supervisory information is confidential, the public would not know about the health of ailing banks unless the banks actually fail. Only the Fed can take advantage from this information. However, the evidence from Peek et al. suggests that such information is not used by the Fed (nor the commercial forecasters). Thus, this information cannot explain the observed difference in the forecast accuracy.

7 Conclusions and Suggestions

In this paper, we re-examine the issue of asymmetric information between the Fed and the public. Previous studies used the consensus of private-sector forecasters to compare with the Fed forecasts. Most of them find that the Fed forecasts are more accurate than the consensus forecasts. This evidence leads to the conclusion that there exists asymmetric information between the Fed and the public. We argue in this paper that such test should be carried out only at the individual level. The fact that the Fed is better than the consensus does not mean that the Fed is better than most of the private-sector forecasters. In this paper, we compare the accuracy of the Fed forecasts to the forecasts from each forecaster. Based on the Diebold-Mariano test and the forecast encompassing test, our results suggest that the Fed produces forecasts more accurately than only half of the private-sector forecasters.

Even with this evidence, one still can not conclude that the Fed has an information advantage over half of the forecasters. There may be other reasons for Fed forecast accuracy; for example, the inefficient use of public information by the private forecasters. We try to find what is the source Fed forecast superiority by conducting the orthogonality tests. The inside information about future policy does not emerge as the main source of Fed superiority, nor the inefficient use of public information by the private-sector forecasters. However, we argue that the source of superiority lies on the mixture of both.

References

Banerjee, Anindya, Juan J. Dolado, John W. Galbraith, and David F. Hendry, *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press, 1993.

Barro, Robert J., "Rational Expectations and the Role of Monetary Policy," *Journal of Monetary Economics*, 1976, *2*, 1–32.

_ **and David B. Gordon**, "Rules, Discretion and Reputation in a Model of Monetary Policy," *Journal of Monetary Economics*, 1983, *12* (12), 101–121.

Batchelor, Roy and Pami Dua, "Blue Chip Rationality Tests," *Journal of Money, Credit and Banking*, November 1991, *23* (4), 692–705.

Bates, J.M. and Clive W.J. Granger, "The combination of forecasts," *Operational Research Quarterly*, 1969, *20*, 451–68.

Bernanke, Ben S. and Jean Boivin, "Monetary Policy in a Data-Rich Environment," *Journal of Monetary Economics*, 2003, *50*, 525–546.

Bonham, Carl S. and Douglas C. Dacy, "In Search of a "Strictly Rational" Forecast," *The Review of Economics and Statistics*, May 1991, *73* (2), 245–253.

_ **and Richard H. Cohen**, "Testing the Rationality of Price Forecast: Comment," *American Economic Review*, March 1995, *85* (1), 284–291.

_ **and** _ , "To Aggregate, Pool, or Neither: Testing the Rational Expectations Hypothesis Using Survey Data," *Journal of Business & Economic Statistics*, July 2001, *19* (3), 278–292.

Canzoneri, Matthew B., "Monetary Policy Games and the Role of Private Information," *American Economic Review*, December 1985, *75* (5), 1056–1070.

Carmona, Carlos C., "Bias in Federal Reserve Inflation Forecasts: Is the Federal Reserve Irrational or Just Cautious?," 2005. Working paper 2005-05, University of California, San Diego.

Clements, Michael P. and David F. Hendry, "On the Limitations of Comparing Mean Square Forecast Errors," *Journal of Forecasting*, December 1993, *12* (8), 617–637.

_ **and** _ , *Forecasting Economic Time Series*, New York: Cambridge University Press, 1998.

Croushore, Dean, "Introducing: The Survey of Professional Forecasters," *Business Review*, The Federal Reserve Bank of Philadelphia, November/December 1993, pp. 3–13.

_ , "The Livingston Survey: Still Usefule After All These Years," *Buiness Review*, The Federal Reserve Bank of Philadelphia, Mar-Apr 1997, pp. 1–12.

_ **and Tom Stark**, "A Funny Thing Happened on the Way to the Data Bank: A Real-Time Data Set for Macroeconomists," *Business Review*, The Federal Reserve Bank of Philadelphia, September/October 2000, pp. 15–27.

_ **and** _ , "A Real-Time Data Set for Macroeconomist," *Journal of Econometrics*, November 2001, *105* (1), 111–130.

Davies, Anthony and Kajal Lahiri, "A New Framework for Analyzing Survey Forecasts Using Three-Dimensional Panel Data," *Journal of Econometrics*, July 1995, *68* (2), 205–227.

Diebold, Francis X. and Roberto S. Mariano, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, July 1995, *3* (3), 253–263.

Figlewski, Stephen and Paul Wachtel, "The Formation of Inflationary Expectations," *Review of Economics and Statistics*, February 1981, *63* (1), 1–10.

Fischer, Stanley, "Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule," *Journal of Political Economy*, February 1977, *85* (1), 191–205.

Granger, Clive W.J. and Paul Newbold, *Forecasting Economic Time Series*, Academic Press, 1977.

Hansen, Lars Peter and Robert J. Hodrick, "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, October 1980, *88* (5), 829–853.

Harvey, David, Stephen Leybourne, and Paul Newbold, "Testing the Equality of Prediciton Mean Squared Errors," *International Journal of Forecasting*, 1997, *13*, 281–291.

_ , _ , and _ , "Tests for Forecast Encompassing," *Journal of Business & Economic Statistics*, April 1998, *16* (2), 254–259.

Holden, K. and D.A. Peel, "On Testing for Unbiasedness and Efficiency of Forecasts," *The Manchester School*, June 1990, *LVII* (2), 120–127.

Jansen, Dennis W. and Ruby Pandey Kishan, "An Evaluation of Federal Reserve Forecasting," *Journal of Macroeconomics*, Winter 1996, *18* (1), 89–109.

Joutz, Frederick L. and Herman O. Stekler, "An Evaluation of the Predictions of the Federal Reserve," *International Journal of Forecasting*, January-March 2000, *16* (1), 17–38.

Karamouzis, Nicholas and Raymond Lombra, "Federal Reserve Policymaking: An Overview and Analysis of the Policy Process," in Karl Brunner and Allan H. Meltzer, eds., *International Debt, Federal Reserve Operations and Other Essays*, Vol. 30 of *Carnegie-Rochester Conference Series on Public Policy* North-Holland Spring 1989, pp. 7–62.

Keane, Michael P. and David E. Runkle, "Testing the Rationality of Price Forecasts: New Evidence from Panel Data," *American Economic Review*, September 1990, *80* (3), 714–735.

Lamont, Owen, "Macroeconomic Forecasts and Microeconomic Forecasters," October 1995. Working paper 5284, National Bureau of Economic Research.

Lombra, Raymond and Michael Moran, "Policy Advice and Policymaking at the Federal Reserve," in Karl Brunner and Allan H. Meltzer, eds., *Monetary Institutions and the Policy Process*, Vol. 13 of *Carnegie-Rochester Conference Series on Public Policy* North-Holland Autumm 1980.

Lovell, Michael, "Tests of the Rational Expectations Hypothesis," *American Economic Review*, March 1986, *76* (1), 110–124.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers, "Disagreement About Inflation Expectations," *NBER Macroeconomics Annual*, 2003, *18*, 209–248.

Mariano, Roberto S., "Testing Forecast Accuracy," in Michael P. Clements and David F. Hendry, eds., *A Companion to Economic Forecasting*, Blackwell Publishers, 2002, chapter 13, pp. 284–298.

McNees, Stephen K., "How Accurate Are Economic Forecasts?," *New England Economic Review*, The Federal Reserve Bank of Boston, Nov-Dec 1974, pp. 2–19.

_ , "Which Forecast Should You Use?," *New England Economic Review*, The Federal Reserve Bank of Boston, Nov-Dec 1985.

Meese, R.A. and K. Rogoff, "Was It Real? The Exchange Rate Interest Differential Relation over the Modern Floating-Rate Period," *Journal of Finance*, 1988, *43*, 933–948.

Mincer, Jacob and Victor Zarnowitz, "The Evaluation of Economic Forecasts," in Jacob Mincer, ed., *Economic Forecasts and Expectataions: Analyses of Forecasting Behavior and Performance*, New York: Columbia University Press, 1969, pp. 3–46.

Muth, John F., "Rational Expectations and the Theory of Price Movements," *Econometrica*, July 1961, *29* (3), 315–335.

Newbold, Paul and David I. Harvey, "Forecast Combination and Encompassing," in Michael P. Clements and David F. Hendry, eds., *A Companion to Economic Forecasting*, Blackwell Publishers, 2002, chapter 12, pp. 268–283.

Newey, Whitney K. and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, May 1987, *55* (3), 703–708.

Nordhaus, William D., "Forecasting Efficiency: Concepts and Applications," *The Review of Economics and Statistics*, Nov 1987, *69* (4), 667–674.

Peek, Joe, Eric S. Rosengren, and Geoffrey M.B. Tootell, "Is Bank Supervision Central to Central Banking?," *Quarterly Journal of Economics*, 1999, *114* (2), 629–655.

_ , _ , and _ , "Does the Federal Reserve Possess an Exploitable Informational Advantage?," *Journal of Monetary Economics*, 2003, *50*, 817–839.

Pesaran, M. Hashem, *The Limits to Rational Expectations*, New York: Blackwell Publisher, 1987.

Romer, Christina D. and David H. Romer, "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, June 2000, *90* (3), 429–457.

_ **and** _ , "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, September 2004, *94* (4), 1055–1084.

Sargent, Thomas J. and Neil Wallace, "Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy*, April 1975, *83* (2), 241–254.

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

Stock, James H. and Mark W. Watson, “Forecasting Inflation,” *Journal of Monetary Economics*, 1999, *44*, 293–335.

Zarnowitz, V. and P. Braun, “Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance,” in J.H. Stock and M.W. Watson, eds., *Business Cycles, Indicators, and Forecasting*, Vol. 28, Chicago: University of Chicago Press, 1993, pp. 11–84.

Zarnowitz, Victor, “Rational Expectations and Macroeconomic Forecasts,” *Journal of Business & Economic Statistics*, October 1985, *3* (4), 293–311.