Federal Reserve Bank of New York
Staff Reports

The Behavior of Uncertainty and Disagreement and Their Roles in Economic Prediction: A Panel Analysis

Robert Rich
Joseph Tracy

Staff Report No. 808
February 2017

The Behavior of Uncertainty and Disagreement and Their Roles in Economic Prediction: A Panel Analysis

Robert Rich and Joseph Tracy

Federal Reserve Bank of New York Staff Reports, no. 808
February 2017
JEL classification: C10, C20, C40

Abstract

This paper examines point and density forecasts from the European Central Bank's Survey of Professional Forecasters. We derive individual uncertainty measures along with individual point- and density-based measures of disagreement. We also explore the relationship between uncertainty and disagreement, as well as their roles in respondents' forecast performance and forecast revisions. We observe substantial heterogeneity in respondents' uncertainty and disagreement. In addition, there is little co-movement between uncertainty and disagreement, and forecast performance shows a more robust inverse relationship with disagreement than with uncertainty. Further, forecast revisions display a more meaningful association with disagreement than with uncertainty: Those respondents displaying higher levels of disagreement revise their point and density forecasts by a larger amount.

Key words: uncertainty, disagreement, ECB-SPF, density forecasts, point forecasts, forecast accuracy, forecast revisions

---

Rich, Tracy: Federal Reserve Bank of New York (correspondence: robert.rich@ny.frb.org). The authors thank Michael Stewart and Michael Fosco for excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

I. Introduction

Individuals' forecast behavior remains an area of considerable research interest because of its importance for understanding individuals' decision-making, as well as for explaining movements in economic and financial variables. The majority of studies have focused on expectations, but there are other dimensions of forecast behavior such as disagreement and uncertainty that play an important role in macroeconomic analysis.[1] Despite their importance, the measurement of disagreement and uncertainty—like the measurement of expectations—is challenging because of the inherent difficulty of observing individuals' subjective magnitudes. While many surveys offer information about point forecasts and their dispersion across respondents, most typically do not provide information about the degree of confidence that respondents attach to their forecasts.

There are, however, a limited number of surveys that provide both point and density forecasts, with the latter taking the form of subjective probability distributions. The availability of density forecasts allows for a more detailed analysis of forecast behavior than can be carried out using point forecasts alone. For example, density forecasts provide a basis to construct measures of individual and aggregate uncertainty. Studies using density forecasts have documented substantial heterogeneity in forecasters' uncertainty. Using metrics to gauge the degree of divergence between two distributions, the notion of disagreement can be extended beyond its conventional association with differences in point forecasts to include, as well, differences in density forecasts. Last, while disagreement and uncertainty are quite distinct from a conceptual standpoint, it is nevertheless instructive to analyze their co-movement, as well as to examine their relationship with and predictive content for variables in which they are viewed as key determinants.

This paper examines point and density forecasts from the European Central Bank's Survey of Professional Forecasters (ECB-SPF) that solicits euro area expectations for a harmonized index of consumer prices (HICP) inflation, real GDP growth and the unemployment rate. In particular, we use the ECB-SPF data for an empirical investigation into five issues related to the measurement and behavior of uncertainty and disagreement. First, we construct <u>individual</u> measures of uncertainty and disagreement, with the latter involving the introduction of new respondent-specific measures of disagreement for both point and density forecasts. The new measures allow us to undertake a parallel analysis between reported point and density forecasts and speak to the robustness of results.

---

[1] For example, it is argued that uncertainty is an important source of business cycle fluctuations [Bloom (2009), Bloom *et al.* (2012)] and a fundamental determinant of asset prices [Campbell (2000)]. Differences in agents' expectations have been cited as an important channel for monetary policy to affect real activity [Woodford (2003), Mankiw and Reis (2002), Lorenzoni (2009)] and a key factor influencing the effect of public information signals [Morris and Shin (2002), Amador and Weill (2010)].

Second, we analyze statistical features of the individual measures of uncertainty and disagreement across the different forecast variables. Specifically, we examine the cross-sectional behavior of uncertainty and disagreement, as well as their movements over time. Moreover, and in contrast to studies conducted at the aggregate level, we consider the roles of respondent and time fixed effects in the behavior of the series which provide a useful background for our formal investigation into the issues of heterogeneity and persistence.

Third, we investigate the relationship between individual uncertainty and disagreement. Numerous empirical studies—lacking a direct measure of aggregate uncertainty—have used aggregate disagreement as a proxy under the maintained assumption of a meaningful positive association between the two variables. Uncertainty and disagreement, however, are distinct concepts. Moreover, while the conceptual distinction between uncertainty and disagreement is discussed at the individual level, direct measurement and testing of their association has, to date, been based entirely on analyses at the aggregate level. Because this aggregation may obscure the true relationship, we evaluate the reliability of disagreement as a proxy for uncertainty at the individual level.

Fourth, we also explore how movements in uncertainty and disagreement relate to the accuracy of a respondent's point and density forecasts. As examples, our measures allow us to consider whether respondents who are ex ante more confident are ex post more accurate, or if respondents who disagree more with other forecasters are less accurate. Further, we discuss and analyze how these results bear upon the practice of deriving heteroskedasticity-based measures of uncertainty.

Fifth, we investigate forecast revisions at the individual level. Most studies focusing on forecast revisions have used surveys in which the target variable remains fixed for some period of time, making it possible to test for statistical properties such as weak/strong efficiency, or to measure the effects of events and news announcements on expectations. While the constant forecast horizon structure of our selected data precludes us from a parallel analysis, we nevertheless can explore the roles that disagreement and uncertainty play in the revisions respondents make to their point and density forecasts across adjacent surveys. Importantly, a novel aspect of the latter analysis is that we examine revisions to the entire density forecast which contrasts with other studies that have only considered revisions to the mean density forecast.

The empirical analysis yields several findings of note. There is substantial heterogeneity in forecasters' uncertainty and disagreement. In addition, individual forecasters' level of uncertainty is associated with very strong respondent fixed effects, while the extent of individual disagreement is associated with very strong time fixed effects. Thus, differences across respondents' measured uncertainty feature a pronounced time-invariant idiosyncratic element which is also consistent with our finding that uncertainty displays higher persistence than disagreement. These different statistical properties also help explain the general absence of an economically or statistically significant co-movement between individual uncertainty and disagreement. Our finding that disagreement is not a reliable proxy for uncertainty corroborates the majority of prior evidence at the aggregate level.

The previous results pertain to the behavior and co-movement of uncertainty and disagreement. When we turn to the relationships that uncertainty and disagreement display with other variables related to forecast behavior, we again find significant differences. Consider first forecast performance. Our results indicate that most of the variation in forecast performance is explained by time effects—that is, the difficulty associated with forecasting accurately changes considerably over time and tends to impact all participants in the same survey. Controlling for these time effects, however, we find that disagreement is more predictive than uncertainty in explaining variation in forecast accuracy. In particular, we find that greater disagreement is associated with lower forecast accuracy, regardless of whether we look at an individual's point or density forecast. However, the nature of the relationship between uncertainty and forecast accuracy is sensitive to the type of reported forecast. While the data indicate that increased uncertainty is associated with lower forecast accuracy using density forecasts, there is no evidence of a similar linkage using point forecasts. Taken together, we view these results as offering a cautionary note about the practice of using ex post predictability as a basis to derive a proxy for ex ante uncertainty.

With regard to forecast revisions, we find that respondents who have a higher degree of disagreement—associated with either their point or density forecasts—tend to revise their subsequent point and density forecasts by a larger amount. On the other hand, the findings either reveal an insignificant or a counter-intuitive role for uncertainty in the forecast revision process. Specifically, our results indicate that uncertainty is not related to the degree of point forecast revisions, while respondents with higher uncertainty tend to revise their density forecast by a smaller amount. The latter result contrasts with the conventional view that greater forecast uncertainty, a priori, would on average be associated with a larger forecast revision as new information arrives.

The paper is organized as follows. The next section provides a summary of the literature that has focused on survey-based measures of uncertainty and disagreement. The third section describes the ECB-SPF data. Section four introduces the individual measures of uncertainty, disagreement and forecast performance, as well as presents a descriptive analysis of the measures of uncertainty and disagreement. The section also details the specification of the regression models and reports the empirical results concerning the relationship between disagreement and uncertainty, their importance as determinants of forecast performance, and their role in respondents' forecast revision process. The final section concludes by discussing the implications of our findings.

II. Literature Review

Our analysis examines several issues that have been explored in other studies. While there are a large number of survey instruments that report point forecasts of respondents, there are only a small group of surveys providing density forecasts. These include the ECB-SPF examined here, as well as the U.S. Survey of Professional Forecasters (US-SPF), the Bank of England Survey of External Forecasters (BOE-SEF) and the Federal Reserve Bank of New York Survey of Consumer Expectations (FRBNY-SCE).

Among studies that have used data from the surveys mentioned above, an issue of central importance is the measurement of disagreement and uncertainty. Disagreement has largely been measured by the dispersion of point forecasts and calculated as the cross-sectional variance/standard deviation, the interquartile range (IQR), or the quasi-standard deviation (qsd).[2] There are a few studies that instead have measured disagreement by the dispersion of the individual density forecast means [Giordani and Soderlind (2003), Boero *et al.* (2008), D'Amico and Orphanides (2008), Bruine de Bruin *et al.* (2011)].

There are several different methods that have been used to measure uncertainty based on the reported density (histogram) forecasts. In some cases, the uncertainty measure is derived from the individual density forecasts assuming that the probability mass within each interval is concentrated at the midpoint of each interval [Rich and Tracy (2010), Kenny *et al.* (2015)], or that the probability mass is distributed uniformly within each interval [Zarnowitz and Lambros (1987), Abel *et al.* (2016)].[3] In other cases, the uncertainty measure is derived by fitting a continuous distribution to the individual density forecasts, where the continuous distribution is a normal distribution [Giordani and Soderlind (2003), Boero *et al.* (2015)], a generalized beta distribution [Bruine de Bruin *et al.* (2011)], or both [Clements (2014a, 2014b)].[4] Alternatively, other studies [Batchelor and Dua (1996), Boero *et al.* (2008)] have focused on the aggregate density forecast and used a decomposition of its estimated variance to derive a measure of uncertainty.

---

[2] In this case, the quasi-standard deviation is equal to half the difference between the 16th and 84th percentiles of the sample of point forecasts appropriately interpolated.

[3] D'Amico and Orphanides (2008) derive measures of uncertainty under both assumptions.

[4] As noted by Bruine de Bruin *et al.* (2011), Clements (2014a, 2014b) and Boero *et al.* (2015), reported non-zero probabilities are needed in at least three bins to allow fitting to proceed. They follow the approach of Engelberg *et al.* (2009) and construct uncertainty measures by fitting triangular distributions when a respondent assigns probabilities to two intervals or less.

Another area of interest concerns the characteristics of disagreement and uncertainty. In addition to examining the behavior of disagreement about point forecasts and density forecast means, studies have explored heterogeneity in forecasters' uncertainty [D'Amico and Orphanides (2008), Boero *et al.* (2008), Bruine de Bruin *et al.* (2011), Clements (2014b), Boero *et al.* (2015)]. The finding of substantial heterogeneity in reported point forecasts and density forecasts has resulted in further exploration into the behavior of disagreement and uncertainty and how they evolve over time. There is evidence of persistent differences in respondents' point forecasts [Patton and Timmermann (2010), Boero *et al.* (2015)], as well as in respondents' forecast uncertainty [Bruine de Bruin *et al.* (2011), Boero *et al.* (2015)].

A number of studies have also explored the relationship between disagreement and uncertainty, although the existing empirical evidence has only been at the aggregate level and has relied mainly on data from the US-SPF. The evidence from the US-SPF, however, has been mixed. Zarnowitz and Lambros (1987) report a modest positive association between disagreement and uncertainty. Giordani and Soderlind (2003) extend the sample period of Zarnowitz and Lambros and report a positive association between disagreement and uncertainty that is both economically and statistically significant, although some studies have argued that their conclusion is problematic.[5] Lahiri and Sheng (2010) examine the US-SPF and find that the disagreement-uncertainty relationship is episodic, with a meaningful co-movement that only emerges during low volatility episodes. Examining matched point and density forecasts, Rich and Tracy (2010) find a very weak relationship between disagreement and uncertainty for the US-SPF. Recent analyses that have examined data from other surveys featuring point and density forecasts, such as Boero *et al.* (2008) for the BOE-SPF and Abel *et al.* (2016) for the ECB-SPF, have tended to find little support for the use of disagreement as proxy for uncertainty.[6]

The availability of point and density forecasts offers a unique opportunity to compare uncertainty measures derived on an <u>ex ante</u> versus <u>ex post</u> basis. Density forecasts allow for the construct of ex ante measures of uncertainty because they only rely on information provided by respondents prior to the realization of the target variable. In contrast, <u>ex post</u> measures of uncertainty are based on a comparison of point predictions to outcomes and therefore depend on forecast accuracy. Because of the limited availability of direct measures of uncertainty, most

[5] Rich and Tracy (2010) and Boero *et al.* (2015) discuss the problematic nature of fitting normal distributions to histograms where respondents place positive probability in less than three bins.

[6] Boero *et al.* (2015) find a strong positive correlation between disagreement and uncertainty when they extend their 2008 analysis to include the recent crisis period. Because the choice and reliability of a proxy for a particular variable of interest is predicated on the unconditional correlation, and not the conditional correlation, between series, the general conclusion remains that disagreement does not appear to be a useful proxy for uncertainty.

empirical measures of uncertainty are calculated on an ex post basis without an opportunity to gauge their reliability. However, density forecasts can be used to measure ex ante uncertainty which provides a benchmark to assess the relationship to ex post uncertainty measures derived from survey forecast errors. Abel *et al.* (2016) examine data at the aggregate level from the ECB-SPF and find little evidence of a link between ex post forecast accuracy and ex ante uncertainty. Clements (2014a) draws a similar conclusion looking at the US-SPF.

The availability of point and density forecasts also allows for a more detailed investigation into the forecast revision process of respondents, with the revisions either reflecting a constant or changing forecast horizon depending on the forecast structure of the survey instrument. Bruine de Bruin *et al.* (2011) examine data from the FRBNY-SCE that involve a constant forecast horizon and find higher forecast uncertainty is associated with a higher variability in individual point forecasts over time. Other studies have focused on data from the US-SPF that involve a time-varying forecast horizon. Lahiri and Sheng (2010) use a Bayesian learning model to investigate the relative importance of the different factors contributing to disagreement as the forecast horizon changes. Patton and Timmermann (2010) also focus on time-varying forecast dispersion and find that heterogeneity in respondents' information sets is relatively unimportant, while heterogeneity in priors plays an important role. Clements (2014a) compares the estimates from a Bayesian learning model for respondents' point forecasts to those for the mean density forecasts and finds notable differences across the two types of forecasts.

In addition to the literature discussed above, there is other work that has some overlap with issues we explore in our analysis. For example, Kenny *et al.* (2014) examine the ECB-SPF to evaluate the informational content of the density forecasts. They find that there is little gain to forecast accuracy from trimming poorly performing forecasters in real time. In another study, Kenny *et al.* (2015) use the ECB-SPF to examine the link between the characteristics of density forecasts and density forecast performance. They find evidence of a downward bias in the variance of the density forecasts, suggesting overconfidence on the part of forecasters. A similar finding was reported in Diebold *et al.* (1999) for the US-SPF, although Clements (2014a) argues that the overconfidence applies to horizons of a year or more and that respondents display under-confident at short horizons. Clements (2004) evaluates density forecasts from the BOE-SEF, but restricts his attention to inflation outcomes close to the target rate of 2½ percent rather than the whole range of values specified in the survey instrument. He reports that the short-horizon density forecasts have higher economic value compared to benchmark unconditional forecasts, but that year-ahead density forecasts perform worse than the benchmark and attribute too much probability mass to relatively high rates of inflation.

III. The European Central Bank's Survey of Professional Forecasters

The ECB-SPF is a quarterly survey of forecasts for the euro area harmonized index of consumer price (HICP) inflation, real GDP growth and the unemployment rate that has been fielded since 1999. The ECB-SPF asks panelists for forecasts at short-, medium- and longer-term horizons including both rolling and calendar year variants. The pool of panelists is drawn from both financial and non-financial institutions throughout the euro area, with 50 panelists responding on average per survey. Further details about the ECB-SPF are provided in Garcia (2003) and Bowles *et al.* (2007).

For the three macroeconomic variables and each horizon, the ECB-SPF asks respondents to submit both a point and a density forecast. For the density forecasts, respondents provide a probability distribution of forecasted outcomes. Specifically, they report their probability distribution as a histogram using a set of intervals provided by the ECB for each macroeconomic variable. The number of closed intervals for the histogram changes over time in an effort to limit the likelihood that respondents place either a significant probability or no probability in either of the two open intervals. However, a common width is used for the closed intervals and this width has remained fixed over time.

In our analysis, we examine matched point and density forecasts for HICP inflation, output growth and the unemployment rate that involve a "rolling" one-year-ahead horizon.[7] Given that the horizon length remains constant through time, the data can be treated as quarterly observations on a set of homogenous series. As Garcia (2003) notes, there is a temporal misalignment between the target variables for inflation, output growth and the unemployment rate because of differences in the data frequency and publication lags of the variables. Specifically, output growth is published quarterly with a two quarter lag, while the unemployment rate and HICP inflation are published monthly with a two month lag and a one month lag, respectively.[8]

With regard to the sample, our study covers the period 1999Q1–2014Q4. For each survey and forecasted variable, we exclude any respondent who does not report matched point and density forecasts or whose probabilities for the density forecast does not sum to unity. In addition, we

[7] There are also forecasts reported for the a one-year/one-year forward horizon and a longer calendar horizon – four-calendar years ahead for surveys conducted in the Q1 and Q2 rounds, and five-calendar years ahead for surveys conducted in the Q3 and Q4 rounds.
[8] For example, the 2010Q1 survey questionnaire asks respondents to forecast one-year-ahead output growth from 2009Q3–2010Q3. For HICP inflation, the corresponding forecast horizon is December 2009–December 2010. For the unemployment rate, the corresponding forecast is for November 2010.

exclude the 2009Q1 GDP density forecast data because many respondents in this survey placed significant probability in the lower open interval of their GDP density forecasts.[9]

Before turning to the empirical analysis, there is a caveat that merits attention relating to the density forecasts reported by the ECB-SPF or similar surveys. As discussed in Boero *et al.* (2015), it is important to recognize that any reported probability distribution will be an imperfect representation of forecast uncertainty. Their argument is partly based on the apparent difficulty of forecasters to give precise numerical values for their subjective probabilities. In addition, the limited number of intervals to which many respondents assign probabilities precludes the specification of a unique probability distribution, with any fitted probability distribution selected by a researcher necessarily expressing uncertainty about the target variable in more detail than the forecaster has conveyed. Consequently, a survey-based measure of uncertainty needs to be viewed as an estimated proxy for the forecaster's unobserved uncertainty, and some caution should be exercised when interpreting results and making inferences about uncertainty and its features based on this proxy.

IV. Variable Definitions, Regression Models and Empirical Results

Disagreement and uncertainty

The traditional survey disagreement measure used in the literature is either the cross-sectional variance or standard deviation of the point forecasts. If we let ${}_i f_t^e$ denote the point forecast from respondent *i* in the survey at date *t,* n_t the number of respondents in the survey, and $\overline{f}_t^e$ the mean of these point forecasts, then the variance-based aggregate disagreement measure is given by:

$$(1) \qquad D_t = \frac{1}{n_t}\sum_{i=1}^{n_t}({}_i f_t^e - \overline{f}_t^e)^2$$

The mean point forecast is also typically referred to as the "consensus" forecast.

[9] For this survey, the significant probability mass at the lower open interval corresponded to a growth rate of "-1% or less" and was due to the survey design of the density forecasts and its inability to provide sufficient coverage for the pessimistic point predictions of output growth. For individuals who either reported point predictions below -1% or wanted to indicate significant downside risk, they assigned most of their probability to the open-ended interval. The narrow spread of the probabilities across intervals results in an artificially low value of the uncertainty measure. See Abel *et al.* (2016) for further discussion.

Generalizing, we can define an individual "average point disagreement" (APD) measure for the j^{th} respondent as the average squared distance between this respondent's point forecast and the other respondents' point forecasts:

$$(2) \qquad {}_{j}APD_{t} = \frac{1}{n_{t}-1}\sum_{i \neq j}({}_{i}f_{t}^{e} - {}_{j}f_{t}^{e})^{2}$$

The traditional variance-based aggregate disagreement measure can then be thought of as the individual APD measure for a representative respondent whose point forecast coincides with the consensus mean.

The APD measure uses the squared norm to measure the distance between two respondents' point forecasts. An alternative is to use the absolute value norm for the individual disagreement measure. This distance norm places less weight on large pair-wise differences. Accordingly, we define the individual average absolute point disagreement (AAPD) measure as:

$$(3) \qquad {}_{j}AAPD_{t} = \frac{1}{n_{t}-1}\sum_{i \neq j}\left|{}_{i}f_{t}^{e} - {}_{j}f_{t}^{e}\right|$$

There is an important robustness consideration associated with the choice of which distance norm to use. Specifically, measures based on the squared norm are more sensitive to outliers compared to the absolute value norm. We also show in subsequent discussion that uncertainty measures can be sensitive to the treatment of the open intervals of the density forecasts, with the sensitivity again being increased by the use of a squared norm. Consequently, the absolute value norm helps to mitigate both of these concerns that could be particularly important for the analysis. Going forward, we will maintain consistency in the construction of variables by adopting the absolute value norm for all distance calculations.

The traditional aggregate disagreement measure and the analogous individual disagreement measures are based on the respondents' point forecasts. For surveys like the ECB-SPF that also report respondents' density forecasts, we can construct additional individual-level disagreement measures that reflect more information than is revealed in the point forecasts. To illustrate this point, consider two respondents *i* and *j*, where we fix the distance between their point forecasts between two scenarios, but vary the nature of their density forecasts. Assume that each respondent's point forecast corresponds to the mid-point of the bin of their density forecast with the highest probability. Consider the contrasting sets of densities in Figure 1. In scenario A, the two densities are skewed in the direction of the other respondent's point forecast. In contrast, scenario B depicts two densities that are skewed away from the other respondent's point forecast. The point forecasts in each scenario are the same, and therefore would indicate the same degree of disagreement based on the point forecasts. However, based on the information conveyed by the density forecasts, it is reasonable to argue that the degree of disagreement is larger in scenario B than it is in scenario A.

What is needed, then, is a metric to convey the degree of divergence between two histograms. Following the choice of the absolute value metric to calculate dispersion of the point forecasts, we use the Wasserstein distance measure for histograms.[10] Let ${}_iF_t^{-1}$ denote the inverse cumulative density function (CDF) for respondent *i* in the survey at date *t*. The Wasserstein disagreement measure between respondent *i* and respondent *j* is given by:

$$(4) \qquad {}_{ij}WD_t = \int_0^1 \left| {}_iF_t^{-1}(z) - {}_jF_t^{-1}(z) \right| dz$$

We then define the individual average absolute density disagreement (AADD) measure for respondent *j* as follows:

$$(5) \qquad {}_jAADD_t = \frac{1}{n_t - 1} \sum_{i \neq j} {}_{ij}WD_t$$

Turning to uncertainty, a popular proxy is the variance of a survey respondent's density forecast. Zarnowitz and Lambros (1987) examine the US-SPF and derive the variance assuming a uniform distribution within each interval. The US-SPF density forecast, like that for the ECB-SPF, contains open intervals on each end of the histogram that must be closed to calculate this variance. A typical—although ad hoc—assumption is that these exterior open intervals have twice the width of the interior closed intervals. After closing off the two open intervals, assume that there are k_t bins associated with the histogram, that the upper and the lower values for the i^{th} bin are given, respectively, by u_i and l_i, and that the probability assigned by respondent j to this bin is ${}_jp_t^i$. The variance measure of individual uncertainty is then given by:

$$(6) \qquad {}_j\sigma_t^2 = \left[\sum_{i=1}^{k_t} {}_jp_t^i \left(\frac{u_{it}^3 - l_{it}^3}{3(u_{it} - l_{it})} \right) \right] - \left[\sum_{i=1}^{k_t} {}_jp_t^i \left(\frac{u_{it}^2 - l_{it}^2}{2(u_{it} - l_{it})} \right) \right]^2$$

[10] See Arroyo and Mate (2009).

To the extent that respondents place any probability in either open interval, the variance measure of uncertainty will be affected by the manner chosen to close off the open intervals. In an attempt to mitigate this concern, our analysis uses the inner quartile range (IQR) to proxy uncertainty because this measure will be robust to the treatment of the open intervals as long as respondents do not place more than 25 percent probability in either open interval.[11] While most forecasters only assign probabilities to the interior closed intervals, there are forecasters who assign probability to the open interval in almost every quarter. Moreover, the number of forecasters assigning probability to the open interval can occasionally be notable. However, forecasters rarely place more than 25 percent probability in an open interval, thereby making the IQR an attractive alternative to equation (6). The IQR uncertainty measure $\left({}_j IQR_t \right)$ can be calculated assuming, as we did for the variance, a uniform distribution within each bin of the histogram.

Figure 3 and Figure 4 plot, respectively, the distributions of our individual disagreement and uncertainty measures for the three forecast variables.[12] The distributions allow us to see the spread in the individual-level measures, as well as their behaviors over time. As a point of reference, the green line in each panel depicts the median value, around which we observe substantial dispersion, implying substantial heterogeneity across respondents in their degrees of disagreement and uncertainty. In addition, the distributions typically display positive skewness. Note that the vertical scale of the disagreement measures is different from that of the uncertainty measures, with the density-based disagreement measures displaying a higher degree of cross-sectional dispersion than the point-based disagreement measures. All of the disagreement measures across the three forecast variables spike during the financial crisis, with a subsequent smaller spike occurring with the onset of the European debt crisis. As another point of reference, the traditional aggregate disagreement measure D_t in equation (1) corresponds roughly to the 10th percentile of the individual disagreement measures. While the individual uncertainty measures in Figure 4 display a greater degree of cross-sectional dispersion compared to the disagreement measures, they display less of a spike during the financial crisis. Instead, the spreads of the uncertainty measures have increased since the financial crisis as the upper-half of the individual uncertainty distributions have moved higher from that time and have not yet reverted to their pre-crisis levels.

[11] That is, the IQR is robust to how the two open intervals are closed as long as the 25th and 75th percentiles fall in the interior closed intervals.

[12] Figure 3 and Figure 4 plot values for the GDP growth forecasts from the 2009Q1 survey for completeness. As noted earlier, these values are excluded from the regression analysis because of problems associated with the reported density forecasts.

While Figure 3 and Figure 4 depict the manner in which the distributions of individual disagreement and uncertainty shift over time, they do not indicate the degree to which respondents move within the distributions over time. In particular, they are not informative about whether there are persistent patterns in individual forecasters' disagreement or uncertainty and, if so, the sources for such persistence. To address the former issue, we draw upon Patton and Timmermann (2010) and provide in Tables 1-3 quarterly transition matrices for the three forecast variables, where all respondents in a survey are ranked according to the quartile in which their measures of disagreement and uncertainty fall in each survey. In the absence of persistence in forecasters' relative disagreement and uncertainty, the entries in the tables should all be approximately one-quarter (0.25). If, however, there is persistence in these features of forecast behavior, then terms on the diagonal should be significantly higher than 0.25, and the off-diagonal terms should be smaller than 0.25. For each table, we provide the sample transition rates and examine if differences in forecasters' relative disagreement and uncertainty persist using the relevant one-tailed test. We also report the chi-square statistic from a joint test for the entire table following a uniform distribution.

Taken together, there is very strong evidence of persistence for all three measures as indicated by the large number of statistical tests rejecting the relevant null hypothesis. For GDP growth, there is more evidence of persistence in the density-based than the point-based disagreement measure. In particular, all of the estimated probabilities of remaining in the same quartile for the density-based disagreement measure are significantly greater than 25 percent. For the individual uncertainty measures, there is even greater persistence as shown by three of the four diagonal probabilities being above 50 percent. Similar conclusions hold for inflation and unemployment. In the case of persistence in individual uncertainty, our results are consistent with those reported by Boero *et al.* (2015) for the BOE-SEF and Bruine de Bruin *et al.* (2011) for the FRBNY-SCE. The evidence of persistence in individual disagreement is a new finding, although evidence of persistence in the relative level of point forecasts has been documented by Patton and Timmermann (2010) for the US-SPF and by Boero *et al.* (2015) again for the BOE-SEF.

The evidence of persistence in the individual disagreement and uncertainty measures motivates our exploration into the sources for the persistence. Here we will consider the role of person effects that reflect any systematic differences across individuals that are unrelated to which surveys they participate in. That is, some respondents may be inherently more or less uncertain, or may consistently display higher or lower disagreement than others. In estimating the contribution of person effects, we will first control for any time effects to account for the fact that not all respondents participate in all surveys. These time effects will reflect the degree to which collectively respondents' uncertainty or disagreement may vary over time.

Table 4 provides variance decompositions for each measure across the three forecast variables. For the two disagreement measures, time effects account for a significant portion of the overall variance, while respondent effects are much less important. Across the three forecast variables, the disagreement time effects are largest for GDP growth. Uncertainty, in contrast, has a much larger component of its variance explained by respondent effects than by time effects, which is consistent with the higher persistence displayed by individual uncertainty in the transition tables.[13] The relative importance of the respondent effects is also fairly uniform for the uncertainty measure across the three forecast variables. Taken together, the results in Table 4 underscore the importance of accounting for person and time effects in our subsequent regression analysis.

To investigate the relationship between disagreement and uncertainty, we consider the following two regression models that allow us to examine the point- and density forecast data at the individual level:

$$(7)\qquad {}_{j}IQR_{t} = \beta_0^P + \beta_1^P\left({}_{j}AAPD_{t}\right) + \alpha_j^P + \mu_t^P + {}_{j}\varepsilon_t^P$$

$$(8)\qquad {}_{j}IQR_{t} = \beta_0^D + \beta_1^D\left({}_{j}AADD_{t}\right) + \alpha_j^D + \mu_t^D + {}_{j}\varepsilon_t^D$$

where α_j denotes a respondent fixed effect, μ_t denotes a time fixed effect, and ${}_{j}\varepsilon_t$ denotes a mean-zero random disturbance term. Because we consider measures of disagreement derived from both the point and density forecasts, the notation in the regression models will differentiate the nature of the disagreement measure through the use of the superscripts 'P' and 'D'. The relationships are estimated on the unbalanced panel data of respondents, and we estimate the standard errors clustering at the respondent level.[14] Given the assumed positive association between disagreement and uncertainty, we conduct a one-sided test of statistical significance for β_1.

[13] Bruine de Bruin *et al.* (2011) also document that there is a strong fixed effect component associated with individual respondents' uncertainty

[14] The ECB-SPF, like other surveys, has experienced exit and entry of respondents over time and occasional non-response to the complete questionnaire. In their analysis of the ECB-SPF at the aggregate level, Abel *et al.* (2016) were concerned that heterogeneity in respondents' forecast behavior and the changing panel could result in incorrect inference being drawn from estimated relationships. They explored the role of compositional effects by considering various subsamples of 'regular' respondents, but found little change in the results across subsamples. The current analysis uses the unbalanced panel structure because the inclusion of respondent fixed effects and time fixed effects allows us to control for possible compositional effects.

Table 5 presents simple regression results for each of the three forecast variables and for our individual point- and density-based disagreement measures. In each case, we show results using the overall variance of each disagreement variable (specifications (1) and (4)), next removing the time-series component of the disagreement variance (specifications (2) and (5)), and then finally removing both the time-series and respondent components of the disagreement variance (specifications (3) and (6)). For these three cases and disagreement measures, we indicate the incremental R^2 value from adding the disagreement measure.

Focusing first on the point-based disagreement measures, there is a positive and significant relationship between individual uncertainty and individual disagreement for GDP, inflation and unemployment using the overall disagreement variance. However, the estimated relationships explain little of the overall variation of individual uncertainty and their features are not robust to including time and person effects. Turning to the density-based disagreement measure, there is a stronger and more robust positive relationship between individual uncertainty and individual disagreement, with disagreement explaining over 8 percent of the overall variation of uncertainty for each of the forecast variables. However, while the statistical significance of the relationship is robust to including time and respondent effects, the incremental explanatory content of disagreement progressively declines as we remove the time- and respondent-specific components of the variance. Consequently, the results for the ECB-SPF at the individual level corroborate the finding of Abel *et al.* (2016) at the aggregate level that there is no meaningful association between uncertainty and disagreement and provide further evidence against the practice of using disagreement as a proxy for uncertainty.

Forecast accuracy and its relationship to uncertainty and disagreement

The previous analysis focused on the statistical properties of disagreement and uncertainty, as well as their joint relationship. We next explore if, and how, disagreement and uncertainty impact other features of forecast behavior. A natural initial candidate for such an investigation is forecast performance. That is, does the degree of alignment between a survey respondent's forecast and other respondents' forecasts or the confidence of a respondent's forecast bear upon his/her relative forecast accuracy?

To explore this issue empirically, we need a measure of forecast accuracy. Because the ECB-SPF provides data on point forecasts, we can use the reported expectations to derive a point accuracy measure for each respondent. Following the approach used to construct the individual-level point disagreement measure, we consider an individual absolute point accuracy (APA) measure given by:

(9) $$_{j}APA_{t+h} = \left| a_{t+h} - {}_{j}f_{t}^{e} \right|$$

where a_{t+h} denotes the realized value of the relevant ECB-SPF target variable in period *t+h*, $_{j}f_{t}^{e}$ again denotes the point forecast from respondent *j* in the survey at date *t,* and *h* denotes the relevant forecast horizon which may exceed the quarterly sampling interval of the data.

Density forecasts also allow for the construct of accuracy measures. As was the case for measuring disagreement, accuracy measures derived from density forecasts can be more informative than those based only on the point forecasts. This feature is illustrated in Figure 2. Assume that we have two respondents who have identical point forecasts. As such, they would also have identical point-based accuracy measures. Similar to the previous analysis using Figure 1, we assume that each respondent's point forecast corresponds to the mid-point of the bin of the density forecast with the highest probability. Respondent *i*'s density is skewed toward the actual outcome, while respondent *j*'s density is skewed away from the actual outcome. Comparing the two density forecasts, a reasonable interpretation is that respondent *i* was more accurate than respondent *j.*

There are a number of possible choices for a density-based accuracy measure. The first one that we consider is an expected accuracy measure. Let l_t denote the lower limit of the histogram and u_t the upper limit after closing off the open intervals. As before, assume that there are k_t bins each with lower and upper limits denoted by (l_{it}, u_{it}) and that the density, $f_t(x)$, is uniform within each bin. We define the expected absolute density accuracy (EADA) measure as follows:

(10) $$_{j}EADA_{t+h} = \int_{l_t}^{u_t} \left| a_{t+h} - x \right|_{j} f_{t}(x)dx$$

A second density-based accuracy measure is the Rank Probability Score (RPS). We need to introduce one additional variable to calculate this score.[15] Let I_{t+h}^{i} denote an indicator variable that takes a value of one if the actual outcome in period *t+h* is in the i^{th} interval of the histogram from

[15] See Kenny *et al.* (2014) and Lopez-Perez (2014) for applications of the RPS.

the survey at date *t*. We adapt the RPS to use the absolute value metric. The Absolute Rank Probability Score for the *j*th respondent in the survey at date *t* is given by:

$$(11) \qquad {}_{j}ARPS_{t+h} = \frac{1}{k_t - 1}\sum_{i=1}^{k_t}\left|\sum_{l=1}^{i} {}_{j}p_t^l - \sum_{l=1}^{i} I_{t+h}^l\right|$$

The ARPS and the EADA both share the property that a respondent receives "credit" by assigning probability in bins close to the bin containing the actual outcome.

Having established that the conceptual distinction between uncertainty and disagreement extends to their empirical counterparts, we now investigate how individual uncertainty and disagreement relate to a respondent's forecast accuracy. For the point forecast data, we can formulate the following regression model that allows us to incorporate a channel of effect of both uncertainty and disagreement on predictive accuracy as well as to control for respondent and time fixed effects:

$$(12) \qquad {}_{j}APA_{t+h} = \delta_0^P + \delta_1^P\left({}_{j}IQR_t\right) + \delta_2^P\left({}_{j}AAPD_t\right) + \alpha_j^P + \mu_{t+h}^P + {}_{j}\varepsilon_{t+h}^P \ ,$$

where ${}_{j}AAPD_t$ denotes the previously defined measure of point forecast disagreement at the individual level and ε_{t+h}^P denotes a mean-zero random disturbance. As an analogue to equation (12), we can consider the following regression model using the density forecast data:

$$(13) \qquad {}_{j}DA_{t+h} = \delta_0^D + \delta_1^D\left({}_{j}IQR_t\right) + \delta_2^D\left({}_{j}AADD_t\right) + \alpha_j^D + \mu_t^D + {}_{j}\varepsilon_{t+h}^D \ ,$$

where the density accuracy measure ${}_{j}DA_{t+h} = \left({}_{j}EADA_{t+h}, {}_{j}ARPS_{th}\right)$, ${}_{j}AADD_t$ denotes the previously defined measure of density forecast disagreement at the individual level, and ε_{t+h}^D again denotes a mean-zero random disturbance.

Before presenting our results, it is worth noting that the relationship between forecast performance, uncertainty and disagreement also bears upon a class of models used to derive measures of uncertainty. Specifically, our findings have implications for time series models of heteroskedasticity and the reliability of using the ex-post error variance to proxy ex-ante uncertainty.[16] As discussed in the Appendix, time series models of heteroskedasticity postulate that forecast accuracy displays a

[16] The most popular example of this modeling strategy is the Autoregressive Conditional Heteroskedasticity (ARCH) model of Engle (1983) in which the conditional variance of a time series is specified as a function of past squared forecast errors.

direct link with uncertainty and no association with disagreement. The implications for equation (12) are that δ_1^P is positive, δ_2^P is zero, and uncertainty has economically significant predictive content for forecast performance. There are analogous implications for the values of δ_1^D and δ_2^D, as well as the explanatory power of the uncertainty measure in equation (13).[17] Accordingly, we conduct a one-sided test of statistical significance for δ_1 and a two-sided test of statistical significance for δ_2.

The results are summarized in Table 6. For each of the forecast variables, we present results for the point accuracy and our two density accuracy measures for specifications that control for time and person effects. In each case, we also indicate the overall R^2 value for the estimated regressions, as well as the incremental R^2 value from first including the uncertainty measure, next adding the disagreement measure, and then adding the respondent component of forecast performance. Given this sequential ordering and recognizing that the explanatory variables are not orthogonalized, we are allowing the highest possible value that can be ascribed to the uncertainty measure in terms of predictive content for the models.

Looking first at the point accuracy measure in specification (1), we find no significant relationship between individual uncertainty and individual accuracy. In contrast, for each of the three forecast variables there is a positive and statistically significant relationship between individual disagreement and individual forecast performance that indicates greater personal disagreement is associated with lower forecast accuracy.

The relationship between individual accuracy and uncertainty changes, however, when we consider a density-based accuracy measure. In specification (2), there is a positive and statistically significant relationship between the expected absolute accuracy measure and uncertainty across the three forecast variables—that is, respondents who are ex-ante more uncertain are ex-post less accurate. This result is also evident in specification (3) where we measure accuracy using the absolute rank probability score, although compared to the expected absolute accuracy the precision of the estimates is lower. These findings run counter to Kenny *et al.* (2015) who find that greater uncertainty is associated with improved forecast accuracy for the ECB-SPF. In addition, the positive relationship between accuracy and disagreement carries over to our density-based measures in almost all cases.[18]

---

[17] Clements (2014b) also explores the issue of ex ante and ex post measures of forecast uncertainty. In contrast to our study, he examines data from the US-SPF and focuses on the term structure of forecast uncertainty.

[18] For GDP, the density disagreement measure is not statistically significant at the 10 percent level for the absolute rank probability score measure of accuracy.

The previous conclusions concerning the linkage between forecast performance, uncertainty and disagreement are based on an evaluation of statistical significance. It is also important, however, to consider economic significance. As shown, the predictive content of uncertainty and disagreement is low. Taken together, the evidence suggests forecast performance displays a limited association with uncertainty and disagreement, although there is a more robust and meaningful inverse relationship with disagreement than with uncertainty.[19] Instead, the results document that most of the variation in forecast performance is explained by time effects. Interestingly, this finding is consistent with the work of D'Agostino *et al.* (2012) who examine the US-SPF and investigate whether some forecasters are better than others. While they observe ex post differences in respondent's accuracy, they find little evidence to suggest they reflect ex ante differences after controlling for variation in the forecasting environment. That is, the forecastability of a variable can be episodic, and taking this consideration into account is important for isolating the effects of other variables and drawing comparisons to the behavior of other respondents.

The findings in Table 6 also allow us to speak to ex-post forecast error variance as a reliable proxy for ex-ante uncertainty and would appear to raise questions about the practice of deriving heteroskedasticity–based measures of uncertainty. We can strongly reject the restriction $\delta_2 = 0$ and the hypothesis that disagreement is not related to forecast performance. Moreover, the absence of a robust and economically meaningful relationship between forecast performance and ex ante uncertainty contradicts a central tenet of these models, with the identified association between forecast accuracy and disagreement making interpretation of these model-based measures of uncertainty problematic. With regard to the latter point, if forecast accuracy is more highly correlated with disagreement than uncertainty, then forecast performance is more informative about the relative position of a prediction than the confidence attached to a prediction.[20]

The forecast revision process

Another interesting aspect of the beliefs formation process focuses on revisions that respondents make to their forecasts. Most analyses have considered surveys in which the target variable remains constant for a period of time, resulting in a sequence of 'fixed-event' forecasts. As first discussed by Nordhaus (1987), the fixed-event forecast structure allows for investigations into

[19] While the incremental R^2 for disagreement is low, it is higher than that for uncertainty and often matches or exceeds that for the person effects.

[20] Clements (2014a) also explores the issue of ex ante and ex post measures of forecast uncertainty. In contrast to our study, he examines data from the US-SPF and focuses on the term structure of forecast uncertainty. Nevertheless, Clements documents substantial heterogeneity across respondents in their forecast accuracy and uncertainty and also finds little evidence of a systematic relationship between the two.

whether respondents make efficient use of information in their forecast revisions. The structure also allows study of the forecast revision process and its relation to macroeconomic news. While most analyses have focused on point forecast revisions, there are studies that have also considered density forecast revisions. Boero *et al.* (2008) examine revisions to point and density forecasts of inflation and GDP growth from the BOE-SEF and report that tests generally do not reject the properties of weak and strong efficiency for the series.[21] In addition, they find that the impact of macroeconomic news—derived as revisions to the survey average point forecasts—is greater the nearer is the forecast target date, while revisions to forecast uncertainty diminish as the forecast target date approaches.

While the ECB-SPF survey structure involves a constant-horizon format, it is still possible to study the forecast revisions of respondents. For example, Bruine de Bruin *et al.* (2011) examine one-year-ahead point and density forecasts of price inflation and wage inflation from the FRBNY-SCE to explore the role of uncertainty in the forecast revision process. Controlling for individual demographic characteristics, their findings indicate that higher uncertainty in one survey is associated with larger absolute revisions in point forecasts from that survey to the next. They interpret the results as being roughly consistent with a model of Bayesian updating by individuals, where a more diffuse prior at one point in time is associated with larger revisions in point forecasts in subsequent periods.

Because the forecast variables examined in our study involve a fixed one-year horizon, our analysis of the forecast revision process is closer in spirit to Bruine de Bruin *et al.* (2011). Similarly, we will consider point as well as density forecast revisions. However, there are aspects of our investigation that do not appear to have a counterpart in previous work. Specifically, we will not restrict our analysis of the density forecast revisions to particular moments such as the mean. Rather, we develop a measure that encompasses the change in the overall density forecast. Consequently, our measure will reflect a revision associated with a mean-preserving spread in the density forecast, whereas a measure based on the density mean will indicate the absence of a forecast revision. In addition, we incorporate a role for both uncertainty and disagreement, with the latter consideration allowing us to determine if the relative alignment of a respondent's forecast bears upon subsequent revisions.

[21] The density forecast means are constructed from the histograms by assuming that the reported probabilities are concentrated at the mid-points of the respective intervals. The concept of weak efficiency implies that forecast revisions are independent of one another, while the concept of strong efficiency implies that forecast revisions are independent of information available at the time of the earlier forecast.

We begin by defining both a point- and density-based measure of a respondent's forecast revision. The point-based forecast revision is simply the absolute value between the respondent's point forecast in the surveys conducted in period *t* and period *t*+1:

$$(14) \qquad {}_jAPR_{t+1,t} = \left| {}_jf^e_{t+1} - {}_jf^e_t \right|$$

Drawing upon our earlier discussion of disagreement for the density forecasts, the density-based forecast revision (DR) is the Wasserstein measure of the difference between the respondent's density forecast in the surveys dated *t* and *t+1:*

$$(15) \qquad {}_jDR_{t+1,t} = {}_jWD_{t,t+1}$$

Note that when the ECB-SPF changes the bin structure between two adjacent surveys, we need to make the interval structures conforming in order to compute the DR measure. For example, if three additional bins are added to the lower end of the histogram, we need to aggregate these back up to the earlier open interval from the prior survey before calculating the density-based forecast revision.

Similar to the analysis of forecast performance, we can investigate the extent to which forecast revisions are related to the respondent's uncertainty and disagreement.

$$(16) \qquad {}_jAPR_{t+1,t} = \gamma_0^P + \gamma_1^P\left({}_jIQR_t\right) + \gamma_2^P\left({}_jAAPD_t\right) + \alpha_j^P + \mu_t^P + {}_j\varepsilon_{t+1,t}^P$$

$$(17) \qquad {}_jDR_{t+1,t} = \gamma_0^D + \gamma_1^D\left({}_jIQR_t\right) + \gamma_2^D\left({}_jAADD_t\right) + \alpha_j^D + \mu_t^D + {}_j\varepsilon_{t+1,t}^D$$

We present the forecast revision results in Table 7. Focusing first on respondents' point forecast revisions in specification (1) to (3), the data find that uncertainty is not significantly related to the degree of point forecast revisions for each of the target variables. In contrast, the data indicate that forecasters who had a higher level of individual disagreement associated with their point forecast tend to revise their subsequent point forecast by a larger amount. Recall our earlier finding that point forecasts for respondents who had a higher level of individual disagreement also tended to be less accurate. Controlling for time effects does not eliminate this relationship between individual disagreement and the point forecast revision, although it reduces the coefficient on disagreement by more than half for GDP forecast revisions. When we control for both time and respondent effects, disagreement remains statistically significant for all three outcome variables.

Looking at respondents' density forecast revisions in specifications (4) to (6), the data still indicate that respondents with higher levels of individual disagreement revise their subsequent density forecast by a greater degree. As for the point forecast revisions, this relationship remains statistically significant for all three target variables even controlling for both time and respondent fixed effects. However, the data now indicate for GDP and unemployment forecasts that individuals who were more uncertain tend to revise their density forecast by a smaller amount. In the case of GDP, this result is robust to controlling for time and respondent fixed effects. This result is counter to the intuition that more uncertainty over a forecast a priori would be associated with larger revision of the forecast on average as new information arrives.

V. Conclusion

This paper provides a detailed exploration into two aspects of forecast behavior—uncertainty and disagreement. We analyze the statistical properties of individual uncertainty and disagreement, as well as assess their roles in respondents' forecast performance and forecast revisions. In terms of motivation and contribution, our study complements the larger literature that has focused on the measurement of expectations. In particular, the expanded scope of our investigation serves as a basis for a better understanding and improved formulation of the beliefs formation process of individuals.

Using data from the ECB-SPF, we derive individual measures of uncertainty and disagreement from reported point and density forecasts. Our empirical analysis indicates substantial heterogeneity in respondents' uncertainty and disagreement. Moreover, there are also notable differences in uncertainty and disagreement. While we find persistence in the relative levels of respondents' uncertainty and personal disagreement, uncertainty displays much stronger persistence. There is also little correlation between uncertainty and disagreement, suggesting movements in the variables are largely independent of each other.

The lack of association between uncertainty and disagreement, however, allows for a relatively straightforward assessment of the relevance of uncertainty and disagreement for forecast performance and forecast revisions. Once again, differences between uncertainty and disagreement emerge from the analysis. While disagreement may not always display economic significance in the estimated relationships, it almost always contains greater predictive content than uncertainty and is generally statistically significant. On the other hand, the evidence does not indicate a robust relationship between the confidence associated with a respondent's forecast and its subsequent accuracy, and is also unable to identify a reasonable linkage between uncertainty and respondents' forecast revisions.

Taken together, our findings lead to suggestions for further research. One issue of interest is identifying differences in the nature or impact of the underlying factors guiding the divergent behavior of uncertainty and disagreement. Another issue relates to the strong persistence displayed by uncertainty during a sample period that encompasses both tranquil and volatile episodes, with the latter including the recent global financial crisis. In addition, there is the issue of why the results speak to a more important role for disagreement than uncertainty for the accuracy and extent of revisions of respondents' forecasts. Developing theoretical models that can account for the features described above would offer a significant advancement in the study of the beliefs formation process.

Appendix: The linkage between heteroskdasticity- and survey-based measures of uncertainty.

Time series models of heteroskedasticity simultaneously model conditional moments of a variable and can be described as follows. If we let X_{t+1} denote a variable of interest, I_t denote the information available in time *t,* and m_t and h_t denote, respectively, the conditional means and variances, then time series models of heteroskedasticity provide measures of:

$$\text{(1A)} \qquad \begin{aligned} m_t &= E\left[X_{t+1} \mid I_t\right] \\ h_t &= E\left[\left(X_{t+1} - m_t\right)^2 \mid I_t\right] \end{aligned}$$

by formulating and estimating a specification for each.[22] By using moment conditions, we can rewrite (1A) in terms of the following regression equations:

$$\text{(2A)} \qquad \begin{aligned} X_{t+1} &= m_t + \eta_{t+1} \\ \left(\eta_{t+1}\right)^2 &= h_t + \varepsilon_{t+1} \end{aligned}$$

where η_{t+1} and ε_{t+1} are mean-zero innovations in (2A) such that $E\left[\eta_{t+1} \mid I_t\right] = E\left[\varepsilon_{t+1} \mid I_t\right] = 0.$

There are two important features of time-series models of heteroskedasticity. First, there is a direct association between conditional variances $\left(h_t\right)$ and forecast accuracy $\left(\eta_{t+1}\right)^2$. Second, heteroskedasticity-based measures of uncertainty $\left(h_t\right)$ are equated to temporal variation in (subjective) forecast uncertainty. Consequently, measures of ex ante uncertainty in this class of models derive from the ex post predictability of a variable.

We can generalize the previous discussion in terms of measures of forecast performance, uncertainty, and disagreement that have been defined in our analysis of the ECB-SPF. In the case of the point forecast data and under the assumption that respondents make efficient use of their information sets, we can modify the system of equations in (2A) to arrive at the following regression model:

$$\text{(3A)} \qquad APA_{t+1} = \delta_0^P + \delta_1^P IQR_t + \varepsilon_{t+1}^p$$

---

[22] The choice of a one-step-ahead forecast horizon is based solely on convenience for expositional and illustrative purposes.

The substitution of APA_{t+1} for $(\eta_{t+1})^2$ reflects our adoption of the absolute value metric to calculate forecast accuracy, with the parameters δ_0 and δ_1 providing a link between the heteroskedsaticity- and survey-based measure of uncertainty. Extending (3A) to the individual level and expanding the forecast horizon beyond one period results in:

$$(4A) \qquad {}_jAPA_{t+h} = \delta_0^P + \delta_1^P\left({}_jIQR_t\right) + {}_j\varepsilon_{t+h}^p$$

Abstracting from respondent fixed effects and time effects, equation (4A) can be seen as a special case of equation (12). Because time series models of heteroskedasticity do not incorporate a role for disagreement, this consideration would imply a value of zero for δ_2^P. With regard to the relationship between forecast accuracy and the uncertainty measure, the assignment of a specific value to δ_1^P is more difficult because it depends on the association between h_t and IQR_t which is an issue under investigation and is therefore currently unknown. Nevertheless, we can still formulate predictions based on properties of time series models of heteoskedasticity and the assumed linkages between forecast accuracy and uncertainty as well as between the model- and survey-based measures of uncertainty. These predictions are that the uncertainty measure not only displays a positive and statistically significant relationship to the forecast accuracy measure, but also has economically significant predictive content for forecast performance.[23] For equation (13), there are analogous implications values for δ_1^D, δ_2^D, and the explanatory power of the uncertainty measure in the regression model.

[23] Recall that a larger value of our forecast performance measure indicates lower forecast accuracy.

References

Abel, Joshua, Robert Rich, Joseph Song, and Joseph Tracy. "The Measurement and Behavior of Uncertainty: Evidence from the ECB Survey of Professional Forecasters." *Journal of Applied Econometrics* 31 (April/May 2016): 533-550.

Arroyo, Javier and Carlos Mate. "Forecasting Histogram Time Series with K-Nearest Neighbours Mehods." *International Journal of Forecasting* 25 (January-March 2009): 192-207.

Amador, Manuel, and Pierre-Oliver Weill. "Learning From Prices: Public Communication and Welfare." *Journal of Political Economy* 118 (October 2010): 866-907.

Batchelor, Roy S., and Pami Dua. "Empirical Measures of Inflation Uncertainty: A Cautionary Note." *Applied Economics* 28 (March 1996): 333-341.

Bloom, Nicholas. "The Impact of Uncertainty Shocks." *Econometrica* 77 (May 2009): 623-685.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Saporata-Eksten, Terry Itay, and J. Stephen. "Really Uncertain Business Cycles." Working Paper No. 18245. National Bureau of Economic Research, July, 2012.

Boreo, Gianna, Jeremy Smith, and Kenneth F. Wallis. "Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters." *Economic Journal* 118 (July 2008): 1107-1127.

Boero, Gianna, Jeremy Smith, and Kenneth F. Wallis. "The Measurement and Characteristics of Professional Forecasters' Uncertainty." *Journal of Applied Econometrics* 30 (November/December 2015): 1029-1046.

Bowels, Carlos, Roberta Friz, Veronique Genre, Geoff Kenny, Aidan Meyler, and Tuomas Rautanen. "The ECB Survey of Professional Forecasters (SPF): A Review After Eight Years' Experience." ECB Occasional Paper No. 59, April 2007.

Bruine de Bruin, Wandi, Charles F. Manski, Giorgio Topa, and Wilbert van der Klaauw. "Measuring Consumer Uncertainty About Future Inflation." *Journal of Applied Econometrics* 26 (April-May 2011): 454-478.

Campbell, John Y. "Asset Pricing at the Millennium." *Journal of Finance* 55 (August 2000): 1515-1567.

Clements, Michael P. "Evaluating the Bank of England Density Forecasts of Inflation." *The Economic Journal* 114 (October 2004): 844-866.

------------. "Probability Distributions or Point Predictions? Survey Forecasts of US Output Growth and Inflation." *International Journal of Forecasting* 30 (January-March 2014a): 99-117.

------------. "Forecast Uncertainty—Ex Ante and Ex Post: U.S. Inflation and Output Growth." *Journal of Business and Economic Statistics* 32 (April 2014b): 206-216.

Diebold, Francis X., Anthony S. Tay, and Kenneth F. Wallis. "Evaluating Density Forecasts of Inflation: The Survey of Professional Forecasters." In *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, edited by Robert F. Engle and Halbert White, 76-90. Oxford and New York, Oxford University Press, 1999.

D'Amico, Stefania, and Athanasios Orphanides. "Uncertainty and Disagreement in Economic Forecasting." Finance and Economics Discussion Series 2008-56. Federal Reserve Board, Washington DC, 2008.

D'Agostino, Antonello, Kieran McQuinn, and Karl Whelan. "Are Some Forecasters Really Better Than Others?" *Journal of Money, Credit and Banking* 44 (June 2012): 715-732.

Engelberg, Joseph, Charles F. Manski, and Jared Williams. "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters." *Journal of Business and economic Statistics* 27 (January 2009): 30-41.

Engle, Robert F. "Estimates of the Variance of U.S. Inflation Based Upon the ARCH Model." *Journal of Money, Credit and Banking* 15 (August 1983): 286-301.

Garcia, Juan A. "An Introduction to the ECB's Survey of Professional Forecasters." ECB Occasional Paper No 8, September 2003.

Giordani, Paolo, and Paul Soderlind. "Inflation Forecast Uncertainty." *European Economic Review* 47 (December 2003): 1037-1059.

Kenny, Geoff, Thomas Kostka, and Federico Masera. "How Informative are the Subjective Density Forecasts of Macroeconomists?" *Journal of Forecasting* 33 (April 2014): 163-185.

------, Thomas Kostka, and Federico Masera. "Density Characteristics and Density Forecast Performance: A Panel Analysis." *Empirical Economics* 48 (May 2015): 1203-1231.

Lahiri, Kajal, and Xuguang Sheng. "Measuring Forecast Uncertainty by Disagreement: The Missing Link." *Journal of Applied Econometrics* 25 (June/July 2010): 514-538.

Lopez-Perez, Victor. "Weighted Averages of Individual Measures of Uncertainty from the European Central Bank's Survey of Professional Forecasters." Working Paper. Universidad Politecnica de Cartagena, December, 2014.

Lorenzoni, Guido. "A Theory of Demand Shocks." *Journal of Political Economy* 99 (December 2009): 2050-2084.

Mankiw, Gregory N., and Ricardo Reis. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (November 2002): 1295-1328.

Morris, Stephen, and Hyun Song Shin. "Social Value of Public Information." *American Economic Review* 92 (December 2002): 1521-1534.

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

Patton, Andrew J., and Allan Timmermann. "Why Do Forecasters Disagree? Lessons From the Term Structure of Cross-Sectional Dispersion." *Journal of Monetary Economics* 57 (October 2010): 803-820.

Rich, Robert W., and Joseph Tracy. "The Relationship Among Expected Inflation, Disagreement, and Uncertainty: Evidence From Matched Point and Density Forecasts." *Review of Economics and Statistics* 92 (February 2010): 200-207.

Woodford, Michael. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by Philippe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford. Princeton, Princeton University Press, 2003.

Zarnowitz, Victor, and Louis A. Lambros. "Consensus and Uncertainty in Economic Predictions." *Journal of Political Economy* 95 (June 1987): 591-621.

Table 1. Transition Probabilities – GDP

a) Average Absolute Point Disagreement

Quartile t		Quartile t+1: 1	2	3	4
	1	**0.39*****	0.24	0.22**	0.15***
	2	0.33	**0.21**	0.27	0.20***
	3	0.31	0.25	**0.24**	0.20***
$\chi^2(16) = 73$	4	0.22*	0.19***	0.27	**0.32*****

b) Average Absolute Density Disagreement

Quartile t		Quartile t+1: 1	2	3	4
	1	**0.40*****	0.28	0.20***	0.12***
	2	0.30	**0.30*****	0.25	0.14***
	3	0.21***	0.23	**0.31*****	0.25
$\chi^2(16) = 277$	4	0.13***	0.18***	0.25	**0.45*****

c) Uncertainty (IQR)

Quartile t		Quartile t+1: 1	2	3	4
	1	**0.64*****	0.27	0.06***	0.02***
	2	0.31	**0.43*****	0.22**	0.04***
	3	0.06***	0.21***	**0.54*****	0.19***
$\chi^2(16) = 1{,}721$	4	0.01***	0.05***	0.21***	**0.73*****

Notes: One-tailed tests for individual diagonal (off-diagonal) elements > (<) 0.25.
Chi-square statistics are for a joint test for uniform distribution for the entire table.
*** significant at the 1% level
** significant at the 5% level
* significant at the 10% level

Table 2. Transition Probabilities − Inflation

a) Average Absolute Point Disagreement

Quartile t	Quartile t+1: 1	2	3	4
1	**0.43*****	0.22**	0.21***	0.14***
2	0.37	**0.23**	0.25	0.14***
3	0.35	0.22**	**0.25**	0.18***
4	0.17***	0.19***	0.22*	**0.42*****

$\chi^2(16) = 206$

b) Average Absolute Density Disagreement

Quartile t	Quartile t+1: 1	2	3	4
1	**0.45*****	0.27	0.17***	0.11***
2	0.27	**0.32*****	0.24	0.16***
3	0.19***	0.26	**0.34*****	0.21***
4	0.12***	0.14***	0.26	**0.48*****

$\chi^2(16) = 419$

c) Uncertainty (IQR)

Quartile t	Quartile t+1: 1	2	3	4
1	**0.64*****	0.26	0.09***	0.02***
2	0.29	**0.45*****	0.23*	0.04***
3	0.08***	0.20***	**0.56*****	0.15***
4	0.02***	0.05***	0.17***	**0.75*****

$\chi^2(16) = 1{,}923$

Notes: One-tailed tests for individual diagonal (off-diagonal) elements > (<) 0.25.
Chi-square statistics are for a joint test for uniform distribution for the entire table.
*** significant at the 1% level
** significant at the 5% level
* significant at the 10% level

Table 3. Transition Probabilities − Unemployment

a) Average Absolute Point Disagreement

Quartile t	Quartile t+1: 1	2	3	4
1	**0.40*****	0.25	0.21***	0.14***
2	0.38	**0.22**	0.25	0.15***
3	0.29	0.22*	**0.30*****	0.18***
4	0.20**	0.17***	0.22*	**0.40*****

$\chi^2(16) = 158$

b) Average Absolute Density Disagreement

Quartile t	Quartile t+1: 1	2	3	4
1	**0.40*****	0.25	0.23	0.12***
2	0.28	**0.33*****	0.22*	0.17***
3	0.20***	0.25	**0.32*****	0.23*
4	0.14***	0.16***	0.24	**0.46*****

$\chi^2(16) = 272$

c) Uncertainty (IQR)

Quartile t	Quartile t+1: 1	2	3	4
1	**0.60*****	0.29	0.09***	0.02***
2	0.32	**0.42*****	0.21**	0.05***
3	0.09***	0.22*	**0.50*****	0.19***
4	0.02***	0.06***	0.18***	**0.73*****

$\chi^2(16) = 1{,}448$

Notes: One-tailed tests for individual diagonal (off-diagonal) elements > (<) 0.25.
Chi-square statistics are for a joint test for uniform distribution for the entire table.
*** significant at the 1% level
** significant at the 5% level
* significant at the 10% level

Table 4. Variance Decompositions

a) Average Absolute Point Disagreement

Source	GDP	Inflation	Unemployment
Time	41.4	27.5	37.0
Person	6.8	10.3	8.3
Residual	50.3	61.9	53.9

b) Average Absolute Density Disagreement

Source	GDP	Inflation	Unemployment
Time	45.3	31.2	32.4
Person	7.2	11.3	10.7
Residual	46.4	57.0	55.6

c) Uncertainty (IQR)

Source	GDP	Inflation	Unemployment
Time	8.4	10.8	11.2
Person	46.3	45.6	40.6
Residual	40.6	39.1	45.3

Notes: ECB-SPF data. Authors calculations.

Table 5. Relationship of Uncertainty (IQR) to Disagreement

a) GDP

	(1)	(2)	(3)	(4)	(5)	(6)
Average Absolute	0.164***	0.077**	0.020			
Point Disagreement	(0.029)	(0.040)	(0.028)			
Average Absolute				0.187***	0.215***	0.120***
Density Disagreement				(0.041)	(0.064)	(0.038)
Constant	0.639***	0.570***	0.639***	0.145***	0.079	0.169***
	(0.028)	(0.041)	(0.031)	(0.026)	(0.049)	(0.042)
Time effects	N	Y	Y	N	Y	Y
Person effects	N	N	Y	N	N	Y
Incremental R^2	0.017	0.002	0.000	0.083	0.039	0.016

b) Inflation

	(1)	(2)	(3)	(4)	(5)	(6)
Average Absolute	0.098**	0.000	0.016			
Point Disagreement	(0.046)	(0.057)	(0.040)			
Average Absolute				0.285***	0.251***	0.204***
Density Disagreement				(0.047)	(0.059)	(0.048)
Constant	0.633***	0.554***	0.670***	0.076***	−0.017	0.095*
	(0.026)	(0.038)	(0.027)	(0.026)	(0.050)	(0.048)
Time effects	N	Y	Y	N	Y	Y
Person effects	N	N	Y	N	N	Y
Incremental R^2	0.004	0.000	0.001	0.123	0.065	0.036

c) Unemployment

	(1)	(2)	(3)	(4)	(5)	(6)
Average Absolute	0.101***	−0.036	−0.033			
Point Disagreement	(0.034)	(0.054)	(0.030)			
Average Absolute				0.244***	0.269***	0.180***
Density Disagreement				(0.051)	(0.072)	(0.050)
Constant	0.611***	0.556***	0.616***	0.062**	−0.060	0.042
	(0.028)	(0.036)	(0.027)	(0.030)	(0.059)	(0.051)
Time effects	N	Y	Y	N	Y	Y
Person effects	N	N	Y	N	N	Y
Incremental R^2	0.005	0.000	0.000	0.120	0.097	0.036

Notes: Standard errors are reported in parentheses and are based on clustering at the respondent level. The incremental R^2 pertains to the disagreement variable. For specifications (1) and (4), the incremental R^2 is identical to the overall R^2.

*** significant at the 1% level

** significant at the 5% level

* significant at the 10% level

Table 6. Forecast Performance

a) GDP	Absolute Point Accuracy	Expected Absolute Accuracy	Absolute Rank Probability Score
	(1)	(2)	(3)
Uncertainty (IQR)	0.004	0.203***	0.238**
	(0.045)	(0.033)	(0.124)
Average Absolute Point Disagreement	0.368***		
	(0.105)		
Average Density Disagreement		0.221***	0.154
		(0.052)	(0.209)
Constant	0.770***	0.691	2.793***
	(0.069)	(0.055)	(0.205)
R^2	0.919	0.937	0.904
Incremental R^2 – uncertainty	0.000	0.001	0.000
Incremental R^2 – disagreement	0.003	0.004	0.000
Incremental R^2 – respondent FEs	0.007	0.005	0.015

b) Inflation	Absolute Point Accuracy	Expected Absolute Accuracy	Absolute Rank Probability Score
	(1)	(2)	(3)
Uncertainty (IQR)	0.056*	0.255***	0.644***
	(0.037)	(0.033)	(0.160)
Average Absolute Point Disagreement	0.664***		
	(0.106)		
Average Density Disagreement		0.304***	0.658**
		(0.051)	(0.285)
Constant	0.316***	0.252***	1.274***
	(0.052)	(0.046)	(0.263)
R^2	0.840	0.868	0.859
Incremental R^2 – uncertainty	0.000	0.007	0.003
Incremental R^2 – disagreement	0.022	0.017	0.005
Incremental R^2 – respondent FEs	0.018	0.013	0.027

c) Unemployment	Absolute Point Accuracy	Expected Absolute Accuracy	Absolute Rank Probability Score
	(1)	(2)	(3)
Uncertainty (IQR)	−0.019	0.171***	0.346**
	(0.041)	(0.041)	(0.163)
Average Absolute Point Disagreement	0.605***		
	(0.119)		
Average Density Disagreement		0.336***	0.549**
		(0.043)	(0.240)
Constant	1.047***	0.956***	−0.520***
	(0.060)	(0.052)	(0.179)
R^2	0.770	0.790	0.870
Incremental R^2 – uncertainty	0.000	0.005	0.001
Incremental R^2 – disagreement	0.027	0.048	0.005
Incremental R^2 – respondent FEs	0.018	0.026	0.017

Notes: Standard errors are reported in parentheses and are based on clustering at the respondent level. Specifications contain respondent and time fixed effects.

*** significant at the 1% level,
** significant at the 5% level,
* significant at the 10% level

Table 7. Forecast Revision

a) GDP

	Absolute Point Revision			Density Revision		
	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty (IQR)	0.012	−0.011	−0.043	−0.254***	−0.195***	−0.201***
	(0.024)	(0.023)	(0.030)	(0.072)	(0.069)	(0.066)
Avg Absolute Point Disagreement	0.807***	0.375***	0.357***			
	(0.069)	(0.087)	(0.095)			
Average Density Disagreement				0.918***	0.375***	0.361***
				(0.074)	(0.103)	(0.110)
Constant	0.069**	0.170***	0.197***	0.112*	0.406***	0.451***
	(0.026)	(0.045)	(0.057)	(0.065)	(0.103)	(0.118)
Time effects	N	Y	Y	N	Y	Y
Person effects	N	N	Y	N	N	Y
R^2	0.174	0.436	0.466	0.177	0.430	0.463

b) Inflation

	Absolute Point Revision			Density Revision		
	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty (IQR)	0.014	0.014	−0.009	−0.094	−0.068	−0.045
	(0.017)	(0.019)	(0.038)	(0.066)	(0.070)	(0.092)
Avg Absolute Point Disagreement	0.508***	0.418***	0.393***			
	(0.038)	(0.039)	(0.042)			
Average Density Disagreement				0.596***	0.498***	0.461***
				(0.041)	(0.046)	(0.054)
Constant	0.052***	0.172***	0.180***	0.054	0.193**	0.171*
	(0.013)	(0.039)	(0.044)	(0.042)	(0.077)	(0.090)
Time effects	N	Y	Y	N	Y	Y
Person effects	N	N	Y	N	N	Y
R^2	0.142	0.203	0.251	0.154	0.226	0.290

c) Unemployment

	Absolute Point Revision			Density Revision		
	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty (IQR)	0.050***	0.015	0.024	−0.163**	−0.111	−0.043
	(0.018)	(0.020)	(0.035)	(0.080)	(0.070)	(0.073)
Avg Absolute Point Disagreement	0.732***	0.364***	0.329***			
	(0.078)	(0.100)	(0.109)			
Average Density Disagreement				0.708***	0.311***	0.249**
				(0.089)	(0.090)	(0.096)
Constant	0.047***	0.201***	0.136**	0.181***	0.544***	0.487***
	(0.022)	(0.053)	(0.055)	(0.068)	(0.106)	(0.102)
Time effects	N	Y	Y	N	Y	Y
Person effects	N	N	Y	N	N	Y
R^2	0.181	0.384	0.412	0.130	0.336	0.371

Notes: Standard errors are reported in parentheses and are based on clustering at the respondent level.
*** significant at the 1% level
** significant at the 5% level
* significant at the 10% level

Figure 1. Motivating a Density Disagreement Measure

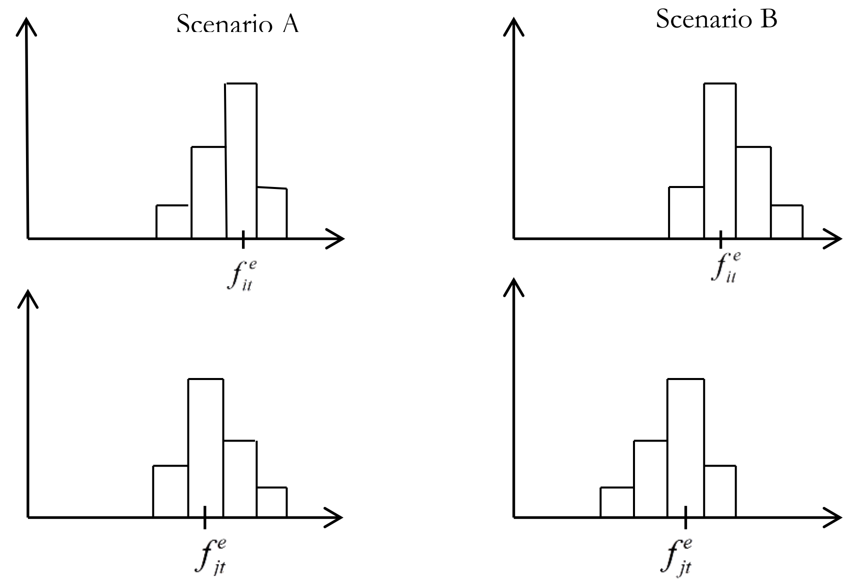


Figure 2. Motivating a Density Accuracy Measure

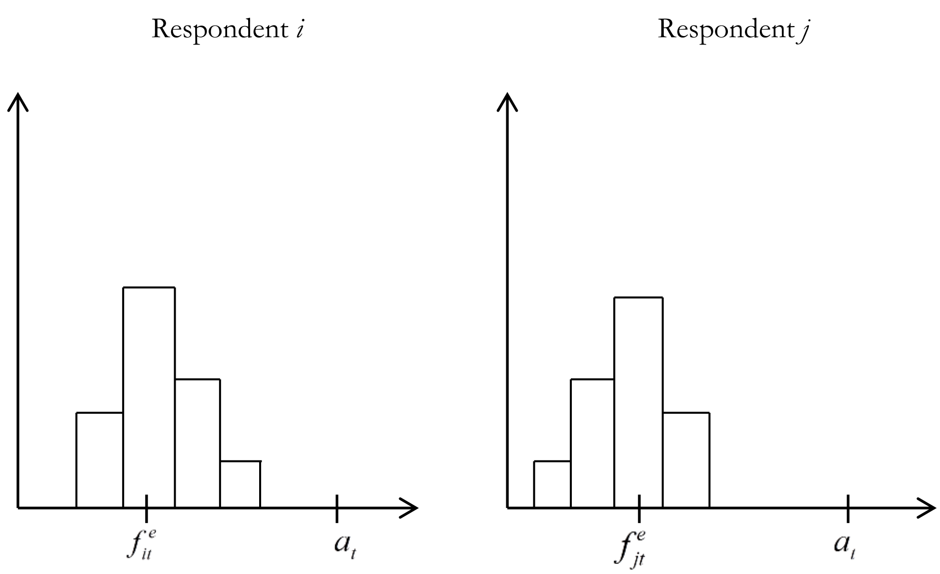
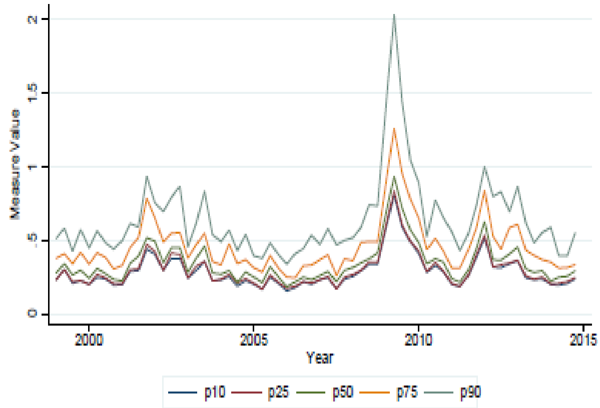


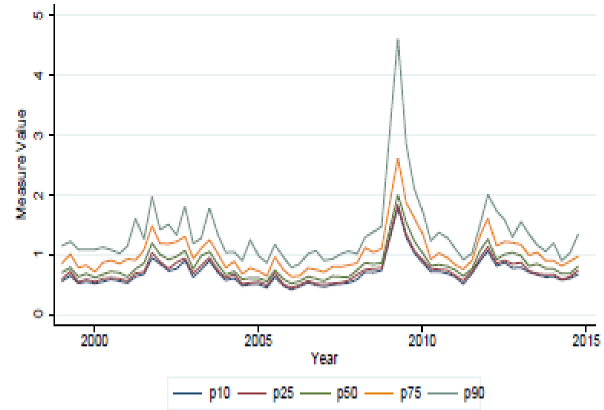
Figure 3. Distribution of Individual Disagreement Over Time

a) GDP

Absolute Point Disagreement

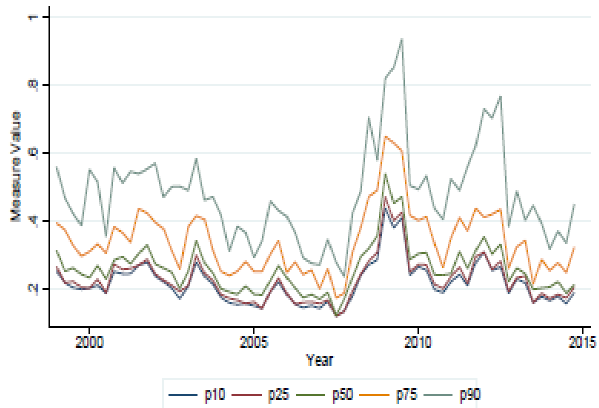


Density Disagreement

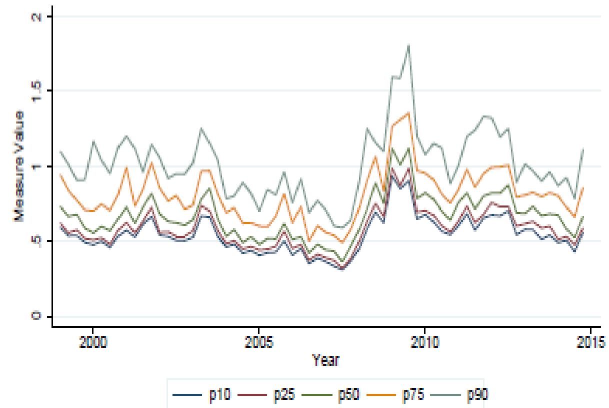


b) Inflation

Absolute Point Disagreement

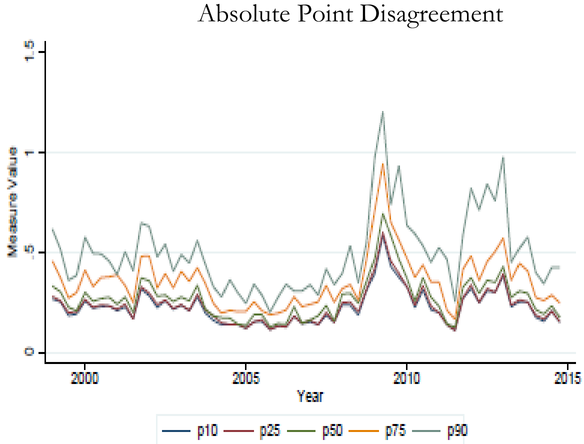


Density Disagreement



c) Unemployment

Absolute Point Disagreement



Density Disagreement

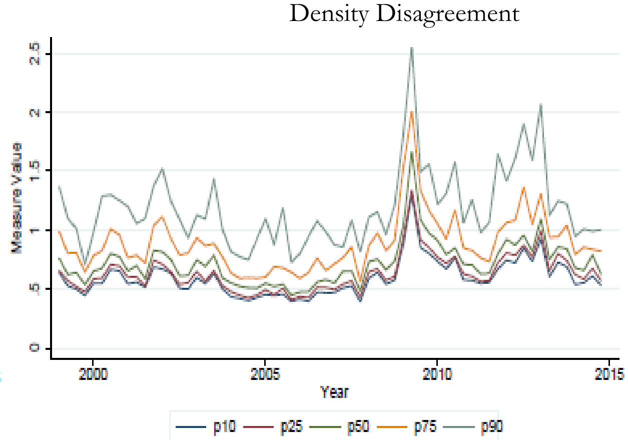


Figure 4. Distribution of Individual Uncertainty (IQR) Over Time

a) GDP

b) Inflation

a) Unemployment