Evaluating the conditionality of judgmental forecasts

Travis J. Berge, Andrew C. Chang, and Nitish R. Sinha

Evaluating the conditionality of judgmental forecasts[*]

Abstract

We propose a framework to evaluate the conditionality of forecasts. The crux of our framework is the observation that a forecast is conditional if revisions to the conditioning factor are faithfully incorporated into the remainder of the forecast. We consider whether the Greenbook, Blue Chip, and the Survey of Professional Forecasters exhibit systematic biases in the manner in which they incorporate interest rate projections into the forecasts of other macroeconomic variables. We do not find strong evidence of systematic biases in the three economic forecasts that we consider, as the interest rate projections in these forecasts appear to be efficiently incorporated into forecasts of other economic variables.

- *Keywords*: Forecast efficiency; macroeconomic forecasting; conditional forecast.
- *JEL Codes*: C22, C53, E17.

Travis J. Berge, Andrew C. Chang, and Nitish R. Sinha
Board of Governors of the Federal Reserve System
20th and C St, NW
Washington, DC 20551
e-mail: travis.j.berge@frb.gov
e-mail: a.christopher.chang@gmail.com
e-mail: nitish.r.sinha@frb.gov

---

1 Introduction

Given the importance of macroeconomic and financial forecasts, it is no surprise that a large literature evaluates forecasts from the Federal Reserve staff and the private sector. Many papers analyze forecast accuracy, while a more recent literature uses forecasts to measure the importance of informational frictions in the economy.[1] However, macroeconomic forecasts are conditioned on a path for future interest rates, and the implications of conditioning forecasts on a future path for interest rates is not as thoroughly explored in the literature. For example, the fact that the Federal Reserve's Greenbook forecasts perform as well as private sector forecasts is surprising considering that, for long periods of time, the Greenbook explicitly provided policymakers with a forecast of the economy consistent with no change in monetary policy and therefore was conditioned on a flat path for the federal funds rate (Reifschneider, Stockton & Wilcox, 1997; Reifschneider & Tulip, 2017).

This paper develops a framework to assess forecast conditionality. The crux of the framework is the following observation: a forecast is conditional if revisions to the conditioning factor (e.g., the expected path of the policy rate) are incorporated into the remainder of the forecast in a coherent manner. We consider two specific properties. First, when the forecast of the conditioning factor changes, other components of the forecast ought to revise as well. Second, *ex-post* forecast errors should not be related to revisions in the conditioning factor. We name these two properties 'absorption' and 'fidelity'. Absorption refers to the idea that changes to a conditioning factor affect, or are absorbed into, other variables that are forecast. Fidelity is a more precise condition that examines whether changes in conditioning factors are *correctly* incorporated into all other elements of the forecast.

We use our framework to compare three important macroeconomic forecasts: the Greenbook, Blue Chip, and the Survey of Professional Forecasters (SPF). We evaluate how these sets of forecasts incorporate three interest rate projections: the federal funds rate, the three-month Treasury yield, and the 10-year Treasury yield. We choose to focus our analysis on consensus forecasts for two

---

[1] Although the forecast accuracy and statistical efficiency of Greenbook and private sector forecasts are sensitive to the sample period used, a typical result is that the Greenbook forecast is as accurate as consensus forecasts from the private sector. See, for example, Romer & Romer (2000), Sims (2002), Clements, Joutz & Stekler (2007), Faust & Wright (2008), Faust & Wright (2009), Tulip (2009), Caunedo, DiCecio, Komunjer & Owyang (2013), Messina, Sinclair & Stekler (2015), Rossi & Sekhposyan (2016), Sinclair, Stekler & Carnow (2017), and Reifschneider & Tulip (2017), among many others.

reasons. First, these consensus forecasts are important inputs to decision making processes, and understanding their statistical properties is valuable. Second, Romer & Romer (2000) argued that the Greenbook forecasts have an informational advantage over the forecasts from the private sector. As we will see, our framework can speak to whether this informational advantage exists or not.

The conditions we consider are special cases of the Nordhaus (1987) tests of informational efficiency, and our work is related to that subsequent literature. Recently, Lahiri & Sheng (2010) examine the relationship between ex-post forecast errors, disagreement and uncertainty. Dovern & Weisser (2011) evaluate an international panel of professional forecasters, while Messina et al. (2015) evaluate the informational efficiency of the Greenbook forecasts. Coibion & Gorodnichenko (2015) use the approach to discern between noisy and sticky information models. When they evaluate the SPF, they reject that the forecasts are informationally efficient, and attribute this deviation to informational rigidities. Andrade & Bihan (2013), Dovern, Fritsche, Loungani & Tamirisa (2015) and Andrade, Crump, Eusepi & Moench (2016) also explore the role of information frictions in professional forecasts. Typically, papers that perform forecast evaluation consider elements of a forecast one at a time, and this is also our approach.[2]

Our analysis is also related to papers that explicitly consider conditional forecasts. While standard Nordhaus-type regression analyses treat forecasts as though they were unconditional, Faust & Wright (2008) and Clark & McCracken (2014) point out that the forecast errors of conditional forecasts contain two elements: the statistical properties of the *unconditional* forecast errors and those of the conditioning of the forecast. Papers that consider conditional forecasts have typically assumed that the conditioning variable is the federal funds rate. However, the federal funds rate need not be the relevant conditioning interest rate. For example, the Federal Reserve staff may not wish to explicitly forecast the actions of the Federal Open Market Committee (FOMC); private forecasters may well take a similar approach. Similarly, the relevant interest rate for macroeconomic forecasts may not be the policy interest rate but some other interest rate. Our framework can shed light on whether a forecast is conditional for a particular interest rate.

We find that forecasts of macroeconomic variables are clearly related to interest rate projections

---

[2] A smaller number of studies have examined multivariate forecast rationality. Komunjer & Owyang (2012) and Caunedo et al. (2013) develop a methodology, applied to the Greenbook forecasts, that derives the weights of a forecaster's loss function. Sinclair et al. (2017) develop multivariate versions of forecast efficiency regressions and apply them to the Greenbook forecasts.

in the sense that they revise together. In addition, we do not find strong evidence that forecast errors of the macroeconomic projections can be predicted using changes in interest rate projections. On the whole, our evidence suggests that forecasters properly incorporate changes in their interest rates projections into their forecasts of macroeconomic variables.

2   A framework for evaluating forecast conditionality

Consider a forecast of K different quarterly variables, denoted by the vector y, at time t. A forecaster produces point forecasts of the paths of each element in y at horizons $h = 0, 1, ..., H$, where $h = 0$ denotes a current-quarter forecast. Associated with the forecast of period $t + h$ produced at period t is a vector of forecast errors, $e_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$. The tools we propose are relevant for any two elements of the forecast vector, however it is convenient to group the elements of y into two non-overlapping sets: macroeconomic variables (y_x) and financial variables (y_f), and their associated forecast errors, $e_{x,t+h|t}$ and $e_{f,t+h|t}$. Of course, all elements of the forecast are interrelated, but this grouping is natural. For example, the Greenbook forecasts are conditioned on an assumed path of future monetary policy (Reifschneider, Stockton & Wilcox (1997); Pagan & Robertson (2004); Reifschneider & Tulip (2007)). Similarly, respondents to the Blue Chip and the SPF provide joint forecasts of both macroeconomic and financial variables. It is natural to expect that macroeconomic forecasts are affected by changes to financial market projections, and vice versa.

Our first criterion is *absorption*, the idea that revisions to one element of the forecast ought to relate to revisions in the other elements. Define $\Delta\hat{y}_{i,t+h|t} = \hat{y}_{i,t+h|t} - \hat{y}_{i,t+h|t-1}$ as the revision to the forecast of variable i in period $t + h$ produced at t. Absorption implies that revisions to the conditioning variable covary with revisions in other elements of the forecast and the rejection of the following null hypothesis:

$$H_0 : E(\Delta\hat{y}_{x,t+h|t}\Delta\hat{y}_{f,t+h|t}|Z_t) = 0, \tag{1}$$

where Z is a vector of controls that we describe in detail later. Of course, the relationship need not be causal, and we do not interpret it as such.

3

Absorption is related to Clements' concept of attentiveness (Clements 2012). He considers the situation where surveys are conducted between new releases of macroeconomic variables, and tests whether forecasters correctly incorporate economic releases into their projections. Absorption differs from attentiveness in that absorption considers how news about one element of a forecast affects the other elements, whereas attentiveness examines whether forecasters correctly incorporate news about a particular variable itself.

Our second concept, *fidelity*, poses the more rigorous question of whether forecast revisions are correctly incorporated into a forecast. Under the null, forecast revisions in one element of the forecast are unrelated to subsequent forecast errors in other elements:

$$H_0 : E(e_{x,t+h|t} \Delta \hat{y}_{f,t+h|t} | Z_t) = 0, \tag{2}$$

where again Z is a vector of relevant controls. The fidelity condition is the heart of what it means for a forecast to be conditional: improper conditioning on the financial variable results in predictable forecast errors of macroeconomic variables.

These two conditions, absorption and fidelity, constitute our framework for understanding how macroeconomic forecasts are conditioned. Absorption explores the behavior of forecasters, although the inference that one can draw from a rejection (or lack of rejection) of the absorption null is limited. Rejection of the *fidelity* null, however, constitutes evidence against the statistical efficiency of a forecast.

3   Revisiting economic forecasts

We evaluate the historical projections of the Federal Reserve Board staff and two consensus forecasts from the private sector: the Blue Chip Economic Indicators survey (Wolters Kluwer Legal and Regulatory Solutions U.S. 2018) and the Survey of Professional Forecasters (Federal Reserve Bank of Philadelphia 2017*b*). Table 1 summarizes the forecasts we consider. The forecasts differ in their history, frequency and the variables forecasted. Taking these limitations into consideration, we evaluate three macroeconomic variables: real gross domestic product (GDP) growth, the unemployment rate, and price inflation as measured the consumer price index (CPI). The interest

rate projections we consider are the federal funds rate, the three-month Treasury bill yield, and the 10-year Treasury bill yield.

Greenbook forecasts are produced by the Federal Reserve Board staff prior to each FOMC meeting, typically in the first and third month of each quarter.[3] Blue Chip surveys about 50 private sector analysts early each month, with results released around the 10th of that month. We consider the consensus Blue Chip forecast, and our sample of Blue Chip forecasts begins in August 1978, although the forecasts of some indicators begin a few years later. Finally, the SPF is a quarterly survey of academic and private-sector forecasters. Surveys are completed around the beginning of the month in the middle of each quarter. The SPF does not ask respondents to produce federal funds rate projections. The exact timing of the Greenbook, Blue Chip, and SPF differ, so comparisons of these forecasts must be viewed with this caveat in mind.

Table 2 introduces our dataset. It shows the root mean squared forecast error (RMSE) for the six indicators we consider.[4] Asterisks on the RMSEs of Blue Chip and the SPF indicate that they are *inferior* to Greenbook forecasts under squared error loss. Since many of the forecasts we consider begin immediately prior to the disinflation of the early 1980's, and because the findings presented in Rossi & Sekhposyan (2016) and Reifschneider & Tulip (2017) indicate that measures of forecast accuracy are sensitive to the sample period, we prefer to focus on the Great Moderation period. We choose to use data from 1984 onwards; this date is arbitrary but approximates the start of the Great Moderation.[5]

Overall, the tables corroborate Reifschneider & Tulip (2017), who conclude that Greenbook forecasts perform as well as the consensus forecasts produced by the private sector. We find some evidence that Greenbook forecasts are more accurate than the forecasts of the private sector, but the superiority is typically limited to the current quarter. The exception is Greenbook forecasts of the 10-year Treasury yield, which are more accurate than the corresponding forecasts from the Blue Chip.[6] That the Greenbook's 10-year yield forecasts are relatively more accurate is peculiar given

---

[3] Greenbook forecasts are from the database maintained by the Federal Reserve Bank of Philadelphia (2017*a*). We augment some of the financial assumptions data on the Philadelphia Fed website using data from Faust & Wright (2009). Greenbook forecasts are made public with a five year lag, and our dataset ends in 2011.

[4] We use the BEA's second release of real GDP growth to calculate GDP forecast errors. Since revisions to the unemployment rate and CPI are negligible, we use current vintage data. Financial market data do not revise.

[5] Results that use the full sample instead are available from the authors upon request.

[6] Although the SPF has a smaller RMSE, its sample size is also smaller, since respondents to the SPF do not forecast the 10-year yield until 1992.

that Greenbook forecasts of the federal funds rate are not more accurate than their private-sector counterparts.

In an influential paper, Romer & Romer (2000) suggest that the Greenbook may contain an informational advantage over private sector forecasts. Since Greenbook forecasts of future short rates are not superior to private sector forecasts, taken at face value table 2 suggests that if the Greenbook contains an informational advantage, then the source of that advantage is not future monetary policy. We explore how the Greenbook incorporates monetary policy changes by considering the reaction to policy actions from the FOMC itself. We regress forecast revisions onto lagged realized policy changes and controls.[7] We are interested in whether realized policy changes affect the forecasts differently. For each survey j forecast of variable i the regression is:

$$\Delta \hat{y}_{i,t+h|t}^{j} = \alpha + \beta \Delta FFR_{t-1} + \gamma_1 NBER_t + \gamma_2 \Delta \hat{y}_{i,t+h-1|t-1}^{j} + \gamma_3 \Delta \hat{y}_{i,t+h|t}^{-j} + e_{i,t+h|t}, \qquad (3)$$

where ΔFFR_{t-1} is the change to the federal funds rate that occurred at the most recent FOMC meeting, $NBER$ is a dummy variable for recessions and $\Delta \hat{y}_{i,t+h|t}^{j}$ denotes the revision to the forecast of variable i in period $t+h$ from forecaster j. Finally, $\Delta \hat{y}_{i,t+h|t}^{-j}$ is the revision to the other forecast. The coefficient on the lagged federal funds rate describes the average reaction of forecasters to an observed 100 basis point change in the federal funds rate.

As shown in table 3, FOMC interest rate actions accompany revisions to nearly all elements of the Blue Chip projection. When the FOMC moves the federal funds rate higher, Blue Chip forecasters revise their inflation and interest rate projections higher and their unemployment rate projections lower. For the SPF, FOMC rate changes produce revisions to short-term interest rates and unemployment rate projections at longer forecast horizons. Greenbook forecasts, however, react differently. Macroeconomic variables are largely unrevised following FOMC actions, with the exception of inflation forecasts at longer horizons. The lack of revision in macroeconomic variables is somewhat peculiar: for large portions of the sample we consider, the Greenbook forecast was

---

[7] We ensure that forecast revisions for all surveys are measured from the survey directly following each FOMC announcement. For the SPF, we cumulate FOMC actions and revisions to the other forecasts that occur between the SPF survey dates. Most importantly, we need to account for the information that has accumulated between the previous FOMC meeting and the current forecast, and so we include as a regressor the revision to the other forecasts we consider. When evaluating the revision to Greenbook forecasts, we control for on the revision to the Blue Chip forecast; when evaluating private sector forecasts, we control for the revision to Greenbook. We also control for National Bureau of Economic Research (NBER) recession dates (Messina et al. 2015), as well as past revisions to account for stickiness in macroeconomic forecasts (Coibion & Gorodnichenko 2015).

conditioned on a random-walk assumption for the policy rate (Reifschneider et al. 1997). The Greenbook projection for the path of 10-year Treasury yields does not strongly respond to FOMC policy.

The differential pattern of revisions across forecasters to FOMC policy suggest that the forecasts are not conditioned in the same way. The next section further explores the conditionality of these forecasts, and evaluates whether the forecasts exhibit informational inefficiencies with respect to these interest rates.

4   Evaluating the conditionality of judgmental forecasts

We evaluate the absorption condition (1) by regressing revisions to macroeconomic variables onto revisions in each interest rate. Since Sinclair, Joutz & Stekler (2010) and Messina et al. (2015) have found that Greenbook forecast errors are systematically related to the state of the business cycle, we include as a control a dummy variable that indicates whether a quarter was defined as recession by the NBER in our regressions.

Table 4 considers the federal funds rate as the conditioning variable. Macroeconomic forecasts revise in a similar way to changes in the federal funds rate as they did to FOMC policy changes (table 3). Tables 5 and 6 show how macroeconomic forecasts incorporate revisions to the three-month and 10-year Treasury rate projections, respectively. The pattern of revisions is similar to the pattern observed for the federal funds rate. When either Treasury rate revises higher, real GDP growth forecasts also revise higher in the very short-term. Revisions to interest rates at longer forecast horizons tend to correspond with downward revisions to output growth. Inflation and other interest rates revise higher when interest rates move up, and the unemployment rate revises lower. The responses for real GDP and inflation tend to be strongest at far-ahead horizons, consistent with the empirical regularity that monetary policy changes act on the real economy with a lag. Unsurprisingly, the federal funds rate and three-month Treasury projections are very closely related: when one of these interest rates revises, the other also revises nearly one-to-one. The elasticity of the Greenbook's 10-year Treasury projections to its short-term interest rate projections is much smaller than the private sector forecasts, while the responses of the macroeconomic variables to any interest rate revisions are broadly the same across the three surveys.

We evaluate the fidelity condition by regressing forecast errors onto interest rate revisions. Again, regressions include NBER-defined recessions. We also include the subsequent forecast error of the conditioning interest rate in the regression to control for the portion of the forecast error that is related to the concurrent error in the conditioning variable. Tables 7–9 give the results and show that forecasts are, for the most part, faithful to revisions in the projected path of interest rates. That is, the null hypothesis of fidelity is generally not rejected.

For real GDP growth, the point estimates from the fidelity regressions tend to be negative but imprecisely estimated. Inflation projections also do not show a regular pattern of biasedness, with the exception of very short-horizon forecasts from the Blue Chip. There is some evidence that short-term interest rate revisions in the Greenbook predict its unemployment rate forecast errors. The point estimates on forecast errors for the unemployment rate projection in tables 7 and 8 are all negative, and statistically significant at the 10 percent level for longer horizons. When the Greenbook federal funds rate projections revise 25 bps higher, the concurrent unemployment rate projection four quarters ahead is roughly 2-3 basis points too high. The pattern of forecast errors is not driven by large negative revisions to short-term interest rates, since the regression controls for NBER recession dates and the subsequent forecast error to the interest rate itself. In addition, auxiliary regressions (not shown) do not reject the null that the response to revisions differs when the revision is positive versus when the revision is negative.

There is also some evidence of inefficiency between revisions to financial variables and the forecast errors of other interest rates. When the Greenbook forecast of the federal funds rate revises higher, the forecast of the three-month Treasury yield tends to be slightly too high, table 7. Given that revisions to the Greenbook's federal funds rate projection and its three-month Treasury projection move about one-for-one with one another, this finding suggests a tension in how the Greenbook links its interest rate projections with one another. The Greenbook's 10-year Treasury projection largely satisfies the fidelity condition; revisions to the federal funds rate do not predict forecast errors in the 10-year rate, although the 10-year Treasury projection is somewhat too low when the Greenbook revises its three-month Treasury rate. For the private sector forecasts the opposite is true: when the consensus forecasts revise short-term interest rates higher, their 10-year projection is subsequently too high, especially for the SPF. These results may reflect the difficulties faced by a central bank when communicating its expected path of policy. For example, this pattern

of revisions would result if a central bank raised interest rates at a modest but persistent pace, but the private sector did not understand this intention. Alternatively, the results could reflect different interpretations between the Federal Reserve staff and private sector forecasters of the secular decline in long-term interest rates during the period. For example, perhaps the Greenbook forecast internalized this secular decline earlier than private sector forecasts.

5   Conclusion

In this paper, we developed a framework that evaluates whether forecasts are correctly conditioned. The framework first examines whether macroeconomic and interest rate projections are related by regressing forecast revisions from one variable onto the revisions of the other, which we call absorption. We then test whether macroeconomic forecasts are properly conditioned by testing the null that *ex-post* forecast errors and interest rate revisions are unrelated. We denote this condition fidelity, since proper conditioning means that the details of the remainder of the forecast are correctly implemented. The tests we propose are straightforward but useful diagnostic tools for macroeconomic forecasts.

Since the economic forecasts from the Greenbook, Blue Chip, and the SPF are used to inform a wide variety of decisions, it is important to understand the statistical properties of these projections. Unsurprisingly, we find that macroeconomic projections tend to revise when interest rate forecasts change, indicating that forecasters believe that these variables are related. We do not find strong evidence of systematic biases in the three economic forecasts that we consider. In most cases, changes in interest rate variables do not predict subsequent forecast errors for other economic forecasts.

References

Andrade, P. & Bihan, H. L. (2013), 'Inattentive professional forecasters', *Journal of Monetary Economics* **60**(8), 967 – 982.

Andrade, P., Crump, R. K., Eusepi, S. & Moench, E. (2016), 'Fundamental disagreement', *Journal of Monetary Economics* **83**, 106 – 128.

Caunedo, J., DiCecio, R., Komunjer, I. & Owyang, M. T. (2013), Federal reserve forecasts: asymmetry and state-dependence, Working papers 2013-12, Federal Reserve Bank of St. Louis.

Clark, T. E. & McCracken, M. W. (2014), Evaluating Conditional Forecasts from Vector Autoregressions, Working Papers 2014-25, Federal Reserve Bank of St. Louis.

Clements, M. P. (2012), 'Do professional forecasters pay attention to data releases?', *International Journal of Forecasting* **28**(2), 297–308.

Clements, M. P., Joutz, F. & Stekler, H. O. (2007), 'An evaluation of the forecasts of the federal reserve: a pooled approach', *Journal of Applied Econometrics* **22**(121-136).

Coibion, O. & Gorodnichenko, Y. (2015), 'Information rigidity and the expectations formation process: A simple framework and new facts', *American Economic Review* **105**(8), 2644–2678.

Diebold, F. X. & Mariano, R. S. (1995), 'Comparing Predictive Accuracy', *Journal of Business & Economic Statistics* **13**(3), 253–263.

Dovern, J., Fritsche, U., Loungani, P. & Tamirisa, N. (2015), 'Information rigidities: Comparing average and individual forecasts for a large international panel', *International Journal of Forecasting* **31**(1), 144 – 154.

Dovern, J. & Weisser, J. (2011), 'Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: an empirical comparison for the g7', *International Journal of Forecasting* **27**(452-465).

Faust, J. & Wright, J. (2008), 'Efficient forecast tests for conditional policy forecasts', *Journal of Econometrics* **146**(2), 293–303.

Faust, J. & Wright, J. H. (2009), 'Comparing greenbook and reduced form forecasts using a large realtime dataset', *Journal of Business & Economic Statistics* **27**(4), 468–479.

Federal Reserve Bank of Philadelphia (2017*a*), 'Greenbook data sets'. Accessed: 2017-10-19.
    URL: *https://www.philadelphiafed.org/research-and-data/real-time-center/greenbook-data*

Federal Reserve Bank of Philadelphia (2017*b*), 'Survey of professional forecasters'. Accessed: 2017-11-28.
    URL: *https://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters*

Komunjer, I. & Owyang, M. T. (2012), 'Multivariate forecast evaluation and rationality testing', *Review of Economics and Statistics* **94**(4), 1066–1080.

Lahiri, K. & Sheng, X. (2010), 'Measuring forecast uncertainty by disagreement: The missing link', *Journal of Applied Econometrics* **25**(4), 514–538.

Messina, J. D., Sinclair, T. & Stekler, H. (2015), 'What can we learn from revisions to the greenbook forecasts?', *Journal of Macroeconomics* **45**(C), 54–62.

Nordhaus, W. D. (1987), 'Forecasting efficiency: Concepts and applications', *The Review of Economics and Statistics* **69**(4), 667–74.

Pagan, A. R. & Robertson, J. (2004), Forecasting for policy, *in* M. P. Clements & D. F. Hendry, eds, 'A Companion to Economic Forecasting', Blackwell Publishing, Malden, MA.

Reifschneider, D. L., Stockton, D. J. & Wilcox, D. W. (1997), 'Econometric models and the monetary policy process', *Carnegie-Rochester Conference Series on Public Policy* **47**(1), 1–37.

Reifschneider, D. & Tulip, P. (2007), Gauging the uncertainty of the economic outlook from historical forecasting errors, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs 2007-60, Federal Reserve Board.

Reifschneider, D. & Tulip, P. (2017), Gauging the uncertainty of the economic outlook using historical forecasting errors: The federal reserve's approach, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs 020, Federal Reserve Board.

Romer, C. D. & Romer, D. H. (2000), 'Federal reserve information and the behavior of interest rates', *American Economic Review* **90**(3), 429–547.

Rossi, B. & Sekhposyan, T. (2016), 'Forecast rationality tests in the presence of instabilities, with applications to the federal reserve and survey forecasts', *Journal of Applied Econometrics* **31**, 507–532.

Sims, C. A. (2002), 'The role of models and probabilities in the monetary policy process', *Brookings Papers on Economic Activity* **33**(2), 1–62.

Sinclair, T. M., Joutz, F. & Stekler, H. (2010), 'Can the fed predict the state of the economy?', *Economics Letters* **108**(1), 28 – 32.

Sinclair, T. M., Stekler, H. O. & Carnow, W. (2017), 'Evaluating a vector of the fed's forecasts', *International Journal of Forecasting* **31**, 157–164.

Tulip, P. (2009), 'Has output become more predictable? changes in greenbook forecast accuracy', *Journal of Money, Credit & Banking* **41**(6), 1217–1231.

Wolters Kluwer Legal and Regulatory Solutions U.S. (2018), 'Blue chip economic indicators'.

	Greenbook	Blue Chip	SPF
Source	FRB-Philadelphia	Wolters Kluwer Legal and Regulatory Solutions U.S.	FRB-Philadelphia
Survey frequency	Typically 2x quarter	Monthly	Quarterly
Period available			
Real GDP	1972m1-2011m12	1978m8-2011m12	1972Q1-2011Q4
CPI	1979m10-2011m12	1980m2-2011m12	1981Q3-2011Q4
Unemployment rate	1972m1-2011m12	1980m3-2011m12	1972Q1-2011Q4
Fed funds rate	1981m1-2011m12	1982m10-2011m12	–
Treasury bill	1980m3-2011m12	1980m10-2011m12	1981Q3-2011Q4
Treasury bond	1980m1-2011m12	1982m1-2011m12	1992Q1-2011Q4

Table 1: Data source and availability.

Source	Horizon	Real GDP		CPI		Un. rate		Fed funds		Treasury bill		Treasury bond	
		N.obs	RMSE	N.obs	RMSE	N.obs	RMSE	N.obs	RMSE	N.obs	RMSE	N.obs	RMSE
Greenbook	0	222	1.53	224	.76	224	.15	224	.30	224	.15	210	.33
	1	220	1.97	222	1.71	222	.30	222	.66	222	.50	208	.60
	2	218	2.17	220	2.09	220	.46	220	1.01	220	.88	206	.79
	3	216	2.25	218	2.09	218	.65	218	1.37	218	1.21	204	.94
	4	214	2.27	216	1.99	216	.82	216	1.67	216	1.51	202	1.05
Blue Chip	0	333	1.76**	336	1.23**	336	.19*	336	.40	336	.24**	336	.71**
	1	330	1.93	333	1.89	333	.34*	333	.75*	333	.60*	333	.94**
	2	327	2.07	330	1.99	330	.50*	330	1.10	330	.95*	330	1.14**
	3	324	2.17	327	2.00	327	.69	327	1.41	327	1.29	327	1.31**
	4	321	2.23	324	2.02	324	.87	324	1.72	324	1.61	324	1.47**
SPF	0	111	1.72	112	1.13	112	.31**	0	—	112	.15	80	.18
	1	110	1.92	111	1.89	111	.38	0	—	111	.52	79	.57
	2	109	2.09	110	1.96	110	.53	0	—	110	.90	78	.77
	3	108	2.20	109	1.98	109	.68	0	—	109	1.24	77	.90
	4	107	2.23	108	2.00	108	.86	0	—	108	1.57	76	1.05

Table 2: Root mean squared forecast errors for Greenbook, Blue Chip, and SPF forecasts, 1984–2011. Asterisks for Blue Chip and SPF forecasts indicate that the private-sector forecasts are *outperformed* by the Greenbook forecast at 10 percent (*) and 5 percent (**) confidence interval, respectively. Tests performed with a one-sided Diebold-Mariano test statistic, assuming squared error loss and using Newey-West standard errors. Sources: Federal Reserve Bank of Philadelphia (2017a), Federal Reserve Bank of Philadelphia (2017b), Wolters Kluwer Legal and Regulatory Solutions U.S. (2018).

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	0.13	0.07	-0.04	0.15	-0.28*
	(0.24)	(0.22)	(0.18)	(0.19)	(0.15)
Blue Chip	0.70**	0.17	-0.11	-0.15**	-0.03
	(0.22)	(0.17)	(0.08)	(0.07)	(0.09)
SPF	0.65**	0.10	0.00	-0.14	-0.20*
	(0.21)	(0.12)	(0.12)	(0.13)	(0.10)
			CPI		
Greenbook	-0.16	-0.28	0.03	0.15**	0.11**
	(0.26)	(0.32)	(0.11)	(0.06)	(0.05)
Blue Chip	0.43**	0.25**	0.20**	0.07	0.13**
	(0.13)	(0.09)	(0.07)	(0.07)	(0.06)
SPF	0.06	0.04	0.06	0.02	0.08*
	(0.23)	(0.09)	(0.04)	(0.04)	(0.05)
			Unemployment rate		
Greenbook	0.07	0.06	0.07	0.05	0.09
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
Blue Chip	-0.13**	-0.17**	-0.18**	-0.14**	-0.15**
	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)
SPF	0.01	-0.05	-0.09*	-0.08*	-0.11**
	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)
			Federal funds rate		
Greenbook	0.08	0.43**	0.59**	0.57**	0.45**
	(0.16)	(0.16)	(0.17)	(0.14)	(0.12)
Blue Chip	0.48**	0.86**	0.98**	0.94**	0.87**
	(0.15)	(0.15)	(0.14)	(0.13)	(0.11)
SPF	–	–	–	–	–
			Three-month yield		
Greenbook	0.11	0.34**	0.44**	0.44**	0.39**
	(0.10)	(0.11)	(0.12)	(0.11)	(0.10)
Blue Chip	0.37**	0.69**	0.71**	0.69**	0.59**
	(0.10)	(0.15)	(0.13)	(0.11)	(0.10)
SPF	0.52**	0.75**	0.75**	0.64**	0.53**
	(0.07)	(0.09)	(0.10)	(0.10)	(0.10)
			10-year yield		
Greenbook	0.10	0.07	0.14	0.15	0.14
	(0.13)	(0.13)	(0.12)	(0.11)	(0.11)
Blue Chip	0.21**	0.24**	0.21**	0.18*	0.19**
	(0.08)	(0.09)	(0.10)	(0.10)	(0.09)
SPF	0.13	0.16	0.18*	0.12	0.14**
	(0.10)	(0.10)	(0.09)	(0.08)	(0.07)

Table 3: Estimates of the response of forecasts to FOMC policy changes, 1984–2011. Table shows estimated β coefficient from regression: $\Delta \hat{y}_{t+h|t} = \alpha + \beta \Delta FFR_{t-1} + \gamma_1 NBER_t + \gamma_2 \hat{y}_{t+h-1|t-1}^j + \gamma_3 \Delta \hat{y}_{t+h|t}^{-j} + u_{t+h}$. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level. See text for details.

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	-0.07	0.26	0.26**	0.02	-0.18**
	(0.24)	(0.16)	(0.12)	(0.09)	(0.08)
Blue Chip	-0.05	-0.01	-0.09	-0.04	-0.11**
	(0.17)	(0.13)	(0.07)	(0.06)	(0.05)
			CPI		
Greenbook	0.41	0.01	0.15**	0.16**	0.10**
	(0.34)	(0.12)	(0.04)	(0.04)	(0.04)
Blue Chip	0.19	0.27**	0.18**	0.15**	0.13**
	(0.17)	(0.06)	(0.03)	(0.03)	(0.02)
			Unemployment rate		
Greenbook	-0.16**	-0.15**	-0.20**	-0.20**	-0.19**
	(0.05)	(0.03)	(0.03)	(0.04)	(0.05)
Blue Chip	-0.16**	-0.19**	-0.22**	-0.24**	-0.23**
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
			Three-month yield		
Greenbook	0.79**	0.93**	0.93**	0.92**	0.93**
	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
Blue Chip	0.72**	0.73**	0.72**	0.71**	0.63**
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)
			10-year yield		
Greenbook	0.37**	0.31**	0.31**	0.30**	0.30**
	(0.08)	(0.06)	(0.05)	(0.04)	(0.04)
Blue Chip	0.77**	0.75**	0.75**	0.74**	0.80**
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)

Table 4: How do other forecasts revise when the federal funds rate projection revises? Table shows estimated β coefficient from regression: $\Delta\hat{y}_{t+h|t} = \alpha + \beta\Delta\hat{y}_{ffr,t+h|t} + \gamma NBER_{t+h} + u_{t+h}$. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level. Sample period is 1984–2011. Note that the SPF does not forecast the federal funds rate. See text for details.

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	0.44	0.46**	0.33**	0.02	-0.17**
	(0.30)	(0.18)	(0.11)	(0.10)	(0.08)
Blue Chip	0.49**	0.27*	-0.03	-0.02	-0.08
	(0.20)	(0.15)	(0.09)	(0.07)	(0.06)
SPF	0.55**	0.29	0.03	-0.07	-0.24**
	(0.26)	(0.22)	(0.11)	(0.12)	(0.08)
			CPI		
Greenbook	0.69	0.09	0.17**	0.17**	0.12**
	(0.51)	(0.13)	(0.04)	(0.04)	(0.04)
Blue Chip	0.54**	0.17	0.23**	0.21**	0.19**
	(0.27)	(0.14)	(0.06)	(0.04)	(0.03)
SPF	0.76*	0.34**	0.24**	0.20**	0.13**
	(0.36)	(0.09)	(0.05)	(0.04)	(0.03)
			Unemployment rate		
Greenbook	-0.26**	-0.18**	-0.23**	-0.22**	-0.21**
	(0.06)	(0.03)	(0.03)	(0.04)	(0.04)
Blue Chip	-0.19**	-0.21**	-0.23**	-0.29**	-0.31**
	(0.04)	(0.04)	(0.06)	(0.05)	(0.05)
SPF	-0.11*	-0.23**	-0.28**	-0.33**	-0.32**
	(0.06)	(0.05)	(0.06)	(0.07)	(0.08)
			Federal funds rate		
Greenbook	1.01**	0.98**	1.02**	1.03**	1.03**
	(0.04)	(0.03)	(0.02)	(0.03)	(0.02)
Blue Chip	1.00**	0.84**	0.85**	0.84**	0.83**
	(0.06)	(0.09)	(0.10)	(0.09)	(0.10)
SPF	–	–	–	–	–
			10-year yield		
Greenbook	0.38**	0.31**	0.32**	0.31**	0.33**
	(0.09)	(0.06)	(0.05)	(0.04)	(0.05)
Blue Chip	0.55**	0.61**	0.62**	0.59**	0.57**
	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)
SPF	0.57**	0.70**	0.69**	0.57**	0.53**
	(0.09)	(0.08)	(0.07)	(0.07)	(0.06)

Table 5: How do other forecasts revise when the three-month Treasury yield projection revises? Table shows estimated β coefficient from regression: $\Delta \hat{y}_{t+h|t} = \alpha + \beta \Delta \hat{y}_{i3m,t+h|t} + \gamma NBER_{t+h} + u_{t+h}$. Sample period is 1984–2011. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level. See text for details.

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	0.42*	0.28	0.01	-0.13	-0.16
	(0.23)	(0.20)	(0.12)	(0.15)	(0.12)
Blue Chip	0.51**	0.18	-0.11	-0.08	-0.10
	(0.20)	(0.15)	(0.12)	(0.09)	(0.06)
SPF	1.32**	0.93**	0.23*	-0.11	-0.18*
	(0.37)	(0.26)	(0.12)	(0.11)	(0.11)
			CPI		
Greenbook	0.52	0.16	0.12	0.06	0.07
	(0.36)	(0.33)	(0.12)	(0.05)	(0.04)
Blue Chip	0.42*	0.10	0.20**	0.19**	0.19**
	(0.25)	(0.10)	(0.05)	(0.04)	(0.04)
SPF	0.65**	0.19**	0.25**	0.23**	0.19**
	(0.23)	(0.06)	(0.05)	(0.05)	(0.05)
			Unemployment rate		
Greenbook	-0.09*	-0.08*	-0.13**	-0.13**	-0.12*
	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)
Blue Chip	-0.14**	-0.18**	-0.19**	-0.21**	-0.20**
	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)
SPF	0.02	-0.14**	-0.21**	-0.28**	-0.33**
	(0.10)	(0.08)	(0.09)	(0.11)	(0.13)
			Federal funds rate		
Greenbook	0.35**	0.32**	0.45**	0.53**	0.52**
	(0.08)	(0.07)	(0.09)	(0.11)	(0.11)
Blue Chip	0.56**	0.57**	0.64**	0.67**	0.72**
	(0.07)	(0.09)	(0.08)	(0.09)	(0.07)
SPF	–	–	–	–	–
			Three-month yield		
Greenbook	0.34**	0.30**	0.42**	0.48**	0.51**
	(0.07)	(0.07)	(0.08)	(0.10)	(0.10)
Blue Chip	0.64**	0.75**	0.79**	0.80**	0.84**
	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
SPF	0.43**	0.58**	0.74**	0.83**	0.92**
	(0.10)	(0.11)	(0.11)	(0.12)	(0.13)

Table 6: How do other forecasts revise when the 10-year Treasury yield projection revises? Table shows estimated β coefficient from regression: $\Delta \hat{y}_{t+h|t} = \alpha + \beta \Delta \hat{y}_{i10y,t+h|t} + \gamma NBER_{t+h} + u_{t+h}$. Sample period is 1984–2011. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level. See text for details.

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	-0.47	-0.14	-0.09	0.18	-0.12
	(0.44)	(0.47)	(0.27)	(0.31)	(0.44)
Blue Chip	-0.11	-0.16	0.30	-0.34	-0.27
	(0.40)	(0.53)	(0.36)	(0.44)	(0.53)
			CPI		
Greenbook	0.07	0.15	-0.25	0.42	0.29
	(0.17)	(0.36)	(0.47)	(0.44)	(0.29)
Blue Chip	0.22	-0.16	-0.59	0.54	0.54
	(0.32)	(0.49)	(0.49)	(0.62)	(0.47)
			Unemployment rate		
Greenbook	-0.02	-0.08	-0.13	-0.18*	-0.23*
	(0.04)	(0.07)	(0.10)	(0.10)	(0.12)
Blue Chip	-0.03	-0.03	-0.13	-0.23	-0.30
	(0.04)	(0.07)	(0.11)	(0.16)	(0.20)
			Three-month yield		
Greenbook	-0.02	-0.11**	-0.11**	-0.06	-0.07
	(0.03)	(0.06)	(0.04)	(0.06)	(0.05)
Blue Chip	-0.03	0.06	0.13	0.17*	0.14*
	(0.05)	(0.08)	(0.12)	(0.09)	(0.08)
			10-year yield		
Greenbook	0.10	0.29	0.16	0.09	0.07
	(0.10)	(0.18)	(0.18)	(0.18)	(0.16)
Blue Chip	0.03	-0.19	-0.25	-0.35*	-0.33
	(0.13)	(0.17)	(0.20)	(0.20)	(0.25)

Table 7: Do forecasts maintain fidelity to the federal funds rate projection? Table shows estimated β coefficient from regression: $e_{i,t+h|t} = \alpha + \beta \Delta \hat{y}_{ffr,t+h|t} + \gamma_1 NBER_{t+h} + \gamma_2 (y_{ffr,t+h} - \hat{y}_{ffr,t+h|t}) + u_{t+h}$. Sample period is 1984–2011. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level. Note the SPF does not forecast the federal funds rate. See text for details.

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	-0.83*	-0.23	-0.03	0.23	-0.13
	(0.43)	(0.42)	(0.29)	(0.33)	(0.49)
Blue Chip	-0.26	-0.23	0.32	-0.77*	-0.69
	(0.51)	(0.50)	(0.41)	(0.45)	(0.59)
SPF	-0.13	-0.47	0.17	-0.23	0.00
	(0.50)	(0.44)	(0.28)	(0.44)	(0.47)
			CPI		
Greenbook	0.28*	0.38	-0.09	0.55	0.41
	(0.16)	(0.32)	(0.52)	(0.52)	(0.32)
Blue Chip	0.78*	-0.02	-0.72	0.46	-0.30
	(0.42)	(0.56)	(0.66)	(0.61)	(0.76)
SPF	0.21	-0.29	-0.15	0.57	0.58
	(0.24)	(0.50)	(0.38)	(0.53)	(0.45)
			Unemployment rate		
Greenbook	-0.03	-0.11	-0.16	-0.19*	-0.25*
	(0.04)	(0.07)	(0.10)	(0.11)	(0.14)
Blue Chip	-0.05	-0.02	-0.11	-0.14	-0.30
	(0.05)	(0.07)	(0.11)	(0.17)	(0.21)
SPF	-0.09*	-0.05	-0.11	-0.17	-0.29*
	(0.05)	(0.05)	(0.08)	(0.12)	(0.15)
			Federal funds rate		
Greenbook	0.38**	0.23**	0.17**	0.12	0.12**
	(0.07)	(0.06)	(0.06)	(0.07)	(0.06)
Blue Chip	0.42**	0.24**	0.18	0.14	0.21**
	(0.07)	(0.09)	(0.14)	(0.12)	(0.09)
SPF	–	–	–	–	–
			10-year yield		
Greenbook	0.28**	0.40**	0.26	0.14	0.13
	(0.13)	(0.17)	(0.18)	(0.18)	(0.17)
Blue Chip	0.13	-0.17	-0.32*	-0.44**	-0.41
	(0.15)	(0.17)	(0.18)	(0.19)	(0.26)
SPF	-0.07**	-0.27**	-0.42**	-0.57**	-0.53**
	(0.03)	(0.10)	(0.12)	(0.13)	(0.19)

Table 8: Do forecasts maintain fidelity to the three-month Treasury yield projection? Table shows estimated β coefficient from regression: $e_{i,t+h|t} = \alpha + \beta\Delta\hat{y}_{i3m,t+h|t} + \gamma_1 NBER_{t+h} + \gamma_2(y_{i3m,t+h} - \hat{y}_{i3m,t+h|t}) + u_{t+h}$. Sample period is 1984–2011. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level.

Forecast horizon	0	1	2	3	4
			Real GDP		
Greenbook	-0.37	0.03	0.61	0.46	-0.14
	(0.38)	(0.36)	(0.40)	(0.39)	(0.51)
Blue Chip	0.37	-0.12	0.66*	-0.38	-1.51**
	(0.41)	(0.47)	(0.38)	(0.45)	(0.49)
SPF	-0.64*	-0.54	0.62	-0.02	-0.27
	(0.35)	(0.50)	(0.45)	(0.60)	(0.64)
			CPI		
Greenbook	-0.09	0.28	-0.15	0.53	0.35
	(0.15)	(0.28)	(0.52)	(0.41)	(0.33)
Blue Chip	0.78**	0.73*	-0.18	0.30	-0.01
	(0.32)	(0.37)	(0.57)	(0.47)	(0.73)
SPF	-0.13	-0.14	-0.63	0.29	1.10
	(0.29)	(0.48)	(0.54)	(0.53)	(0.70)
			Unemployment rate		
Greenbook	0.01	0.01	-0.05	-0.11	-0.28*
	(0.03)	(0.05)	(0.09)	(0.13)	(0.15)
Blue Chip	-0.11**	-0.11	-0.14	-0.20	-0.17
	(0.05)	(0.08)	(0.13)	(0.19)	(0.25)
SPF	-0.08	0.02	-0.07	-0.14	-0.30
	(0.07)	(0.06)	(0.11)	(0.17)	(0.25)
			Federal funds rate		
Greenbook	0.16**	0.14	0.15	0.34	0.67**
	(0.05)	(0.13)	(0.17)	(0.26)	(0.29)
Blue Chip	0.30**	0.38**	0.59**	0.84**	1.00**
	(0.10)	(0.15)	(0.20)	(0.25)	(0.39)
SPF	–	–	–	–	–
			Three-month yield		
Greenbook	0.07*	0.08	0.11	0.25	0.60**
	(0.03)	(0.08)	(0.15)	(0.23)	(0.26)
Blue Chip	0.13**	0.16	0.33*	0.60**	0.62*
	(0.06)	(0.14)	(0.19)	(0.26)	(0.36)
SPF	0.02	0.06	0.25**	0.47**	0.72**
	(0.03)	(0.05)	(0.08)	(0.13)	(0.28)

Table 9: Do forecasts maintain fidelity to the 10-year Treasury yield projection? Table shows estimated β coefficient from regression: $e_{i,t+h|t} = \alpha + \beta \Delta \hat{y}_{i10y,t+h|t} + \gamma_1 NBER_{t+h} + \gamma_2(y_{i10y,t+h} - \hat{y}_{i10y,t+h|t}) + u_{t+h}$. Sample period is 1984–2011. Newey-West standard errors in parentheses, asterisks denote statistical significance at 10 (*) and 5 percent (**) confidence level.