Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Forecast Disagreement Among FOMC Members

Chanont Banternghansa
and
Michael W. McCracken

Working Paper 2009-059A
http://research.stlouisfed.org/wp/2009/2009-059.pdf

December 2009

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Forecast Disagreement Among FOMC Members *

Chanont Banternghansa
Federal Reserve Bank of St. Louis

Michael W. McCracken
Federal Reserve Bank of St. Louis

December 2009

Abstract

This paper presents empirical evidence on the disagreement among Federal Open Market Committee (FOMC) forecasts. In contrast to earlier studies that analyze the range of FOMC forecasts available in the *Monetary Policy Report to the Congress*, we analyze the forecasts made by each individual member of the FOMC from 1992 to 1998. This newly available dataset, while rich in detail, is short in duration. Even so, we are able to identify a handful of patterns in the forecasts related to i) forecast horizon; ii) whether the individual is a Federal Reserve Bank president, governor, and/or Vice Chairman; and iii) whether individual is a voting member of the FOMC. Additional comparisons are made between forecasts made by the FOMC and the Survey of Professional Forecasters.

JEL Nos.: E52, E58, C53
Keywords: FOMC, forecast accuracy, forecast disagreement

* *Banternghansa*: Research associate, Research Division; Federal Reserve Bank of St. Louis; P.O. Box 442; St. Louis, MO 63166; chanont.banternghansa@stls.frb.org. *McCracken (corresponding author)*: Research officer, Research Division; Federal Reserve Bank of St. Louis; P.O. Box 442; St. Louis, MO 63166; michael.w.mccracken@stls.frb.org. We are grateful to Todd Clark, Bob Rasche, and participants of the Saint Louis Brown Bag seminar for valuable comments. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

1 Introduction

In 1979, legislation was passed requiring the Federal Reserve to report economic forecasts to Congress. After an initial release in July of that year, forecasts have been provided in February and July of each year thereafter.[1] Before each of these releases, each member of the Federal Open Market Committee (FOMC) makes a forecast of end-of-year nominal and real gross domestic product (GDP) growths, inflation, and the unemployment rate. The February forecasts are for the current calendar year. In July, two sets of forecasts are given: an updated forecast for the current calendar year and a longer-horizon forecast for the next calendar year. Once these forecasts have been collected from each member of the FOMC, the maximum, minimum, and a trimmed range (based on dropping the three highest and three lowest values) of each of the four variables are reported in *Monetary Policy Report to the Congress* (*MPR*).

In this paper, we use a newly available dataset (as described by Romer, 2009) to document the disagreement among forecasts made by individual members of the FOMC between February 1992 and July 1998. Until now, the only publicly available information consisted of the aggregated information (i.e., the maximum, minimum, and the trimmed range) contained in the *MPR*. In contrast, this new dataset provides not only the individual forecasts for each economic variable, but it also associates the forecasts with every member of the FOMC other than the Chairman.

To date, the dataset is the richest source of information on the FOMC forecasts that is available to the public. This richness allows us to construct a variety of measures of disagreement among FOMC members. With these measures in hand, our goal is to identify any patterns in the disagreement among the forecasts. Examples of potential patterns include seasonal effects related to the forecast horizon, as well as treatment effects related to whether the member is a regional bank president, governor, or Vice Chairman, and whether the individual is a voting member of the FOMC. In addition, we link disagreement to the accuracy of the forecasts and whether a voting member of the FOMC dissented at the time of his/her submission to the *MPR*.

Even so, the dataset is very limited in its duration. Although FOMC forecasts have been made since 1979, the documentation of the individual forecasts is limited. Under the guidance of David Small, the Board of Governors has constructed a complete series of the

[1] Starting in 2008, this process was expanded to include forecasts made in April and October.

forecasts starting only as far back as February 1992. A complete series of forecasts exists through the present day, but a 10-year release window has been enacted that limits the most recent forecasts publicly available. After pruning any individual forecasts missing one of the four variables of interest, our data consists of a total of 358 individual forecasts each containing forecasts for the four variables, over three distinct forecast horizons, over a 7-year span, made by each regional bank president and each governor.

We are not the first to assess forecast disagreement among FOMC members, but the literature is limited by the lack of availability of the detailed data. Mankiw, Reis, and Wolfers (2003) note that the range of FOMC inflation forecasts is positively correlated with the interquartile range of similar forecasts made in the Livingston Survey. McNees (1995) notes that the average range of the FOMC forecasts increases with the forecast horizon. More often than not, the literature on FOMC forecasts has focused on the accuracy and efficiency of the FOMC forecasts (as proxied by the midpoint of either the full or trimmed range). Examples include Gavin (2003), Gavin and Mandal (2003), and Gavin and Pande (2008). While not directly related to forecast disagreement, Meade and Sheets (2005) as well as Chappell and McGregor (2000) discuss the related issue of dissent in the voting patterns of FOMC members.

Our results differ from all previous in at least two respects. First, we emphasize the degree of disagreement by each individual member of the FOMC and not the aggregate level of disagreement. Second, although we discuss disagreement in the context of forecasts for each of the four variables, we also address forecast disagreement among the vectors of forecasts themselves. Our logic for doing so is based on an assumption that the FOMC members construct their vectors of forecasts in a congruent fashion that jointly describes their view of the economy rather than construct their forecasts irrespective of the other elements. For example, those who believed in a Phillips curve relationship would likely adjust their forecasts of inflation and unemployment in an inverse fashion as their information set changes across time.

With these caveats in mind, our main results are as follows. First, there is disagreement among the members of the FOMC, but the degree of disagreement is small relative to the degree of disagreement among a universe of forecasters exemplified by the Survey of Professional Forecasters (SPF). Second, the Vice Chairman tends to have the most centrally located forecasts among all members of the FOMC. Third, while on aggregate there is little evidence that the level of disagreement varies with a regional bank's voting status, for some regional banks disagreement does vary with voting status. In particular, the Cleveland Federal Reserve Bank tends to be more consensus oriented when voting while the Dallas Federal Reserve Bank tends to be less consensus oriented when voting. Fourth, both the Cleveland and St. Louis Feds tend to be in greater disagreement than all other members of the FOMC. Finally, consumer price index (CPI) forecasts in general seem to be constructed for reasons other than accuracy as measured by quadratic loss.

This last point is important and should be kept in mind when interpreting our results on both disagreement and accuracy. As noted by Faust and Wright (2008), the FOMC (and Greenbook) forecasts are conditional rather than unconditional forecasts. The distinction between the two types of forecasts is that the conditional forecasts are constructed based on a hypothetical future path of monetary policy (i.e., a future path of the Federal Funds rate). Federal Reserve Bank of St. Louis president James Bullard (2009) makes this distinction clear when he states that "The FOMC members forecasts are made under appropriate monetary policy." In this framework, "appropriate monetary policy" is left to the discretion of the individual FOMC member constructing their own forecast. As argued by Ellison and Sargent (2009), this induces disagreement among the members irrelevant of whether the members are working from the same information sets (or even the same baseline models). As such, our results on disagreement and accuracy capture not only variation in the information sets and models the FOMC members are working with but also the variation in beliefs on what appropriate monetary policy should be. Not surprisingly, we find that this variation reveals itself most clearly in the forecasts of nominal GDP and inflation

The paper proceeds as follows. Section 2 describes the data and the methods used for our analysis. In Section 3 we characterize the degree of disagreement among FOMC forecasts. Section 4 describes the relationship between disagreement and the accuracy of FOMC forecasts. Section 5 links disagreement with voting dissent at the most recent FOMC meeting. Section 6 concludes.

2 Data and Methods

Before presenting our results, we first provide a brief description of the data and methods used in our analysis.

2.1 FOMC Data

As described in the introduction, we use the FOMC data provided by Romer (2009).[2] The FOMC data contain forecasts of each of the FOMC members from 1992 to 1998. These forecasts are made in February and July of each year. The forecasts include annual fourth quarter to fourth quarter (Q4 to Q4) growth rates of nominal GDP, real GDP, and the CPI as well as unemployment rate forecasts for the fourth quarter of the relevant year. Forecasts made in February are for the current calendar year; the forecasts made in July are for both the current and following calendar year. Thus each year has a total of three forecast sets: a 10 month-ahead forecast submitted in February, a 5 month-ahead forecast submitted in July for that year, and a 17 month-ahead forecasts for the next year. The dataset also contains the name of every FOMC member, affiliation (regional bank president, governor, or Vice Chairman), and whether the individual is a voting member of the FOMC for that particular year.

Due to the limited time frame of our dataset, we use the term "individual" interchangeably with "institution." As a result, our analysis treats each regional bank—not the bank president themselves—as the smallest unit. Similarly, the Vice Chairman is defined by the individual's title, not the person. Finally, we treat the governors on average rather than by person. As a result, between 1992 and 1998 we have 21 individual forecasts for each regional bank (except Cleveland which has 19 individual forecasts) and 108 individual forecast for the governor, of which 19 individual forecasts are made by the Vice Chairman.

2.2 SPF Data

To get a general sense of how the FOMC forecasts compare to the universe of professional forecasters, we also consider disagreement and accuracy of the participants in the SPF as collected by the Federal Reserve Bank of Philadelphia.[3] The surveys are released four times a year: February, May, August, and November. For our disagreement comparisons, we used data only released in February of each year because for this forecast, the information sets associated with the SPF are most closely aligned to those of the FOMC members (whereas the August SPF forecasts are released a full month after the July FOMC forecasts).

The unemployment rate forecast is for the fourth quarter of the current calendar year

[2] The data set is titled "A New Data Set on Monetary Policy Report: The Economic Forecasts of Individual Members of the FOMC" and is available at http://elsa.berkeley.edu/~dromer/.

[3] SPF data can be obtained at www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/historical-data/individual-forecasts.cfm.

while the CPI forecasts are Q4 to Q4 growth rates. In contrast to the FOMC forecasts, the SPF nominal and real GDP forecasts are in levels. We translate these into Q4 to Q4 growth rates using the Bureau of Economic Analysis (BEA's) preliminary estimates of nominal and real GDP for the fourth quarter of the previous year. Calculating the growth rates in this fashion is possible in real time because the BEA releases these estimates at the end of January while the SPF forecasts are submitted in mid-February. The only exception is in 1996 when the BEA's estimates of real and nominal GDP for 1995:Q4 were first released on February 23 instead of at the end of January. We assume the SPF forecasters used this BEA's release in the February survey to calculate growth and inflation rates in 1996.

2.3 Methods

A measure of dispersion must be chosen to evaluate disagreement among the FOMC forecasts. In choosing a metric, our first goal was to select one that was internally consistent regardless of the dimension of the forecast—that is, choose a metric that was not only well defined when analyzing the level of disagreement for each of the four individual variables but was also well defined when evaluating the level of disagreement among the vector-valued forecasts themselves. Our second goal was to choose a metric that accounted for any correlations across the individual variables when we measured the multivariate level of disagreement.

Figure 1 shows why this second point is a concern. Here we simulated 18 distinct bivariate standard normals with a correlation coefficient of 0.9. As expected, the pairs essentially lie along a line through the origin with slope equal to 0.9. Now consider points A, B, C, and D on the circle centered at the origin. Because each of these four points is equidistant from the origin, if we used Euclidean distance, they might be considered to be equally in "disagreement." In contrast, if one adjusts for the fact that the two variables are correlated, it is clear that points A and C are in greater "disagreement" with the bivariate sample as a whole than points B and D. In our four-variate sample of forecasts, we expect such an issue to arise since, for example, in so far as CPI-based inflation is highly correlated with GDP deflator-based inflation, a coherent forecast would roughly satisfy the property that the growth of nominal GDP would be the sum of the growth rate in real GDP and inflation.

The Mahalanobis distance satisfies each of our two requirements and is our baseline measure of disagreement. Let $x_{i,t,\tau} = (x_{i,t,\tau}^{(1)}, ..., x_{i,t,\tau}^{(4)})'$ denote the vector-valued forecast generated by individual i, at time t, for forecast horizon τ and let $\bar{x}_{t,\tau}$ denote the average of the time t forecasts for horizon τ. If we then define the sample covariance matrix associated with the forecasts as $S_{t,\tau} = n_{t,\tau}^{-1} \sum_{i=1}^{n_{t,\tau}} (x_{i,t,h} - \bar{x}_{t,\tau})(x_{i,t,h} - \bar{x}_{t,\tau})'$ the Mahalanobis distance of $x_{i,t,\tau}$ is

$$D(x_{i,t,\tau}) = \sqrt{(x_{i,t,h} - \bar{x}_{t,\tau})' S_{t,\tau}^{-1} (x_{i,t,h} - \bar{x}_{t,\tau})}. \tag{1}$$

In the scalar case $j = 1, ..., 4$ this simplifies to

$$D(x_{i,t,\tau}^{(j)}) = \frac{|x_{i,t,h}^{(j)} - \bar{x}_{t,\tau}^{(j)}|}{s_{t,\tau}^{(j)}} \tag{2}$$

where $s_{t,\tau}^{(j)}$ is the sample standard deviation of the forecasts.[4]

At some level we have tied our hands by wanting our measure of distance to be applicable for both multivariate and univariate comparisons. Were we to focus exclusively on the scalar case, we could have chosen the interquartile range as used in Mankiw, Reis, and Wolfers (2003) and Capistrán and Timmermann (2008). Instead, as a check of the robustness of our results, we also consider a variant of absolute deviations from the median as our measure of distance.

First consider the scalar case. If we let $m_{t,\tau}^{(j)}$ denote the median forecast at time t for horizon τ, then a simple outlier-robust measure of distance for individual i is the absolute deviation from the median $|x_{i,t,\tau}^{(j)} - m_{t,\tau}^{(j)}|$. Given these individual deviations, an outlier-robust measure of the "typical" deviation can be constructed using the concept of median absolute deviation (MAD) defined as $MAD_{t,\tau}^{(j)} = med_{1 \leq i \leq n_{t,\tau}} \{|x_{i,t,\tau}^{(j)} - m_{t,\tau}^{(j)}|\}$. When useful, a normalized distance metric could then be constructed as

$$D(x_{i,t,\tau}^{(j)}) = \frac{|x_{i,t,h}^{(j)} - m_{t,\tau}^{(j)}|}{MAD_{t,\tau}^{(j)}}. \tag{3}$$

The multivariate case is more difficult. The first complication is that while it is simple to generalize from the sample mean of scalars to a sample average of vectors, it is significantly more complex to generalize from a median of scalars to a median of vectors. A second complication is that the concept of MAD is an inherently scalar concept which, when constructed element by element, does not account for the correlations across the variables as discussed relative to Figure 1.

[4] Lahiri and Sheng (2008) also use squared deviations from the mean to measure disagreement.

Fortunately, a multivariate concept of a median does exist. The *Tukey median* of a collection of vectors is defined as the point within the cloud formed by the sample of vectors that has the greatest *ldepth*. Here we eschew a detailed discussion of *ldepth* and Tukey medians but instead provide an analogy using the scalar case.[5] Consider a collection of points on a line and begin with the left-most point. If we draw a vertical line through that point, the *ldepth* associated with that point is the minimum of (i) the number of points on, or to, its left and (ii) the number of points on, or to it's right. Since this number is 1 that point has an *ldepth* of 1. Now we repeat this process for every point in the sample. If there is a point with a unique largest *ldepth*, it is the median. If the point is not unique, then the median can be any point on or between those two points. In the multivariate case, the algorithm is similar but instead of drawing lines through points, we draw half-spaces. Throughout, we use the publicly available Fortran code provided by Struyf and Rousseeuw (1998) to calculate the Tukey median.[6]

Given this multivariate measure of central tendency, we now generalize the concept of MAD by analogy. Let $m_{t,\tau} = (m_{t,\tau}^{(1)}, ..., m_{t,\tau}^{(4)})'$ denote the four-variate Tukey median and let $|x_{i,t,\tau} - m_{t,\tau}|$ denote the vector of absolute deviations from the median. The (Tukey) median absolute deviation is then $MAD_{t,\tau} = (MAD_{t,\tau}^{(1)}, ..., MAD_{t,\tau}^{(4)})' = med_{1 \leq i \leq n_{t,\tau}}\{|x_{i,t,\tau} - m_{t,\tau}|\}$, which produces an outlier-robust measure of the vector of "typical" deviations from the median. Note that this approach does capture at least some of the effect of comovements across the individual elements. A normalized distance metric can then be constructed as

$$D(x_{i,t,\tau}) = \sum\nolimits_{j=1}^{4} \frac{|x_{i,t,h}^{(j)} - m_{t,\tau}^{(j)}|}{MAD_{t,\tau}^{(j)}}. \tag{4}$$

3 Disagreement

In this section we present the aggregate level of disagreement among the forecasters and attempt to discern any patterns in the degree of disagreement by the individual members of the FOMC. In each instance, we begin with an overview focused on the disagreement among the vector-valued forecasts. We then provide some discussion on disagreement related to the individual elements of the forecasts. Given our very limited data, we do not pursue identifying treatment effects using the panel data methods proposed by Davies and Lahiri

[5] See Rousseeuw and Struyf (1998) for a detailed discussion of the Tukey median.
[6] ftp://ftp.win.ua.ac.be/pub/software/agoras/newfiles/ldeptha.gz

(1995) or Lahiri and Sheng (2008). Instead we use HAC-robust t-tests of equal means between the relevant groups.

Figure 2 provides the sample paths of aggregate disagreement among the FOMC forecasts using the square root of the determinant of $S_{t,\tau}$ as the relevant metric. The plot consists of three lines, one for each of the forecast horizons. There is little clear evidence of any patterns among the lines, but one could certainly argue that in aggregate, forecast disagreement is lowest at the shortest (5-month ahead) forecast horizon.

Figure 3 provides the same plots of disagreement but subdivided by element (and hence the plots are of $s_{t,\tau}$). In most cases, there is little clear evidence of any patterns among the lines. But again, there is some indication that forecast disagreement is lowest at the shortest forecast horizon. This is particularly true for the CPI forecasts for which the degree of disagreement is monotone increasing in the forecast horizon at every forecast origin. Recall that McNees (1995) documents that the average range of the FOMC forecasts is increasing in the forecast horizon.

This is a somewhat surprising result since, when using standard OLS regression methods for constructing forecasts, all forecasts are expected to eventually converge to the (historical) sample mean. Hence, regardless of whether the "models" used by members of the FOMC are different, eventually one would expect the forecasts to disagree less. Our contrasting observation (again, especially for CPI inflation) suggests that the forecasts are not being constructed in a minimum mean square error (MSE) sense but are being constructed for other reasons.[7] While other statistical loss functions could explain this result (e.g. Capistrán and Timmermann, 2008), Ellison and Sargent (2009) argue that the FOMC members are being strategic when putting their forecasts together. In particular, in the context of a model of robust decision making, they argue that the forecasts are a strategic tool for convincing the other members of their policy view. As such, the members have incentives to (say) raise their inflation forecasts if they think policy should be tighter or lower their inflation forecasts if they think policy should be looser—regardless of what they think the actual level of inflation will truly be.

If there are horizon-driven disagreement effects, and we admit they are difficult to identify given our limited dataset, they will have to be accounted for when we try to identify

[7] This observation also lends some criticism to Romer and Romer's (2008) suggestions that the FOMC forecasts would be more "accurate" if they were to adapt the Greenbook forecasts instead of their own. Such a suggestion presumes that the Board of governors staff has the same loss function as the members of the FOMC.

other treatment effects. Fortunately, our baseline measure of disagreement $D(x_{i,t,\tau})$, at least in part, mitigates the issue by rescaling the nominal distances separately for each time period. To see whether the rescaling does in fact remove all horizon-based effects, in Table 1 we report the mean levels of disagreement for each horizon at both the multivariate level and individually for each of the four variables. In each case, a simple t-test of equal means among the three horizons fails to reject the null of equal disagreement. As such, we proceed with our analysis treating the normalized distance measures as having fully accounted for the horizon effects.

3.1 Voting Status

Table 2 provides the mean levels of disagreement by horizon and across all horizons for voting and non-voting members of the FOMC. Columns 2 and 3 are the values including all members of the FOMC. In each instance, there is no statistically significant difference in the level of disagreement based on voting status. In column 4 we make the same comparison but exclude the New York Fed and the governors from the analysis because they always vote. In broad terms, the results are unchanged, though in one instance we find that for real GDP forecasts at the 10-month ahead horizon, the regional banks tend to have a greater level of disagreement when they are not voting than when they are voting. Even so, there are 40 t-tests in Table 2 and even if the null hypothesis held in each case, we would expect some spurious rejections simply due to multiple testing.

3.2 Regional Bank and Governor

Tables 3 and 4 decompose the mean levels of disagreement to a finer level. In column 1 of both tables, each regional bank, governor, or Vice Chairman is listed. Associated with these FOMC participants, the second column provides the average level of disagreement at the multivariate level in Table 3 and by individual variable in Table 4. The third and fourth columns further distinguish the level by voting status. The remaining three columns provide p-values associated with simple t-tests for equal means for comparisons based on voting status, comparisons with the governors, and comparisons with the Vice Chairman.

We begin by simply noting the biggest and smallest values of average disagreement. At the multivariate level, the Cleveland Fed has a high level of disagreement with the other FOMC members, but the St. Louis Fed has—by a substantial margin—the highest level of disagreement. In contrast, the Vice Chairman exhibits the lowest level of disagreement among the FOMC members. Column 7 of Table 3 formally tests whether the individuals exhibit significantly different mean levels of disagreement from the Vice Chairman. We reject the null of equal disagreement between the Vice Chairman and the overall group with a p-value of 2.3%. We also reject the null of equal disagreement between the Vice Chairman and both the Cleveland and St. Louis Feds individually. Interestingly, we also reject the null when comparing the average among the other governors with that of the Vice Chairman. Strengthening the argument that the Cleveland and St. Louis Feds forecasts are outliers is the comparison, made in column 6, between the regional banks and the governors as a whole. Again we find that only the St. Louis and Cleveland Feds exhibit significantly different levels of mean disagreement. This is despite the fact that, as seen in column 2, the mean degree of disagreement by the governors is the fourth highest behind only Cleveland, Minneapolis, and St. Louis!

Continuing with the multivariate comparisons in Table 3, in column 5 we provide p-values associated with a t-test for equal mean disagreement based on the voting status for each regional bank (except New York, which always has a vote). Cleveland, Dallas, and to a lesser extent Minneapolis and Philadelphia, seem to exhibit different levels of disagreement based on their voting status. Whereas the Dallas Fed seems to be in greater disagreement when voting, the Cleveland Fed seems to be more consensus oriented when voting. That said, the Cleveland Fed should by no means be seen as a consensus builder. When voting they still have the third-highest level of disagreement behind only the Dallas and St. Louis Feds. It is also interesting to note that we *fail to reject* the null that the St. Louis Fed varies its level of disagreement by voting status. The implication is that the St. Louis Fed has arguably the highest degree of disagreement among the FOMC members regardless of its voting status.

In each panel of Table 4, the same comparisons are made but decomposed into nominal and real GDP growths, CPI inflation, and the unemployment rate. In each case, the St. Louis Fed has one of the two highest values of mean disagreement. Moreover, when comparing these values with those of the Vice Chairman or the governors as a whole, we reject the null of equal mean disagreement 7 of 8 instances at the 10% level.

Looking across the four panels, it appears that there is a wider range of degrees of disagreement on the nominal side than on the real side. In the panels associated with real GDP growth and the unemployment rate, only 2 of 26 tests have statistically significant differences in mean disagreement between the Vice Chairman and either the regional banks or other governors. Moreover, there are only 3 of 24 instances between both panels in which a regional bank has significantly different levels of mean disagreement from the governors. If we exclude the St. Louis Fed from our tally, there would be no significant difference between the disagreement in the Vice Chairman and other regional banks or other governors and only two significant differences in disagreement between the governors and a regional bank.

This is in contrast to the results in the panels associated with nominal GDP growth and inflation. Here we find that the Cleveland, Minneapolis, and St. Louis Feds, and to a lesser extent the governors, exhibit statistically significant different degrees of disagreement with the Vice Chairman. Moreover, these and several other regional banks exhibit significant levels of differences in disagreement with the governors. Among these instances, the Minneapolis, Cleveland, and St. Louis Feds exhibit greater disagreement while the Atlanta, Philadelphia, San Francisco, and Chicago Feds exhibit less disagreement than the governors.

When separated by voting status, there are a few instances of statistically significant differences in the level of disagreement. Richmond, Cleveland and the Minneapolis Feds appear to become more consensus-oriented in their inflation forecasts when voting, whereas Chicago appears less consensus-oriented when voting. And while the Philadelphia Fed seems to be more consensus-oriented with their unemployment forecasts when voting, the Richmond Fed tends to exhibit greater disagreement when voting—we reject the null of equal disagreement with p-value of 1.6%. Interestingly, while at the multivariate level the Dallas Fed exhibits a significant increase in disagreement when voting, it does not exhibit any significant differences in any one of the individual subcases. This is also marginally the case for Philadelphia Fed in which it exhibits significant decrease in disagreement at the multivariate level with a p-value of 10.5% when voting even though three out of four times its voting pattern did not change in the scalar case—thus highlighting our view that multivariate comparisons of disagreement provide additional information not contained in any of the scalar cases.

3.3 Relative to the SPF

In this section we look at disagreement among the FOMC members but couched in a larger world of forecasts made by other professional forecasters. Specifically, we imbed the SPF forecasts with those made by the FOMC for each February forecast. We restrict attention to these forecasts because the information sets, while not perfectly timed, are significantly better timed than those associated with the July forecasts. By doing so we add 216 more individual-year observations to the population of forecasts made in February.

The purpose of this exercise is to get a feel for whether or not the degree of disagreement among the FOMC members is "large" or "small." In order to reach such a conclusion, we need other forecasts to serve as a baseline and the SPF is a well known and timely collection of publicly available forecasts. Even so, we admit that there is a sense in which we are mixing apples and oranges: The FOMC forecasts are conditional while those from the SPF are unconditional.

With this caveat in mind, Figure 4 provides a box-and-whisker plot of each individual's measure of vector-valued disagreement. The red asterisks denote disagreements associated with the SPF while the blue circles are those associated with the FOMC members. One immediately notices that the dots associated with the FOMC are on the left side of the plot while the SPF's asterisks are more likely to be on the right side and hence, at least visually, it appears that members of the SPF exhibit far higher levels of disagreement than members of the FOMC.[8] For the sake of comparison, we also include the levels of disagreement based on the Greenbook forecasts associated with the January FOMC meeting. The green squares associated with these forecasts appear to be centrally located relative to the FOMC and SPF forecasts.

Table 5 provides the detailed measures of disagreement among the SPF, the FOMC members, and the various subgroups of the FOMC in our analysis. As noted in Figure 4, the most obvious result is simply that the degree of disagreement among the SPF is much larger than any disagreement among the FOMC and any of its subgroups. Though not reported here, all t-tests for equal mean disagreement between the SPF and members of the FOMC (i.e., SPF vs. FOMC, SPF vs. Voters) are statistically significant at a very high level. Also not reported, when the FOMC is couched in this larger universe of forecasters, we find no evidence of statistically significant levels of disagreement among the FOMC members—a result driven by the fact that any disagreement among the FOMC members is swamped by the aggregate degree of disagreement including that from the SPF.

[8] In this plot, there is no substance to the vertical axis. The "height" associated with any point is chosen at random simply to prevent the dots from piling on top of one another in the graph.

3.4 Robustness

One criticism of our results is that our preferred measure of disagreement is fundamentally based on means rather than medians, and hence outliers may be unduly influencing the measure of central tendency from which we base the degree of disagreement. As noted previously, we replicated the results in Tables 1 through 5 using medians as the outlier-robust measure of central tendency and used MAD as the outlier-robust measure of a "typical" distance (these tables are available on request).

Although the nominal measurements of disagreement are very different across the two metrics, in most instances our characterizations of "significant" outliers are unchanged. The outlier-robust variant of Table 2 fails to reject the null of voter status effects among the regional banks in each instance except for the very same one case relating to real GDP growth at the 10-month horizon. The outlier-robust variants of Table 3 and 4 are slightly less similar but still very highly correlated. Of the 9 instances in which Table 3 reports a p-value less than 10%, the outlier-robust variant matches 7 times. The remaining few instances indicate some differences relating to the metric. The mean-based metric finds that at the multivariate level, Cleveland, Dallas, and Philadelphia show different degrees of disagreement when voting than when not voting while the median-based metric fails to reject the null of equal level of disagreement when voting or not. The outlier-robust variant of Table 4 has a similar rate of success matching with Table 4: 22 of 28 times. Of these that do not match, many differ by basis points around the 10% threshold.

When we imbed the SPF with the FOMC members, an outlier-robust variant of Table 5 continues to show the same patterns. The SPF has a much higher degree of disagreement than the FOMC members. Tests of equal disagreement between the SPF and the FOMC members are highly significant. Again, in the greater universe of forecasting agents, we fail to reject the null of equal disagreement among the FOMC subgroups.

To be fair, we should make clear that we are not interpreting the similarity of results as support of our main conclusions. Rather, we interpret the similarity as at least not contrasting with our observations using the means-based metric. Our caution stems from some dissimilarities between Table 1 and its outlier-robust variant. Recall that in Table 1 we fail to reject the null of any remaining seasonal effects induced by horizon after scaling. In contrast, the outlier-robust variant still finds strong evidence of horizon-based effects in average disagreement at the 5-month horizon for both real GDP growth and CPI inflation.

That is, we reject the null of equal disagreement for the 10-month vs. 5-month and 17-month vs. 5-month comparisons for both real GDP and CPI.

4 Accuracy and Disagreement

In this section we describe the accuracy of the forecasts provided by the FOMC with an eye toward any linkages with disagreement. For each of the individual members of the FOMC this is straightforward because we have the actual forecasts. For the FOMC in aggregate, recall that there is no single "forecast" reported in the *MPR*; the *MPR* reports only the range and trimmed range. While others have chosen to use the midpoint of the range or the trimmed range as the FOMC "forecast," we use the trimmed mean constructed as the simple average of the sample after dropping the three highest and three lowest values of the variable. Finally, although other loss functions could be used to characterize forecast accuracy (Capistrán and Timmermann, 2008) we restrict attention to the most commonly used quadratic loss function noting, however, that there is no evidence suggesting that the members of the FOMC construct their forecasts with this loss function in mind.

4.1 The Trimmed Mean Forecast

Before presenting our results on the accuracy of the forecasts it is useful to take a closer look at how disagreement affects the behavior of the trimmed mean forecast. Since this forecast is constructed by first dropping the three lowest and three highest forecasts of that variable and then taking the simple average of the remaining forecasts, by definition this implies that individuals with greater degrees of disagreement are less likely to have their forecasts explicitly incorporated in the trimmed mean forecast.

In Table 6, for each regional bank, Vice Chairman, and the governors as a group we provide the percentage of forecasts excluded from the trimmed mean forecast for each variable. Panel A shows all horizons while in the remaining panels this exclusion is subdivided by each of the three forecast horizons.[9,10] Not surprisingly given our previous results on disagreement among the FOMC, in columns 2 through 5 of Panel A we find that, averaging

[9] For the regional banks and the Vice Chairman there are a maximum of 21 forecasts, with a maximum of 7 for each horizon. For the governors the maximum is larger because we aggregate across all of the governors.

[10] In some instances there is a tie for the third-highest or third-lowest value of the forecast. While it is irrelevant which value is dropped for the nominal value of the trimmed mean forecasts it does affect our percentages of times a value was dropped by individual. When a tie exists, we randomize among the choices using equal weights across the individuals. As such, these percentages should be viewed as approximations.

across all horizons, the St. Louis Fed had its forecast dropped from the trimmed mean either the most or second most often for each of the four variables. In fact, the St. Louis Fed forecast is dropped more than half of the time for each variable and is dropped roughly 86% of the time for the nominal GDP growth forecast. Moreover, at the 17-month horizon the nominal GDP growth forecast is dropped from the trimmed mean 100% of the time!

Note, however, that since this is done variable by variable, some individual members of the FOMC do have their forecast explicitly incorporated into (say) the trimmed mean nominal GDP growth forecast but not the trimmed mean CPI inflation forecast. But if the vector-valued forecasts are constructed in the congruent fashion we expect them to be (that is, taking account of the linkages across the variables), there is a sense in which the vector of trimmed mean forecasts is still including forecasts from individuals who are "multivariate outliers." For example, consider the St. Louis Fed 17-month ahead forecasts. While it is true that their nominal GDP growth forecasts are always excluded, at times the real GDP growth, CPI inflation, and unemployment forecasts are included in the trimmed mean despite the fact that the St. Louis Fed forecast as a whole has an outlier mentality relative to the majority of the FOMC.

In the final column of each panel we therefore consider a slightly different approach to constructing the trimmed mean forecasts that is based on our multivariate measure of disagreement. Specifically, for each time period and horizon, we construct the measure of disagreement for each vector-valued forecast and "trim" those with the 6 largest levels of disagreement—analogous to the present approach that drops the 3 largest and smallest values of the forecast. This approach omits those forecasts that, considered as a vector, are least in agreement with the FOMC as a whole. Using this trimming rule, over all the horizons, in panel A we see that the St. Louis Fed is dropped more than 80% of the time and at the 17-month horizon it is dropped 100% of the time. In contrast, the Atlanta and Richmond Feds are rarely dropped; in fact, at the 17-month horizon they never are.

4.2 Mean Square Errors

We now proceed to documenting the accuracy of the FOMC forecasts. In our approach we calculate the mean square errors (MSEs) associated with each of the regional banks, the Vice Chairman, and the governors separately for each forecast horizon and for each of the four variables. In addition, we evaluate the accuracy of the trimmed mean forecast as our proxy for the FOMC forecast as reported in the *MPR*. For comparison we also report a few other forecasts that could have been constructed using the forecasts from the FOMC: the equally weighted average of the forecasts without trimming, the equally weighted average formed using those forecasts that were trimmed (i.e., the average of the highest 3 and lowest 3 forecasts), and the trimmed mean forecast using the concept of multivariate trimming considered in the previous section. As a further source of comparison, we also include the forecast associated with the median of the SPF and the Greenbook forecasts.

Table 7 reports these MSEs. Specifically, the first row provides the MSEs of the trimmed mean forecasts by variable and horizon. The remaining elements of the rows provide the ratio of the MSEs for that row relative to that for the trimmed mean. A number smaller (larger) than 1 indicates that the individual associated with the row was on average more (less) accurate than the trimmed mean forecast. Before proceeding, we should note that due to the extremely small sample sizes in each of the cells (which are typically based on 7 observations) we make no attempts to test for statistical significance across the MSEs by group. Whereas we felt that our normalizations removed the "horizon-based" effects in our analysis of disagreement (and hence we were willing to aggregate across horizons after normalizing), we feel much less comfortable doing so when measuring accuracy. As such, all of our observations should be interpreted keeping the small sample sizes in mind.

With that caveat, we begin by first noting that in nearly all cases, the MSEs of the forecasts decreases as the event horizon shrinks. For example, the trimmed mean forecast of nominal GDP growth has MSEs of 1.187, 0.929, and 0.211 for the 17-month, 10-month and 5-month horizons respectively. In general, the trimmed mean forecast tends to perform better than the individual members in terms of MSE—and even more so for nominal GDP growth and inflation than for real GDP growth and unemployment. For the nominal variables, the trimmed mean is better than 9 to 10 of the individuals at each horizon. Only the Philadelphia and Richmond Feds produce forecasts of the nominal variables that are more accurate than the trimmed mean for more than half the horizons. But for the real variables, the trimmed mean does better than only 6 to 9 of the individuals. Atlanta, Chicago, Dallas, Minneapolis, Richmond, and St. Louis each are more accurate than the trimmed mean for more than half the horizons. Overall, the New York and San Francisco Feds were least likely (on average) to be more accurate than the trimmed mean, outperforming it only once and twice, respectively. In contrast, the Richmond and Philadelphia Feds were more accurate than the trimmed mean 9 and 8 times, respectively. Interestingly, the governors and Vice Chairman were among the worst forecasters relative to the trimmed mean with one very notable exception: They nearly always did better forecasting inflation.

The bottom portion of each panel reports MSEs for the SPF, Greenbook, and some alternative model averaging–type forecasts that could have been constructed with the FOMC forecasts. In 9 of 12 comparisons, the simple average of the FOMC forecasts has a lower MSE than the trimmed mean forecast—though, admittedly, in most instances the relative gains in accuracy are small. Our alternative trimmed forecast, based on trimming vectors as a whole, had a lower MSE in 8 of 12 comparisons relative to the trimmed mean. In those instances where it did worse, the relative losses were very small but in some of those instances in which it did better, the gains were a substantial 10% or more.

Amusingly, in 9 of 12 instances, the simple average of the forecasts that were "trimmed" by the FOMC did better than the trimmed mean forecast itself. And as was the case for our multivariate trimmed forecast, when it did worse, the relative losses were small while in instances where it did better, the gains were 10% and even 20%.

4.3 Linking Disagreement and Accuracy

Here we attempt to identify any empirical connections between disagreement and accuracy. Our approach is partially motivated by Lahiri and Sheng (2009) who provide a theoretical link between disagreement among forecasters and aggregate forecast uncertainty. To do so, first let $\hat{u}_{i,t,\tau}^{(j)^2}$ denote the squared forecast error of variable j, associated with forecasts made at time t, with horizon τ, made by individual i. If we then let $D(x_{i,t,\tau}^{(j')})$ denote an individual's level of disagreement on variable j', and let RB and V denote dummy variables for regional bank and voting respectively, we estimate the following pooled regression (pooled across i and t) separately for each variable j, j', and horizon τ:

$$\hat{u}_{i,t,\tau}^{(j)^2} = \alpha_1 V_{i,t} + \alpha_2 RB_i + \alpha_3 V_{i,t} \cdot RB_i + D(x_{i,t,\tau}^{(j')})(\beta_1 V_{i,t} + \beta_2 RB_i + \beta_3 V_{i,t} \cdot RB_i) + \epsilon_{i,t,\tau}.$$

The first three predictors—those associated with the α's—are controls. The latter three predictors—those associated with the β's—are the ones on which we focus our attention. We use these predictors to parse out any effects the level of disagreement may have on the accuracy of the forecasts. In particular, the goal is to identify any disagreement effects driven by whether that individual is a governor ($H_0 : \beta_1 = 0$), voting regional bank president ($H_0 : \beta_1 + \beta_2 + \beta_3 = 0$), or non-voting regional bank president ($H_0 : \beta_2 = 0$).

Table 8 reports the results of the pooled regressions linking the accuracy of CPI inflation forecasts to disagreement among either CPI inflation forecasts (columns 2, 4, and 6) or nominal GDP growth forecasts (columns 3, 5, and 7). We find no significant evidence that the degree of disagreement by the governors affects the accuracy of their forecasts (i.e., $\beta_1 = 0$) at any horizon. Similarly, in the bottom panel we find no evidence that the degree of disagreement by voting regional banks affects the accuracy of their forecasts (i.e., $\beta_1 + \beta_2 + \beta_3 = 0$). This, in turn, is supported by no evidence of differences in disagreement effects between the governors and the voting regional banks (i.e., $\beta_1 - (\beta_1 + \beta_2 + \beta_3) = 0$).

However, at the two longest horizons—those most associated with policy decisions—the level of disagreement on nominal variables among non-voting members of the regional banks has a negative impact on the accuracy of their corresponding inflation forecasts (i.e., $\beta_2 > 0$). This is reinforced in the bottom panel by the fact that for these same horizons we find a significant difference in the effect of disagreement by governors and non-voting regional banks (i.e., $\beta_1 - \beta_2 < 0$) with p-values all less than 10%. We also witness some significant differences in disagreement between voting and non-voting regional bank (i.e., $\beta_1 + \beta_2 + \beta_3 - [\beta_2] < 0$).

In contrast, in unreported results, we find little evidence of any relationship between disagreement among forecasts of real variables and the accuracy of CPI inflation forecasts. In addition, we find little evidence of any measure of disagreement and the accuracy of forecasts of real GDP growth, unemployment, and nominal GDP growth. This lack of significance supports the notion that the CPI inflation forecasts play a special role among the FOMC members above and beyond simply being a forecast of an unknown future event. While we can only conjecture what this role may be, one interpretation is that as a non-voting member of the FOMC, these regional banks are regularly reporting their inflation (and nominal GDP growth) forecasts not so much as an indicator of what they expect future values of inflation to be but as a worst-case scenario (Ellison and Sargent, 2009) designed to influence the present voting members of their view of monetary policy.

5 Dissent and Disagreement

This section briefly looks at whether an individual's level of disagreement is related to whether that individual casts a dissenting vote at the corresponding FOMC meeting. Unfortunately, trying to make such a connection is seriously limited by the available data. For example, not only do we have data that span a mere 7 years, our forecasts are associated only with the February and July FOMC meetings, which are only a portion of the FOMC meetings in a given year. Making the situation even harder is the fact that during the time frame for which we have data, not a single voting member of the FOMC dissented in any of the February meetings.

For the July meetings, there were a total of 7 dissenting votes cast, among a total of 143 votes. To see if these dissenting votes are related to an individual's forecast disagreement, Table 9 reports the mean levels of disagreement among voters based on whether the member dissented. The top panel contains the results for disagreement among the 17-month ahead forecasts and the lower panel contains those for the 5-month ahead forecasts. In each, the first row is related to vector-valued disagreement while the remaining rows relate to disagreement for the individual forecasts.

A quick look at columns 2 and 3 indicates that in most cases, disagreement was on average higher among those who cast dissenting votes. This is particularly true at the shorter time horizon. Columns 5 reports p-values associated with t-tests of equal mean disagreement between those who did not dissent and those who dissented. Column 6 does the same but between non-voters and dissenters. Given our very small sample sizes, there is very little evidence of statistically significant differences in mean disagreement among these groups.

6 Conclusion

Using a novel dataset, we characterize the degree of disagreement and accuracy of the FOMC forecasts used in the *Monetary Policy Report* to Congress. While the time duration of the dataset is very limited, we feel that a handful of patterns related to the forecast horizon, related to whether the member is a regional bank president, governor, or Vice Chairman, and whether the individual is a voting member of the FOMC are fairly clear.

Although it is difficult to parse out explicitly from our limited dataset, we believe that underlying many of our results is the fact that, as noted by Bullard (2009) as well as Ellison and Sargent (2009), the members of the FOMC construct their forecasts for reasons other than accuracy as measured by MSEs. Since we most clearly observe this in the CPI-based inflation forecasts one can infer that is where the battle lines were typically drawn over the time frame of our dataset.[11] Reinforcing that argument is the empirical observation that it is the nominal variables for which there are the most significant deviations in mean disagreement between the regional banks and the Vice Chairman, whom we find to be one of the most consensus-oriented members of the FOMC.

[11] In other words, there is a reason the terms "inflation hawk" and "inflation dove" are common decriptions of FOMC members. Put differently, one never hears a member of the FOMC described as an "unemployment hawk" or "unemployment dove."

References

Bullard, James (2009), "Discussion of Ellison and Sargent: What Questions are Staff and FOMC Forecasts Supposed to Answer?," Powerpoint Presentation at the 10th EABCN Workshop on Uncertainty over the Business Cycle, European Central Bank, Frankfurt, March 30, 2009.

Capistrán, Carlos and Allan Timmermann (2009), "Disagreement and Biases in Inflation Expectations," *Journal of Money, Credit and Banking*, 41(2-3), 365-396.

Chappell Jr., Henry W. and Rob Roy McGregor (2000), "A Long History of FOMC Voting Behavior," *Southern Economic Journal*, 66, 906-922.

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

Ellison, Martin and Thomas J. Sargent (2009), "A Defence of the FOMC," CEPR/EABCN Working Paper 7510.

Faust, Jon and Jonathan H. Wright (2008), "Efficient Forecast Tests for Conditional Policy Forecasts," *Journal of Econometrics*, 146, 293-303.

Gavin, William T. (2003), "FOMC Forecast: Is All the Information in the Central Tendency?" *Federal Reserve Bank of St. Louis Review*, May/June, 85, 27-48.

Gavin, William T. and Mandal, Rachel J. (2003), "Evaluating FOMC Forecasts," *International Journal of Forecasting*, 19, 655-67.

Gavin, William T. and Geetanjali Pande (2008), "FOMC Consensus Forecasts," *Federal Reserve Bank of St. Louis Review*, May/June, 90 (Part 1), 146-163.

Lahiri, Kajal and Xuguang Sheng (2008), "Evolution of Forecast Disagreement in a Bayesian Learning Model," *Journal of Econometrics*, 144, 325-340.

Lahiri, Kajal and Xuguang Sheng (2009), "Measuring Forecast Uncertainty by Disagreement: The Missing Link," *Journal of Applied Econometrics*, forthcoming.

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

McNees, Stephen K. (1995), "An Assessment of the 'Official' Economic Forecasts," *New England Economic Review*, July/August, 13-23.

Meade, Ellen E. and D. Nathan Sheets (2005), "Regional Influences on FOMC Voting Patterns," *Journal of Money, Credit, and Banking*, 37, 661-677.

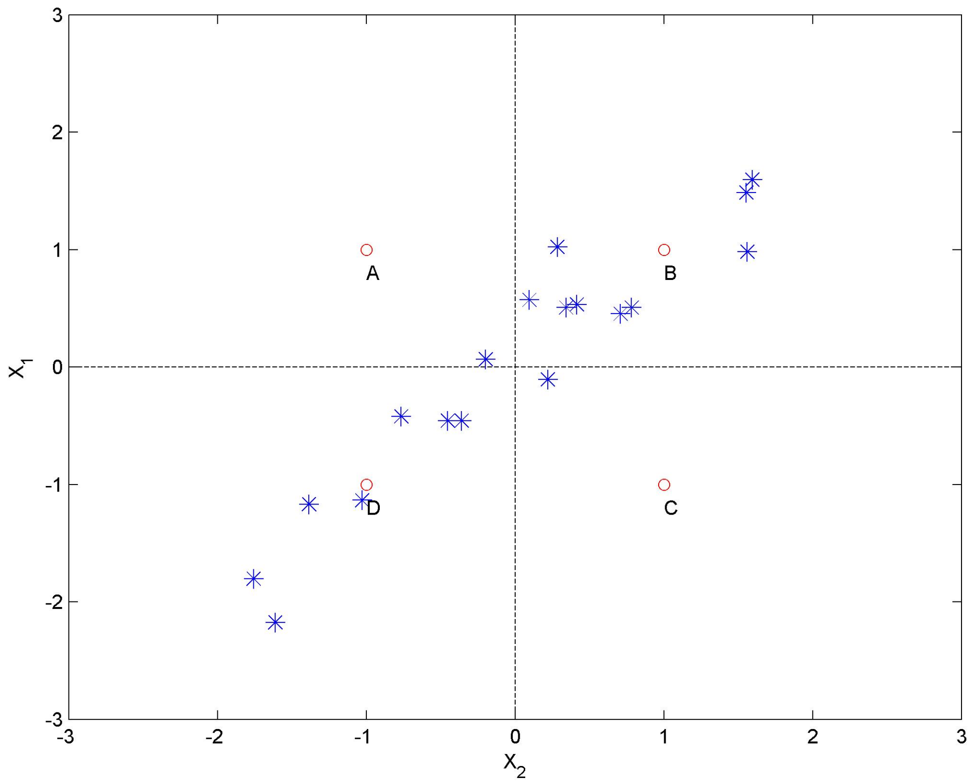
Romer, David H. (2009), "A New Data Set on Monetary Policy: The Economic Forecasts of Individual Members of the FOMC," NBER Working Paper 15208.

Romer, Christina D. and David H. Romer (2008), "The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?" *American Economic Review*, 98, 230-235.

Rousseeuw, Peter J. and Struyf, A. (1998), "Computing Location Depth and Regression Depth in Higher Dimensions," *Statistics and Computing*, 8, 192-203.

Rousseeuw, Peter J. and Struyf, A. (2000), "High-dimensional computation of the deepest location," *Computational Statistics and Data Analysis*, 34, 415-426.

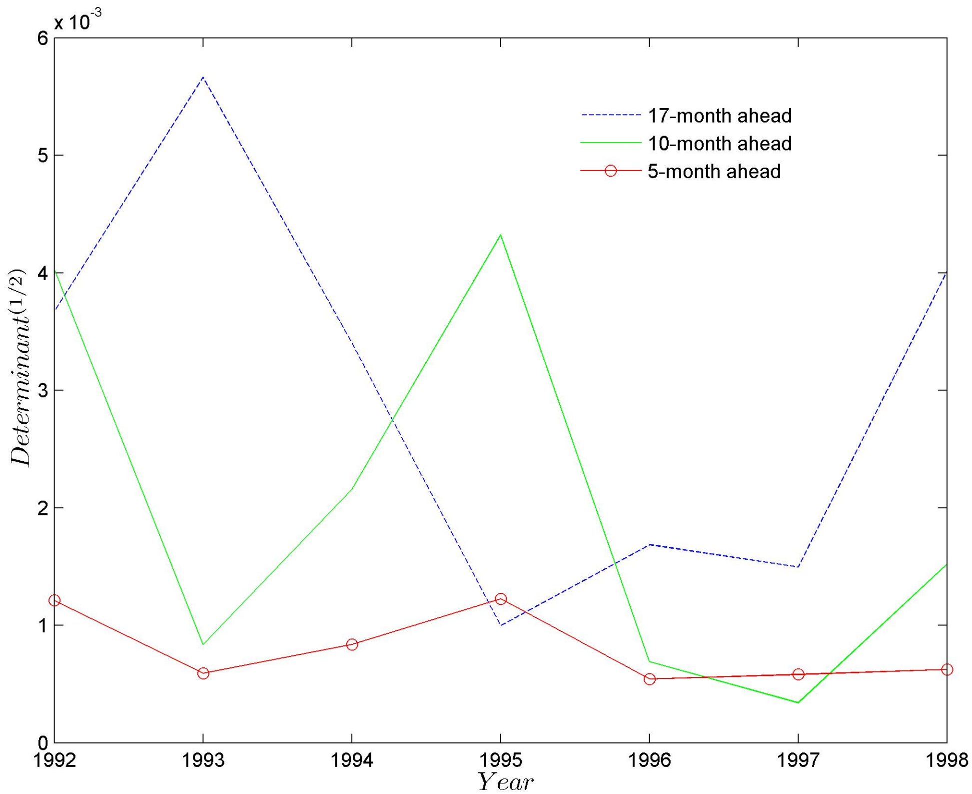
Figure 1: Bivariate Disagreement



Note:

(i) Data were generated as bivariate N(0,1) with correlation 0.90. The Euclidean distance from the mean for points A,B,C and D are all 1.41 whereas the estimates of the Mahalanobis distances for those points are 5.59, 1.23, 5.59, and 1.23, respectively.

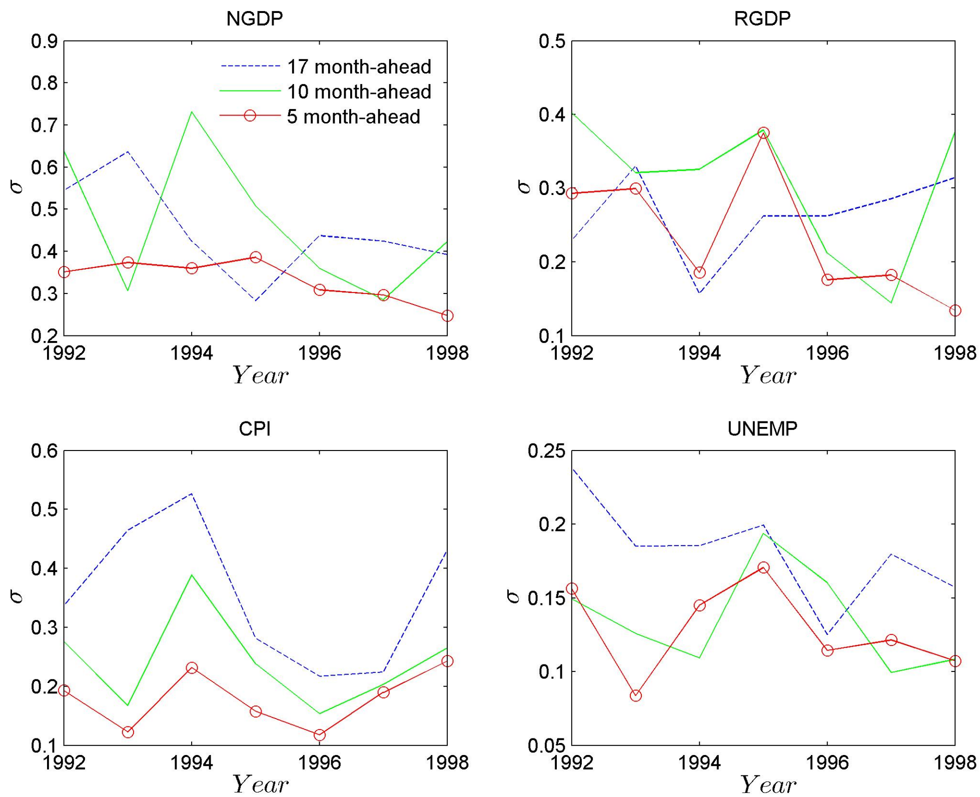
Figure 2: FOMC Multivariate Disagreement



Note:

(i) The lines consist of seven points. Each point is the square root of the determinant of the sample covariance of the vectors of forecasts.

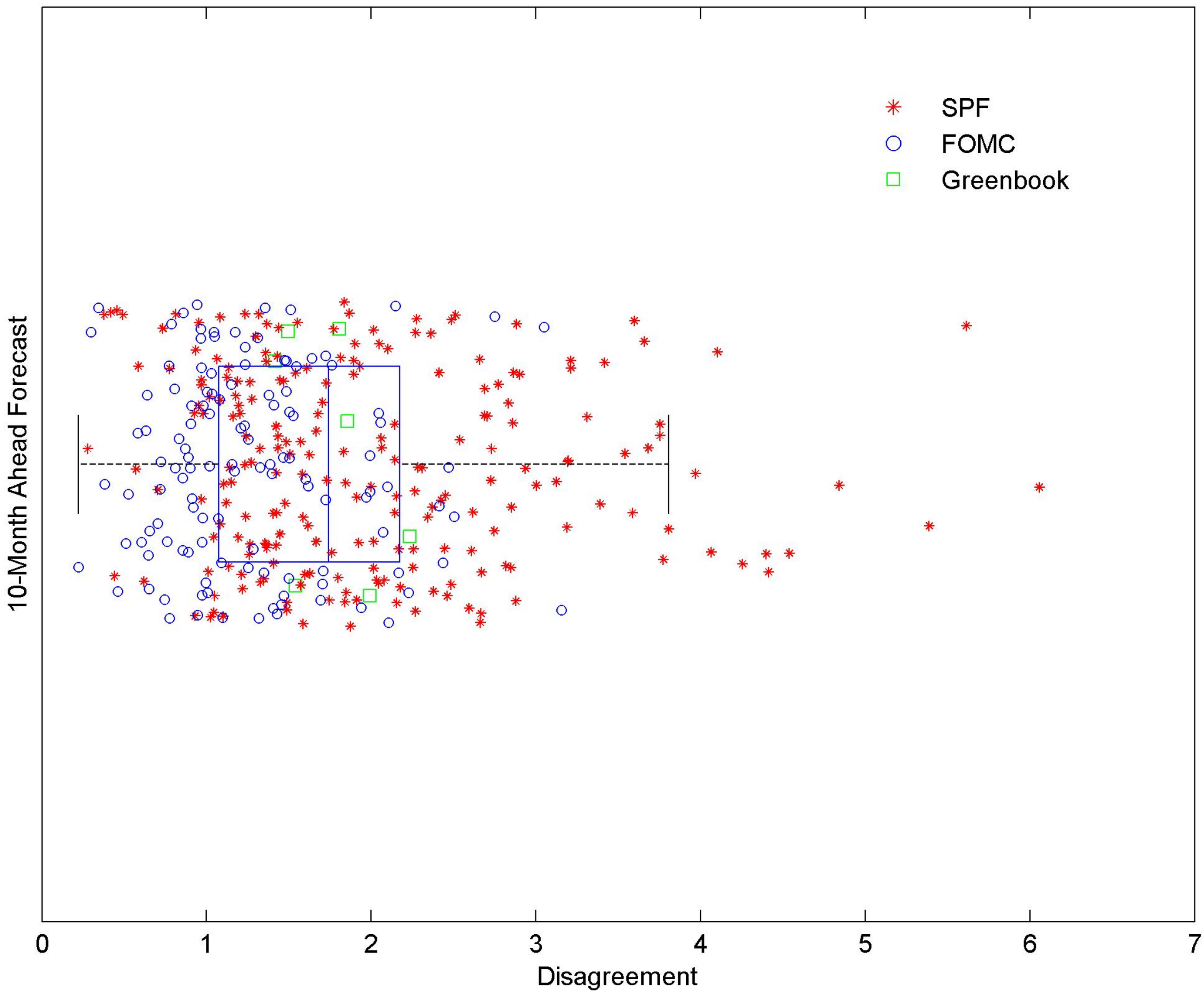
Figure 3: FOMC Scalar Disagreement



Notes:

(i) The lines consist of 7 points. Each point is the sample standard deviation of the scalar forecasts.

Figure 4: Multivariate Disagreement of FOMC and SPF



Notes:

(i) Values are calculated using equation (1). Box corresponds to the interquartile range (IQR). The mean is represented by the vertical line in the box. The whisker on the left(right) is 1.5 times less(more) than the IQR. Observations beyond the whiskers are considered outliers.

(ii) In this plot, there is no substance to the vertical axis. The "height" associated with any point is chosen at random simply to prevent the dots from piling on top of one another in the graph.

Table 1: **Mean Disagreement by Horizon**

	Mean			**p-value**		
	17-month	10-month	5-month	17m vs. 10m	10m vs. 5m	17m vs. 5m
Vector	1.840	1.844	1.832	0.962	0.885	0.927
NGDP	0.772	0.801	0.767	0.721	0.671	0.947
RGDP	0.750	0.780	0.744	0.705	0.647	0.948
CPI	0.736	0.734	0.807	0.983	0.362	0.409
UNEMP	0.784	0.753	0.758	0.698	0.950	0.738

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.
(ii) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (1). The scalar measures of disagreement were constructed similarly using equation (2).

Table 2: **Mean Disagreement by Voting Status**

	Mean			**p-value**	
	Voters	Nonvoters	Voters excl. NY/G	Voters vs. Nonvoters	Voters excl. NY/G vs. Nonvoters
Vector					
17-month	1.822	1.867	1.823	0.684	0.715
10-month	1.855	1.828	1.896	0.815	0.606
5-month	1.771	1.922	1.725	0.197	0.179
Total	1.816	1.872	1.816	0.395	0.461
Nominal GDP					
17-month	0.722	0.847	0.772	0.244	0.571
10-month	0.793	0.812	0.833	0.861	0.874
5-month	0.745	0.799	0.830	0.574	0.819
Total	0.753	0.820	0.812	0.272	0.917
Real GDP					
17-month	0.698	0.827	0.633	0.270	0.163
10-month	0.841	0.688	0.975	0.145	0.031**
5-month	0.709	0.797	0.688	0.446	0.467
Total	0.750	0.771	0.768	0.747	0.974
CPI					
17-month	0.745	0.724	0.884	0.862	0.281
10-month	0.688	0.804	0.704	0.334	0.480
5-month	0.812	0.800	0.853	0.910	0.653
Total	0.748	0.776	0.812	0.673	0.645
Unemployment					
17-month	0.773	0.799	0.765	0.786	0.779
10-month	0.792	0.694	0.648	0.391	0.709
5-month	0.750	0.770	0.709	0.863	0.674
Total	0.772	0.755	0.706	0.790	0.516

Notes:

(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.

(ii) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (1). The scalar measures of disagreement were constructed similarly using equation (2).

Table 3: **Mean Multivariate Disagreement by Individual**

	Mean			**p-value**		
	Aggregate	Voters	Non-voters	Voter vs. Non-voters	Agg. vs. Gov.	Agg. vs. Vice-chair
Atlanta	1.613	1.370	1.710	0.215	0.148	0.837
Boston	1.721	1.681	1.751	0.703	0.485	0.421
Chicago	1.820	1.806	1.831	0.911	0.983	0.172
Cleveland	2.176	1.936	2.589	0.017**	0.019**	0.002***
Dallas	1.726	2.173	1.503	0.010***	0.525	0.418
Kansas City	1.766	1.668	1.839	0.584	0.670	0.269
Minneapolis	1.834	1.468	1.980	0.053*	0.947	0.176
New York	1.780	1.780			0.809	0.329
Philadelphia	1.780	1.511	1.887	0.105	0.764	0.260
Richmond	1.720	1.676	1.737	0.801	0.434	0.393
San Francisco	1.727	1.727	1.727	0.999	0.462	0.370
St Louis	2.512	2.555	2.480	0.699	0.000***	0.000***
Governor	1.823	1.823				0.033**
Vice-chair	1.577	1.577				
Total	1.839	1.816	1.872	0.461	0.781	0.023**

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.
(ii) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (1).

Table 4: **Mean Scalar Disagreement by Voting Individuals**

	Mean			p-value				Mean			p-value		
	Aggregate	Voters	Non-voters	Voter vs. Non-voters	Agg. vs. Gov.	Agg. vs. Vice-chair		Aggregate	Voters	Non-voters	Voter vs. Non-voters	Agg. vs. Gov.	Agg. vs. Vice-chair
NGDP							CPI						
Atlanta	0.789	0.583	0.871	0.206	0.726	0.427	Atlanta	0.432	0.542	0.388	0.425	0.001***	0.303
Boston	0.657	0.587	0.710	0.574	0.458	0.990	Boston	0.600	0.634	0.574	0.769	0.203	0.779
Chicago	0.653	0.795	0.546	0.290	0.435	0.986	Chicago	0.555	0.728	0.426	0.026**	0.040**	0.959
Cleveland	0.896	0.894	0.900	0.983	0.368	0.231	Cleveland	1.446	1.266	1.754	0.096*	0.000***	0.000***
Dallas	0.751	0.979	0.636	0.194	0.943	0.562	Dallas	0.794	1.010	0.686	0.229	0.650	0.136
Kansas City	0.639	0.600	0.668	0.744	0.338	0.906	Kansas City	0.738	0.704	0.763	0.839	0.990	0.232
Minneapolis	1.144	0.819	1.275	0.098*	0.019**	0.017**	Minneapolis	0.936	0.487	1.116	0.011**	0.162	0.030**
New York	0.580	0.580			0.141	0.604	New York	0.543	0.543			0.085*	0.895
Philadelphia	0.559	0.473	0.593	0.323	0.056*	0.471	Philadelphia	0.674	0.682	0.670	0.963	0.630	0.474
Richmond	0.639	0.728	0.604	0.711	0.510	0.930	Richmond	0.723	0.430	0.840	0.034**	0.917	0.314
San Francisco	0.751	0.736	0.757	0.917	0.926	0.500	San Francisco	0.514	0.515	0.513	0.993	0.022**	0.706
St Louis	1.515	1.506	1.521	0.952	0.000***	0.000***	St Louis	1.334	1.360	1.314	0.898	0.006***	0.002***
Governor	0.741	0.741				0.386	Governor	0.736	0.736				0.074*
Vice-chair	0.655	0.655					Vice-chair	0.562	0.562				
Total	0.780	0.753	0.820	0.917	0.400	0.227	Total	0.759	0.748	0.776	0.645	0.651	0.045**
RGDP							UNEMP						
Atlanta	0.700	0.552	0.760	0.295	0.751	0.581	Atlanta	0.608	0.398	0.692	0.046**	0.145	0.551
Boston	0.780	0.594	0.919	0.234	0.802	0.932	Boston	0.776	0.803	0.756	0.867	0.896	0.691
Chicago	0.741	0.795	0.700	0.786	0.961	0.749	Chicago	0.780	0.713	0.830	0.498	0.896	0.628
Cleveland	0.981	0.841	1.221	0.372	0.186	0.327	Cleveland	0.923	0.762	1.200	0.309	0.439	0.236
Dallas	0.816	1.062	0.693	0.344	0.693	0.898	Dallas	0.602	0.625	0.591	0.887	0.199	0.573
Kansas City	0.667	0.679	0.659	0.930	0.584	0.447	Kansas City	0.767	0.813	0.732	0.755	0.778	0.640
Minneapolis	0.740	0.594	0.799	0.265	0.956	0.736	Minneapolis	0.550	0.569	0.543	0.884	0.012**	0.247
New York	0.686	0.686			0.614	0.453	New York	0.902	0.902			0.474	0.243
Philadelphia	0.749	0.960	0.665	0.162	0.990	0.767	Philadelphia	0.709	0.341	0.856	0.014**	0.481	0.961
Richmond	0.558	0.501	0.581	0.715	0.134	0.110	Richmond	0.790	1.103	0.665	0.016**	0.944	0.520
San Francisco	0.626	0.610	0.632	0.885	0.374	0.284	San Francisco	0.539	0.474	0.565	0.690	0.029**	0.270
St Louis	1.128	1.095	1.153	0.863	0.012**	0.048**	St Louis	1.078	0.919	1.198	0.353	0.165	0.096*
Governor	0.748	0.748				0.629	Governor	0.798	0.798				0.328
Vice-chair	0.792	0.792					Vice-chair	0.701	0.701				
Total	0.758	0.750	0.771	0.974	0.833	0.700	Total	0.765	0.772	0.755	0.516	0.495	0.519

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.
(ii) The scalar disagreement is calculated by averaging over all individuals' level of disagreement using equation (2).

Table 5: **Mean Disagreement of SPF and FOMC**

	SPF	FOMC	Voters	Nonvoters	Regional	Governors
Vector	2.000	1.272	1.281	1.257	1.255	1.311
NGDP	0.845	0.549	0.550	0.548	0.522	0.612
RGDP	0.787	0.446	0.485	0.387	0.430	0.483
CPI	0.826	0.620	0.568	0.697	0.640	0.571
UNEMP	0.889	0.495	0.524	0.452	0.476	0.541

Notes:

(i) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (1). The scalar measures of disagreement were constructed similarly using equation (2).

Table 6: **Percent of Forecasts Trimmed by Individual**

	NGDP	RGDP	CPI	UNEMP	Vector
Panel A: All Horizon					
Atlanta	0.524	0.286	0.095	0.143	0.143
Boston	0.286	0.429	0.238	0.333	0.286
Chicago	0.238	0.333	0.143	0.381	0.286
Cleveland	0.474	0.579	0.789	0.368	0.579
Dallas	0.333	0.429	0.333	0.143	0.381
Governor	0.296	0.361	0.352	0.407	0.361
Kansas City	0.333	0.190	0.524	0.476	0.238
Minneapolis	0.571	0.333	0.429	0.333	0.333
New York	0.286	0.238	0.333	0.381	0.381
Philadelphia	0.143	0.333	0.238	0.238	0.286
Richmond	0.190	0.190	0.238	0.381	0.190
San Francisco	0.286	0.190	0.238	0.190	0.286
St Louis	0.857	0.667	0.667	0.571	0.810
Vicechair	0.263	0.421	0.105	0.211	0.263
Total	0.352	0.352	0.352	0.352	0.352
Panel B: 17-month ahead					
Atlanta	0.571	0.286	0.000	0.286	0.000
Boston	0.429	0.571	0.143	0.429	0.286
Chicago	0.143	0.286	0.286	0.571	0.286
Cleveland	0.500	0.667	0.667	0.333	0.833
Dallas	0.286	0.429	0.429	0.000	0.143
Governor	0.306	0.333	0.306	0.361	0.417
Kansas City	0.286	0.143	0.571	0.429	0.286
Minneapolis	0.571	0.143	0.429	0.143	0.286
New York	0.286	0.286	0.429	0.286	0.429
Philadelphia	0.000	0.429	0.286	0.143	0.000
Richmond	0.143	0.143	0.286	0.429	0.000
San Francisco	0.286	0.286	0.286	0.286	0.429
St Louis	1.000	0.714	0.714	0.857	1.000
Vicechair	0.429	0.143	0.000	0.143	0.143
Total	0.353	0.353	0.353	0.353	0.353

	NGDP	RGDP	CPI	UNEMP	Vector
Panel C: 10-month ahead					
Atlanta	0.571	0.143	0.143	0.143	0.143
Boston	0.143	0.286	0.286	0.286	0.143
Chicago	0.286	0.429	0.143	0.429	0.571
Cleveland	0.143	0.429	0.857	0.429	0.429
Dallas	0.571	0.571	0.286	0.286	0.571
Governor	0.278	0.361	0.306	0.472	0.306
Kansas City	0.571	0.429	0.571	0.429	0.429
Minneapolis	0.857	0.429	0.571	0.429	0.571
New York	0.143	0.143	0.429	0.429	0.286
Philadelphia	0.143	0.286	0.143	0.429	0.429
Richmond	0.286	0.143	0.143	0.143	0.143
San Francisco	0.143	0.000	0.286	0.000	0.143
St Louis	0.714	0.857	0.571	0.143	0.571
Vicechair	0.200	0.600	0.400	0.000	0.400
Total	0.350	0.350	0.350	0.350	0.350
Panel D: 5-month ahead					
Atlanta	0.429	0.429	0.143	0.000	0.286
Boston	0.286	0.429	0.286	0.286	0.429
Chicago	0.286	0.286	0.000	0.143	0.000
Cleveland	0.833	0.667	0.833	0.333	0.500
Dallas	0.143	0.286	0.286	0.143	0.429
Governor	0.306	0.389	0.444	0.389	0.361
Kansas City	0.143	0.000	0.429	0.571	0.000
Minneapolis	0.286	0.429	0.286	0.429	0.143
New York	0.429	0.286	0.143	0.429	0.429
Philadelphia	0.286	0.286	0.286	0.143	0.429
Richmond	0.143	0.286	0.286	0.571	0.429
San Francisco	0.429	0.286	0.143	0.286	0.286
St Louis	0.857	0.429	0.714	0.714	0.857
Vicechair	0.143	0.571	0.000	0.429	0.286
Total	0.353	0.353	0.353	0.353	0.353

Notes:
(i) For each panel, values in columns 2-5 are the percent of forecasts omitted from the central range in the Monetary Policy Report.
(ii) For each panel, values in column 6 are the percent of forecasts trimmed by dropping the forecasts associated with 6 largest levels of multivariate disagreement.

Table 7: **Forecast Accuracy by Variable**

NGDP	17-month	10-month	5-month
Trimmed Mean	1.187	0.929	0.211
Atlanta	1.308	1.033	1.307
Boston	1.568	0.683	1.872
Chicago	1.045	0.969	0.754
Cleveland	1.325	1.405	2.712
Dallas	0.705	1.073	1.377
Kansas City	1.326	1.210	0.789
Minneapolis	0.889	1.211	0.576
New York	0.894	1.000	1.584
Philadelphia	0.747	0.993	1.107
Richmond	0.940	0.857	0.880
San Francisco	1.221	1.281	1.836
St Louis	1.780	1.558	2.646
Governor	1.072	1.566	1.363
Vicechair	1.044	1.718	0.953
Untrimmed Mean	0.962	0.980	0.922
Mean of Trimmed	0.908	0.978	0.812
Vect Trimmed Mean	0.976	0.889	0.972
Greenbook	1.432	1.076	1.664
SPF		0.911	

CPI	17-month	10-month	5-month
Trimmed Mean	0.509	0.240	0.134
Atlanta	1.235	1.125	0.968
Boston	1.055	1.232	1.667
Chicago	1.131	1.453	1.053
Cleveland	0.704	1.124	0.983
Dallas	1.645	1.048	1.446
Kansas City	1.145	0.905	1.032
Minneapolis	1.490	3.024	1.606
New York	1.241	1.279	1.071
Philadelphia	0.763	0.726	0.978
Richmond	0.988	1.703	1.733
San Francisco	1.187	1.131	0.808
St Louis	3.413	2.379	2.357
Governor	0.818	0.840	0.919
Vicechair	0.954	0.975	1.106
Untrimmed Mean	0.960	1.036	0.974
Mean of Trimmed	0.899	1.098	0.949
Vect Trimmed Mean	1.003	0.914	0.893
Greenbook	1.139	0.697	0.978
SPF		1.365	

RGDP	17-month	10-month	5-month
Trimmed Mean	2.201	1.535	0.574
Atlanta	0.956	0.891	1.353
Boston	1.442	0.965	1.362
Chicago	1.004	0.878	0.896
Cleveland	0.799	1.064	1.929
Dallas	0.722	0.861	0.868
Kansas City	1.091	1.483	1.067
Minneapolis	0.795	0.677	0.680
New York	1.206	1.160	1.397
Philadelphia	0.674	1.046	0.891
Richmond	1.085	0.728	0.778
San Francisco	1.279	1.151	1.024
St Louis	0.834	0.773	0.722
Governor	1.009	1.253	1.001
Vicechair	0.972	1.628	1.181
Untrimmed Mean	0.997	0.992	0.964
Mean of Trimmed	0.995	0.982	0.909
Vect Trimmed Mean	0.999	0.951	1.041
Greenbook	1.161	1.277	1.504
SPF		1.048	

UNEMP	17-month	10-month	5-month
Trimmed Mean	0.410	0.289	0.070
Atlanta	0.668	0.844	0.811
Boston	1.407	0.992	1.337
Chicago	0.818	1.358	0.952
Cleveland	0.981	1.128	1.389
Dallas	0.988	0.943	1.013
Kansas City	1.187	1.205	0.831
Minneapolis	1.103	0.795	1.621
New York	1.321	1.239	1.550
Philadelphia	0.986	1.239	1.241
Richmond	0.724	0.874	0.547
San Francisco	1.128	0.992	1.682
St Louis	0.922	0.849	1.378
Governor	1.317	1.039	1.277
Vicechair	1.065	0.622	1.398
Untrimmed Mean	1.008	1.024	0.979
Mean of Trimmed	1.027	1.067	0.947
Vect Trimmed Mean	0.967	1.010	1.000
Greenbook	1.420	1.061	1.763
SPF		0.959	

Notes:
(i) Values associated with the trimmed mean are mean square errors. The remaining values are ratios of MSEs relative to that of the trimmed mean.

Table 8: **Impact of Individual's Disagreement on Forecast Error Square of CPI**

	17-month		10-month		5-month	
	Disagreement in		Disagreement in		Disagreement in	
	CPI	NGDP	CPI	NGDP	CPI	NGDP
V (α_1)	0.389***	0.450***	0.225***	0.278***	0.115***	0.078**
	(0.145)	(0.120)	(0.060)	(0.087)	(0.034)	(0.036)
RB (α_2)	0.276*	0.160	0.037	0.061	0.110***	0.179***
	(0.161)	(0.141)	(0.097)	(0.086)	(0.040)	(0.055)
RB × V (α_3)	−0.147	0.149	0.027	−0.104	−0.143*	−0.151*
	(0.280)	(0.281)	(0.141)	(0.138)	(0.075)	(0.078)
V × D (β_1)	0.042	−0.048	−0.033	−0.093	0.010	0.066
	(0.138)	(0.115)	(0.038)	(0.071)	(0.041)	(0.052)
RB × D (β_2)	0.541**	0.598***	0.406***	0.373***	0.094	0.008
	(0.256)	(0.198)	(0.155)	(0.132)	(0.057)	(0.056)
RB × V × D (β_3)	−0.364	−0.626**	−0.336**	−0.177	0.003	0.001
	(0.305)	(0.289)	(0.164)	(0.178)	(0.098)	(0.085)
Regional Voter	0.219	−0.075	0.038	0.103	0.107	0.074
p-value ($\beta_1 + \beta_2 + \beta_3 = 0$)	0.175	0.740	0.618	0.342	0.185	0.116
R.Voter vs. R.Non-Voter	−0.322	−0.673	−0.368	−0.270	0.013	0.067
p-value ($\beta_1 + \beta_2 + \beta_3 - [\beta_2] = 0$)	0.238	0.012	0.023	0.101	0.882	0.324
Gov. vs. R.Voter	−0.177	0.027	−0.070	−0.196	−0.097	−0.008
p-value ($\beta_1 - [\beta_1 + \beta_2 + \beta_3] = 0$)	0.405	0.914	0.408	0.133	0.287	0.906
Gov. vs. R.Non-Voter	−0.499	−0.646	−0.439	−0.465	−0.084	0.058
p-value ($\beta_1 - \beta_2 = 0$)	0.089	0.006	0.007	0.002	0.237	0.449
Adj. R^2	0.462	0.461	0.459	0.396	0.508	0.489
N	119	119	120	120	119	119

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively. Panel Newey-West standard errors in parentheses.
(ii) Each column represents one regression model.

Table 9: **Mean Disagreement by Dissent**

	Mean			**p-value**	
	Assenters	Dissenters	Non-voters	Assenters vs. Dissenters	Non-voters vs. Dissenters
17-month					
Vector	1.821	1.835	1.867	0.957	0.904
NGDP	0.702	0.900	0.847	0.509	0.859
RGDP	0.659	1.050	0.827	0.135	0.393
CPI	0.745	0.740	0.724	0.988	0.961
UNEMP	0.784	0.673	0.799	0.596	0.539
5-month					
Vector	1.732	2.127	1.922	0.280	0.579
NGDP	0.712	1.043	0.799	0.199	0.329
RGDP	0.686	0.916	0.797	0.573	0.772
CPI	0.792	0.987	0.800	0.324	0.383
UNEMP	0.699	1.218	0.770	0.029**	0.079*

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.
(ii) No dissent in 10-month forecast (February).

Appendix Tables

Table 1-A: **Median Disagreement by Horizon**

	Mean			**p-value**		
	17-month	10-month	5-month	17m vs. 10m	10m vs. 5m	17h vs. 5m
Vector	4.800	5.069	4.844	0.540	0.600	0.914
NGDP	1.335	1.285	1.364	0.753	0.638	0.870
RGDP	1.143	1.259	1.838	0.439	0.023**	0.007***
CPI	1.434	1.451	1.075	0.940	0.043**	0.048**
UNEMP	1.300	1.144	1.217	0.297	0.661	0.621

Notes:

(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.

(ii) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (4). The scalar measures of disagreement were constructed similarly using equation (3).

Table 2-A: **Median Disagreement by Voting Status**

	Mean			**p-value**	
	Voters	Nonvoters	Voters ex. NY/G	Voters vs. Nonvoters	Voters ex. NY/G vs. Nonvoters
Vector					
17-month	4.585	5.118	4.706	0.281	0.507
10-month	4.994	5.182	5.025	0.773	0.829
5-month	4.633	5.158	4.780	0.308	0.536
Total	4.738	5.153	4.839	0.196	0.401
Nominal GDP					
17-month	1.222	1.503	1.317	0.222	0.485
10-month	1.250	1.337	1.324	0.706	0.961
5-month	1.282	1.486	1.636	0.370	0.665
Total	1.252	1.442	1.424	0.147	0.917
Real GDP					
17-month	1.072	1.249	0.888	0.404	0.162
10-month	1.303	1.192	1.628	0.591	0.083*
5-month	1.813	1.875	1.509	0.887	0.416
Total	1.396	1.439	1.345	0.807	0.626
CPI					
17-month	1.448	1.414	1.714	0.905	0.376
10-month	1.326	1.639	1.363	0.331	0.484
5-month	0.996	1.191	1.174	0.317	0.936
Total	1.257	1.414	1.416	0.319	0.991
Unemployment					
17-month	1.228	1.406	1.262	0.416	0.587
10-month	1.231	1.012	0.947	0.273	0.775
5-month	1.148	1.319	1.071	0.533	0.447
Total	1.202	1.246	1.092	0.749	0.324

Notes:

(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.

(ii) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (4). The scalar measures of disagreement were constructed similarly using equation (3).

Table 3-A: **Median Multivariate Disagreement by Individual**

	Mean			**p-value**		
	Aggregate	Voters	Non-voters	Voter vs. Non-voters	Agg. vs. Gov.	Agg. vs. Vice-chair
Atlanta	3.930	3.082	4.269	0.058*	0.142	0.912
Boston	4.156	3.589	4.582	0.101	0.271	0.637
Chicago	4.429	4.715	4.215	0.585	0.492	0.326
Cleveland	7.125	6.426	8.324	0.323	0.018**	0.003***
Dallas	4.517	5.337	4.107	0.192	0.690	0.362
Kansas City	4.499	4.167	4.749	0.532	0.594	0.283
Minneapolis	5.621	3.623	6.420	0.035**	0.346	0.068*
New York	4.117	4.117			0.225	0.672
Philadelphia	4.335	3.761	4.564	0.292	0.310	0.354
Richmond	4.229	5.041	3.904	0.335	0.404	0.608
San Francisco	3.715	3.332	3.868	0.340	0.009***	0.764
St Louis	9.042	7.957	9.857	0.309	0.000***	0.000***
Governor	4.780	4.780				0.024**
Vice-chair	3.859	3.859				
Total	4.905	4.738	5.153	0.401	0.611	0.007***

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.
(ii) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (4).

Table 4-A: **Median Scalar Disagreement by Voting Individuals**

	Mean			**p-value**				**Mean**			**p-value**		
	Aggregate	Voters	Non-voters	Voter vs. Non-voters	Agg. vs. Gov.	Agg. vs. Vice-chair		Aggregate	Voters	Non-voters	Voter vs. Non-voters	Agg. vs. Gov.	Agg. vs. Vice-chair
NGDP							CPI						
Atlanta	1.228	1.222	1.231	0.977	0.738	0.461	Atlanta	0.688	0.833	0.630	0.557	0.011**	0.673
Boston	1.011	0.873	1.114	0.559	0.504	0.914	Boston	0.950	1.150	0.801	0.297	0.180	0.512
Chicago	1.144	1.466	0.902	0.112	0.895	0.674	Chicago	0.905	1.222	0.667	0.095*	0.087*	0.618
Cleveland	1.561	1.680	1.356	0.576	0.249	0.185	Cleveland	2.786	2.578	3.143	0.505	0.004***	0.001***
Dallas	1.421	1.423	1.420	0.995	0.399	0.283	Dallas	1.329	1.500	1.244	0.633	0.687	0.089*
Kansas City	0.985	0.792	1.130	0.308	0.283	0.812	Kansas City	1.277	1.147	1.375	0.704	0.806	0.088*
Minneapolis	2.279	1.249	2.691	0.024**	0.009***	0.010***	Minneapolis	1.726	0.583	2.183	0.021**	0.167	0.023**
New York	0.992	0.992			0.418	0.856	New York	0.809	0.809			0.054*	0.947
Philadelphia	0.910	0.761	0.970	0.461	0.210	0.621	Philadelphia	1.301	1.370	1.274	0.892	0.819	0.202
Richmond	1.382	2.278	1.024	0.237	0.640	0.498	Richmond	1.268	1.000	1.375	0.186	0.825	0.076*
San Francisco	1.161	1.111	1.181	0.872	0.977	0.637	San Francisco	0.805	0.472	0.938	0.210	0.052*	0.960
St Louis	2.719	2.560	2.838	0.645	0.000***	0.000***	St Louis	2.661	2.415	2.845	0.666	0.017**	0.005***
Governor	1.166	1.166				0.498	Governor	1.218	1.218				0.024**
Vice-chair	1.042	1.042					Vice-chair	0.792	0.792				
Total	1.328	1.252	1.442	0.917	0.081*	0.136	Total	1.320	1.257	1.414	0.991	0.364	0.004***
RGDP							UNEMP						
Atlanta	1.198	1.056	1.254	0.617	0.409	0.818	Atlanta	1.071	0.500	1.300	0.055*	0.573	0.690
Boston	1.536	1.107	1.857	0.208	0.953	0.563	Boston	1.213	0.981	1.387	0.531	0.966	0.538
Chicago	1.251	1.019	1.425	0.542	0.469	0.921	Chicago	1.459	1.460	1.458	0.997	0.393	0.093*
Cleveland	1.637	1.429	1.993	0.560	0.746	0.386	Cleveland	1.530	1.292	1.939	0.444	0.368	0.113
Dallas	1.413	2.010	1.114	0.277	0.790	0.744	Dallas	0.704	0.429	0.842	0.177	0.019**	0.350
Kansas City	0.940	0.881	0.984	0.809	0.076*	0.300	Kansas City	1.183	1.315	1.083	0.636	0.831	0.368
Minneapolis	1.413	1.095	1.540	0.372	0.781	0.735	Minneapolis	1.056	1.389	0.922	0.430	0.491	0.703
New York	1.004	1.004			0.092*	0.366	New York	1.506	1.506			0.380	0.108
Philadelphia	1.539	2.452	1.173	0.135	0.941	0.512	Philadelphia	1.060	0.389	1.329	0.011**	0.418	0.646
Richmond	0.957	0.833	1.007	0.669	0.124	0.376	Richmond	1.246	1.667	1.078	0.164	0.948	0.274
San Francisco	1.229	0.806	1.398	0.176	0.461	0.881	San Francisco	0.802	0.833	0.789	0.928	0.072*	0.595
St Louis	2.354	2.063	2.571	0.553	0.146	0.078*	St Louis	1.783	1.315	2.134	0.206	0.196	0.071*
Governor	1.512	1.512				0.431	Governor	1.231	1.231				0.097*
Vice-chair	1.288	1.288					Vice-chair	0.946	0.946				
Total	1.413	1.396	1.439	0.626	0.551	0.590	Total	1.220	1.202	1.246	0.324	0.910	0.100

Notes:
(i) ***, **, and * denote statistical signifiance at the 1%, 5%, and 10% levels, respectively.
(ii) The scalar disagreement is calculated by averaging over all individuals' level of disagreement using equation (3).

Table 5-A: **Median Disagreement of SPF and FOMC**

	SPF	FOMC	Voters	Nonvoters	Regional	Governors
Vector	6.032	3.691	3.735	3.617	3.689	3.695
NGDP	1.510	0.974	0.962	0.993	0.939	1.061
RGDP	1.903	1.008	1.040	0.957	1.004	1.020
CPI	1.502	1.064	1.005	1.162	1.144	0.863
UNEMP	1.670	0.881	0.953	0.762	0.849	0.961

Notes:

(i) Mean multivariate disagreement is calculated by averaging over all individuals' level of disagreement using equation (4). The scalar measures of disagreement were constructed similarly using equation (3).