Empir Econ (2014) 47:365–388
DOI 10.1007/s00181-013-0736-z

Evaluating FOMC forecast ranges: an interval data approach

Henning Fischer · Marta García-Bárzana · Peter Tillmann · Peter Winker

Received: 29 February 2012 / Accepted: 3 June 2013 / Published online: 25 July 2013

Abstract The Federal Open Market Committee (FOMC) of the U.S. Federal Reserve publishes the range of members' forecasts for key macroeconomic variables, but not the distribution of forecasts within this range. To evaluate these projections, previous papers compare the midpoint of the range with the realized outcome. This paper proposes an alternative approach to forecast evaluation that takes account of the interval nature of projections. It is shown that using the conventional Mincer–Zarnowitz approach to evaluate FOMC forecasts misses important information contained in the width of the forecast interval. This additional information plays a minor role at short forecast horizons but turns out to be of sometimes crucial importance for longer-horizon forecasts. For 18-month-ahead forecasts, the variation of members' projections contains information that is more relevant for explaining future inflation than information embodied in the midpoint. Likewise, when longer-range forecasts for real

Electronic supplementary material The online version of this article (doi:10.1007/s00181-013-0736-z) contains supplementary material, which is available to authorized users.

H. Fischer (✉) · P. Winker
Department of Economics, University of Giessen, Licher Str. 64, 35394 Giessen, Germany
e-mail: henning.fischer@wirtschaft.uni-giessen.de

P. Winker
e-mail: peter.winker@wirtschaft.uni-giessen.de

M. García-Bárzana
Department of Statistics and Operational Research and Mathematics Didactics, University of Oviedo, C/ Calvo Sotelo s/n, 33007 Oviedo, Spain
e-mail: garciabarmarta.uo@uniovi.es

P. Tillmann
Department of Economics, University of Giessen, Licher Str. 66, 35394 Giessen, Germany
e-mail: peter.tillmann@wirtschaft.uni-giessen.de

GDP growth and the unemployment rate are considered, the width of the forecast interval comprises information over and above the one given by the midpoint alone.

Keywords Forecast evaluation · Interval data · Federal Reserve · Monetary policy

JEL Classification C53 · E37 · E58

1 Introduction

Forecasts of key macroeconomic aggregates are an important input for monetary policy decisions. The Federal Open Market Committee (FOMC), the decision-making body of the U.S. Federal Reserve, regularly produces projections for inflation, real gross domestic product (GDP) growth, and unemployment for different forecast horizons. Fed Chairman Bernanke (2011) recently argued that

> the committee's economic projections provide important context for understanding today's policy action as well as the committee's general policy strategy.

Moreover, policy rules estimated by Orphanides and Wieland (2008) and Wieland and Wolters (2011) show that FOMC forecasts have more explanatory power for actual interest rate decisions than the observed outcomes. Thus, evaluating forecasts that are formulated by monetary policymakers is an important element in the analysis of monetary policy decisions.

While each committee member submits her own set of projections, the set of forecasts made available to the public only consists of a forecast range constructed from individual forecasts. They are communicated both as the so-called `full range (FR)` of all individual forecasts and as the truncated `central tendency (CT)` interval, the latter eliminating the three highest and the three lowest individual forecasts. Fed watchers do not receive information about the distribution of forecasts within these intervals.[1] The forecast ranges published by the FOMC are notably different from probability forecasts produced, among others, by the Bank of England, which are meant to project the range of possible realizations of the variable based on a given coverage probability.[2]

This paper applies a procedure to evaluate these forecasts, which has recently been developed in the interval estimation literature. To gauge the quality of FOMC projections, the literature typically compares the midpoint, i.e., the mean of the upper and lower bound of either the `FR` or the `CT`, with the actual realization of the forecast variable, e.g., Gavin (2003), Gavin and Mandal (2003), and Gavin and Pande (2008). This

[1] The Federal Reserve Board's staff members produce their own set of forecasts collected in the Greenbook. These projections are point forecasts and are available to each FOMC member prior to the meeting. A separate strand of the literature analyzes the quality of Greenbook forecasts, see e.g. Romer and Romer (2000), Sims (2002), D'Agostino and Whelan (2008), Capistrán (2008), Gamber and Smith (2009), and Sinclair et al. (2010).

[2] The field of density forecasts in economics is still far from having reached a mature state. For a brief overview of the relevant literature on predictive densities and the problem of how to evaluate their accuracy when the true density cannot be observed cf. Kascha and Ravazzolo (2010) and the references therein.

approach, however, is viable only to the extent that all individual projections are drawn from the same underlying distribution. Hence, a general objection against using the mean forecast to assess forecast accuracy stems from the fact that the point forecasts collected from individual members represent the modes of each member's individual forecast density.[3] They represent each member's projection of the most likely outcome. Treating these modes as alternative draws from a single distribution is not an innocuous assumption as FOMC members entertain a variety of monetary policy preferences and models used to generate forecasts. Moreover, all forecasts are supposed to be conditional on each member's own judgment of the "appropriate policy" path over the forecast horizon. In case this path differs across members, the midpoint of either forecast range is not particularly informative. Furthermore, the individual mode forecasts could be the result of asymmetric densities with the degree of asymmetry varying across members.

We explicitly acknowledge the interval nature of individual FOMC projections and propose an alternative way to forecast evaluation that does not rely on the midpoint of either forecast range only. The approach draws on recently developed methods to estimating interval data regressions, see Blanco-Fernández et al. (2012) and Fischer et al. (2013). In the present context, the interval data approach is shown to collapse to a particularly straightforward extension of the conventional Mincer–Zarnowitz (1969) regression. In a regression of the realization of the forecast variable on the midpoint of the forecast range, the spread between the midpoint and the bounds of the forecast range enters as an additional regressor.

To the extent the width of the forecast interval measures the FOMC members' disagreement about the future, our paper also adds to the literature on disagreement among macroeconomic forecasters. Mankiw et al. (2004) and Capistrán and Timmermann (2009) provide evidence of a positive correlation between the dispersion in inflation beliefs and both the level and volatility of the inflation rate. Finding the amount of disagreement in inflation expectations to be varying over time together with a host of other macroeconomic aggregates, Mankiw et al. (2004) even conjecture that "disagreement may be a key to macroeconomic dynamics." Giordani and Söderlind (2003) and Lahiri and Sheng (2010) argue that forecast disagreement is a good proxy for forecast uncertainty. We show that the inclusion of a measure of forecast disagreement arises naturally in a Mincer–Zarnowitz regression as a consequence of the interval nature of the forecast data.

Our results suggest that these enhanced forecast regressions play a particularly important role in the evaluation of long-run forecasts. While the spread, and thus the interval nature of forecasts, is less important at short horizons, the spread contains important information about the eventual realization 12 or 18 months ahead. This is consistent with the notion that forecast uncertainty matters most for longer forecast horizons. If members entertain different—potentially asymmetric—loss functions and update forecasts differently in light of incoming information, the level of disagreement at a given forecast meeting might contain information relevant for the eventual future outcome.

[3] See also Reifschneider and Tulip (2007) and Rudebusch (2008) for this point.

The remainder of the paper is organized as follows. Section 2 introduces the interval data approach to forecast evaluation. Section 3 briefly discusses the data set and presents the key results of the paper. The final section offers some tentative conclusions.

2 Forecast evaluation: point versus interval data

Since neither all individual forecasts nor the mean or the median of their distribution is released, a common way to evaluate the informational content of the FOMC forecasts is to compare the actual outcome of the respective variable with the midpoint of either the `FR` or the `CT` serving as the FOMC's "consensus" forecast.[4] To do so, one can run a Mincer–Zarnowitz-type regression, see Mincer and Zarnowitz (1969), as given by

$$y_t = \alpha + \beta_1 \cdot \operatorname{mid} X_t + \epsilon_t, \tag{1}$$

where y_t is the outcome eventually realized, $\operatorname{mid} X_t$ is its "consensus" forecast represented by the midpoint of the `FR` or the `CT` interval, and $\epsilon_t \sim N(0, \sigma^2)$, $t = 1, \ldots, T$. An unbiased forecast implies that the null hypothesis $\alpha = 0$ and $\beta_1 = 1$ is met.

This conventional point data approach implicitly assumes that each forecast submitted by individual FOMC members is drawn from the same underlying probability distribution. In this case, using the mean forecast would indeed be informative about the most likely outcome. As discussed in the introduction, however, it is plausible to assume that the individual forecasts are not drawn from the same distribution, see the test results in Dowd (2004). By limiting oneself to considering the midpoint of the interval of forecasts only, one effectively discards the information about how much uncertainty surrounds this "consensus" forecast as given by the dispersion in the FOMC members' views. Hence, it might prove beneficial to also include the range between the highest and the lowest value of the `FR` or the `CT`, respectively, when assessing the accuracy of the FOMC's forecasts.

In general, the interval data approach specifies an interval Y_t as a linear function of another interval X_t, i.e., $Y_t = f(X_t)$. Let that model be the novel model "$\mathrm{M_G}$" studied in detail in Blanco-Fernández et al. (2012), which is also being used in an empirical application for stock market volatility by Fischer et al. (2013), as given by

$$\begin{aligned} Y_t &= \gamma_1 \cdot X_t^{\mathrm{Mid}} + \gamma_2 \cdot X_t^{\mathrm{Spr}} + \gamma_3 \cdot X_t^{\mathrm{Spr}_2} + \gamma_4 \cdot X_t^{\mathrm{Mid}_2} + \mathcal{E}_t \\ &= \gamma_1 \cdot \left[\operatorname{mid} X_t, \operatorname{mid} X_t\right] + \gamma_2 \cdot \left[-\operatorname{spr} X_t, \operatorname{spr} X_t\right] \\ &\quad + \gamma_3 \cdot \left[\operatorname{spr} X_t, \operatorname{spr} X_t\right] + \gamma_4 \cdot \left[-|\operatorname{mid} X_t|, |\operatorname{mid} X_t|\right] + \mathcal{E}_t. \end{aligned} \tag{2}$$

Here, $Y_t = \left[\inf Y_t, \sup Y_t\right] = \left[\operatorname{mid} Y_t \pm \operatorname{spr} Y_t\right]$ is the interval-valued response variable, where $\operatorname{mid} Y_t = (\sup Y_t + \inf Y_t)/2$ is its midpoint and $\operatorname{spr} Y_t = (\sup Y_t - \inf Y_t)/2$ is its spread. $X_t = \left[\inf X_t, \sup X_t\right] = \left[\operatorname{mid} X_t \pm \operatorname{spr} X_t\right]$ is an interval-valued regressor, whose so-called canonical decomposition $X_t = \left[\operatorname{mid} X_t, \operatorname{mid} X_t\right] +$

[4] Gavin and Pande (2008) find that the midpoint of the `CT` closely matches both the mean and the median of the distribution of all individual forecasts, which are the conventional measures of consensus among policymakers.

$\left[-\text{spr } X_t, \text{spr } X_t\right] = X_t^{\text{Mid}} + X_t^{\text{Spr}}$ permits the flexible model setup as given in (2) with seemingly four distinct interval regressors, yet only considering mid X_t and spr X_t, the two basic characteristics of the interval X_t. Finally, $\mathcal{E}_t$ is an interval error with $E(\mathcal{E}_t|X_t) = \left[\alpha_1 \pm \alpha_2\right]$. Details on how to estimate the interval data model in order to obtain the regression coefficients can be found in the Appendix.

The interval data model (2) eventually results in the following two linear relationships between the midpoints and spreads of both intervals and hence between point data variables again:

$$\text{mid } Y_t = \gamma_1 \cdot \text{mid } X_t + \gamma_3 \cdot \text{spr } X_t + \alpha_1 + \varepsilon_{1,t}, \quad E(\varepsilon_{1,t}) = 0 \tag{3}$$

$$\text{spr } Y_t = |\gamma_2| \cdot \text{spr } X_t + |\gamma_4| \cdot |\text{mid } X_t| + \alpha_2 + \varepsilon_{2,t}, \quad E(\varepsilon_{2,t}) = 0 \tag{4}$$

This framework lends itself to evaluate the forecast ranges for inflation, output growth, and unemployment published by the FOMC.

Before turning to the empirical application, it has to be noted that a "genuine" interval version of the Mincer–Zarnowitz-type regression—as given in Eq. (1) for the point data case—requires both an interval-valued regressor X_t and an interval-valued response variable Y_t. But if one wishes to evaluate the accuracy of the FOMC's `FR` or `CT` projections as given by the interval X_t, the benchmark Y_t is typically a "degenerated" interval, since the actual outcome realized and published later on is a point variable such that $Y_t = [\text{mid } Y_t \pm \text{spr } Y_t] = [\text{mid } Y_t \pm 0]$ is a zero-spread interval with $\text{mid } Y_t = \inf Y_t = \sup Y_t = y_t$. Hence, the empirical application of the interval data model (2) undertaken here inevitably leads to zero values for the coefficient estimates γ_2 and γ_4 in the linear relationship (4) since $\text{spr } Y_t = 0 \; \forall \, t$.

With Eq. (3) remaining as the only relevant relationship resulting from approach (2), one can see that in this case of a "degenerated" zero-spread response interval $Y_t = \left[y_t, \; y_t\right]$, the interval data model basically collapses to an enhanced point data specification similar to approach (1), but employing spr X_t (i.e., half the width of the forecast range) as an additional regressor. This point might become clearer when rewriting the pivotal Eq. (3) in this situation as

$$y_t = \alpha + \beta_1 \cdot \text{mid } X_t + \beta_2 \cdot \text{spr } X_t + \epsilon_t, \tag{5}$$

which is easily seen to nest the typical Mincer–Zarnowitz regression as given by Eq. (1). The collapse might appear as foregoing much of the interval data model's flexibility; however, it still constitutes an interval theory-based suggestion that also the dispersion in the FOMC participants' views might contain information about the actual outcome to be realized later on.

Hence, we will use the enhanced point data model (5) to evaluate both the `FR` and the `CT` of FOMC projections. A significant β_2 coefficient indicates that the width of the forecast interval contains important information to explain the eventual realization. Put differently, a significant β_2 is not necessarily a sign of biased forecasts, but might be considered a natural consequence of the interval nature of forecasts pointing to the importance of both the midpoint and the spread of the respective FOMC forecast range.[5] We compare the models (1) and (5) by means of root mean squared error (RMSE) measures as discussed below.

3 Data

Within the last 30 years, the FOMC published its economic projections typically twice a year in its Monetary Policy Report (MPR) to the Congress. As part of the preparation, each FOMC member prepares a set of forecasts to be presented and discussed at the FOMC meeting. Prior to the forecast meeting, members have full access to the Greenbook forecasts prepared by the staff of the Federal Reserve Board. Forecasts are submitted by board members as well as voting and nonvoting presidents of the twelve regional Federal Reserve Banks. The published report, however, contains information about the range of individual forecasts only. Individual forecasts are not published.[6] They are communicated both as the so-called `full range (FR)` of all individual forecasts and as the truncated `central tendency (CT)` interval, in which the three highest and the three lowest individual forecasts are eliminated. The public has no information about the distribution of individual forecasts within these ranges.

In the February report, the FOMC prepares forecasts for the inflation rate and the growth rate of real output, both from the fourth quarter of the previous year to the fourth quarter of the current year, and for the average civilian unemployment rate in the fourth quarter of the current year.[7] We refer to these projections as 12-month-ahead forecasts for simplicity. The July report does not only contain updates of the February forecasts for the current year, which we call 6-month-ahead forecasts, but also preliminary predictions for all three variables (i.e., two Q4-on-Q4 growth rates and one average rate for Q4) that are to be realized in the fourth quarter of next year, which we label 18-month-ahead forecasts.[8]

[5] The spread might thus be considered an omitted variable in the usual Mincer–Zarnowitz regression as given by Eq. (1). Since the spread as a measure of uncertainty in the distribution of forecasts can be viewed as a nontrivial function of the midpoint, this notion is also in line with Ramsey (1969) who suggested adding nonlinear functions of the regressors as additional explanatory variables in order to test for specification errors.

[6] Recently, individual forecasts are made available for a short sample period with a publication lag of ten years, see Romer (2010). Tillmann (2011) uses this new data set to uncover strategic forecasting behavior of FOMC members. Based on that data set, Banternghansa and McCracken (2009) study the degree of forecast disagreement among FOMC members.

[7] Both the Bureau of Economic Analysis (BEA) and the Fed used GNP as the measure of real aggregate output until 1992, when they switched to GDP. As regards the inflation rate, the FOMC switched among several price indices in the past. From 1979 to 1988, the inflation rate forecasts were based on the change in the GNP deflator. In 1989, however, the committee switched to inflation based on the consumer price index (for all urban consumers), which was then replaced by the price index of (overall) personal consumption expenditures in 2000. The latter was interchanged with the price index of core personal consumption expenditures in July 2004. The FOMC started reporting inflation rate projections based on both the overall and the core PCE price index with the February 2008 MPR.

[8] Since 2005, the forecasts in the February report also pertain to the next calendar year (24-month-ahead forecasts). Following the October 2007 meeting, the FOMC changed the frequency of forecasts, lengthened the forecast horizon to around 3 years, and raised the number of variables to be forecast. In addition, members are asked for their perception of forecast uncertainty. See Reifschneider and Tulip (2007) and Rudebusch (2008) for a discussion of these changes.

In this paper, we use these forecasts for a period from 1983 (when the first CT interval was published in the February MPR) to 2011. Since the first seven MPRs (issues February 1979 to July 1982) only exhibit range projections with varying language, sometimes suggesting they are FR and sometimes they are CT projections, we follow Romer and Romer (2008) in using the 18-month-ahead forecasts published in the July 1982 MPR as both the FR and the CT projection for the year 1983. As regards the series of mixed inflation rate projections, from 2008 on, we switch back to using the FOMC forecasts based on the overall PCE price index again. All projections are taken from the respective MPRs available online back to the July 1996 issue at the Federal Reserve Board's Web site (http://www.federalreserve.gov/boarddocs/hh/) and back to the February 1979 issue through the Federal Reserve Archival System for Economic Research (FRASER, at http://fraser.stlouisfed.org/publication/?pid=671). This amounts to 29 observations per variable for each of the three forecast horizons considered.

To evaluate the forecast accuracy, we compare these forecasts with actual realizations. We measure outcomes using real-time data rather than the latest vintage data available at the time of writing. Following Romer and Romer (2008) and Reifschneider and Tulip (2007), for variables in the National Income and Product Accounts (NIPA)—such as real GNP/GDP and the two PCE price indices—we define the actual data to be the BEA's so-called final estimates. These slightly revised numbers are released in late March or early April, so roughly three months after the end of the quarter being forecast, and correspond most closely to what the FOMC was trying to forecast. As regards non-NIPA variables, such as the unemployment rate or inflation measured by changes in the CPI, which are hardly affected by immediate revisions, we measure the outcomes using the data as first released. In particular, these non-NIPA series are typically Q4 estimates originally reported in January or early February. For further details about the data and their sources, please see the Appendix. The data themselves are available as supplementary material with the online version of this article on the journal Web page (10.1007/s00181-013-0736-z).

Figures 1, 2, and 3 show both FOMC forecast interval types as well as the actual outcome for each of the three key economic figures considered, respectively. In all figures, the respective years on the x-coordinate refer to the points in time for which the respective forecasts were formed (which happened roughly 6, 12, or 18 months in advance) and when the actual outcome of the variable was eventually realized. A visual inspection reveals that for all three variables, the width of both forecast interval types increases monotonously in the forecast horizon. Whereas the unemployment rate features the smallest forecast intervals throughout, the intervals typically are widest for the real GDP growth projections. The only exception to this is the 18-month horizon, for which the inflation rate projections span the widest intervals. Interestingly, for real GDP, the average interval width increases only marginally from the 12- to the 18-month-forecast horizon, whereas the most substantial rise in disagreement among FOMC members becomes apparent concerning the inflation rate which is likely to realize more than 1 year ahead in the future.[9]

[9] Gavin and Mandal (2003) attribute the relatively low degree of disagreement on a short-term point forecast for inflation to the FOMC members' perceived lack of control over the inflation rate over horizons shorter than 18 months.

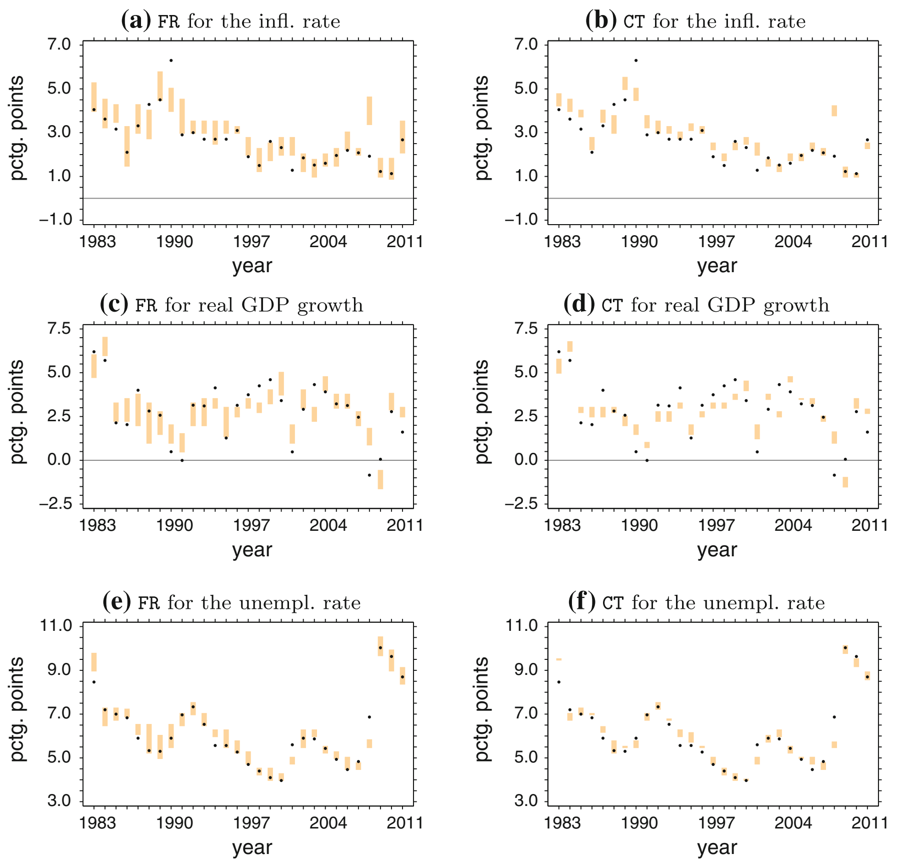

Fig. 1 Six-month-ahead interval projections and realizations. The three figures in the *left* (right) *column* depict the `Full Range` (`Central Tendency`) of all FOMC members' projections together with the corresponding actual point realization

While the FOMC performed very poorly—even at the 6-month horizon—when predicting real GDP growth, the figures show that the committee members did better for the (mixed) inflation rate and the unemployment rate, both being less volatile. The FOMC had a tendency to overpredict inflation in the 1980s and 1990s, but the predictions improved in the new millennium, with less disagreement to be found among the members at the same time. Table 1 reports two descriptive statistics that support the observations made above: the average width of the `FR` and `CT` intervals (given in percentage points) and the percentage of times that the actual value fell inside the respective forecast interval. Even vast forecast intervals did not prevent the real GDP growth projections from resulting in the lowest inclusion rate among all variables for all forecast horizons.

The projections made after the onset of the financial crises in fall 2008 finally show that the turmoils experienced recently on financial markets and in the real economy did not only lead to a rise in disagreement about the appropriate monetary and fiscal

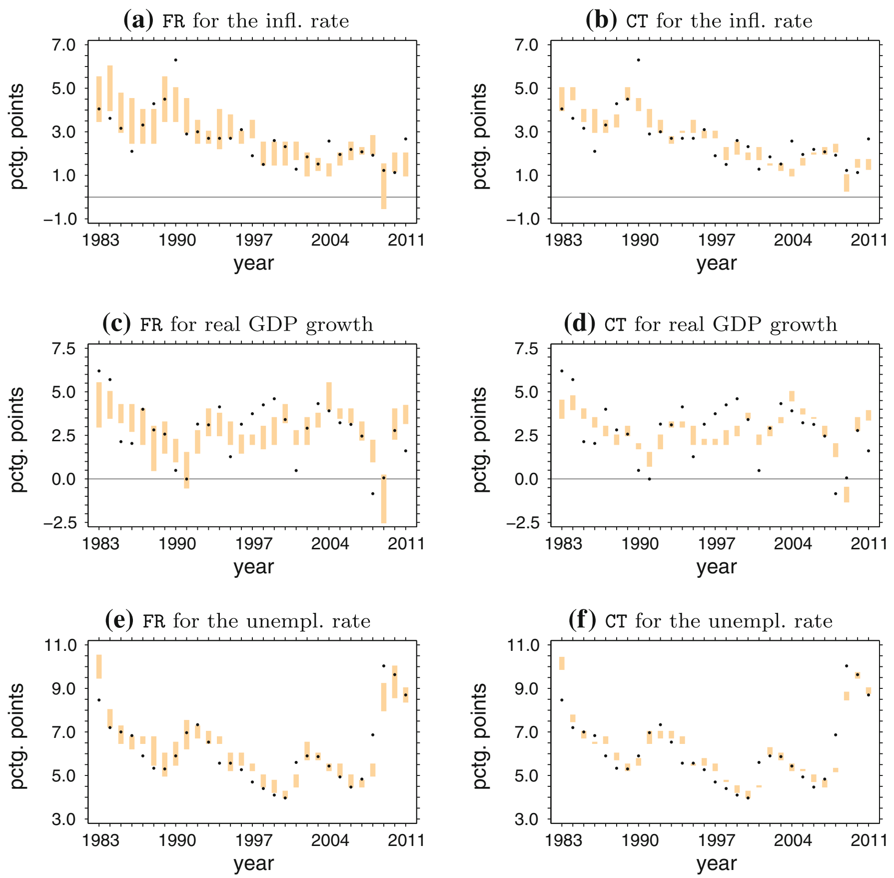

Fig. 2 Twelve-month-ahead interval projections and realizations. The three figures in the *left* (right) *column* depict the `Full Range` (`Central Tendency`) of all FOMC members' projections together with the corresponding actual point realization

policies to be taken, but also to less consensus on the future outcome of the policies being reflected once again in wider FOMC forecast intervals.

4 Results

4.1 The "simple" approach to forecast accuracy evaluation

Conducting the standard point data approach for evaluating the FOMC's forecast accuracy, one would regress the actual (point) realization y_t of the variable under study on a constant and the midpoint mid X_t of the interval of the respective forecast as in regression model (1). The results from ordinary least squares (OLS) regressions for the three variables, the `FR` and the `CT` as well as the three alternative forecast horizons

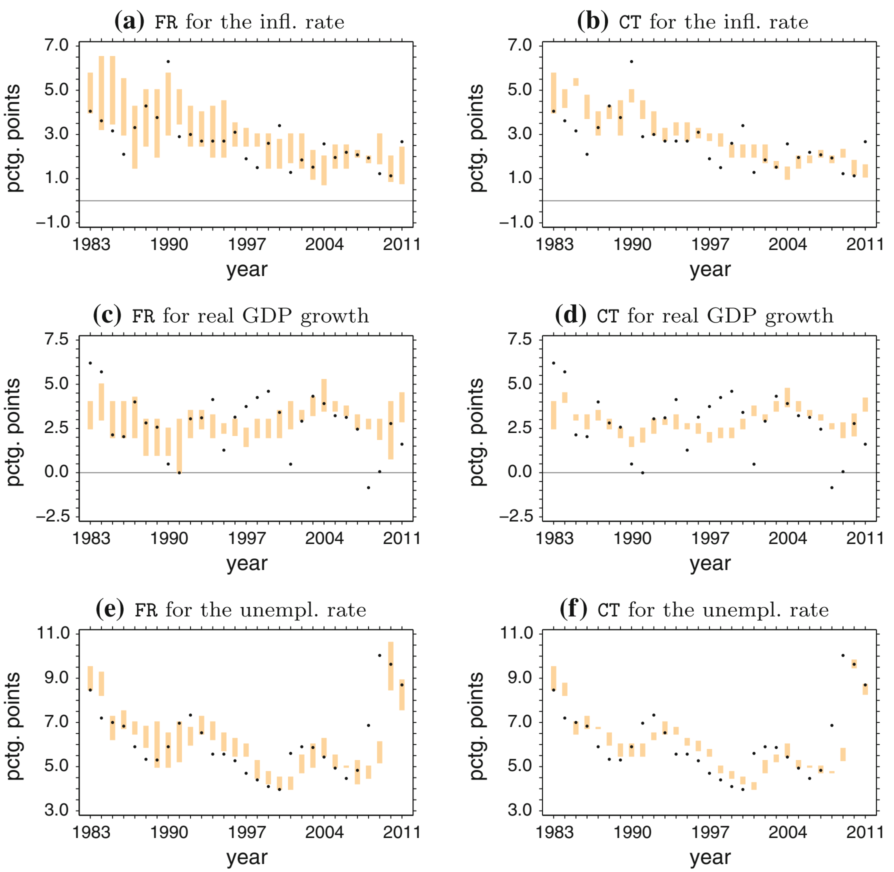

Fig. 3 Eighteen-month-ahead interval projections and realizations. The three figures in the *left* (right) *column* depict the `Full Range` (`Central Tendency`) of all FOMC members' projections together with the corresponding actual point realization

are presented in Tables 2, 3, and 4 in the columns entitled "simple".[10] If the estimates $(\hat{\alpha}, \hat{\beta}_1)$ are statistically not significantly different from (0, 1), the historical forecasts are unbiased. The joint null hypothesis $(\alpha, \beta_1) = (0, 1)$ is tested by the reported F statistic. The RMSE employed as a measure of forecast accuracy is defined as

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{\epsilon}_t)^2}, \tag{6}$$

and its version corrected for the degrees of freedom, i.e., the standard error of regression (SER), is calculated as

[10] The columns headed "enhanced" refer to the the OLS regressions of model (5). We will turn to these results after analyzing those of the "simple" models.

Table 1 Descriptive statistics for the FR and the CT forecast intervals

	Inflation rate	Real GDP growth	Unemployment rate
6 months			
FR average interval width	0.90	1.01	0.54
FR inclusion rate (%)	62	24	59
CT average interval width	0.35	0.38	0.20
CT inclusion rate (%)	17	7	34
12 months			
FR average interval width	1.14	1.37	0.64
FR inclusion rate (%)	52	31	38
CT average interval width	0.41	0.51	0.22
CT inclusion rate (%)	21	10	21
18 months			
FR average interval width	1.65	1.41	0.89
FR inclusion rate (%)	62	34	34
CT average interval width	0.60	0.57	0.36
CT inclusion rate (%)	34	10	24

The average interval width gives the mean distance between the upper and the lower bound of the respective forecast interval, averaged over all $T = 29$ observations, measured in percentage points. The inclusion rate measures the percentage of times that the actual value fell inside the respective forecast interval

$$\mathrm{SER} = \sqrt{\frac{1}{T-k}\sum_{t=1}^{T}\left(y_t - \hat{y}_t\right)^2} = \sqrt{\frac{1}{T-k}\sum_{t=1}^{T}\left(\hat{\epsilon}_t\right)^2}, \tag{7}$$

where $\hat{y}_t$ denotes the fitted values from the respective OLS regression, T is the sample size, and k is the total number of coefficients to be estimated. The smaller the both measures are, the smaller the errors and, consequently, the more accurate the forecasts are.

One unsurprising finding is clear-cut across all three variables to be forecast (for both the FR and the CT case, respectively): The accuracy of the midpoint forecasts is decreasing in the length of the forecast horizon. This becomes obvious when considering the respective values of the adjusted R^2, RMSE, or SER. Comparing all three key economic statistics that shall be forecast, the unemployment rate stands out as the figure whose midpoint projections match the actual outcome most closely at all forecast horizons as suggested by the adjusted R^2. This is most likely due to the high degree of persistence in the time series of the actual unemployment rate.

In almost all "simple" model regressions, the p value of the F statistic is quite large such that one would not reject the unbiasedness hypothesis at conventional significance levels. However, an exception to this uniform finding is the 6- and 18-month-ahead inflation rate forecasts, for which the interval midpoints seem to represent biased forecasts for their subsequent realizations according to the reported F test results. For the short-range midpoint forecasts, the unbiasedness hypotheses can be rejected at the 5 % significance level, whereas the F statistic is very close to being significant at the 10 % level for the long-range midpoint forecasts, a result holding true for both the FR and the CT cases. The rather high F statistics in the regressions for the 18-month-ahead

Table 2 Evaluating the FOMC's inflation rate forecasts using simple and enhanced regressions

	`Full Range` case						`Central Tendency` case					
	6 month		12 month		18 month		6 month		12 month		18 month	
	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced
Constant	0.0295	0.0807	0.6461	0.6256	0.8281	0.7840	0.0804	−0.0179	0.5866	0.6608	0.8152	0.7824
	(0.2947)	(0.2874)	(0.3264)	(0.3515)	(0.4384)	(0.4535)	(0.2880)	(0.2934)	(0.3215)	(0.3064)	(0.4506)	(0.4049)
	[0.9209]	[0.7812]	[0.0580]	[0.0868]	[0.0697]	[0.0957]	[0.7822]	[0.9518]	[0.0791]	[0.0404]	[0.0816]	[0.0643]
mid X_t	0.9084	0.9210	0.7466[a]	0.7292[a]	0.6312[b]	0.3707[c]	0.8923	0.8316	0.7754	0.8299	0.6325[b]	0.6057[a]
	(0.1286)	(0.1355)	(0.1483)	(0.1469)	(0.1717)	(0.1543)	(0.1248)	(0.1500)	(0.1399)	(0.1571)	(0.1746)	(0.2125)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0010]	[0.0237]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0012]	[0.0084]
spr X_t		−0.1939		0.1185		0.9786		1.5589		−1.0676		0.3694
		(0.4079)		(0.3911)		(0.3749)		(1.9774)		(0.9615)		(0.7118)
		[0.6384]		[0.7642]		[0.0148]		[0.4376]		[0.2770]		[0.6081]
F statistics	3.6961	2.4126	2.1009	1.5725	2.4181	8.1966	3.5058	3.2384	1.7270	2.0608	2.4579	3.0399
	[0.0381]	[0.0895]	[0.1419]	[0.2199]	[0.1081]	[0.0005]	[0.0443]	[0.0383]	[0.1969]	[0.1300]	[0.1046]	[0.0468]
$\overline{R^2}$	0.6699	0.6825	0.5277	0.5101	0.3643	0.4194	0.7059	0.7018	0.5796	0.5751	0.4285	0.4096
RMSE	0.6329	0.6320	0.7457	0.7452	0.8434	0.7909	0.5974	0.5903	0.7036	0.6940	0.7997	0.7976
SER	0.6559	0.6675	0.7728	0.7871	0.8741	0.8353	0.6191	0.6234	0.7291	0.7330	0.8288	0.8423
TU_{SER}	1.0177		1.0184		0.9557		1.0069		1.0053		1.0163	

Numbers in tables are estimated OLS regression coefficients for the respective explanatory variables, their corresponding standard errors given in round, and their corresponding p values (for the respective t test of the null hypothesis stating that the coefficient value is zero) given in square brackets. The superscripts a, b and c are used to indicate when the estimated coefficient of mid X_t is significantly different from unity at the 10, 5, or 1 % level, respectively. In order to take the potential conditional heteroskedasticity and serial correlation in the error into account, standard errors and p values are calculated based on heteroskedasticity and autocorrelation consistent (HAC) estimators of the covariance matrix of the parameter estimates, for which Bartlett kernel weights as described in Newey and West (1987) are used. The F statistic reported corresponds to the joint null hypothesis $H_0\colon (\alpha, \beta_1) = (0, 1)$ in all columns regarding the simple model as given by Eq. (1), whereas it tests the joint null hypothesis $H_0\colon (\alpha, \beta_1, \beta_2) = (0, 1, 0)$ whenever the enhanced model as given by Eq. (5) is considered. TU_{SER} is calculated as the ratio of the enhanced model's SER to the simple model's SER, with values less than one indicating the superiority of the enhanced model

Table 3 Evaluating the FOMC's real GDP growth rate forecasts using simple and enhanced regressions

	`Full Range` case						`Central Tendency` case					
	6 month		12 month		18 month		6 month		12 month		18 month	
	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced
Constant	−0.0709	−0.1150	0.3211	−1.1388	0.0148	−0.3712	−0.0078	0.3148	0.0815	−0.1355	0.1213	−0.5099
	(0.3847)	(0.6450)	(0.5194)	(1.2768)	(1.2394)	(1.7706)	(0.3878)	(0.5694)	(0.6135)	(1.1253)	(1.6622)	(1.8253)
	[0.8553]	[0.8599]	[0.5415]	[0.3806]	[0.9906]	[0.8356]	[0.9842]	[0.5851]	[0.8953]	[0.9051]	[0.9424]	[0.7822]
mid X_t	1.0052	1.0064	0.8943	1.0284	0.9686	1.0123	0.9823	0.9735	0.9706	0.9874	0.9097	0.8551
	(0.1118)	(0.1120)	(0.1764)	(0.2075)	(0.4070)	(0.4539)	(0.1124)	(0.1196)	(0.2037)	(0.2320)	(0.5291)	(0.5389)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0247]	[0.0346]	[0.0000]	[0.0000]	[0.0001]	[0.0002]	[0.0970]	[0.1247]
spr X_t		0.0812		1.5868		0.3712		−1.5625		0.6710		2.7885
		(0.9432)		(0.9443)		(0.8109)		(1.9536)		(2.3953)		(1.7405)
		[0.9320]		[0.1049]		[0.6509]		[0.4311]		[0.7816]		[0.1212]
F statistics	0.0590	0.0402	0.1980	3.3637	0.0330	0.0916	0.0883	0.2491	0.0104	0.0630	0.1379	1.3815
	[0.9428]	[0.9890]	[0.8215]	[0.0338]	[0.9676]	[0.9640]	[0.9158]	[0.8613]	[0.9896]	[0.9789]	[0.8718]	[0.2705]
$\overline{R}^2$	0.6755	0.6631	0.3887	0.4324	0.1261	0.0988	0.6662	0.6607	0.3801	0.3587	0.0910	0.1204
RMSE	0.9196	0.9195	1.2622	1.1935	1.5085	1.5032	0.9327	0.9228	1.2710	1.2686	1.5385	1.4851
SER	0.9531	0.9711	1.3081	1.2604	1.5634	1.5876	0.9666	0.9746	1.3173	1.3398	1.5945	1.5685
TU_{SER}	1.0189		0.9636		1.0155		1.0083		1.0171		0.9837	

Numbers in tables are estimated OLS regression coefficients for the respective explanatory variables, their corresponding standard errors given in round, and their corresponding p values (for the respective t test of the null hypothesis stating that the coefficient value is zero) given in square brackets. The superscripts a, b and c are used to indicate when the estimated coefficient of mid X_t is significantly different from unity at the 10, 5, or 1 % level, respectively. In order to take the potential conditional heteroskedasticity and serial correlation in the error into account, standard errors and p values are calculated based on heteroskedasticity and autocorrelation consistent (HAC) estimators of the covariance matrix of the parameter estimates, for which Bartlett kernel weights as described in Newey and West (1987) are used. The F statistic reported corresponds to the joint null hypothesis $H_0: (\alpha, \beta_1) = (0, 1)$ in all columns regarding the simple model as given by Eq. (1), whereas it tests the joint null hypothesis $H_0: (\alpha, \beta_1, \beta_2) = (0, 1, 0)$ whenever the enhanced model as given by Eq. (5) is considered. TU_{SER} is calculated as the ratio of the enhanced model's SER to the simple model's SER, with values less than one indicating the superiority of the enhanced model

Table 4 Evaluating the FOMC's unemployment rate forecasts using simple and enhanced regressions

	Full Range case						Central Tendency case					
	6 month		12 month		18 month		6 month		12 month		18 month	
	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced	Simple	Enhanced
Constant	0.1403	0.1632	0.1221	0.1983	1.3081	1.3964	0.2481	0.1223	0.4374	0.4522	1.3960	1.4438
	(0.2895)	(0.2894)	(0.6834)	(0.6553)	(0.8704)	(0.8221)	(0.3226)	(0.2495)	(0.6328)	(0.5502)	(0.8560)	(0.9126)
	[0.6320]	[0.5777]	[0.8595]	[0.7646]	[0.1445]	[0.1013]	[0.4485]	[0.6282]	[0.4953]	[0.4187]	[0.1145]	[0.1257]
mid X_t	0.9660	0.9916	0.9660	0.9121	0.7912	0.6931[b]	0.9535	0.9502	0.9189	0.9767	0.7825[a]	0.7210[a]
	(0.0447)	(0.0457)	(0.1180)	(0.1191)	(0.1239)	(0.1289)	(0.0531)	(0.0461)	(0.1093)	(0.0837)	(0.1177)	(0.1466)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
spr X_t		−0.6766		0.8184		1.1579		1.4332		−3.4455		1.8325
		(0.4444)		(0.7943)		(0.8471)		(0.9519)		(1.9655)		(1.5943)
		[0.1399]		[0.3124]		[0.1834]		[0.1442]		[0.0914]		[0.2609]
F statistics	0.7295	2.5753	0.2053	0.7068	1.5177	2.6878	0.4404	1.1173	0.2906	1.1315	1.8753	1.2908
	[0.4914]	[0.0755]	[0.8157]	[0.5567]	[0.2373]	[0.0672]	[0.6483]	[0.3601]	[0.7501]	[0.3546]	[0.1727]	[0.2985]
$\overline{R}^2$	0.9357	0.9364	0.8234	0.8226	0.4691	0.4727	0.9328	0.9340	0.8188	0.8288	0.4637	0.4551
RMSE	0.3820	0.3728	0.6330	0.6225	1.0974	1.0732	0.3904	0.3797	0.6411	0.6116	1.1030	1.0910
SER	0.3959	0.3937	0.6560	0.6574	1.1373	1.1334	0.4046	0.4010	0.6645	0.6459	1.1431	1.1522
TU_{SER}	0.9946		1.0021		0.9966		0.9911		0.9721		1.0080	

Numbers in tables are estimated OLS regression coefficients for the respective explanatory variables, their corresponding standard errors given in round, and their corresponding p values (for the respective t test of the null hypothesis stating that the coefficient value is zero) given in square brackets. The superscripts a, b, and c are used to indicate when the estimated coefficient of mid X_t is significantly different from unity at the 10, 5, or 1 % level, respectively. In order to take the potential conditional heteroskedasticity and serial correlation in the error into account, standard errors and p values are calculated based on heteroskedasticity and autocorrelation consistent (HAC) estimators of the covariance matrix of the parameter estimates, for which Bartlett kernel weights as described in Newey and West (1987) are used. The F statistic reported corresponds to the joint null hypothesis $H_0: (\alpha, \beta_1) = (0, 1)$ in all columns regarding the simple model as given by Eq. (1), whereas it tests the joint null hypothesis $H_0: (\alpha, \beta_1, \beta_2) = (0, 1, 0)$ whenever the enhanced model as given by Eq. (5) is considered. TU_{SER} is calculated as the ratio of the enhanced model's SER to the simple model's SER, with values less than one indicating the superiority of the enhanced model unemployment rate forecasts may also be interpreted as being suggestive of biased midpoint forecasts, given that the β_1 estimate is also significantly different from unit value, at least for the CT case.

Interestingly, the informational content of the midpoint of long-range real GDP growth forecast intervals—be it the FR or the CT type—is markedly low, as revealed by the $\overline{R^2}$. Another interesting observation is that for the inflation rate forecasts, the midpoint of the CT is a more accurate predictor than the midpoint of the FR, which has previously also been noted by McCracken (2010). The reason for this finding might be that members intentionally submit extreme forecasts to influence policy decisions. Eliminating these outliers may then improve the midpoint's forecast accuracy.

The results mentioned for predominantly long-horizon forecasts suggest that factors other than the midpoint of the forecast intervals could provide additional informational content for the actual outcomes. Since so far only the midpoints of the FOMC forecast intervals have been considered, a natural extension is to raise the question whether the dispersion in the FOMC members' individual forecasts contains additional useful information about the uncertain future outcome. This issue will be explored by the interval data theory-based approach now.

4.2 The "enhanced" approach to forecast accuracy evaluation

Remember that regressing $Y_t = \left[y_t,\ y_t\right]$ on X_t in the interval data framework is in our analysis equivalent to running the point data regression of y_t on a constant, mid X_t, and spr X_t as given by (5), which can be thought of as an augmented Mincer–Zarnowitz regression. Thus, an F statistic can be used again to test the joint null hypothesis $(\alpha, \beta_1, \beta_2) = (0, 1, 0)$ which postulates that the midpoint of the forecast interval is an unbiased estimate for the actual outcome while its spread has no additional explanatory power. Testing this type of joint null hypothesis for Eq. (5) can also be considered a special type of encompassing test in the spirit of Chong and Hendry (1986). When the null has to be rejected, the midpoint forecast does not encompass the spread forecast; instead, the latter contains additional information about the point outcome.

The results from estimating the regression as given by Eq. (5) by OLS for both the FR and the CT cases, again for all three variables and all three forecast horizons, are for the sake of comparison also reported in Tables 2, 3, and 4. They can be found in the columns entitled "enhanced" and corroborate the presumptions made before about the importance of the disagreement among the forecasters, which will be elucidated in the following.

4.2.1 Inflation rate

Analyzing the dispersion in the FOMC's inflation forecasts, one can see that it does not do a good job in providing additional information about the outcome at both the 6- and the 12-month horizon. The corresponding results of the interval data regressions show that for both the FR and the CT, the interval spread is not significant when used as a second regressor supplementing the interval midpoint. The SER increases compared to the simple regressions, too.

When it comes to exploiting the full information contained in the 18-month-ahead forecast intervals, however, the spread of the respective interval plays a remarkable role. Compared to the simple regressions, the adjusted R^2 increases from 36 to 42 % in the FR case when the spread enters as a regressor, while the RMSE and also the SER decline. The value of the midpoint's coefficient reduces nearly by half after controlling for the spread, being far away from unit value. The F statistics confirm these findings and clearly approve the rejection of the null hypothesis $(\alpha, \beta_1, \beta_2) = (0, 1, 0)$, thereby indicating that the 18-month-ahead midpoint forecast does not encompass the corresponding forecast spread. Hence, the variation in members' expectations seems to contain information that is considerably more relevant for explaining the inflation rate realized 18 months later than the one embodied in the midpoint of the FR of individual projections.

For the truncated CT case, the F statistics advocating the rejection of the encompassing hypothesis also underline the information content of the width of the forecast range, although the point estimate for the spread coefficient might not be statistically significant.[11] The findings are consistent with the view that forecast disagreement that arises endogenously if members report their mode forecasts taken from different underlying distributions embodies different interpretations of new information. The dispersion of forecasts has predictive power if, e.g., members' degree of disagreement is positively related to business cycle turning points or structural changes in forecast variables.

4.2.2 Real GDP growth rate

For the GDP growth rate forecasts, the enhanced model specification does not seem to result in additional explanatory power in contrast to the simpler specification. The partial influence of the interval spread is not significantly different from zero. Moreover, adding the forecast dispersion as a second regressor neither yields higher values of the adjusted R^2, nor does it reduce the RMSE or the SER in general. Only the spread of the FR of 12-month-ahead forecasts is close to being statistically significant at the 10 % level and seems to add some explanatory power above the one contained in the interval midpoint. The adjusted R^2 increases from 39 to 43 % when exploiting the information enclosed in the spread. Besides, the SER also turns out to be smaller in comparison with the simpler specification with only one regressor. Finally, according to the F statistic, we can reject the null hypothesis postulating that the dispersion in the FOMC members' forecasts does not contain useful information over and above the one given by the supposedly unbiased midpoint forecast, so we can conclude that the latter does not encompass the former. This, in turn, speaks for a combination of both basic characteristics of the middle-range forecast interval in order to obtain a forecast superior to the single midpoint forecast.

[11] Since the FR interval cannot have a smaller width than the truncated CT interval by construction, the FR spread automatically represents a degree of dispersion in the FOMC members' individual views, which is at least as high as the one given by the CT spread. Hence, differences between the FR and the CT results can be attributed to the relevance of eliminating extreme views whereby the forecast dispersion is in general decreased beforehand. Note, however, that also the consensus forecast given by the midpoint might be altered when eliminating the six outliers.

Similar results can be found for the longer-range 18-month-ahead CT forecast interval, even if the F statistic might not be significantly different from zero. Here, the midpoint loses its significance after controlling for the spread, while $\overline{R}^2$ increases and the SER falls. Since the same results cannot be observed for the FR of 18-month-ahead projections, this suggests that the full dispersion in individual forecasts contains useful information at the 12-month horizon, while censoring extreme views is appropriate in order to distill the relevant information embodied in the interval width when considering 18-month-ahead real GDP growth projections.

4.2.3 Unemployment rate

The dispersion of unemployment forecasts does in general not exhibit a significant influence on the realizations at conventional significance levels, except for the 12-month-ahead CT forecasts where it helps to increase $\overline{R}^2$ and decrease the SER, respectively. Interestingly, however, for the FR of both the 6- and the 18-month-ahead forecasts, the F statistic signals that the encompassing hypothesis $(\alpha, \beta_1, \beta_2) = (0, 1, 0)$ has to be rejected. For the short-range horizon, this F test result emerges despite the coefficient of the midpoint being virtually identical to unity, whereas for the long-range horizon, the test result seems to follow from the midpoint's coefficient being significantly different from unity at the 5 % level, disqualifying the midpoint as an unbiased, fully information efficient forecast to be used on its own. This again suggests that the interval width reflects information that should not be discarded when evaluating longer-range FOMC forecasts.

5 Concluding remarks

The evaluation of macroeconomic forecasts published by the FOMC proves to be difficult due to the fact that only a range of forecasts is made available to the public. Often the midpoint of these ranges is used to represent the average forecast. This paper proposed an alternative that explicitly recognizes the interval nature of these projections. We found that the width of the forecast range contains important information that significantly affects the evaluation of longer-horizon forecasts. For the 18-month-ahead horizon, the variation of all members' projections contains information that is more relevant for explaining future inflation than information embodied in the midpoint of the range. The dispersion in individual members' views also appears to be a noteworthy factor not to be neglected when considering 12- or 18-month forecasts for real GDP growth or the unemployment rate, partly in the condensed form of the CT omitting the most deviant opinions.

The results derived in this paper imply that the midpoint alone is often not a useful indicator for the Fed's long-run outlook on the economy. Given that monetary policy actions typically affect the real economy with a time lag of about four to eight quarters, this is the decisive horizon for market participants and the broader public. In order to uncover the whole information content of all FOMC members' long-run forecasts, Fed watchers ought to consider not only the midpoint of the forecast range, since it is shown to often be a biased estimate for the eventual realization 12 or 18 months ahead. Instead, they should pay attention to the interval width as well. The latter represents the disagreement among the FOMC members, which serves as the best available proxy for the uncertainty surrounding the "consensus" midpoint forecast, and it is shown to also be a crucial factor significantly affecting the actual realization. Hence, when uncertainty becomes a major issue, i.e., for long-range forecast horizons, even a simple linear combination of both the midpoint and the width of each forecast range can already lead to a superior point forecast compared to the potentially biased one given by the single midpoint alone.

It follows that steps toward higher transparency in the sense of releasing information about either the dispersion of forecasts within the committee or the publication of fan charts along the lines of the Bank of England's communication strategy could potentially play an important role in guiding public expectations. Future research will assess whether the measures taken since 2007 to reform the forecasting process have significantly changed the information content of FOMC forecasts.

Acknowledgements Financial support from Spain's Ministerio de Ciencia e Innovacion (MICINN) through *Acciones Integradas Hispano-Alemanas 2012*, cofinanced by the German Academic Exchange Service (DAAD) through *PPP Spanien 2012*, Research Grant 54367957, is kindly acknowledged. We are grateful to three anonymous referees for helpful comments on an earlier version of this paper. Any remaining errors are our own.

Data appendix: measuring the outcomes

We measure the outcomes for the three variables being forecast using real-time data, closely following the procedure employed by Romer and Romer (2008) in a related study that comprises the period 1979–2001. The exact computations and data sources are described in the following.

Real output growth rate

Real-time actuals are 4Q-on-4Q growth rates calculated using BEA's "final (third)" Q4 estimates of real GNP/GDP, typically released in March and first published in the March or April issue of the Survey of Current Business (SCB). The data were downloaded as monthly vintages from the Federal Reserve Bank of Philadelphia's Real-Time Data Research Center (RTDSM) at http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/ROUTPUT/. We compute percentage changes using numbers from the same SCB issue (i.e., the same mid-April vintage from the real-time data set). For instance, our figure for real GDP growth in 1999 is computed as the percentage change in the estimates of real GDP from 1998Q4 to 1999Q4 that are contained in the NIPA tables of the April 2000 SCB issue.

The formerly mentioned real-time actuals are used throughout for all years from 1983 until 2011 for all three forecast horizons under study, except for the years 1991 and 1992. This is due to the fact that the MPRs until (including) July 1991 contained forecasts for growth in real GNP, not real GDP, whereas the real-time data series as available in the RTDSM report values for real GDP only starting with the 1991M12 vintage. We follow Romer and Romer (2008) in solving this issue as follows: Because the 6-, 12-, and 18-month projections for 1991Q4 all forecasted growth in real GNP, but the actual time series mentioned above gives real-time values for real GDP only, the actual for all three forecast horizons for 1991Q4 is calculated using BEA's "final (third)" Q4 estimates for real GNP in 1990Q4 and 1991Q4 taken from the respective SBC issue 3/92. Similarly, the actual for the 18-month projections for 1992Q4 is calculated using BEA's "final (third)" Q4 estimates for real GNP in 1991Q4 and 1992Q4 as given in the respective SBC issue 3/93. Archived digital copies of hardcover SCB issues can be found online on the BEA homepage at http://www.bea.gov/scb/date_guide.asp from 1921 till today.

Our outcome measures match the measures being forecast as closely as possible. One issue remaining where we cannot completely reconcile the actual with the forecast concerns the change in base years. This involves the switches in reporting standards from 1990 to 1991 and from 1994 to 1995, as described in detail in the data appendix of Romer and Romer (2008) available online at http://www.aeaweb.org/articles.php?doi=10.1257/aer.98.2.230.

Inflation rate

Real-time actuals are 4Q-on-4Q growth rates calculated using estimates from several sources, depending on whether the measure of inflation is based on a variable in the NIPA tables or not.

For NIPA variables (i.e., inflation measured by the implicit GNP deflator, the PCE chain-type price index, or the PCE core chain-type price index), we use BEA's "final (third)" Q4 estimates typically released in March and first published in the March or April issue of the SCB. We compute percentage changes using numbers from the same SCB issue. For instance, we compute our figure for the inflation rate in 1986 as the percentage change in the implicit GNP deflator using the estimates of real and nominal GDP from 1985Q4 to 1986Q4 that are all contained in the NIPA tables of the March 1987 SCB issue. These "final (third)" estimates can be downloaded as monthly vintages from the Philadelphia Fed's RTDSM at http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/ROUTPUT/ and http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/NOUTPUT/, respectively. Our figure for the inflation rate in 2002 is computed as the percentage change in the estimates of the PCE chain-type price index from 2001Q4 to 2002Q4, which are contained in the NIPA tables of the April 2003 SCB issue. These "final (third)" Q4 estimates can be downloaded as monthly vintages from the Federal Reserve Bank of St. Louis's ALFRED database at http://alfred.stlouisfed.org/series?seid=PCECTPI&cid=21. As regards the real-time actuals for (percentage changes in) the PCE core chain-type price index, "final (third)" Q4 estimates can be downloaded as vintages either from the Philadelphia Fed RTDSM at http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/PCONX/ or from the ALFRED database at http://alfred.stlouisfed.org/series?seid=JCXFE&cid=21. Both sources in general provide the same quarterly estimates taken from the respective

SCB issues, the RTDSM offering monthly vintages starting in February 1996 and the ALFRED database offering monthly vintages starting in July 1999. Note that all real-time actuals can also be found in the original SCB issues available online as scans or digital issues via the BEA homepage at http://www.bea.gov/scb/date_guide.asp.

For the CPI as a non-NIPA variable that is not subject to immediate revisions, we use the 4Q-on-4Q percentage changes as first reported by the Bureau of Labor Statistics (BLS) in January. Since the BLS publishes monthly figures, but does not construct its own quarterly averages, following Romer and Romer (2008), we use the figures for actual Q4-to-Q4 CPI inflation from the first Greenbook prepared after the release of the December data. This is always the Greenbook prepared in late January or the very beginning of February. Historical Greenbooks are published online by the Board of Governors of the Federal Reserve System at http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm. The 5-year publication lag is not relevant for our study, since we need actual CPI data for the period 1989–2000 only. In contrast, the Philadelphia Fed's RTDSM provides real-time CPI data in monthly frequency as either monthly vintages starting in November 1998 or quarterly vintages starting in 1994Q3 only.

Unemployment rate

Real-time actuals for the civilian unemployment rate are averages for the fourth quarter of the relevant year. We obtain real-time data from the Philadelphia Fed's RTDSM available as quarterly vintages at http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files/RUC/. This data represents the values from the BLS's February, May, August, and November issues of *Employment and Earnings*. We use the monthly estimates available in mid-February of each year for the previous October, November, and December and calculate the arithmetic mean of these three values. The unemployment rate is hardly revised; for the time period considered here, in particular, the estimates from the February vintages employed by us do not differ at all from those of the May vintages available three months later. Comparisons with other potential real-time unemployment rates reveal that the mean values computed by us for the fourth quarter of each year are virtually identical to the Q4 estimates from the first Greenbook prepared after the release of the BLS December data (i.e., the Greenbook available in late January or the very beginning of February). The monthly real-time figures used by us do not differ from the respective monthly values contained in the January issue of the BLS's *Monthly Labor Review*, either, the latter being archived at http://www.bls.gov/opub/mlr/archive.htm for editions starting in 2000.

Technical appendix: estimating the interval data model

This appendix is intended to provide details on the estimation of the interval model "M_G" as given by Eq. (2) for the general case in which the response variable Y_t is not necessarily an interval degenerated into a real number. More profound information on the model itself and the estimation process can be found in Blanco-Fernández et al. (2012), who also present the model's cognate predecessors being the simple, less flexible linear regression models for interval data of Blanco-Fernández et al. (2011) and González-Rodríguez et al. (2007). The reader may also refer to these papers for some preliminary concepts of the interval data framework including the basics of interval arithmetic.

The interval data model "$\mathrm{M_G}$" as given in Eq. (2) can also be written as

$$Y_t = \boldsymbol{x}_t' \cdot \boldsymbol{\gamma} + \mathcal{E}_t, \tag{8}$$

where $\boldsymbol{x}_t' = (X_t^{\mathrm{Mid}}, X_t^{\mathrm{Spr}}, X_t^{\mathrm{Spr_2}}, X_t^{\mathrm{Mid_2}})$, and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)'$; hence, the equivalent matrix expression—which consolidates all T random interval pairs $\{(X_t, Y_t)\}_{t=1,\ldots,T}$ of the sample considered—is given by

$$\boldsymbol{y} = \boldsymbol{X} \cdot \boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{9}$$

where $\boldsymbol{y} = (Y_1, Y_2, \ldots, Y_T)'$ and $\boldsymbol{\varepsilon} = (\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_T)'$ are $(T \times 1)$ column vectors collecting the T interval-valued response variables and interval-valued errors, respectively, and $\boldsymbol{X} = (\boldsymbol{x}^{\mathrm{Mid}}, \boldsymbol{x}^{\mathrm{Spr}}, \boldsymbol{x}^{\mathrm{Spr_2}}, \boldsymbol{x}^{\mathrm{Mid_2}})$ is a $(T \times 4)$ matrix containing the four $(T \times 1)$ interval-valued column vectors $\boldsymbol{x}^{\mathrm{Mid}} = (X_1^{Mid}, X_2^{\mathrm{Mid}}, \ldots, X_T^{\mathrm{Mid}})'$, $\boldsymbol{x}^{\mathrm{Spr}} = (X_1^{\mathrm{Spr}}, X_2^{\mathrm{Spr}}, \ldots, X_T^{\mathrm{Spr}})'$, $\boldsymbol{x}^{\mathrm{Spr_2}} = (X_1^{\mathrm{Spr_2}}, X_2^{\mathrm{Spr_2}}, \ldots, X_T^{\mathrm{Spr_2}})'$, and $\boldsymbol{x}^{\mathrm{Mid_2}} = (X_1^{\mathrm{Mid_2}}, X_2^{\mathrm{Mid_2}}, \ldots, X_T^{\mathrm{Mid_2}})'$.

In order to estimate the coefficient vector $\boldsymbol{\gamma}$ in (9), a *least-squared-error*-type minimization problem is solved, which features two important characteristics setting it apart from the "standard" unrestricted least-squares method. The first distinctive feature worth mentioning is one restricting the search space for two of the four model parameters of interest. Note that for the two explanatory intervals of positive width given in Eq. (2), by definition, the identities $X_t^{\mathrm{Spr}} = -X_t^{\mathrm{Spr}}$ and $X_t^{\mathrm{Mid_2}} = -X_t^{\mathrm{Mid_2}}$ hold, which leads to the interval model as specified in Eq. (2) being not unique. Hence, there are four equivalent representations for model "$\mathrm{M_G}$", which allows—without loss of generality—to consider the parameters γ_2 and γ_4 to be nonnegative. The second peculiarity to be noted regards the class of nonempty, closed, and bound intervals in $\mathbb{R}$ employed in the linear regression, denoted by $\mathcal{K}_c(\mathbb{R})$. It is a semi-linear space as, in general, the existence of a symmetric element with respect to the addition is not guaranteed. For this reason, it is useful to consider the so-called Hukuhara difference (Hukuhara 1967), which is the most common difference when dealing with intervals, even though it does not always exist.[12] Taking into account that the model-implied estimated interval errors given by the corresponding Hukuhara differences $Y_t -_H \boldsymbol{x}_t' \cdot \widehat{\boldsymbol{\gamma}}$ have to exist (i.e., the residual has to be a well-defined interval whose supremum is not smaller than its infimum, for each observation $t = 1, 2, \ldots, T$), the minimization problem becomes a constrained one, where the set of constraints is devoted to assure the existence of the Hukuhara differences and hence the residuals.

[12] González-Rodríguez et al. (2007, p. 69) provide a formal definition of the Hukuhara difference, together with an example of the condition guaranteeing its existence.

Thus, the constrained minimization problem at hand is

$$\min_{\boldsymbol{\gamma}} \left(\boldsymbol{y}^* - \boldsymbol{X}^* \cdot \boldsymbol{\gamma}\right)' \left(\boldsymbol{y}^* - \boldsymbol{X}^* \cdot \boldsymbol{\gamma}\right) \quad \text{s.t.} \quad \boldsymbol{X}_s \cdot \boldsymbol{\gamma} \leq \mathbf{spr}\boldsymbol{Y}, \tag{10}$$

for which the real-valued vectors and matrices are defined as

$$\begin{aligned}
\boldsymbol{X}^* &= \begin{pmatrix} \boldsymbol{X}_m^* \\ \boldsymbol{X}_s^* \end{pmatrix} \quad \text{is a } (2T \times 4) \text{ matrix, wherein} \\
\boldsymbol{X}_m^* &= \boldsymbol{X}_m - \mathbf{1} \cdot \overline{\boldsymbol{x}_m}', \text{ wherein} \\
\boldsymbol{X}_m &= \left(\mathbf{mid}\boldsymbol{X}, \mathbf{0}, \mathbf{spr}\boldsymbol{X}, \mathbf{0}\right), \text{ with} \\
\mathbf{mid}\boldsymbol{X} &= \left(\text{mid } X_1, \text{mid } X_2, \ldots, \text{mid } X_T\right)', \\
\mathbf{spr}\boldsymbol{X} &= \left(\text{spr } X_1, \text{spr } X_2, \ldots, \text{spr } X_T\right)', \\
&\quad \mathbf{0} \quad \text{being a } (T \times 1) \text{ vector of zeros,} \\
\overline{\boldsymbol{x}_m} &= \left(\overline{\text{mid } X}, 0, \overline{\text{spr } X}, 0\right)', \text{ with} \\
\overline{\text{mid } X} &= \frac{1}{T} \sum_{t=1}^{T} \text{mid } X_t, \\
\overline{\text{spr } X} &= \frac{1}{T} \sum_{t=1}^{T} \text{spr } X_t, \\
&\quad \mathbf{1} \quad \text{being a } (T \times 1) \text{ vector of ones,} \\
\boldsymbol{X}_s^* &= \boldsymbol{X}_s - \mathbf{1} \cdot \overline{\boldsymbol{x}_s}', \text{ wherein} \\
\boldsymbol{X}_s &= \left(\mathbf{0}, \mathbf{spr}\boldsymbol{X}, \mathbf{0}, |\mathbf{mid}\, \boldsymbol{X}|\right), \text{ with} \\
\mathbf{spr}\boldsymbol{X} &= \left(\text{spr } X_1, \text{spr } X_2, \ldots, \text{spr } X_T\right)', \\
|\mathbf{mid}\, \boldsymbol{X}| &= \left(|\text{mid } X_1|, |\text{mid } X_2|, \ldots, |\text{mid } X_T|\right)', \\
&\quad \mathbf{0} \quad \text{being a } (T \times 1) \text{ vector of zeros,} \\
\overline{\boldsymbol{x}_s} &= \left(0, \overline{\text{spr } X}, 0, \overline{|\text{mid } X|}\right)', \text{ with} \\
\overline{\text{spr } X} &= \frac{1}{T} \sum_{t=1}^{T} \text{spr } X_t, \\
\overline{|\text{mid } X|} &= \frac{1}{T} \sum_{t=1}^{T} |\text{mid } X_t|, \\
&\quad \mathbf{1} \quad \text{being a } (T \times 1) \text{ vector of ones,}
\end{aligned}$$

$$\boldsymbol{\gamma} = \left(\gamma_1,\ \gamma_2,\ \gamma_3,\ \gamma_4\right)',\ \text{with}\ \gamma_1, \gamma_3 \in \mathbb{R}\ \text{and}\ \gamma_2, \gamma_4 \in \mathbb{R}_0^+,$$

$$\begin{aligned}
\boldsymbol{y}^* &= \begin{pmatrix} \boldsymbol{y}_m^* \\ \boldsymbol{y}_s^* \end{pmatrix} \qquad \text{is a } (2T \times 1) \text{ vector, wherein} \\
\boldsymbol{y}_m^* &= \mathbf{mid}\boldsymbol{Y} - \overline{\text{mid}\ Y} \cdot \mathbf{1},\ \text{with} \\
\mathbf{mid}\boldsymbol{Y} &= \left(\text{mid}\ Y_1, \text{mid}\ Y_2,\ \ldots\ , \text{mid}\ Y_T\right)', \\
\overline{\text{mid}\ Y} &= \frac{1}{T} \sum_{t=1}^{T} \text{mid}\ Y_t, \\
\mathbf{1} &\quad \text{being a } (T \times 1) \text{ vector of ones}, \\
\boldsymbol{y}_s^* &= \mathbf{spr}\boldsymbol{Y} - \overline{\text{spr}\ Y} \cdot \mathbf{1},\ \text{with} \\
\mathbf{spr}\boldsymbol{Y} &= \left(\text{spr}\ Y_1, \text{spr}\ Y_2, \ldots, \text{spr}\ Y_T\right)', \\
\overline{\text{spr}\ Y} &= \frac{1}{T} \sum_{t=1}^{T} \text{spr}\ Y_t, \\
\mathbf{1} &\quad \text{being a } (T \times 1) \text{ vector of ones.}
\end{aligned}$$

As the objective function has quadratic shape and the inequality constraints are linear, some standard routines from numerical analysis can be used to solve the minimization problem. We implement a MATLAB code based on the Karush–Kuhn–Tucker conditions to obtain the regression coefficients.

References

Banternghansa C, McCracken M (2009) Forecast disagreement among FOMC members. Work paper 2009–059A, Fed Reserve Bank, St Louis

Bernanke B (2011) Transcript of Chairman Bernanke's press conference, 27 April 2011. Board of Governors of the Fed Reserve System, Washington DC

Blanco-Fernández Á, Corral N, González-Rodríguez G (2011) Estimation of a flexible simple linear model for interval data based on set arithmetic. Comput Stat Data Anal 55(9):2568–2578

Blanco-Fernández Á, García-Bárzana M, Colubi A, Kontoghiorghes E (2012) Extensions of linear regression models based on set arithmetic for interval data. Working paper (arXiv:1210.5881 [math.ST])

Capistrán C (2008) Bias in Federal Reserve inflation forecasts: is the Federal Reserve irrational or just cautious? J Monet Econ 55(8):1415–1427

Capistrán C, Timmermann A (2009) Disagreement and biases in inflation expectations. J Money Credit Bank 41(2/3):365–396

Chong Y, Hendry D (1986) Econometric evaluation of linear macro-economic models. Rev Econ Stud 53(4):671–690

D'Agostino A, Whelan K (2008) Federal Reserve information during the great moderation. J Eur Econ Assoc 6(2–3):609–620

Dowd K (2004) FOMC forecasts of macroeconomic risks. Occasional paper series no 2004–2012. University of Nottingham, Nottingham

Fischer H, Blanco-Fernández Á, Winker P (2013) Predicting stock return volatility: can we benefit from regression models for return intervals?, MAGKS joint discussion paper series in economics, forthcoming

Gamber E, Smith J (2009) Are the Fed's inflation forecasts still superior to the private sector's? J Macroecon 31(2):240–251

Gavin W (2003) FOMC forecasts: is all the information in the central tendency? Fed Reserve Bank St Louis Rev 85(3):27–46
Gavin W, Mandal R (2003) Evaluating FOMC forecasts. Int J Forecast 19(4):655–667
Gavin W, Pande G (2008) FOMC consensus forecasts. Fed Reserve Bank St Louis Rev 90(3, part 1):149–164
Giordani P, Söderlind P (2003) Inflation forecast uncertainty. Eur Econ Rev 47(6):1037–1059
González-Rodríguez G, Blanco A, Corral N, Colubi A (2007) Least squares estimation of linear regression models for convex compact random sets. Adv Data Anal Classif 1:67–81
Hukuhara M (1967) Intégration des applications mesurables dont la valeur est un compact convexe. Funkcialaj Ekvacioj 10:205–223
Kascha C, Ravazzolo F (2010) Combining inflation density forecasts. J Forecast 29(1–2):231–250
Lahiri K, Sheng X (2010) Measuring forecast uncertainty by disagreement: the missing link. J Appl Econ 25(4):514–538
Mankiw N, Reis R, Wolfers J (2004) Disagreement about inflation expectations. In: Gertler M, Rogoff K (eds) NBER macroeconomics annual 2003, vol 18. MIT Press, Cambridge, pp 209–248
McCracken M (2010) Using FOMC forecasts to forecast the economy. Federal Reserve Bank of St. Louis Economic Synopses, vol 5
Mincer J, Zarnowitz V (1969) The evaluation of economic forecasts. In: Mincer J (ed) Economic forecasts and expectations: analysis of forecasting behavior and performance. National Bureau of Economic Research, New York, pp 1–46
Newey W, West K (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55(3):703–708
Orphanides A, Wieland V (2008) Economic projections and rules of thumb for monetary policy. Fed Reserve Bank St Louis Rev 90(4):307–24
Ramsey J (1969) Tests for specification errors in classical linear least-squares regression analysis. J R Stat Soc B (Methodol) 31(2):350–371
Reifschneider D, Tulip P (2007) Gauging the uncertainty of the economic outlook from historical forecasting errors. Finance and Economics Discussion Series, vol 60, Divisions of Research and Statistics and Monetary Affairs, Fed Reserve Board, Washington, DC
Romer D (2010) A new data set on monetary policy: the economic forecasts of individual members of the FOMC. J Money Credit Bank 42(5):951–957
Romer C, Romer D (2000) Federal reserve information and the behavior of interest rates. Am Econ Rev 90(3):429–457
Romer C, Romer D (2008) The FOMC versus the staff: where can monetary policymakers add value? Am Econ Rev 98(2):230–235
Rudebusch G (2008) Publishing FOMC economic forecasts. FRBSF Economic Letter. Fed Reserve Bank of San Francisco, San Francisco
Sims C (2002) The role of models and probabilities in the monetary policy process. Brookings Pap Econ Act 2002(2):1–40
Sinclair T, Joutz F, Stekler H (2010) Can the fed predict the state of the economy? Econ Lett 108(1):28–32
Tillmann P (2011) Strategic forecasting on the FOMC. Eur J Polit Econ 27(3):547–553
Wieland V, Wolters M (2011) Forecasting and policy making. In: Elliott G, Timmermann A (eds) Handbook of economic forecasting, vol 2. Elsevier, Amsterdam