EUROPEAN CENTRAL BANK

EUROSYSTEM

In 2014 all ECB publications feature a motif taken from the €20 banknote.

WORKING PAPER SERIES

NO 1635 / FEBRUARY 2014

DOES THE FEDERAL RESERVE STAFF STILL BEAT PRIVATE FORECASTERS?

Makram El-Shagi, Sebastian Giesen and Alexander Jung

Acknowledgements
The authors thank Joseph Haubrich, Mikael Apel, Marianna Blix Grimaldi, Aaron Jackson, Axel Lindner, Claus Michelsen, Barbara Rossi, Oreste Tristani, three anonymous referees, and the participants of the Pacific Rim conference in Tokyo and those of the Annual Western Economic Association Conference in Seattle for useful comments and suggestions. The views expressed by the authors are their own and do not necessarily reflect those of the Eurosystem. The authors remain responsible for any errors or omissions.

Makram El-Shagi
Halle Institute for Economic Research and Department of Economics California State University-Long Beach;
e-mail: makram.el-shagi@iwh-halle.de

Sebastian Giesen
Deutsche Bundesbank; e-mail: sebastian.giesen@bundesbank.de

Alexander Jung
European Central Bank; e-mail: alexander.jung@ecb.europa.eu

Address	Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address	Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Internet	http://www.ecb.europa.eu
Fax	+49 69 1344 6000

ISSN	1725-2806 (online)
EU Catalogue No	QB-AR-14-009-EN-N (online)

Abstract

The aim of this paper is to assess whether the findings of Romer and Romer (2000) on the superiority of staff forecasts are still valid today. The paper uses both latest available econometric techniques as well as conventional tests. Several tests for forecast rationality show that a necessary condition for good forecast performance is satisfied both for Greenbook and private forecasts, as measured by the Survey of Professional Forecasters (SPF). Tests for forecast accuracy and the encompassing test confirm the superiority of Greenbook forecasts for inflation and output using an extended sample (1968 to 2006). The relative forecast performance is, however, not robust in the presence of large macroeconomic shocks such as the Great Moderation and oil price shocks. Other econometric tests show that a relative better forecast performance by staff is observed when there is increased uncertainty. Staff's better knowledge about the Fed's future interest rate path also plays an important role in this respect.

JEL Codes: C53, E37, E52, E58

Keywords: Forecast performance, forecast rationality, forecast stability, Greenbook forecasts, Survey of Professional Forecasters.

Non-technical summary

Today, most central bank watchers invest vast resources in making good forecasts of inflation and output. Why are they doing this? It helps them to better assess the monetary policy stance in real time and to form expectations about the likely future interest rate path. In a seminal paper Romer and Romer (2000) demonstrated that the Federal Reserve's (Fed) Greenbook forecasts outperform private forecasts. Hence, if private forecasters had knowledge of the Fed's internal forecasts, they could have made better forecasts of output and inflation in the United States.

The aim of this paper is to assess whether the findings of Romer and Romer (2000) on the superiority of staff forecasts are still valid today. We examine this question using both conventional tests as well as latest available econometric techniques. Several tests for forecast rationality show that a necessary condition for good forecast performance is satisfied both for Greenbook and private forecasts, as measured by the Survey of Professional Forecasters (SPF). Tests for forecast accuracy and the encompassing test confirm the superiority of Greenbook forecasts for inflation and output using an extended sample (1968 to 2006). The relative forecast performance is, however, not robust in the presence of large macroeconomic shocks such as the Great Moderation and oil price shocks. Other econometric tests show that a relative better forecast performance by staff is observed when there is increased uncertainty. Staff's better knowledge about the Fed's future interest rate path also plays an important role in this respect.

Several explanations for the Fed's information advantage have been proposed in the literature. Three explanations are prominent: (i) the Fed's thorough forecasting process including a vast range of resources devoted to forecasting macroeconomic variables, (ii) the Fed's knowledge of its own likely policy actions and the Fed's comparative advantage in collecting detailed information about current and recent movements in the economy, and (iii) the Fed's privileged access to confidential data based on its bank supervisory authority. Among them, Romer and Romer (2000) reject inside information by staff on the future interest rate path, the early access to government statistics and the better knowledge about data revisions as possible explanations.

This paper suggests that further qualifications to previous findings by Romer and Romer (2000) have to be made. On the one hand, this paper confirms their finding that the Fed has a significant information advantage concerning inflation and output forecasts for an extended sample (1968 to 2006). The result is robust to a possible timing advantage of the Fed's staff relative to private forecasters. Further evidence suggests superiority of Greenbook forecasts in particular when uncertainty is high. On the other hand, this paper differs from Romer and Romer (2000) when it comes to the assessment of the driving factors explaining the information advantage. First, this paper finds that the Fed's staff access to better information on the future fed funds rate explains a different relative performance concerning inflation forecasts. Second, the finding on the Fed's superiority is sensitive to the presence of large macroeconomic shocks such as the Great Moderation and oil price shocks. In this context, an interesting question is whether the non-availability of Greenbook forecasts in real time ultimately explains their superiority relative to private forecasts. This question could be addressed in further research by comparing the evidence with that for other central banks, which publish their staff forecasts in real-time.

1. INTRODUCTION

Today, most central bank watchers invest vast resources in making good forecasts of inflation and output. Why are they doing this? It helps them to better assess the monetary policy stance in real time and to form expectations about the likely future interest rate path. Can private forecasters learn something from central bank forecasts of these key macroeconomic variables? Romer and Romer (2000) have demonstrated that the Federal Reserve's (Fed) Greenbook forecasts outperform private forecasts of output and inflation in the United States. But, have relationships changed since then or has the US economy become more predictable, as suggested by Tulip (2009)? Moreover, the presence of information cascades (see Bikhchandani, Hirshleifer and Welch (2008)) appears to have contributed to a better sharing of information and to a reduction in the dispersion of private forecasts.

At least for the following three reasons it can be questioned whether the superiority of staff forecasts still holds today. First, the Fed and other main central banks have achieved a high level of transparency, thereby reducing their relative information advantage with the aim to enhance the effectiveness of monetary policy (see Woodford, 2005). In this context, the Fed only publishes its Summary of Economic Projections in real time, while its Greenbook forecasts are released with a lag of five years. Second, given several large macroeconomic shocks (the Great Moderation, oil price shocks, and financial crises) which contributed to changes in volatility patterns of macroeconomic time series, it is conceivable that the relative forecast performance between staff and private forecasters has changed. Because both groups of forecasters have been subject to profound uncertainty in the presence of these shocks, however, it could be the case that their performance is different. Third, the superiority of Greenbook forecasts is at odds with another paper by Romer and Romer (2008) suggesting that forecasts made by the Federal Reserve Open Market Committee (FOMC) are less informative than the Greenbook forecasts. These forecasts by FOMC policy-makers are informed by the internal Greenbook forecasts and should not differ fundamentally from the latter. While Ellison and Sargent (2012) have challenged this view in their defence of the FOMC policy-makers' forecasts, the two other points remain open for a more detailed investigation.

The aim of this paper is to assess the validity of the findings of Romer and Romer (2000)

on the superiority of staff forecasts. The paper uses both latest available econometric techniques as well as conventional tests. Several tests for forecast rationality show that a necessary condition for good forecast performance is satisfied both for Greenbook and private forecasts, as measured by the Survey of Professional Forecasters (SPF). Tests for forecast accuracy and the encompassing test confirm the superiority of Greenbook forecasts for inflation and output using an extended sample (1968 to 2006). The relative forecast performance is, however, not robust in the presence of large macroeconomic shocks such as the Great Moderation and oil price shocks. Other econometric tests show that a relative better forecast performance by staff is observed when there is increased uncertainty. Staff's better knowledge about the Fed's future interest rate path also plays an important role in this respect.

The paper is organized as follows. Section 2 briefly reviews the literature. Section 3 explains the data used for this study. Section 4 look into both forecast performance and relative forecast performance over the entire sample, essentially replicating the analysis of Romer and Romer (2000) with more data. Building on this analysis, Section 5 analyses the dynamics of forecast performance, testing the changes of forecast rationality over time, the changes in relative forecast performance, and – most importantly - identifying the driving factors underlying fluctuations of relative forecast performance. Section 6 concludes.

2. A BRIEF REVIEW OF THE LITERATURE

Greenbook forecasts are thought to provide the FOMC with an information advantage relative to private forecasters in the following sense (see Romer and Romer, 2000). First, Greenbook forecasts are more accurate that is they have lower root mean square errors (RMSE) than private forecasts. Second, given the Fed's Greenbook forecast, private sector forecasts have little or no additional explanatory power for inflation. The performance of private sector forecasts relative to staff forecasts has been the subject of a series of empirical studies. For reasons related to data availability, most studies examining this issue have been made for the US. For different samples ranging from the late 1960s to the mid-1990s several studies looking into forecast accuracy support the finding on the information advantage of the Fed (see Sims, 2002; Gavin and Mandal, 2003; Peek, Rosengren and Tootell, 2003; D'Agostino

and Whelan, 2008). While Reifschneider and Tulip (2007) and Gamber and Smith (2009) find that Greenbook forecasts are not more accurate than private forecasts (at least since the mid-1980s), Table 1 shows that this is still the case for inflation and output forecasts for the US (with a horizon of up to 4 quarters ahead). Moreover, Figure 1 shows that the nowcasts for inflation and output from the Greenbook and the SPF as well as the (corresponding) nowcast errors are highly correlated. A similar observation applies to forecasts and forecast errors for up to four quarters ahead. Hence, the gap in relative forecast performance may have narrowed over time (we analyse this question in more detail in Section 5). Table 1 reports the root mean square errors (RMSE) for the forecasts plotted in the figures for several horizons. By contrast, this comparison of forecast accuracy shows that the SPF forecasts for inflation and output are inferior to the Greenbook forecasts for the whole sample and at all horizons considered.

The Fed releases the FOMC forecast made by policy-makers once a quarter, but its Greenbook staff forecasts for each FOMC meeting are only published with a lag of about five years. These staff forecasts are therefore not available to the public when assessing the Fed's monetary policy stance and the economic outlook. Since July 1979, the Fed has regularly published summary statistics of FOMC policy-makers' economic projections twice a year (in February and July). Since October 2007, it has published a Summary of Economic Projections four times a year (this is done in connection with the FOMC's policy meetings in January, April, June, and November). [2]

Why should staff forecasts be superior to private forecasts? In fact, this is a puzzling proposition, because the level of data and model uncertainty is profound for all forecasters. A priori it is therefore

[2] In a separate study, Romer and Romer (2008) look into the controversial question whether the published FOMC policy-maker forecasts are inferior or similar to internal Greenbook forecasts prepared by staff, which are not available in real-time. They find that FOMC forecasts do not provide useful information relative to the Greenbook forecast even though FOMC members know the staff forecast when making their individual forecast. A study by Ellison and Sargent (2012) suggests that an inferiority of FOMC forecasts is at odds with evidence documenting that differences between FOMC and Greenbook forecasts are very small. We cannot resolve this debate here, but emphasise that the following important differences between both types of forecasts could have a bearing on the relative forecast performance. First, the FOMC forecast is made by individual FOMC members, it is not a staff forecast and it has not yet been established as a consensus forecast that is based on identical interest rate assumptions across FOMC members (currently work is under way in that direction). The accuracy of FOMC forecasts could be negatively influenced by specific factors. Strategic motives of individual members and a non-harmonised interest rate assumption (see McCracken, 2010; Tillmann, 2011) are examples of such factors. Second, it has a larger dispersion around the mean, because the range of possible outcomes considered by all individual policy-makers given multiple uncertainties may be more disperse than what Fed staff forecasts report.

not clear whether forecasts by central bank staff (or by international organizations such as IMF and OECD) are similar accurate than those produced by the private sector (see Batchelor, 2000). One of the main reasons here is the high level of transparency achieved by main central banks over the last decade. Thereby, they have deliberately reduced their relative information advantage with the aim to enhance the effectiveness of monetary policy (see Woodford, 2005). In parallel to central banks' efforts to increase transparency about their economic assessment, the private sector (mainly banks and other agencies) has increased the amount of resources it invests in making their own forecasts of these variables. It has allowed them to scrutinize the central banks' forward-looking assessment, when forming expectations, and it has contributed to better predictions of future monetary policy decisions (see Brand, Buncic and Turunen, 2010; Blattner, Catenaro, Ehrmann, Strauch and Turunen, 2008). Lange, Sack and Whitesell (2003) find increased predictability of FOMC decisions because of improved transparency. In a similar spirit, Swanson (2006) suggests that since the late 1980s increases in Fed transparency have been instrumental to the ability of both US financial markets and the private sector to forecast the federal funds rate at horizons of several months.

Nevertheless, in the literature it is widely assumed that central banks have an information advantage relative to the private sector, which could contribute to a better forecast accuracy in relative terms. First, they have more timely and complete knowledge of official statistics and may have access to first estimates of data releases. Second, they have more insight into their likely reaction to future shocks and, in the absence of forward guidance they should have better knowledge on their own intentions on future interest-rate setting.[3]

As institutions funded by the government, central banks could possibly be subject to political pressure in their response to economic shocks and this may diminish the accuracy of their forecasts. At the very least, political pressure could imply a deterioration of the quality of forecasts and would become visible in frequent revisions of first estimates. Today, most central banks are independent in the

[3] In the case of international organizations, this information advantage may be attenuated by the fact that forecasters are not based in the countries they are forecasting. They may benefit from consultations with central bank staff during country missions, but may not have access to various informal pieces of information which is available to local forecasters regardless of whether they work for the central bank or the private sector (see Abreu, 2011).

pursuit of their monetary policy goals, though to a varying extent (see Alesina and Summers, 1993 and the more recent evidence provided in Moutot, Jung and Mongelli, 2008). Central bank independence ensures in most instances that central banks are in a position to make an independent assessment of the monetary policy stance and the underlying assumptions.

As concerns inflation forecasts by central bank staff (and possibly for output forecasts, if they are made consistent with the inflation forecast), a further argument is that in the pursuit of a price stability goal central banks could have the incentive to align forecasts at the policy horizon with their numerical inflation target. While this could help them to stabilise private expectations and thereby enhance the effectiveness of monetary policy, it would result in a deterioration of forecast accuracy in absolute terms and relative to other forecasters. For example, assessments of inflation forecasts by the Bank of England and the Swedish Riksbank provide some evidence on the presence of this kind of bias, when real-time forecasts are compared with their outcomes (see e.g. Jung, 2013).

Most central banks are only responsible for preparing statistics on money and interest rates, and other government institutions are in charge of preparing the National Accounts and the fiscal data. Even, if central bank independence is considered to be a strong argument in this discussion, it still could be that the other government institutions, which provide important inputs to forecasting exercises, are subject to political pressure (see Frankel and Schreger, 2012). These institutions may have incentives to provide data which cast a favourable light on the government and thereby contribute to worsen forecast accuracy of those who use these data. At the same time, it must be acknowledged that statistical agencies all over the world make continuous efforts to improve the timeliness and quality of their statistics. Moreover, private forecasters often have no other sources when making their macroeconomic forecasts, but given knowledge about quality problems of official statistics, they may more easily justify the use of other non-official sources as inputs. This could have a positive impact on their forecast accuracy.

3. DATA

This paper uses quarterly forecasts for inflation (GDP deflator) and (real) output from the

Greenbook forecast, and the Survey of Professional Forecasters (SPF) for the US. It includes Greenbook forecasts and outcomes for the sample 1968Q4 to 2006Q4 from the real-time database of the Federal Reserve Bank of Philadelphia.[4] For a more detailed analysis of GDP and inflation forecasts, the database also includes Greenbook forecasts and outcomes of the GDP components (i.e. real consumption, real fixed business investment, real residential investment, real federal government consumption, real local and state government consumption) as well as those for nominal GDP and CPI inflation. The Greenbook projections are prepared independently by the research staff at the Board of Governors for each FOMC meeting without interference from the Board. Greenbook forecasts are available for five or six quarters into the future, though the horizon of the forecast varies over time and with the date of the FOMC meeting. They generally report forecasts in terms of real GDP growth, but before 1992 the reported data are for real GNP. Likewise for inflation, longer series are typically reported for the GDP (GNP) deflator. Other measures for inflation are available at somewhat shorter horizons, namely the CPI since 1979 and the PCE since 2000. These measures have been used more prominently in policy debates.

The SPF is the oldest quarterly survey of macroeconomic forecasts in the US. It has been conducted by the American Statistical Association and the National Bureau of Economic Research. The Bank of Philadelphia took over the survey in 1990 and as of third quarter of 1990 the Bank has transformed it into a real-time survey. Forecasts for core CPI inflation, PCE inflation, and core PCE inflation were only added as late as the first quarter of 2007. We therefore use the GDP (GNP) deflator (and, where meaningful we also report tests for CPI inflation). For real GDP (GNP) observations are fully comparable with the Greenbook forecasts. In this context, the present analysis uses the mean survey responses which are based on the implied forecast for each panellist (for details see the website of the Federal Reserve of Philadelphia). These forecasts are available for the long time series and their calculation takes into account that the sample composition has changed and that for some periods there have been no forecasts reported by individual forecasters. Using the mean also takes on board possible

[4] All results reported in this paper use real-time data, i.e. first releases. For the US GDP data are subject to deep revisions and this could have a bearing on the results. In fact, when using final vintages the results for GDP are not fully robust, whereas for inflation for which revisions have been less substantial the results are robust.

pooling advantages among private forecasters since the mean forecasts gives little weight to extreme values.

Staff forecasters have a slight timing advantage over the SPF and private forecasts have no knowledge of the Greenbook forecast until five years later. We use the Philadelphia Fed's Greenbook dataset which matches the timing of the SPF forecasts. The Greenbook forecasts are released to the FOMC members prior to each meeting (the time of the month when the forecast is made also varies, because the date of the FOMC meeting varies). For the purpose of this study, we use the observations that become available in the following months: March, June, September and December. For the SPF the timing is such that new observations are released to the public by middle or end of the following months: February, May, August and November.[5]

4. TESTS OF FORECAST PERFORMANCE FOR THE FULL SAMPLE

In this section we analyse the forecasting performance of Fed staff forecasters and private forecasters for inflation and output, both for each type of forecasters individually and in relative terms. We report empirical results for the US for the sample 1968 to 2006. First, we test for each group of forecasters whether their forecasts are rational. Forecast rationality is a criterion which can be regarded as a necessary condition for good forecast performance. We report results based on regressions for the full sample and for different subsamples. Second, we test whether these staff forecasts actually have contained information that would have helped private forecasters to improve their predictions in real-time.

4.1 Forecast rationality

When examining forecast performance, it is common to first check whether forecasts are "rational" in the sense of Muth (1961). Forecast rationality tests should check whether a forecast is equal to the mathematical expectation conditional on all (available) information. Forecast rationality requires forecasts to be both unbiased and efficient estimates of future outcomes (see e.g. Keane and

[5] For details see documentation of the Federal Reserve of Philadelphia: http://www.phil.frb.org/research-and-

Runkle, 1990). Since it is not always clear which set of exogenous variables should be included in these tests, it is common to test for weak efficiency, which requires that forecast errors are not correlated over time. Both, the unbiasedness and the (technical) weak efficiency assumptions have been challenged in the literature. For example, a biased forecast might be optimal in the case of asymmetric loss function (Holden and Peel, 1990), but such forecasts may not be useful seen from the perspective of professional forecasters. Likewise, it has been shown that weak efficiency will falsely be rejected for rational (and efficient) forecasts under a broad range of models.[6] Therefore, while testing for both unbiasedness and weak efficiency, we focus on the results related to unbiasedness, when discussing rationality.

We consider two alternative tests to assess unbiasedness of forecasts. First, we employ a standard Mincer-Zarnewitz regression (which is frequently applied, see e.g. Romer and Romer, 2000; Rossi and Sekhpoysan, 2011).

$$A_{h,t} = \alpha + \beta \hat{F}_{h,t} + \varepsilon_{h,t}, \qquad (1)$$

where $A_{h,t}$ denotes the outcome (either inflation rate or real GDP growth rate) at time t plus h-steps ahead, and $\widehat{F}_{h,t}$ is the corresponding h-step ahead forecast. The null hypothesis of forecast rationality is H_0: $\alpha = 0$ and $\beta = 1$ jointly. If the forecasting horizon *h* is larger than one period ahead, the residuals of this equation will exhibit moving average behaviour of an order of *h*-1. This, however, does not indicate inefficiency, because it only reflects that the same shock to the variable of interest shows up repeatedly in the forecast errors since the periods covered by subsequent forecasts overlap. To deal with the arising problem of serial correlation in the forecast errors when estimating equation (1), we calculate robust standard errors (HAC) for all regressions. Second, we employ a test for unbiasedness as proposed by Holden and Peel (1990). While the condition $\alpha = 0$ and $\beta = 1$ in the test equation (1) is sufficient for unbiasedness, Holden and Peel have shown that it is unnecessarily restrictive. Instead, the

---

data/real-time-center/survey-of-professional-forecasters/spf-documentation.pdf.

[6] This holds for a data generating process producing frequent persistent level shifts with uncertain timing such as a regime switching error correction model (where the equilibrium level of the dependent variable is observed, but uncertainty prevails about different adjustment regimes). Rational inflation forecasts would then produce correlated errors. While models with such features are frequently applied to a variety of macroeconomic indicators including inflation, their properties may make the validity of a so defined efficiency criterion questionable. This argument has been applied to the rationality of inflation forecasts. See for example Johnson, 1997; Evans and Lewis, 1995 and El-Shagi, 2011.

condition $\alpha/\mu A + \beta = 1$ (where μA is the mean of A), which is necessary and sufficient for unbiasedness should be used. Whether this condition holds, can implicitly be tested through the equation:

$$A_{h,t} - \widehat{F}_{h,t} = \alpha + \eta_{h,t}, \tag{2}$$

where η_t is a moving average process of an order equal to the effective forecasting horizon minus one. Holden and Peel (1990) propose to test the null hypothesis H_0: $\alpha = 0$. In small samples the Holden and Peel (1990) test could be somewhat problematic. The reduction in degrees of freedom due to the MA terms in the already small subsamples substantially reduces the power of the test, thereby falsely accepting unbiasedness too often. This problem is aggravated by the fact that dynamic estimation is typically subject to a small sample bias. When comparing the results from the traditional test (equation 1) and the Holden-Peel test it should be borne in mind that the latter test is too restrictive, whereas the former test is too forgiving.

We test for (weak) efficiency by estimating the augmented form of equation (1), as proposed by Keane and Runkle (1990):

$$A_{h,t} = \alpha + \beta\hat{F}_{h,t} + \gamma(A_{h,t-1} - \hat{F}_{h,t-1}) + \nu_{h,t}, \tag{3}$$

where v follows a moving average (MA) process of appropriate order. The null hypothesis of weak efficiency is H_0: $\gamma = 0$, i.e. forecast errors have no predictive power for the dependent variable. If the null hypothesis is valid, any dynamic behaviour in the forecast errors should come from the MA process, which is caused by overlapping forecast horizons. In principle, it would be possible to test for unbiasedness ($\alpha = 0$ and $\beta = 1$) and efficiency ($\gamma = 0$) simultaneously in this setup. Like for the Holden-Peel (1990) test, an augmented small sample problem may arise due to fewer degrees of freedom after accounting for the AR and MA behaviour. Since this would reduce the power of the test, we do not interpret the results obtained from equation (3) as additional check for unbiasedness. When interpreting the results on efficiency, we note that the presence of a small sample bias would imply a bias of the AR term towards zero. Consequently, the test would detect inefficiency less frequently, but it would not falsely detect inefficiencies where these are absent.

In terms of forecast rationality, our results for the US using an extended sample of almost 40 years (see Table 2) resemble those obtained earlier on by Romer and Romer (2000). We find that the

null hypothesis of forecast unbiasedness cannot be rejected at conventional levels (of 5 per cent), neither for the Greenbook nor for the SPF forecasts of inflation and output. That is, both Greenbook and private forecasts contain important information about future inflation and output developments. This finding can be extended to other private forecasters, namely the Blue Chip forecasts (see Rossi and Sekhpoysan, 2011). We provide additional results on forecast efficiency. For output, the test results support weak efficiency for Fed staff and SPF forecasts. But, for inflation the tests reject efficiency for both types of forecasters at all horizons, except for the nowcast. Uncertainty concerning the timing of future shifts in inflation is a possible factor which could be responsible for autocorrelated forecast errors. Efficiency is not rejected for nowcasts, since such uncertainty mostly affects forecasts that are based on structural (or structurally inspired) models. At the same time, expectations of an upcoming shift in inflation may turn out to be well founded, but with a different timing. Hence, the detected "inefficiency'" of inflation forecasts with longer horizons should not be interpreted as an outright rejection of their rationality.

4.2 Testing for additional information

It is widely believed that central banks have superior information when assessing the economic situation and the future course of monetary policy. If so, could private forecasters improve their forecasts by learning from the central bank staff forecast? In the following, as proposed by Romer and Romer (2000), the paper provides estimates of a regression which allows testing whether Greenbook forecasts contain information about the current and future economic environment that is potentially helpful for private forecasters to improve their forecasts of inflation and output. This test implicitly assumes that both types of forecasters have the same information set. The regression for the encompassing test compares forecasts from staff and private forecasters which are made in the same quarter. Since the SPF and Greenbook forecasts have somewhat different release dates, Fed staff in practice has a slight timing advantage of a few weeks and may use this information to improve its forecasts in real-time. To check whether the findings are robust to the timing issue, the paper controls for a possible timing advantage of Fed staff, following the approach proposed by Romer and Romer

(2000). To test for additional information, the following equation is estimated:

$$A_{h,t} = \delta + \gamma^P \hat{F}_{h,t}{}^P + \gamma^S \hat{F}_{h,t}{}^S + \nu_{h,t}, \quad (4)$$

where $A_{h,t}$ denotes outcomes at time t (inflation rate or real GDP growth) h-steps ahead, and $\hat{F}_{h,t}$ is the corresponding h-step ahead forecast from the central bank staff (superscript *S*) or the private forecaster (superscript *P*). The existence of additional information by central bank staff would require that γ^S is positive and significantly different from zero. When assessing the possible impact of timing on the relative forecast performance, we repeat the tests from equation 4 with an important modification. We put central bank staff at a timing disadvantage of one quarter, i.e. we use their forecasts from the previous quarter and check whether these forecasts are still informative for private forecasters. We estimate:

$$A_{h,t} = \delta + \gamma^P \hat{F}^P_{h,t} + \gamma^S \hat{F}^S_{h+1,t-1} + \nu_{h,t}, \quad (5)$$

Table 3 reports estimation results of equation (4) and (5) for the US. We confirm findings by Romer and Romer (2000) on the potential usefulness of Greenbook forecasts for the private sector. Our results indicate that for an extended sample Greenbook forecasts possess additional information on inflation and output which is not contained in the SPF forecasts. All estimates of γ^S are significantly positive for all forecasting horizons considered, and the estimates of γ^P are mostly insignificant and close to zero. Only, for the nowcast on inflation and the one-period-ahead forecast of output the SPF forecast contains valuable information. Hence, including Greenbook forecasts would have improved private forecasts.

When giving Fed staff a timing disadvantage of one quarter and extending the forecast horizon by one in the test (see equation 5), we find that Greenbook forecasts would no longer provide additional information for short-term forecasts. But, Greenbook forecasts for inflation and output are still useful for private forecasters at longer forecast horizons. This might indicate that there actually is a deeper understanding behind the structural causes of inflation (and GDP growth) at central banks, because this is required for making good forecasts at longer horizons. Contrarily, the access to the most recent information is essential to perform well in nowcasts.

5. ANALYSING THE DYNAMICS OF FORECAST PERFORMANCE

Owing to changes in volatility patterns and given the presence of extraordinary uncertainty, the forecast performance of central bank forecasters and of private forecasters can change over time. Anecdotal evidence supports this point. First, Fed Chairman Bernanke mentioned in a press conference (on 12 December 2012) that the Fed overestimated real GDP growth in past years. Second, in the presence of persistent oil price shocks, several inflation-targeting central banks appear to have underestimated inflation for some time. The Stockton (2012) Report suggests that the Bank of England's recent forecast performance has deteriorated and was somewhat worse than that of private forecasters. Third, Kenny and Morgan (2011) document the predictive failure of macroeconomic tools and expert judgement of forecasters more broadly during the financial crisis for both short and medium-term horizons. In this context, performance assessments by central banks are indicative that wrong assumptions concerning oil prices and fiscal policies were at the root of the forecast errors.

In order to examine this issue in more depth, Section 5.1 presents the results from the rolling window rationality tests, Section 5.2 introduces the tests for stability of the relative forecast performance and Section 5.3 presents the results from the conditional predictive ability test.

5.1 Rolling window rationality tests

Rossi (2005) and Rossi and Sekhpoysan (2011) argue that the tests described in Section 4 for forecast rationality are invalid in the presence of parameter instability.[7] Changes in paradigms of US monetary policy may imply structural breaks in the relationship. In 1979, the Fed embarked on a disinflationary monetary policy. In an unusual announcement chairman Volcker broke with past traditions and made it clear that the Fed would take responsibility for inflation (see Goodfriend, 1997, p. 12). This was an important clarification, because it implied that in the aftermath the Fed would give more weight to price stability within the dual mandate. The Volcker disinflation led to a regime shift towards lower inflation in the US. We show that this change has also had implications for the forecast rationality of both central bank staff and private forecasters. In order to check for the existence of a

[7] This issue will be addressed in Section 5.2 where we apply fluctuation tests (see Giacomini and Rossi, 2010).

break in the relationship we conduct a break point test.[8] These tests show that a break has likely occurred at the beginning of the 1980s, i.e. when the Volcker disinflation started (see last column of Table 2). To account for those changes we run our rationality tests for a moving window with a bandwidth of 25 quarters.

The rolling window estimates (see Figure 2 and 3) show occasional or even prolonged departures from the unbiasedness property by Greenbook and by SPF forecasts. Efficiency is rejected for most periods and for all forecast horizons for inflation and output. However, the above mentioned phenomenon of autoregressive behaviour of forecast errors obtained from rational models is particularly severe in small samples as shown by Evans and Lewis (1995), which would render the test invalid.

5.2 Testing for forecast stability

In this section we consider that differences in the relative performance of staff and private forecasts could vary over time. It is conceivable that for certain episodes there are no differences in the performance, while for other periods such differences are significant. We examine the stability of relative forecast performance using a recently developed fluctuation test by Giacomini and Rossi (2010). The null hypothesis of the test is forecast stability, i.e. that the detected difference in the relative forecast performance is not time-varying:

$$H_0: E[\Delta L_t(\hat{f}^S_{t-h,R}, \hat{f}^P_{t-h,R})] = 0\,, \qquad \text{for all t= R+h, …, T,} \tag{6}$$

where $\hat{f}_{t-h,R}$ denotes the h-step ahead forecast errors at time t by Fed staff (superscript *S*) and the private forecasters (superscript *P*). *L* denotes the corresponding loss function. The test statistics are computed using rolling (out-of-sample) windows of a given size *R*. We choose the window size to equal 15% of the sample, as suggested by Giacomini and Rossi (2010).

Figure 4 shows the test results for inflation, output and Figure 5 shows the results for the GDP components for the US. The null hypothesis of forecast stability is rejected, if the test statistics hits one of the confidence bounds shown in this Figure at least once (one-time reversal test). Inflation exhibits a

[8] We use the Andrews, Lee and Ploberger (1996) procedure to test for breaks at unknown time with a trimming

clear tendency towards instability in the relative performance. In order to check for robustness, we also make tests for the CPI as proxy for inflation. With the exception of the nowcast, the test also detects instability in the relationship. Since the CPI series provided by the Fed is substantially shorter than the GDP deflator series, the results for the CPI are not comparable here. Most importantly, the results for the CPI do not capture the Great Moderation. For (real) output the relative performance was also not stable.

Although fluctuations in performance do not suffice to reject stability, visual inspection indicates that the time series show some marked fluctuations. Prior to the mid-1980s, the fluctuation statistics indicates for most variables superiority of the Fed Greenbook forecasts.[9] The Greenbook forecasts significantly outperformed the private forecasts during the 1970s when the US economy had to face severe oil price shocks. Coinciding with the Great Moderation, which reduced overall volatility, the forecast performance of both types of groups became more similar. The fluctuation test statistics shows that since the mid-1980s there were overall no meaningful differences in the relative forecasting performance for inflation and output.

Furthermore, when analysing the GDP components (see Figure 5), the tests do not detect any signs of instability in the relative forecast performance. The detected difference in behaviour compared to the aggregate data mainly relates to the use of different samples. It is mostly due to missing observations for GDP components, since the analysis of the components only starts from the mid-1980s. As was shown for the aggregate data, changes in the relative forecast performance were most pronounced before the Great Moderation.

5.3 Testing for conditional predictive ability

The relative performance of different forecast groups may be explained by differences in the underlying assumptions. In this context, it has been argued that central bank staff has more timely and complete knowledge of official statistics and may have earlier access to data releases than private

---

parameter of 15%.

forecasters. In addition, in the absence of forward guidance, central bank staff should have better knowledge of the central bank's reaction to future shocks and the implied future interest-rate path. A test by Giacomini and White (2006) allows accounting for the possibility that the forecast performance is related to specific factors. We examine predictive ability conditional on three key factors (data revisions, interest rate path, oil and commodity prices) and we test for the influence of increased uncertainty as measured by a volatility index.

The test provides information on whether changes in the relative forecasting performance are linked to developments in specific exogenous variables. The null hypothesis is that given the information set Ω_t it is not possible to distinguish which forecast group has a lower forecast error at horizon τ. It can be written as:

$$H_0: E[L(\hat{f}^S_{t+\tau}) - L(\hat{f}^P_{t+\tau}) | \Omega_{t_t}] = 0 . \quad (7)$$

First, we examine the relative forecasting performance for inflation and real GDP growth, given uncertainties in the economic environment. They are proxied by the cross sectional dispersion for the quarterly forecasts (i.e. the dispersion of inflation, real GDP, industrial production, and housing starts).[10] Since the dispersion measures are obtained from the SPF, they could report an information disadvantage that is specific to professional forecasters, because uncertainty perceived by individual forecasters cannot necessarily be deduced from their mutual disagreement. To check for the robustness of the results, we use the predicted variance of inflation obtained from a simple GARCH(1,1) model as alternative proxy for economic uncertainties.

Second, we check whether data revisions had an impact on the relative forecasting performance. We use revisions in the variables inflation and real GDP growth.[11] Third, to account for the Fed's better knowledge of its interest rate policy, we test for the impact of upcoming interest rate changes on relative forecast performance. As proxy, we use the absolute quarter-on-quarter change in the federal funds rate at the corresponding forecast horizon. Fourth, we condition on oil prices and the commodity

[9] Negative (positive) values of the test statistics indicate the superiority (inferiority) of Greenbook forecasts.
[10] The dispersion measure equals the 75th percentile minus the 25th percentile of the forecasts for quarter on quarter variables. These variables are available from the Federal Reserve Bank of Philadelphia.

price index.

Table 5 reports the results. First, if we condition the relative forecast performance for uncertainties in the economic environment, we find that Fed staff generally made better inflation forecasts than private forecasters (SPF) during times of elevated economic uncertainties. The same holds for the nowcasts of output. However, for the nowcasts of inflation and for output forecasts for longer horizons this advantage cannot be detected.

Second, if we condition the relative forecast performance on data revisions, we find that the relative forecasting performance is only significantly affected in the very short term. Surprisingly, revisions in inflation cause improvements in the relative forecasting performance for real GDP, and revisions in real GDP cause improvements in the relative forecasting performance for inflation. Hence, for most horizons the tests support the argument by Romer and Romer (2000) that Fed staff makes better forecasts for reasons which are not related to Fed's staff earlier access to government statistics.

Third, if we condition the relative forecast performance on future changes in the federal funds rate, the relative performance of Fed staff inflation forecasts is better at a longer horizon of four quarters ahead. This test suggests that the Fed likely made better inflation forecasts when interest rate changes were looming, but no forward guidance was applying.[12] In that sense, the Fed's Greenbook inflation forecasts seem to have benefited from the staff's better knowledge of the Fed's future interest rate path.

Fourth, if we condition relative forecast performance on oil prices and the HWWA index for energy, oil and raw materials, we find that these factors have no significant influence on the horserace between Fed staff and private forecasters.[13] These test results are consistent with common knowledge that both types of forecasters face an even challenge when attempting to predict the consequences of changes in oil and commodity prices for inflation and output.

---

[11] This set of revisions is obtained from the real-time data set from the Federal Reserve of Philadelphia.

[12] Since interest rate changes may be frequent in times of higher economic uncertainty, we also check for correlation between these variables that may drive our results and, hence, distort the interpretation of our results. We find only some correlation of the variables (0.30), implying that the results are mainly attributable to the separate effect of the interest rate.

[13] This finding is robust to using different measures for oil. Since the HWWA index also compromises commodity

6. CONCLUSIONS

Several explanations for the Fed's information advantage have been proposed in the literature. Three explanations are prominent: (i) the Fed's thorough forecasting process including a vast range of resources devoted to forecasting macroeconomic variables, (ii) the Fed's knowledge of its own likely policy actions and the Fed's comparative advantage in collecting detailed information about current and recent movements in the economy, and (iii) the Fed's privileged access to confidential data based on its bank supervisory authority. Among them, Romer and Romer (2000) reject inside information by staff on the future interest rate path, the early access to government statistics and the better knowledge about data revisions as possible explanations.

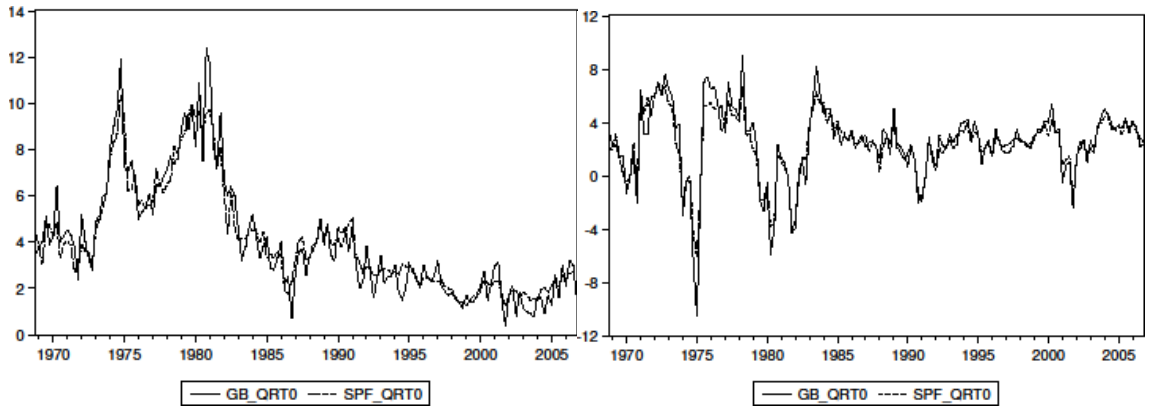
This paper suggests that further qualifications to previous findings by Romer and Romer (2000) have to be made. On the one hand, this paper confirms their finding that the Fed has a significant information advantage concerning inflation and output forecasts for an extended sample 1968 to 2006. The result is robust to a possible timing advantage of the Fed's staff relative to private forecasters. Further evidence suggests superiority of Greenbook forecasts in particular when uncertainty is high. On the other hand, this paper differs from Romer and Romer (2000) when it comes to the assessment of the driving factors explaining the information advantage. First, this paper finds that the Fed's staff access to better information on the future fed funds rate explains a different relative performance concerning inflation forecasts. Second, the finding on the Fed's superiority is sensitive to the presence of large macroeconomic shocks such as the Great Moderation and oil price shocks. In this context, an interesting question is whether the non-availability of Greenbook forecasts in real time ultimately explains their superiority relative to private forecasts. This question could be addressed in further research by comparing the evidence with that for other central banks, which publish their staff forecasts in real-time.

---

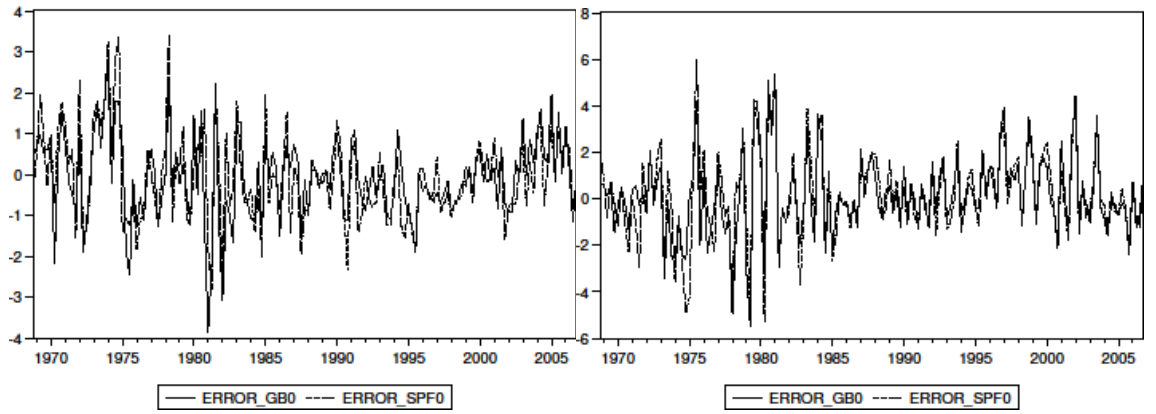
prices we only report the results for this measure in Table 5.

FIGURE 1: Nowcasts and nowcast errors for the Greenbook and the SPF

(a) Nowcast for inflation (GDP deflator) **(b) Nowcast for real GDP**



(c) Nowcast error for inflation **(d) Nowcast error for real GDP**



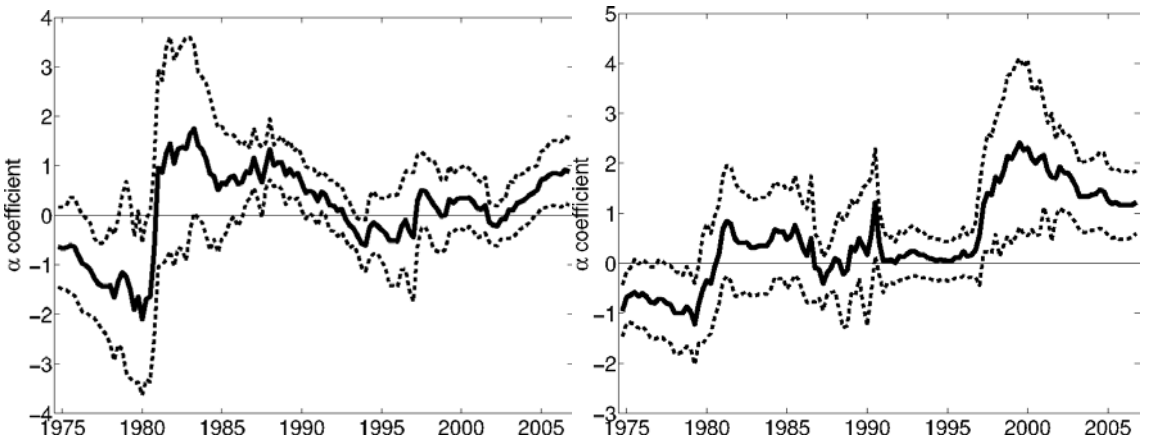
Notes: GB_QRT0 is the Greenbook nowcast; SPF_QRT0 is the SPF nowcast; ERROR_GB0 is the Greenbook nowcast error and ERROR_SPF0 is the SPF nowcast error.

Source: see data section.

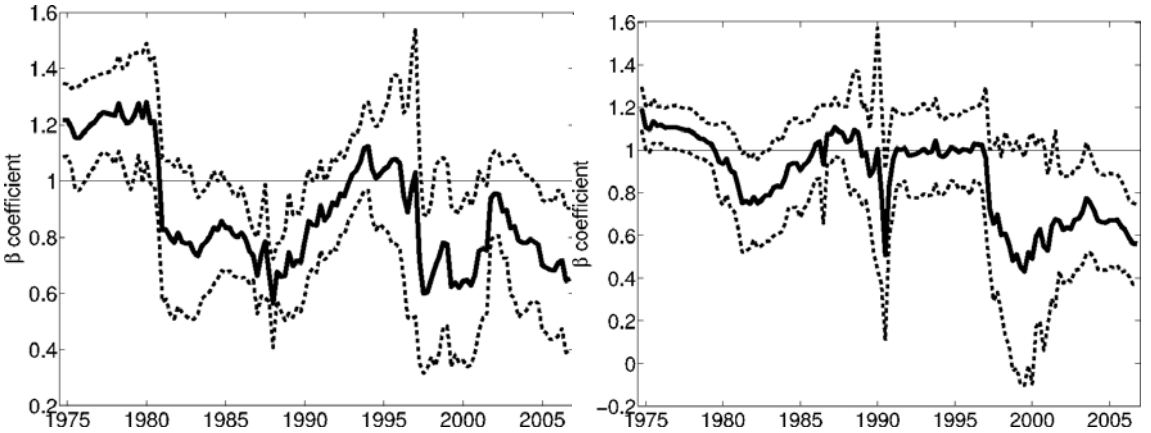
FIGURE 2: Rolling window estimates: Fed Greenbook forecasts

Inflation **Real GDP**

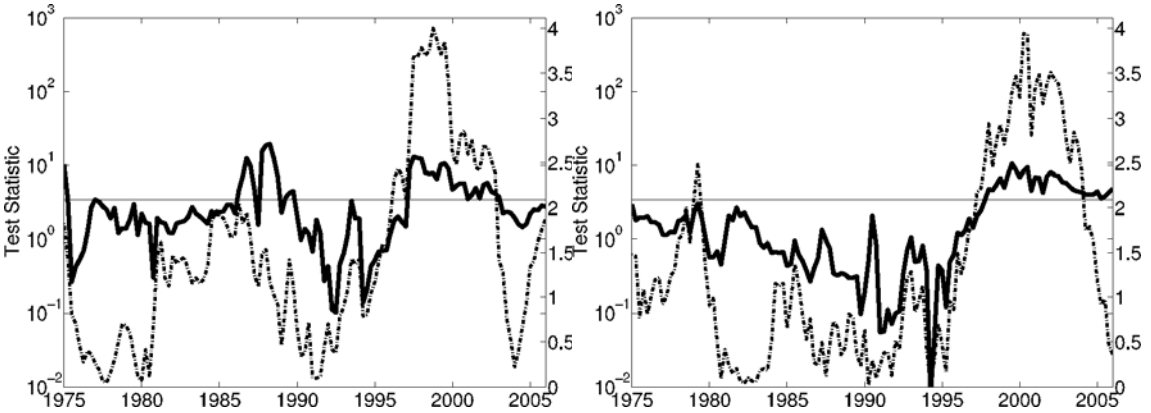
(a) coefficient α



(b) coefficient β



(c) Rolling F-test and HP-test



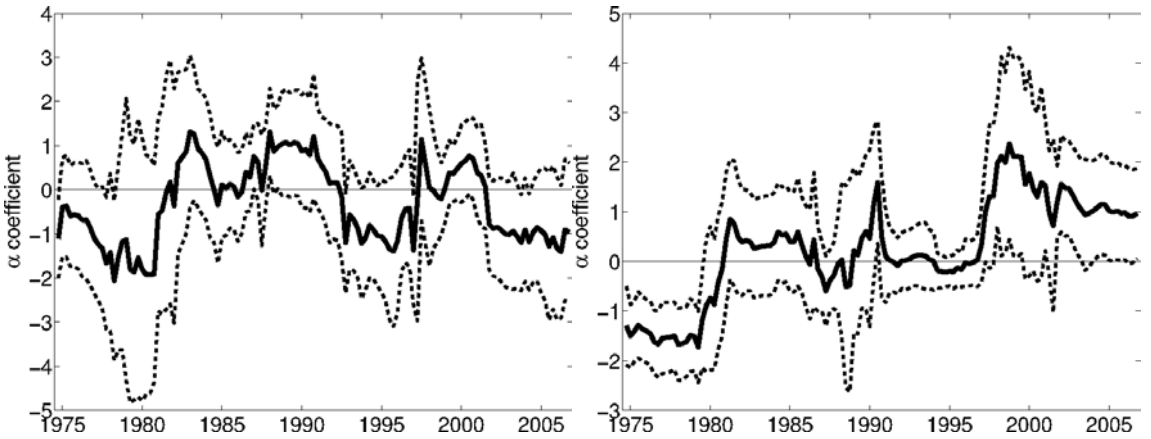
Note: The top four pictures show rolling window estimates for the individual α and β coefficients (solid lines) with corresponding 95% confidence bounds (dashed lines). The window size for the estimation comprises 25 observations. The two bottom pictures show the corresponding evolution of the F-statistic (for the joint hypothesis α=0 and β=1; see solid line) and the HP-test: Holden-Peel test (see dotted line). The axis is scaled so that the horizontal line represents the critical value for both tests.

Source: see data section.

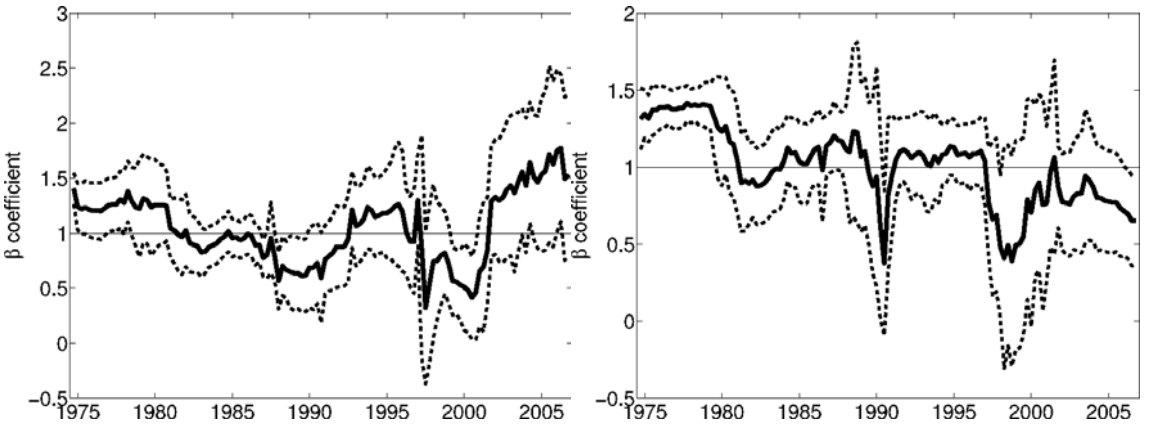
FIGURE 3: Rolling window estimates: SPF forecasts

Inflation **Real GDP**

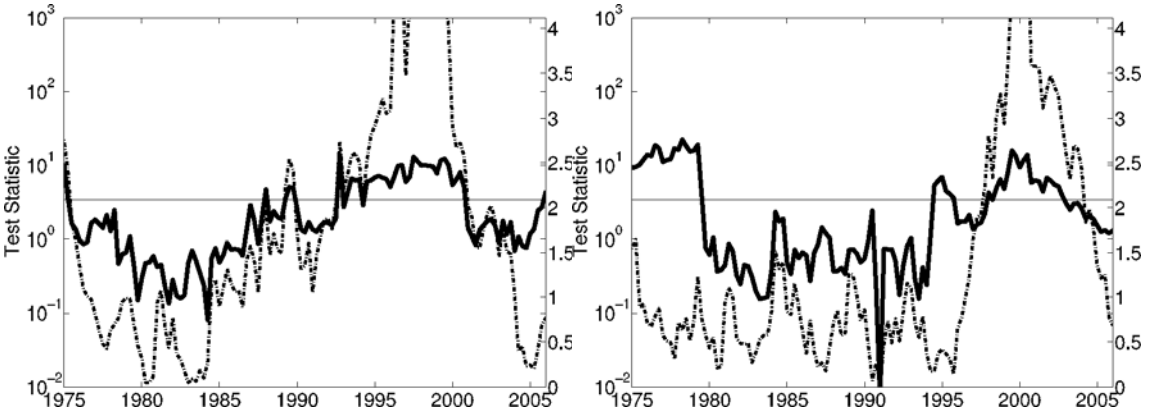
(a) coefficient α



(b) coefficient β



(c) Rolling F-test and HP-test



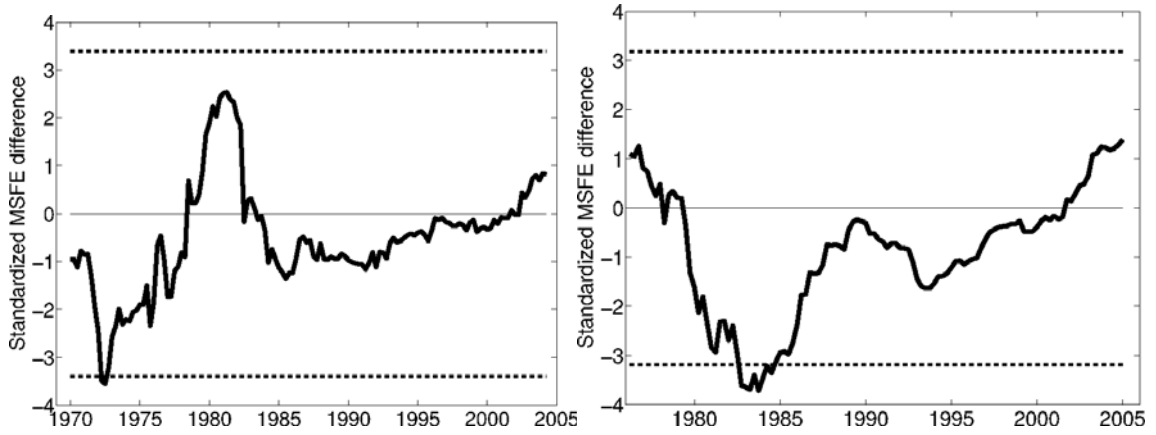
Note: The top four pictures show rolling window estimates for the individual α and β coefficients (solid lines) with corresponding 95% confidence bounds (dashed lines). The window size for the estimation comprises 25 observations. The two bottom pictures show the corresponding evolution of the F-statistic (for the joint hypothesis α=0 and β=1; see solid line) and the HP-test: Holden-Peel test (see dotted line). The axis is scaled so that the horizontal line represents the critical value for both tests.

Source: see data section.

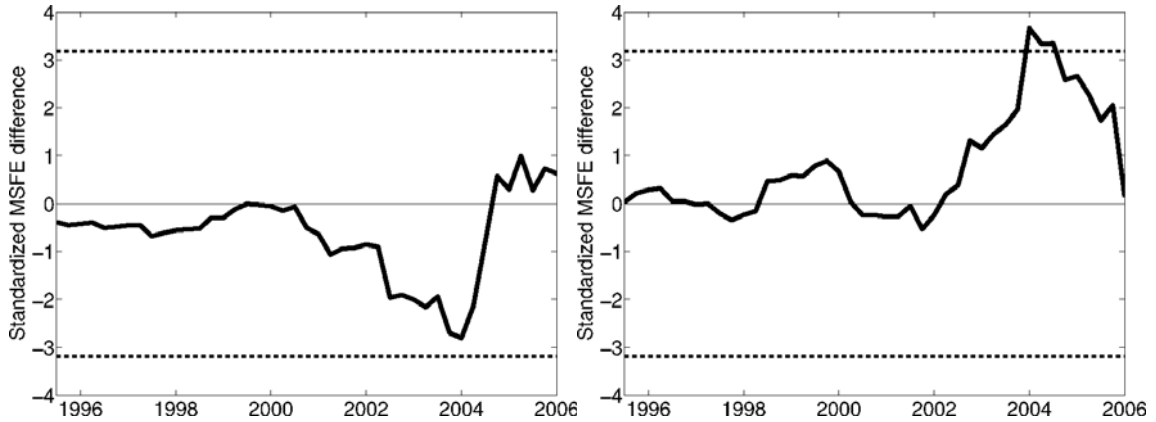
FIGURE 4: Fluctuation test statistic for inflation and output

Nowcast **Forecast four quarters ahead**

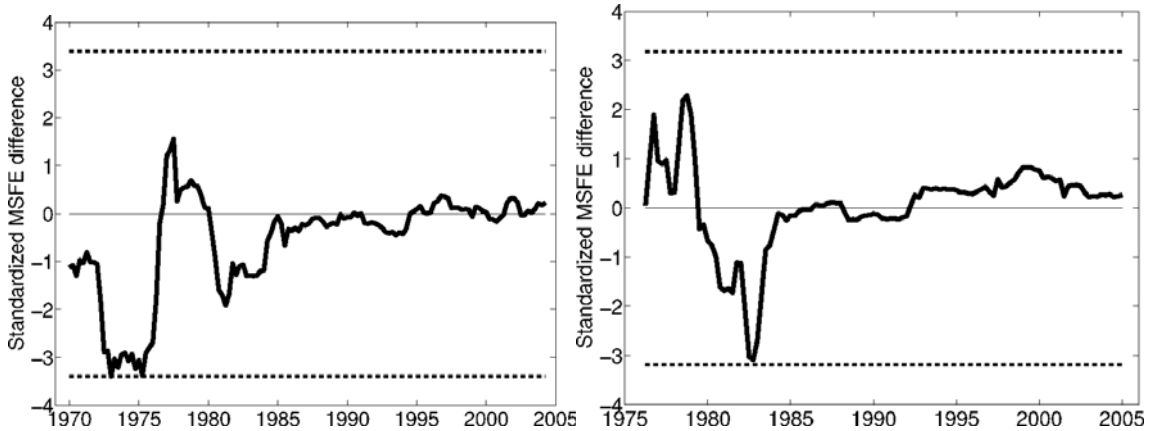
(a) GDP deflator



(b) CPI inflation



(c) Real GDP



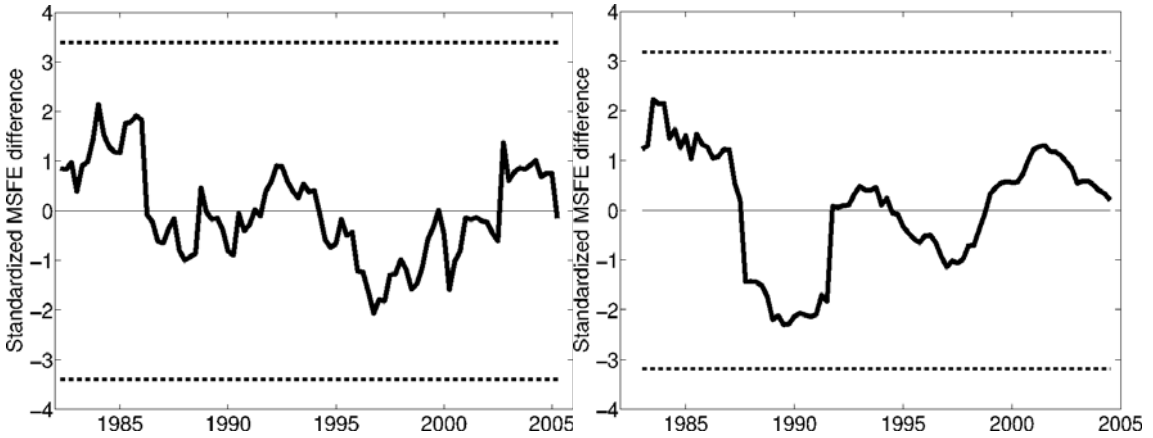
Note: The solid line shows the fluctuation test statistic and the dashed lines represent the corresponding critical values. MSFE: mean squared forecast error.

Source: see data section.

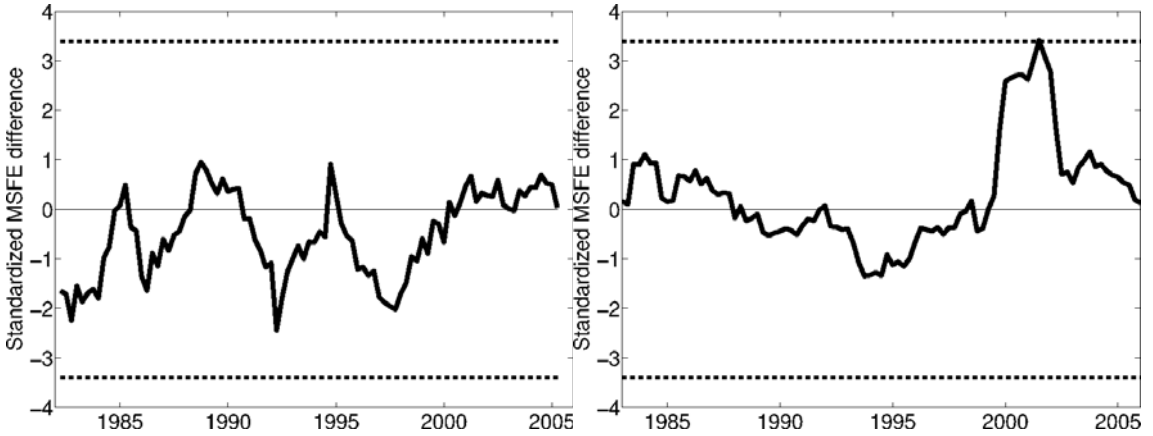
FIGURE 5: Fluctuation test statistic for GDP components

Nowcast **Forecast four quarters ahead**

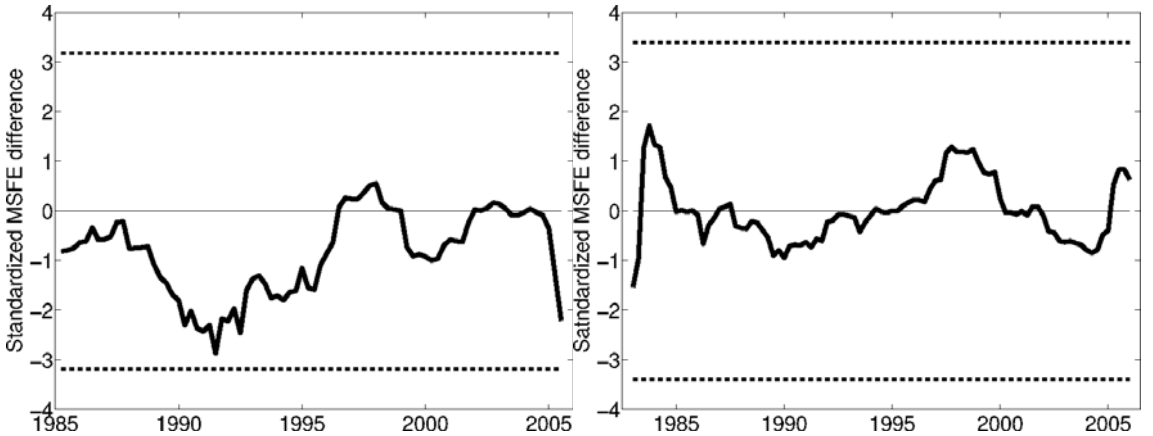
(a) Real consumption



(b) Real fixed business investment



(c) Residential investment



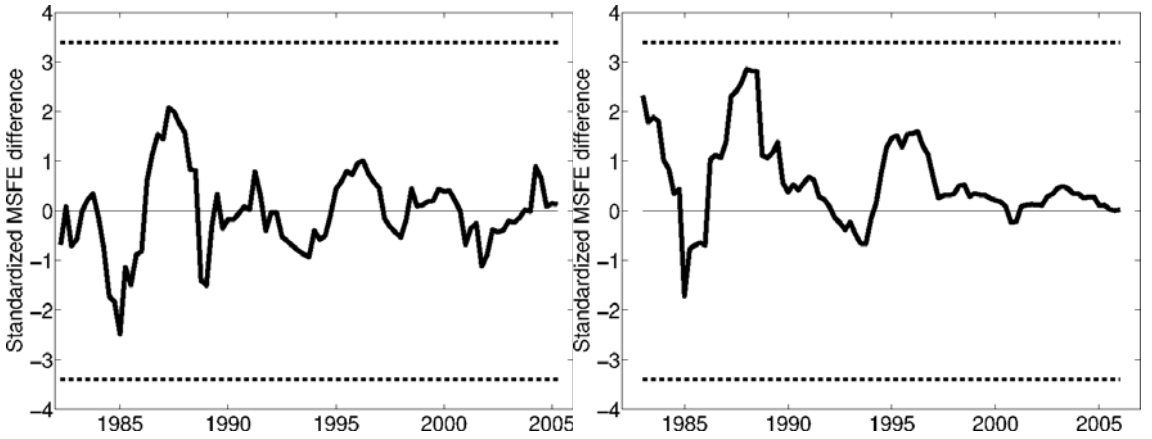
Note: The solid line shows the fluctuation test statistic and the dashed lines represent the corresponding critical values. MSFE: mean squared forecast error.

Source: see data section.

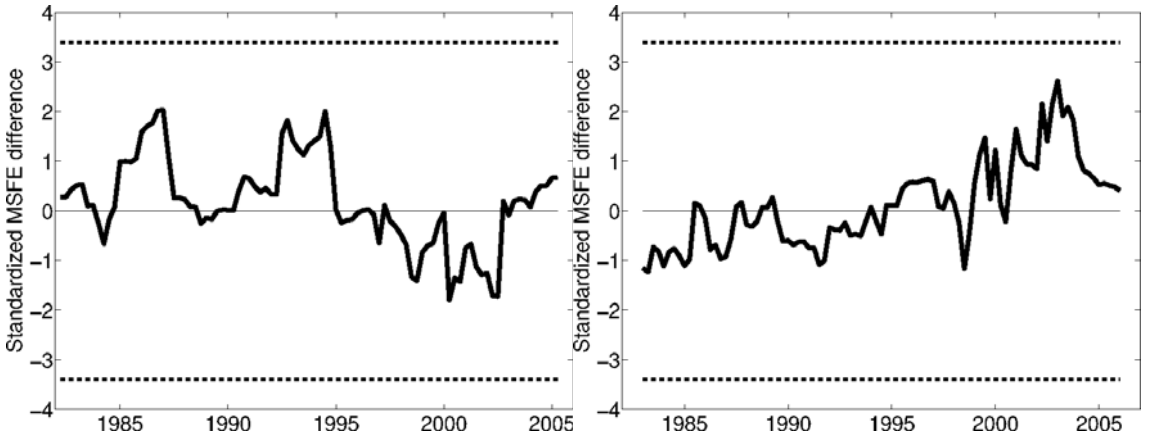
FIGURE 5 (cont.): Fluctuation test statistic for GDP components

Nowcast **Forecast four quarters ahead**

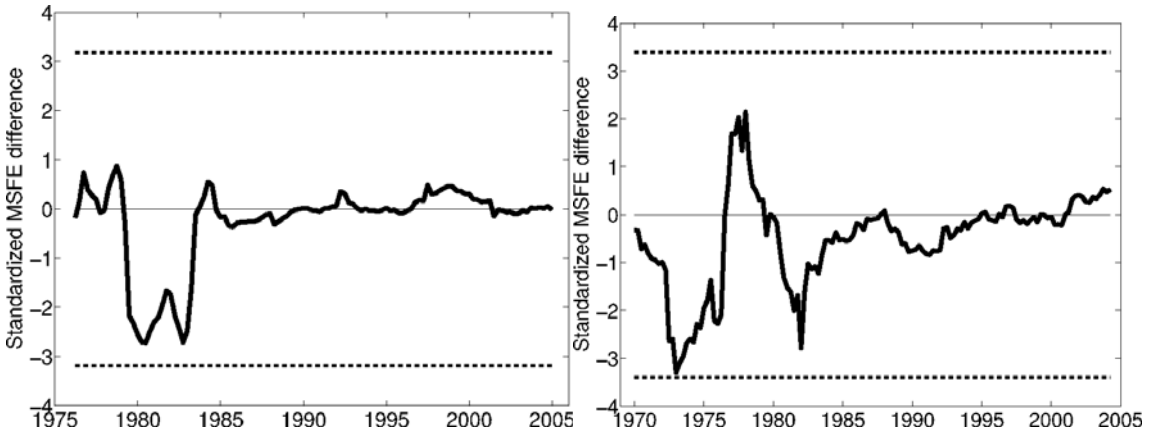
(d) Federal government consumption



(e) Local and State government consumption



(f) Nominal GDP



Note: The solid line shows the fluctuation test statistic and the dashed lines represent the corresponding critical values. MSFE: mean squared forecast error.

Source: see data section.

TABLE 1: Descriptive statistics on forecast accuracy

Horizon	Mean Absolute Error		Root Mean Square Error		Mean Error	
	Greenbook	**SPF**	**Greenbook**	**SPF**	**Greenbook**	**SPF**
Inflation (GDP deflator)						
Nowcast	0.80	0.89	1.04	1.13	-0.09	-0.07
1 quarter ahead	1.01	1.14	1.32	1.51	0.07	-0.03
2 quarters ahead	1.08	1.24	1.50	1.69	0.12	-0.03
3 quarters ahead	1.06	1.37	1.53	1.85	0.08	-0.04
4 quarters ahead	1.01	1.40	1.55	1.99	0.03	-0.06
Real output (GDP)						
Nowcast	1.34	1.50	1.80	1.96	0.13	0.19
1 quarter ahead	1.94	1.85	2.66	2.63	-0.17	-0.12
2 quarters ahead	2.04	2.02	2.85	2.91	-0.32	-0.33
3 quarters ahead	1.99	2.17	3.00	3.19	-0.34	-0.53
4 quarters ahead	1.83	2.06	2.75	3.09	-0.23	-0.61

Notes: Sample 1968Q4 to 2006Q4. SPF: Survey of Professional Forecasters.

TABLE 2: Tests for forecast rationality

Horizon	α		β		R^2		Wald test		HP test		KR test		BP test	
	GB	SPF	GB	SPF	GB	SPF	GB	GB	SPF	SPF	GB	SPF	GB	SPF
Inflation (GDP deflator)														
Nowcast	0.07 (0.22)	-0.30 (0.21)	0.96 (0.06)	1.05 (0.06)	0.67	0.82	0.85	0.30	0.17	0.51	0.51	0.59	80Q4	81Q1
1 quarter ahead	0.06 (0.29)	-0.21 (0.31)	1.00 (0.08)	1.05 (0.09)	0.75	0.68	0.90	0.78	0.65	0.91	0	0	80Q4	80Q4
2 quarters ahead	0.08 (0.30)	-0.20 (0.36)	1.01 (0.09)	1.04 (0.10)	0.68	0.59	0.80	0.86	0.56	0.91	0	0	80Q3	80Q4
3 quarters ahead	0.04 (0.33)	-0.16 (0.47)	1.01 (0.10)	1.03 (0.13)	0.65	0.51	0.92	0.94	0.74	0.87	0	0	80Q2	81Q2
4 quarters ahead	0.01 (0.38)	-0.02 (0.55)	1.01 (0.12)	0.99 (0.14)	0.61	0.43	0.99	0.98	0.86	0.89	0	0	80Q1	81Q1
Real output (GDP)														
Nowcast	0.24 (0.20)	-0.11 (0.23)	0.96 (0.05)	1.12 (0.07)	0.64	0.64	0.49	0.10	0.39	0.22	0.78	0.78	79Q3	79Q3
1 quarter ahead	0.30 (0.34)	-0.21 (0.39)	0.83 (0.10)	1.03 (0.12)	0.33	0.33	0.16	0.83	0.53	0.65	0.19	0.14	79Q2	79Q2
2 quarters ahead	0.10 (0.44)	-0.32 (0.39)	0.85 (0.13)	1.00 (0.17)	0.19	0.19	0.20	0.37	0.32	0.31	0	0	80Q1	79Q1
3 quarters ahead	0.92 (0.54)	0.33 (0.81)	0.56 (0.16)	0.72 (0.22)	0.06	0.06	0.01	0.07	0.23	0.15	0.71	0.22	79Q4	79Q4
4 quarters ahead	0.83 (0.61)	0.05 (0.87)	0.63 (0.19)	0.79 (0.26)	0.06	0.06	0.09	0.04	0.38	0.14	0.56	0.87	92Q4	82Q1

Notes: Sample: 1968Q4 to 2006Q4 (about 150 observations were included). GB: Greenbook forecasts. SPF: Survey of Professional Forecasters. HAC standard errors in brackets. HP test: test by Holden and Peel (1990); KR test: test by Keane and Runkle (1990); BP test: Andrews, Lee and Ploberger (1996) breakpoint test.

TABLE 3: Encompassing test

Horizon	δ	γ^P	γ^S	Adj. R^2
Inflation (GDP deflator)				
Nowcast	-0.11 (0.19)	0.32*** (0.11)	0.69*** (0.13)	0.85
1 quarter ahead	0.11 (0.33)	-0.09 (0.17)	1.08*** (0.16)	0.75
2 quarters ahead	0.26 (0.36)	-0.26 (0.35)	1.23*** (0.36)	0.68
3 quarters ahead	0.21 (0.47)	-0.21 (0.35)	1.19** (0.32)	0.65
4 quarters ahead	0.58 (0.52)	-0.77* (0.45)	1.68*** (0.43)	0.64
Real output (GDP)				
Nowcast	0.18 (0.22)	0.11 (0.22)	0.87*** (0.18)	0.69
1 quarter ahead	-0.07 (0.39)	0.54* (0.29)	0.44* (0.23)	0.35
2 quarters ahead	-0.20 (0.54)	0.29 (0.30)	0.67*** (0.23)	0.23
3 quarters ahead	1.01 (0.81)	-0.06 (0.37)	0.59** (0.25)	0.08
4 quarters ahead	0.09 (0.83)	0.16 (0.31)	0.69*** (0.23)	0.11

Notes: Sample: 1968Q4 to 2006Q4 (about 150 observations were included).
Asterisks mark significance at the one (***), five (**) and ten (*) per cent level.

TABLE 4: Timing test

Horizon	δ	γ^P	γ^S	Adj. R^2
Inflation (GDP deflator)				
Nowcast	-0.29 (0.19)	0.99*** (0.13)	0.06 (0.13)	0.82
1 quarter ahead	-0.17 (0.26)	0.50*** (0.18)	0.55*** (0.17)	0.69
2 quarters ahead	-0.14 (0.28)	0.25 (0.15)	0.80*** (0.13)	0.67
3 quarters ahead	-0.10 (0.30)	0.11 (0.14)	0.93*** (0.12)	0.66
Real output (GDP)				
Nowcast	0.14 (0.25)	1.36*** (0.11)	-0.30** (0.12)	0.66
1 quarter ahead	-0.21 (0.41)	1.04*** (0.22)	-0.01 (0.22)	0.33
2 quarters ahead	-0.38 (0.54)	0.67*** (0.24)	0.35* (0.18)	0.21
3 quarters ahead	0.32 (0.70)	-0.06 (0.24)	0.87*** (0.12)	0.29

Notes: Sample: 1968Q4 to 2006Q4 (about 150 observations were included).
Asterisks mark significance at the one (***), five (**) and ten (*) per cent level.

TABLE 5: Tests for conditional predictive ability

Horizon	Uncertainty					Data revisions		Interest rates	Commo-dity prices
	Inflation dispersion	Real output dispersion	Industrial production dispersion	Housing starts dispersion	ARCH measure	Inflation revisions	Real output revisions	Fed funds rate changes	HWWA-Index
Inflation (GDP deflator)									
Nowcast	0.359	0.226	0.264	0.478	0.270	0.400	0.254	0.112	0.919
1 quarter ahead	**0.025**	**0.022**	**0.002**	**0.049**	**0.027**	0.422	**0.025**	0.151	0.380
2 quarters ahead	**0.085**	**0.080**	**0.012**	0.193	0.135	0.814	0.148	0.108	0.367
3 quarters ahead	**0.052**	**0.018**	**0.016**	**0.000**	**0.001**	0.524	0.652	0.306	0.443
4 quarters ahead	**0.001**	**0.001**	**0.001**	**0.000**	**0.000**	0.556	0.178	**0.009**	0.494
Real output (GDP)									
Nowcast	**0.050**	**0.015**	**0.017**	**0.024**	**0.080**	**0.008**	0.295	**0.081**	0.307
1 quarter ahead	0.241	0.258	0.156	0.293	0.219	0.561	0.966	0.655	0.300
2 quarters ahead	0.878	0.757	0.950	0.820	0.850	0.231	0.595	0.808	0.304
3 quarters ahead	0.798	0.555	0.802	0.810	0.678	0.232	0.470	0.861	0.502
4 quarters ahead	0.767	0.906	0.728	0.519	0.822	0.357	0.342	0.975	0.287

Notes: P-values for the conditional predictive ability test are computed following Giacomini and White (2006). P-values below 10 per cent are marked bold. ARCH measures the volatility in past inflation rate and is used as an alternative uncertainty measure. Fed funds rate changes refers to quarter-on-quarter changes of the Fed's funds rate at the corresponding forecast horizon and HWWA is an index comprising world market prices for energy, oil and raw materials.

REFERENCES

Abreu, Ideberta. (2011). "International Organisations' vs. private analysts' forecasts: an evaluation." Technical Report 20, Banco de Portugal.

Alesina, Alberto, and Lawrence H. Summers. (1993). "Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence." *Journal of Money, Credit and Banking* 25(2), 151-162.

Andrews, Donald W., Inpyo Lee, and Werner Ploberger. (1996). "Optimal changepoint tests for normal linear regression." *Journal of Econometrics* 70(1), 9-38.

Batchelor, Roy. (2000). "The IMF and OECD versus Consensus Forecasts." *Applied Economics* 33(2), 225-235.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. (2008). "Information Cascades." In: Steven N. Durlauf and Lawrence E. Blume (eds.), The New Palgrave Dictionary of Economics, Second Edition, Palgrave Macmillan,U.K.

Blattner, Tobias, Marco Catenaro, Michael Ehrmann, Rolf Strauch, and Jarkko Turunen. (2008)."The Predictability of Monetary Policy." ECB Occasional Paper Series 83.

Brand, Claus, Daniel Buncic, and Jarkko Turunen. (2010). "The Impact of ECB Monetary Policy Decisions and Communication on the Yield Curve." *Journal of the European Economic Association* 8(6), 1266-1298.

D'Agostino, Antenello, and Karl Whelan. (2008). "Federal Reserve information during the Great Moderation." *Journal of the European Economic Association* 6(2-3), 609-620.

El-Shagi, Makram (2011). "Inflation expectations: Does the market beat econometric forecasts?" *The North American Journal of Economics and Finance* 22(3), 298-319.

Ellison, Martin, and Thomas J. Sargent. (2012). "A Defense of the FOMC." *International Economic Review* 53(4), 1047-1065.

Evans, M. D. and Lewis, Karen K. (1995). "Do Expected Shifts in Inflation Affect Estimates of the Long-Run Fisher Relation?" *Journal of Finance* 50, 225-253.

Frankel, Jeffrey A., and Jesse Schreger. (2012). "Over-optimistic official forecasts in the Euro- zone and

fiscal rule." Technical Report 18283, NBER.

Gamber, Edward N. and Julie K. Smith, J. K. (2009). "Are the Fed's inflation forecasts still superior to the private sector's?" *Journal of Macroeconomics* 31(2), 240-251.

Gavin, William T., and Rachel J. Mandal. (2003). "Evaluating FOMC forecasts." *International Journal of Forecasting* 19(4), 655-667.

Giacomini, Raffaella, and Barbara Rossi. (2010). "Forecast comparisons in unstable environments." *Journal of Applied Econometrics* 25(4), 595-620.

Giacomini, Raffaella, and Halbert White. (2006). "Tests of conditional predictive ability." *Econometrica* 74(6), 1545-1578.

Goodfriend, Marvin. (1997). "Monetary policy comes of age: a 20th century odyssey." *Federal Reserve Bank of Richmond Economic Quarterly* 83(1), 1-22.

Holden, Karl, and David A. Peel. (1990). "On testing for unbiasedness and efficiency of forecasts." *The Manchester School* 58(2), 120-127.

Johnson, David R. (1997). "Expected inflation in Canada 1988-1995: an evaluation of Bank of Canada credibility and the effect of Inflation Targets." *Canadian Public Policy* 23(3), 233-258.

Jung, Alexander. (2013). "Policy-makers' interest rate preferences: recent evidence for three monetary policy committees." *International Journal of Central Banking* 9(3), 145-192.

Keane, Michael P., and David E. Runkle. (1990). "Testing the rationality of price forecasts: new evidence from panel data." *American Economic Review* 80(4), 714-735.

Kenny, Geoff, and Julian Morgan. (2011). "Some lessons from the financial crisis for the economic analysis." ECB Occasional Paper Series 130.

Lange, Joe, Brian P. Sack, and William C. Whitesell. (2003). "Anticipations of Monetary Policy in Financial Markets." *Journal of Money, Credit and Banking* 35(6), 889–909.

McCracken, Michael W. (2010). "Disagreement at the FOMC: the dissenting votes are just part of the story." *The Regional Economist*, 10-16.

Moutot, Philippe, Alexander Jung, and Francesco P. Mongelli. (2008). "The Workings of the Eurosystem, Monetary Policy Preparations and Decision-Making - Selected Issues." ECB Occasional

Paper Series 79.

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

Peek, Joe, Eric S. Rosengren, and Geoffrey M.B. Tootell. (2003). “Does the Federal Reserve possess an exploitable informational advantage?” *Journal of Monetary Economics* 50(4), 817-839.

Reifschneider, David, and Peter Tulip. (2007). “Gauging the Uncertainty of the Economic Outlook from Historical Forecasting Errors.” FEDS Papers 2207-60, Board of Governors of the Federal Reserve System.

Romer, Christina D., and David H. Romer. (2000). “Federal Reserve information and the behavior of interest rates.” *American Economic Review* 90(3), 429-457.

Romer, Christina D., and David H. Romer. (2008). “The FOMC versus the Staff: Where can monetary policymakers add value?” *American Economic Review: Papers and Proceedings* 98(2), 230-235.

Rossi, Barbara. (2005). “Optimal tests for nested model selection with underlying parameter instabilities.” *Econometric Theory* 21(5), 962-990.

Rossi, Barbara, and Tatevik Sekhposyan. (2011). “Forecast Optimality Tests in the Presence of Instabilities.” Technical report. Working Paper.

Rudebusch, Glen D., and John C. Williams. (2008). “Revealing the secrets of the temple: the value of publishing central bank interest rate projections.” In: John Y. Campbell (ed.) Asset Prices and Monetary Policy, University of Chicago Press.

Sims, Chris A. (2002). “The role of models and probabilities in the monetary policy process.” *Brookings Papers on Economic Activity* 2002(2), 1-40.

Stockton, David. (2012). “Review of the Monetary Policy Committee’s Forecasting Capability.” Technical report, Bank of England.

Swanson, Eric T. (2006). “ Have Increases in Federal Reserve Transparency Improved Private Sector Interest Rate Forecasts?” *Journal of Money, Credit, and Banking* 38(3), 791-819.

Tillmann, Peter. (2011). “Strategic forecasting on the FOMC.” *European Journal of Political Economy* 27(3), 547–553.

Tulip, Peter. (2009). “Has the economy become more predictable? Changes in Greenbook forecast accuracy.” *Journal of Money, Credit and Banking* 41(6), 1217-1231.

Woodford, Michael. (2005). “Central bank communication and policy effectiveness.” Proceedings, Federal Reserve Bank of Kansas City, August, 399-474.