

Conference on

Forecasting and Monetary Policy

Berlin, 23-24 March 2009

Kajal Lahiri
SUNY at Albany

„Measuring Forecast Uncertainty by Disagreement“

Measuring Forecast Uncertainty by Disagreement: The Missing Link

Kajal Lahiri and Xuguang Sheng
University at Albany - SUNY

Since the seminal work of Zarnowitz and Lambros (1987), economists have studied but disagreed on whether disagreement is a good proxy for uncertainty. See, for example, studies by Bomberger (1996, 1999), Rich and Butler (1998), Giordani and Söderlind (2003), Lahiri and Liu (2005), Boreo, Smith and Wallis (2007) and many others.

Interestingly, there has been a parallel but largely independent research in the accounting and finance literature on whether disagreement among financial or market analysts can be used as a proxy for uncertainty about future earnings. See Zhang (2006), Barry and Jennings (1992), Abarbanell et al. (1995), Barron et al. (1998), Diether et al. (2002), Johnson (2004) and others .

MAIN RESULTS

Using a standard decomposition of forecasts errors into common and idiosyncratic shocks, we show that aggregate forecast uncertainty can be expressed as the disagreement among the forecasters plus the perceived variability of future aggregate shocks. Thus, the reliability of disagreement as a proxy for uncertainty will be determined by the stability of the forecasting environment, and the length of the forecast horizon. Using density forecasts from the *Survey of Professional Forecasters* over 1969-2007, we find direct evidence in support of our hypothesis.

In recent Accounting and Finance literature, squared errors in consensus forecasts have been used to proxy for the variance of future aggregate shocks as a component of forecast uncertainty. Our results suggest that adding squared mean forecast error to disagreement can make the estimated uncertainty worse than the use of disagreement alone. If one wants to construct a robust *ex ante* measure of uncertainty, our suggestion is to use the sum of the observed disagreement from the survey and the variance of future aggregate shocks generated by GARCH-type models that use a moving average squared errors over past few years as one of the covariates.

Time permitting, we will also look into the implications of our result for

- i) the recent forecast combination literature, cf. Pesaran et al (2009); and
- ii) providing benchmark forecast uncertainty by FOMC, cf. Reifschneider and Tulip (2007).

THE ECONOMETRIC MODEL

For N individuals, T target years, H forecast horizons, let F_{ith} be the forecast of the variable of interest made by agent i , for the target year t and h -quarter ahead to the end of the target year, and A_t be the actual value of variable. The individual forecast error (e_{ith}) is defined as: $e_{ith} = A_t - F_{ith}$.

Following Davies and Lahiri (1995), we write e_{ith} as the sum of a component common to all forecasters (λ_{th}) and idiosyncratic errors (ε_{ith}):

$$e_{ith} = \lambda_{th} + \varepsilon_{ith}, \quad \lambda_{th} = \sum_{j=1}^h u_{tj}.$$

The common component (λ_{th}) represents the cumulative effect of all shocks that occurred from h -quarter ahead to the end of target year t . Equation specifies λ_{th} as the accumulation of all quarterly aggregate shocks (u_{tj}) over the forecast horizon.

We make the following simplifying assumptions:

Assumption 1:

$$E(u_{tj}) = 0; \text{ var}(u_{tj}) = \sigma_{u|tj}^2 \text{ for any } t \text{ and } j; E(u_{tj}u_{ts}) = 0 \text{ for any } t \text{ and } j \neq s;$$

$$E(u_{th}u_{t-k,h}) = 0 \text{ for any } t, h \text{ and } k \neq 0.$$

Assumption 2:

$$E(\varepsilon_{ith}) = 0; \text{ var}(\varepsilon_{ith}) = \sigma_{\varepsilon|ith}^2 \text{ for any } i, t \text{ and } h; E(\varepsilon_{ith}\varepsilon_{jth}) = 0 \text{ for any } t, h \text{ and } i \neq j.$$

Assumption 3:

$$E(\varepsilon_{ith}u_{t-k,j}) = 0 \text{ for any } i, t, h, k \text{ and } j.$$

The observed disagreement (d_{th}) among forecasters is the variance of their point forecasts which can be expressed as:

$$d_{th} \equiv \frac{1}{N-1} \sum_{i=1}^N (F_{ith} - F_{.th})^2 = \frac{1}{N-1} \sum_{i=1}^N (\varepsilon_{ith} - \frac{1}{N} \sum_{i=1}^N \varepsilon_{ith})^2,$$

where $F_{.th} = \frac{1}{N} \sum_{i=1}^N F_{ith}$. Note that the sample variance d_{th} is a random variable prior to observing forecasts. Taking expectations, we get the non-random disagreement, denoted by D_{th} :

$$\begin{aligned} D_{th} \equiv E(d_{th}) &= \frac{1}{N-1} \sum_{i=1}^N E(\varepsilon_{ith} - \frac{1}{N} \sum_{i=1}^N \varepsilon_{ith})^2 \\ &= \frac{1}{N-1} \sum_{i=1}^N (\sigma_{\varepsilon|ith}^2 + \frac{1}{N^2} \sum_{i=1}^N \sigma_{\varepsilon|ith}^2 - \frac{2}{N} \sigma_{\varepsilon|ith}^2) \\ &= \frac{1}{N} \sum_{i=1}^N \sigma_{\varepsilon|ith}^2. \end{aligned}$$

Thus, not surprisingly, we find that D_{th} is determined by the average variance of idiosyncratic errors.

The uncertainty associated with a forecast of any specific individual is measured by the variance of individual forecast error, and can be expressed as

$$U_{ith} \equiv \text{Var}(A_t - F_{ith}) = \text{Var}(\lambda_{th} + \varepsilon_{ith}) = \sigma_{\lambda|th}^2 + \sigma_{\varepsilon|ith}^2.$$

Individual forecast uncertainty above is comprised of two components: perceived uncertainty associated with forthcoming common shocks, $\sigma_{\lambda|th}^2$ and idiosyncratic shocks, $\sigma_{\varepsilon|ith}^2$. Following Zarnowitz and Lambros (1987), we measure overall forecast uncertainty (U_{th}) as the average of the individual forecast error variances

$$U_{th} \equiv \frac{1}{N} \sum_{i=1}^N U_{ith},$$

which measures the confidence an outside observer will have in a randomly drawn typical individual forecast from the panel of forecasters. Given our model, U_{th} can be expressed as a function of the model parameters as:

$$U_{th} = \sigma_{\lambda|th}^2 + \frac{1}{N} \sum_{i=1}^N \sigma_{\varepsilon|ith}^2.$$

After substituting, we get

$$U_{th} = \sigma_{\lambda|th}^2 + D_{th}.$$

Given the model assumptions, aggregate forecast uncertainty is simply the disagreement plus the variance of the accumulated aggregate shocks over the forecast horizon. Thus, the wedge between uncertainty and disagreement will be determined partly by the size of the forecast horizon over which the aggregate shocks accumulate – the longer is the forecast horizon the bigger will be the difference on the average. It also suggests that the robustness of the relationship between two will depend on the variability of aggregate shocks over time. In relatively stable time periods where the variability of aggregate shocks is small, whether the variability of these shocks were correctly perceived or not, disagreement will be a good proxy for the unobservable aggregate forecast uncertainty.

Bomberger (1996) examined the dependence of the variance of consensus forecast errors (called “consensus uncertainty”) on forecast disagreement using Livingston’s survey data on inflation expectations. Certainly, a positive relationship between the two during periods of economic instability will ensure that disagreement will continue to be positively correlated with the overall forecast uncertainty. However, since the difference between uncertainty and disagreement is the variance of unanticipated aggregate shocks (as will be explained later, this is approximately the same as the “consensus uncertainty”), theoretically it is not clear why disagreement will be able to predict it. Certainly, to the contrary to a statement in Bomberger (1996, p.385), it is *not* necessary that “if disagreement is to be a good proxy for individual uncertainty, it must also track consensus uncertainty”.

In our current framework, we model the variance of forecast errors without modeling forecasters' expectation formation process. Actually, it is easy to connect our model with Bayesian learning framework that models individuals' forecasting behavior. Suppose that each forecaster is endowed with two signals: one public signal, represented by

$$l_{it} = A_t + \eta_{it}, \quad \eta_{it} \sim N(0, 1/\sigma_{\eta|it}^2),$$

and one private signal, represented by

$$s_{it} = A_t + \zeta_{it}, \quad \zeta_{it} \sim N(0, 1/\sigma_{\zeta|it}^2).$$

The private signal is assumed to be independent of the public signal and also independent of other private signals, which are standard assumptions in the literature, cf. Lahiri and Sheng (2007). Each forecaster then combines these two sources of information, via Bayes rule, to derive the conditional expected value of A_t as

$$F_{it} \equiv E(A_t | l_{it}, s_{it}) = (\sigma_{\eta|it}^2 l_{it} + \sigma_{\zeta|it}^2 s_{it}) / (\sigma_{\eta|it}^2 + \sigma_{\zeta|it}^2),$$

and the conditional variance of A_t as

$$U_{it} \equiv \text{Var}(A_t | l_{it}, s_{it}) = 1 / (\sigma_{\eta|it}^2 + \sigma_{\zeta|it}^2).$$

The individual forecast uncertainty defined above reflects the uncertainty in both the public and private information, which is similar to what we derived before where the individual forecast uncertainty is comprised of perceived uncertainty associated with forthcoming common shocks and idiosyncratic shocks. Then we measure overall forecast uncertainty (U_{it}) as the average of the individual

forecast uncertainties $U_{it} \equiv \frac{1}{N} \sum_{i=1}^N U_{it}$. Given the Bayesian learning model, U_{it} can

be expressed as:

$$U_{it} = \frac{1}{N} \sum_{i=1}^N \frac{1}{(\sigma_{\eta|it}^2 + \sigma_{\zeta|it}^2)}.$$

Note that overall forecast uncertainty above, derived in the context of Bayesian learning framework, provides the justification that the aggregate uncertainty should be defined as the simple average of individual uncertainties, and not the uncertainty of the average forecast.

The disagreement among forecasters can be measured by the expected dispersion of F_{ith} . To examine the effect of new information on the disagreement, we consider the so-called pre-posterior variance of opinions across forecasters. For any given information system represented by $\sigma_{\eta|th}^2$ and $\sigma_{\zeta|ith}^2$, the pre-posterior variance is the variance based on the distribution of the signals l_{th} and s_{ith} for $i=1,2,\dots,N$. The disagreement among forecasters can then be measured as

$$\begin{aligned} D_{th} &\equiv E\left[\frac{1}{N-1}\sum_{i=1}^N(F_{ith} - \frac{1}{N}\sum_{i=1}^N F_{ith})^2\right] \\ &= \frac{1}{N}E\left[\sum_{i=1}^N F_{ith}^2\right] - \frac{1}{N(N-1)}E\left[\sum_{i=1}^N\sum_{j\neq i}^N F_{ith}F_{jth}\right]. \end{aligned}$$

After substituting for F_{ith} , we get

$$D_{th} = \left[\frac{1}{N}\sum_{i=1}^N \frac{1}{(\sigma_{\eta|th}^2 + \sigma_{\zeta|ith}^2)}\right] - \left[\frac{1}{N(N-1)}\sum_{i=1}^N\sum_{j\neq i}^N \frac{\sigma_{\eta|th}^2}{(\sigma_{\eta|th}^2 + \sigma_{\zeta|ith}^2)(\sigma_{\eta|th}^2 + \sigma_{\zeta|jth}^2)}\right].$$

Note that the first term on the right-hand side of the equation above is forecast uncertainty, U_{th} and the second term is the average covariance among forecast errors, C_{th} , where

$$C_{th} = \frac{1}{N(N-1)}\sum_{i=1}^N\sum_{j\neq i}^N Cov(A_t - F_{ith}, A_t - F_{jth}).$$

Barry and Jennings (1992) derived a similar relationship among uncertainty, disagreement and the average covariance in forecasts. Their result justifies forecast disagreement as one component of forecast uncertainty, which has, unfortunately, been unnoticed in the economics literature. Given our model, we can simplify the expression for the average covariance among forecast errors as

$$C_{th} = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j \neq i}^N E[(\lambda_{th} + \varepsilon_{ith})(\lambda_{th} + \varepsilon_{jth})] = \sigma_{\lambda|th}^2,$$

which can be easily interpreted as the uncertainty shared by all forecasters due to their exposure to common shocks. Thus, we simplify the results in Barry and Jennings (1992) and Barron et al. (1998) greatly.

EMPIRICAL TEST OF THE RELATIONSHIP BETWEEN UNCERTAINTY AND DISAGREEMENT

Data

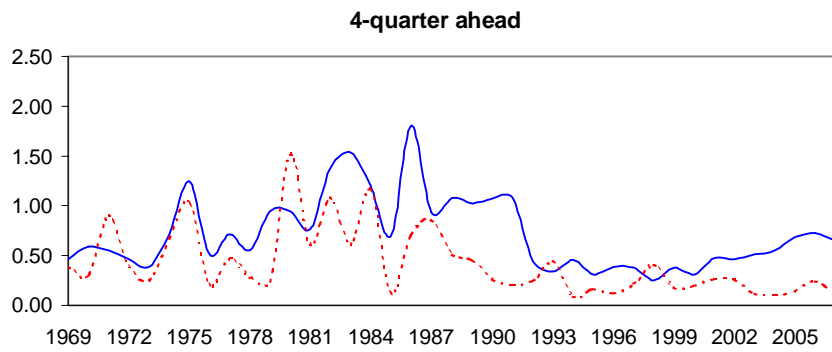
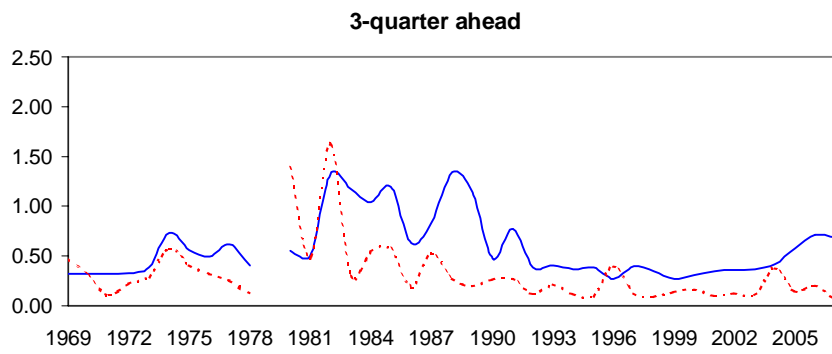
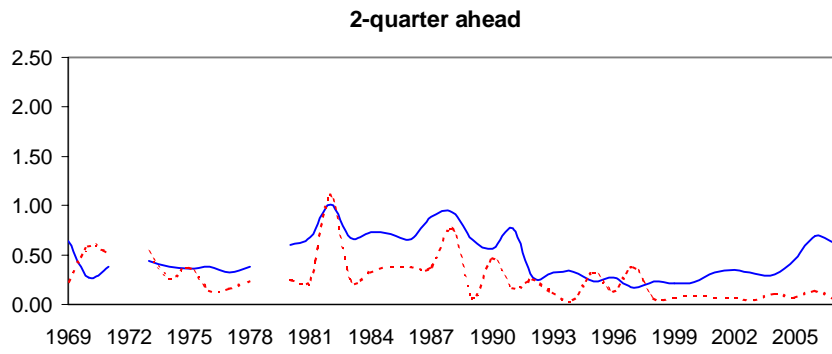
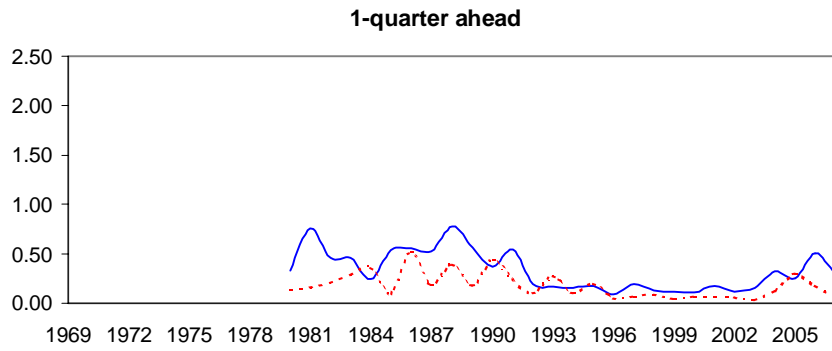
The data in our study are taken from Survey of Professional Forecasters (SPF) that is provided by the Federal Reserve Bank of Philadelphia. A unique feature of SPF data is that forecasters are also asked to provide density forecasts for output growth and inflation, which is the focus of this paper. We focus on the density forecasts for the change from year $t-1$ to t that were issued in the four consecutive surveys from the first quarter through the fourth quarter of year t . The actual horizons for these four forecasts are approximately $3\frac{1}{2}$, $2\frac{1}{2}$, $1\frac{1}{2}$, and $\frac{1}{2}$ quarters but we shall refer to them simply as horizons 4, 3, 2, and 1 quarter. After deleting observations with missing values, we obtain a total of 4,986 observations for inflation over 1969:Q1 to 2007:Q4 and 3,312 observations for output growth over 1981:Q3 to 2007:Q4. For the purpose of estimation, we eliminate observations for infrequent respondents. We focus on the “regular” respondents who participated in at least 25 surveys in inflation forecasts and at least 17 surveys in output growth forecasts – approximately 15% in both cases. This leaves us with a total of 2,787 observations for inflation forecasts and 2,342 observations for output growth forecasts. We use real-time data available from the Federal Reserve Bank of Philadelphia web site.

Test of the relationship between uncertainty and disagreement

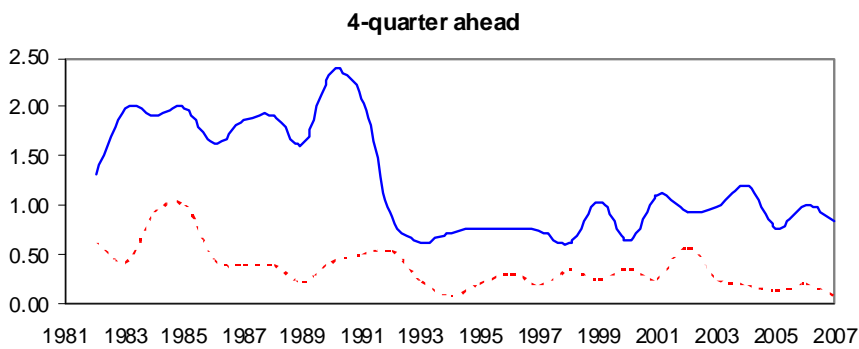
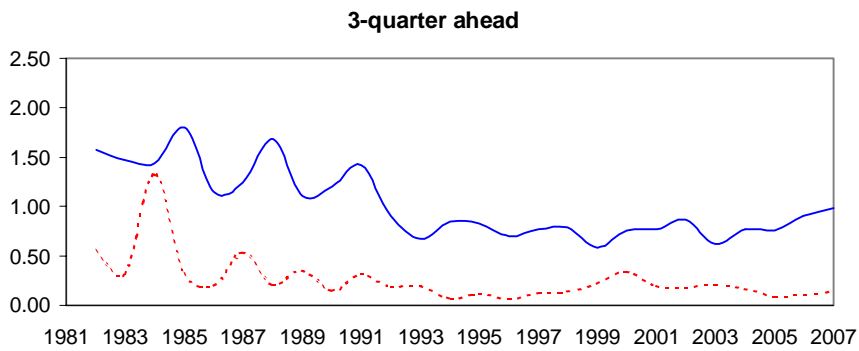
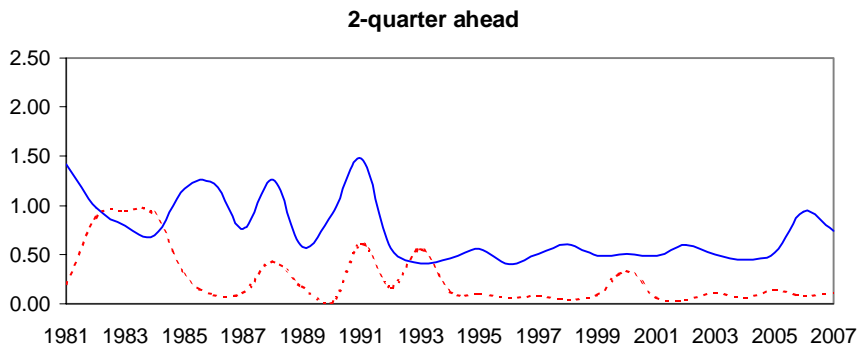
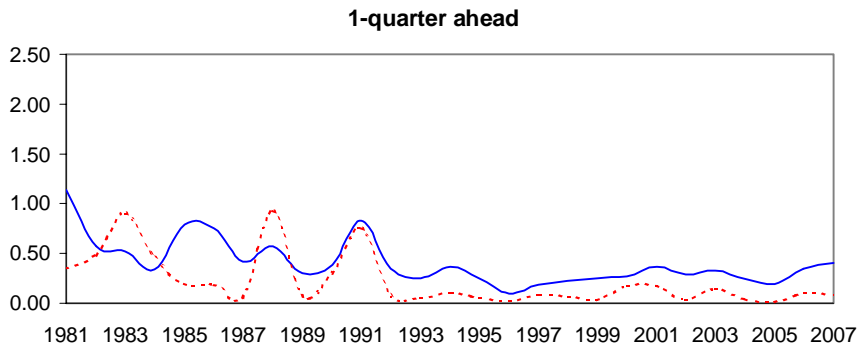
We should note that our equation specifies a relationship between uncertainty, disagreement and the variance of aggregate shocks based on unconditional expectations before observing any forecast or actual. However, the SPF forecast density data were obtained sequentially in the real time. Thus, we should develop a corresponding relationship in terms of expectations conditional on observing the individual forecasts (and hence disagreement d_{it}) at time $t-h$, but before the actual value A_t was realized.

For typical values of N and $\sigma_{\varepsilon|it}^2$ in our context, an extra term involving $(1/N^2)$ comes in but will be very close to zero for large N and can be ignored. Thus, the difference between the reported *ex ante* forecast uncertainty and disagreement will give approximately estimates of *ex ante* variance of aggregate shocks in real time before the actual values were realized. Estimates of uncertainty are plotted in Figures 1 to 2. Their average values are given in Table 1. Note that the difference between uncertainty and disagreement is an estimate of the variance of *ex ante* aggregate shock in real time.

Uncertainty (solid line) and disagreement (dotted line) in inflation forecasts



Uncertainty (solid line) and disagreement (dotted line) in real GDP forecasts



Uncertainty and disagreement averaged over time

	SPF inflation forecast (1969-2007)				SPF GDP forecast (1981-2007)			
	1Q ahead	2Q ahead	3Q ahead	4Q ahead	1Q ahead	2Q ahead	3Q ahead	4Q ahead
Uncertainty	0.33	0.48	0.58	0.69	0.41	0.74	1.02	1.25
Disagreement	0.18	0.26	0.32	0.42	0.22	0.25	0.26	0.37
Difference	0.15	0.22	0.26	0.27	0.19	0.49	0.76	0.88

Several points are worth noting in these graphs and table.

Disagreement and uncertainty typically move together but the former is almost always smaller than the latter in both series, which is in line with the evidence that the former tends to underestimate the latter (cf. Zarnowitz and Lambros, 1987; Lahiri et al. 1988). Also, the difference between uncertainty and disagreement (i.e., the variance of *ex ante* aggregate shocks) in both series becomes larger, as forecast horizon gets longer from 1 quarter to 4 quarters, which provides evidence in support of the horizon effect.

Second, somewhat unexpectedly, in some quarters disagreement exceeds uncertainty, especially for inflation. One explanation can be the well-documented structural break of the early 80s. As is well known, inflation rose sharply and unexpectedly during 1979-1981, and is characterized by a break in the inflation process. This may lead to forecasters adopting disparate forecasting functions and as a result, their predictions may generate extraordinary disagreement. As Mankiw et al. (2003) have argued, the distribution of point forecasts may become multi-modal as forecasters learn about the new regime and react differently. More specifically, this may lead to significant correlations – some negative and some positive – in idiosyncratic shocks between the forecasters, violating one of the assumptions in our forecast error decomposition. Note that for the validity of relation, we need the average pair-wise covariances in forecast errors (net of common shocks) to be zero. Using the 4-quarter ahead inflation forecasts, we however, found that this condition was not violated even during the tumultuous period 1979-83.¹⁷ Thus, on balance, we feel that the estimated negative variance of aggregate shocks on few occasions in our sample can be attributed to the imprecision in the estimation of disagreement or uncertainty due to small samples and outliers.

Now we can test formally the implications of our relationship between uncertainty and disagreement which depends on the variance of aggregate shocks over time and across horizons. We thus divide the sample of inflation forecasts into two periods: the unstable period (1969-1983) and the stable period (1984-2007). To study the relationship between uncertainty and disagreement, we run the following regression:

$$U_{it} = \beta D_{it} + \rho_1 H_{1i} + \rho_2 H_{2i} + \rho_3 H_{3i} + \rho_4 H_{4i} + \varepsilon_{it},$$

where $H_i = 1$ if the forecast is made at horizon i for $i = 1, 2, 3, 4$, and 0 otherwise. Table below shows the estimation results.

Regression of survey measure of uncertainty on disagreement

	SPF inflation forecast		SPF GDP forecast
	1969-1983	1984-2007	1984-2007
Disagreement	0.43*	0.76*	0.72*
	(0.08)	(0.19)	(0.25)
H1	0.39*	0.17*	0.24*
	(0.13)	(0.03)	(0.04)
H2	0.34*	0.31*	0.56*
	(0.04)	(0.04)	(0.05)
H3	0.36*	0.42*	0.81*
	(0.06)	(0.04)	(0.06)
H4	0.53*	0.39*	0.96*
	(0.06)	(0.07)	(0.09)
Adj. R ²	0.34	0.39	0.53

The estimated coefficient on disagreement is 0.43 for inflation forecasts during 1969-83. The same coefficient during 1984-2007 is estimated to be 0.76 and 0.72 for inflation and GDP forecasts, respectively. Thus the evidence from SPF density forecasts supports our model implication that disagreement is a good proxy for uncertainty when the variance of aggregate shocks is small, and is consistent with the empirical results presented by Bomberger (1996) and Giordani and Söderlind (2003). As is also clear in Table above, the difference between uncertainty and disagreement, which is an estimate of *ex ante* variance of aggregate shocks, is larger, as forecast horizon gets longer. For example, as the horizon increases from 1 quarter to 4 quarters, the difference increases monotonically from 0.24 to 0.96 in output growth forecast. This pattern is also observed for inflation forecasts during the stable period at all horizons with the exception of 4-quarter ahead forecasts, which means that the additional variability due to the shocks that fell during the first quarter of the current year (on the average during 1984-2007) compared to the remaining quarters is not significant. This is caused by the relatively high disagreement in 4-quarter ahead forecasts during the 1986-1989 period compared to other forecasts. Furthermore, all horizon dummies are estimated to be statistically significant at the 5% level. On balance, the empirical evidence above shows that the variance of aggregate shocks accumulates systematically over horizons, as predicted by our model. This finding is important since most of studies have focused on their relationship over time, without specifying the underlying forecast horizons.

Should squared error of mean forecast be used as a proxy for σ_λ^2 ?

An influential paper in the accounting literature by Barron et al. (1998) suggested “one can infer uncertainty and consensus from observable forecast dispersion, error in the mean forecast and the number of forecasts” (Barron et al. 1998, p. 427). Their suggestion has been extensively used to study the information environment in analysts’ earning forecasts.

Barron et al. (1998) argued that one can use the squared error in the mean forecast as a proxy for $\sigma_{\lambda|th}^2$ to empirically estimate forecast uncertainty as in the following equation:

$$\hat{U}_{th} = (A_t - F_{.th})^2 + (1 - \frac{1}{N})d_{th}.$$

Because forecast errors are known to respondents only after the announcement of actual values, above indeed yields a measure of *ex post* uncertainty. Its reliability as a proxy for *ex ante* uncertainty faced by individual forecasters at the time of forecast is questionable. With density forecasts at our disposal, we can compare them directly. The general message is that, compared to survey measure of uncertainty, *ex post* uncertainty is considerably more volatile. The *ex post* uncertainty overstates the survey measure of uncertainty whenever a forecast is followed by a large unanticipated forecast error.

One may think that it may be an acceptable practice to use mean squared forecast error as a proxy for its *ex ante* counterpart because Barron et al. (1998) are looking at forecast uncertainty retrospectively. Even in this historical context, squared forecast error can give very misleading indication of the uncertainty environment in real time in a past sample, as shown by the extra variability in *ex post* uncertainty during periods that are characterized by large *ex post* forecast errors.

Construction of an ex ante measure of uncertainty

Because uncertainty is essentially an *ex ante* concept attached to a forecast before the actual outcome is known, it must be constructed using data available in real time. To form a measure of forecast uncertainty, we should use the observed disagreement from the survey, d_{it} and the variance of aggregate shocks generated conditionally by GARCH-type models, $\hat{\sigma}_{\lambda|it}^2$ (cf. Engle, 1982; Bollerslev, 1986) to estimate U_{it} :

$$\hat{U}_{it} = \hat{\sigma}_{\lambda|it}^2 + d_{it}.$$

In order to generate GARCH-type estimates of the variability of aggregate shocks, we first filter the mean forecast errors for possible autocorrelation. The order of autocorrelation present in a given mean forecast error series is found by fitting moving average models of varying order, the preferred model being chosen by the use of Schwarz information criterion. We then estimated $\sigma_{\lambda|t}^2$ using different GARCH-type models with various distributional assumptions on the filtered mean forecast errors. For convenience, these models are labeled as Model 1 through Model 8. In Model 1, we estimated the standard GARCH (1, 1) model with the following specification:

$$e_t \sim N(0, \sigma_{\lambda|t}^2), \quad \sigma_{\lambda|t}^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 \sigma_{\lambda|t-1}^2,$$

where e_t is the serially uncorrelated mean forecast error. This equation has been estimated using the quasi-maximum likelihood for the 1984-2007 subsample and for each horizon. Consistent with many earlier studies, in Model 2 we estimated the equation using the t -distribution with 6 degrees of freedom.

As an alternative specification, we replaced the lagged mean squared forecast error with the average of mean squared errors over the last ten years. In Model 3, we estimated $\sigma_{\lambda|t}^2$ using the following model specification:

$$e_t \sim N(0, \sigma_{\lambda|t}^2), \quad \sigma_{\lambda|t}^2 = \beta_0 + \beta_1 \left(\sum_{s=1}^{10} MSE_{t-s} / 10 \right) + \beta_2 \sigma_{\lambda|t-1}^2.$$

Model 4 estimated this equation using the t -distribution with 6 degrees of freedom.

Models 5 through 8 correspond to Models 1 through 4, except that we modeled the standard deviation instead of the variance in the GARCH-type models.

Forecast uncertainty should be generated by the sum of the estimated variance of aggregate shocks $\hat{\sigma}_{\lambda|th}^2$ from GARCH-type models and the disagreement from the survey. Table below shows the correlations between survey and other measures of uncertainty. Several points stand out. First, the GARCH estimates of uncertainty with the average squared errors over the last ten years (in place of the last period forecast error) help to capture the variation in the survey measure of uncertainty (Models 3, 4, 7 and 8) fairly well. Compared to the simple correlation with the disagreement alone (the first row), the correlations between the survey uncertainty and the uncertainty generated by Models 3, 4, 7 and 8 increase by about 5% for 1- and 2-quarter ahead inflation forecasts, and by more than 15% and 10% for 3- and 4-quarter ahead GDP forecasts, respectively. Second, models with t -distributions (Models 2, 4, 6 and 8) match survey measure of uncertainty better. In general, Models 2, 4, 6 and 8 using t -distribution with 6 degree of freedom perform better to capture the variation in survey uncertainty than Models 1, 3, 5 and 7 using normal distribution. Third, modeling the standard deviation instead of the variance tends to do a better job in representing the variation in

survey measure of uncertainty. For output growth forecasts, the best model seems to be Model 8 that performs even better at longer horizons. For inflation forecasts, the best model is Model 8 at shorter horizons and Model 6 at longer horizons. In addition, when we add squared errors to disagreement (Model 0), its predictive power to proxy survey uncertainty decreases across almost all horizons for both variables!

In summary, the GARCH-type models are very successful in modeling the variability of future aggregate shocks to the economy in the sense that when added to disagreement, this composite measure of *ex ante* forecast uncertainty explains the corresponding survey measure better than disagreement alone.

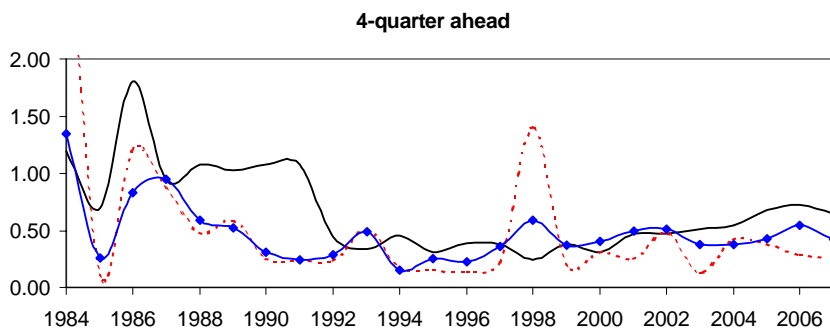
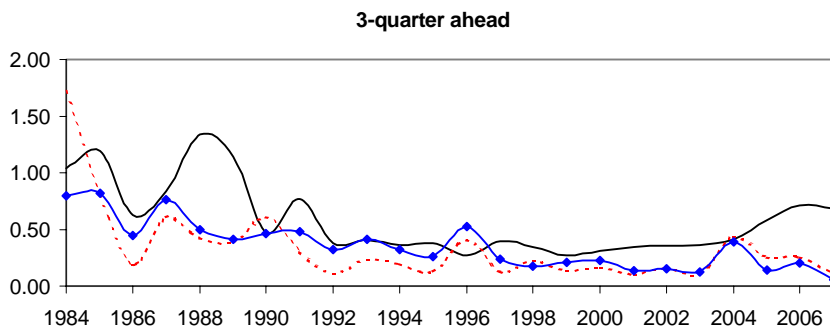
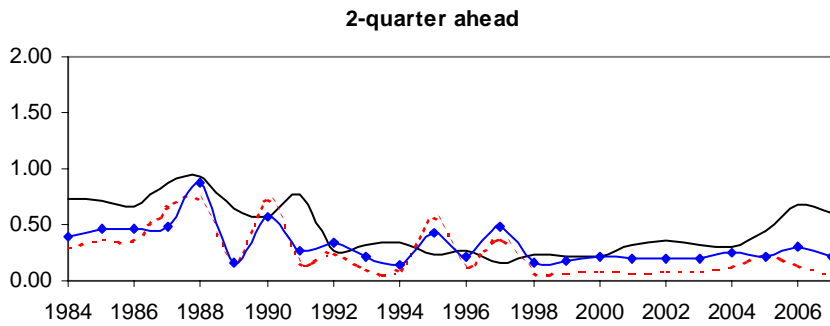
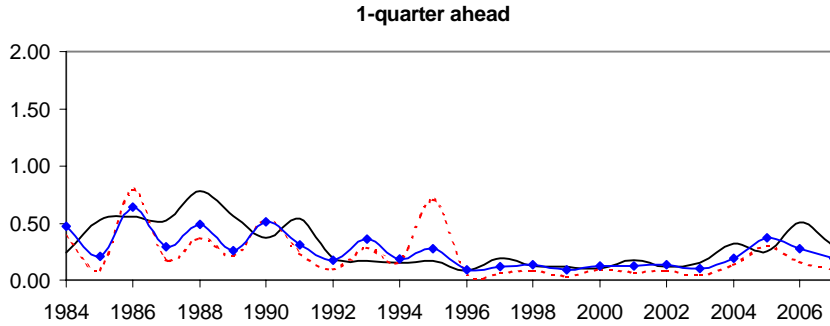
We plot the evolution of uncertainty generated from the best models in inflation and output growth forecasts over time. Compared to the uncertainty constructed using the squared error in the mean forecast, the uncertainty from GARCH-type models is less volatile and thus matches better the survey measure of uncertainty. This underscores the important point that *ex ante* uncertainty has to be generated conditionally based on the information known to survey respondents when making their forecasts, which is exactly what GARCH-type models do. We should, however, note that the error-based measures of uncertainty including the GARCH have failed to signal the slowly creeping uncertainty in inflation and output growth forecasts since 2002 as indicated by the density forecasts. This is because the corresponding forecast errors have continued to be small despite the slow but steady increase in uncertainty due to unusual financial market developments and political instability in recent years. Uncertainty estimates based on density forecasts have an obvious advantage in this regard.

Measures of uncertainty in inflation forecasts:

Survey measure of uncertainty (solid line)

Uncertainty using squared error of mean forecast (dotted line)

Uncertainty from GARCH-type model (line with diamond)

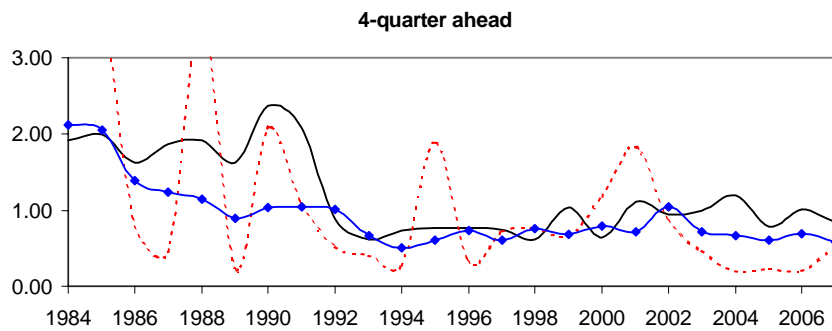
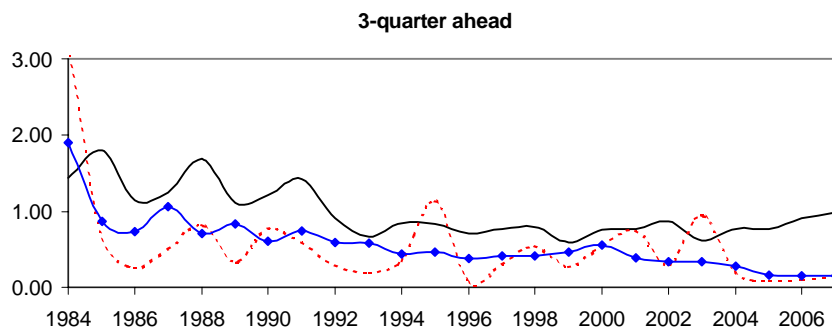
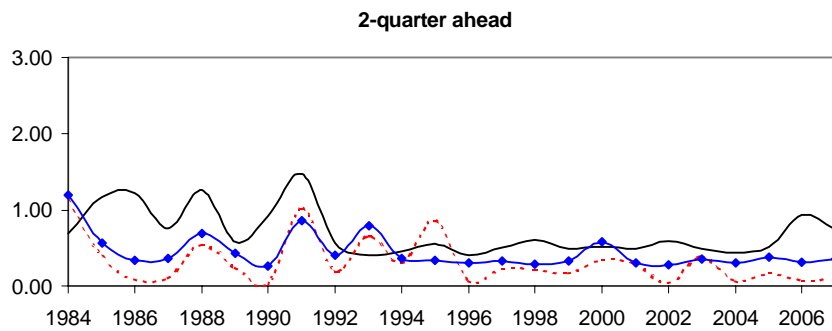
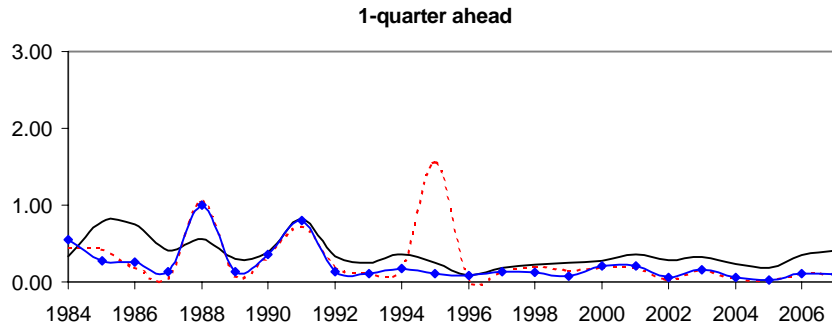


Measures of uncertainty in output growth forecasts:

Survey measure of uncertainty (solid line)

Uncertainty using squared error of mean forecast (dotted line)

Uncertainty from GARCH-type model (line with diamond)



Correlation between survey uncertainty and alternative measures of uncertainty

	SPF inflation forecast (1984-2007)				SPF GDP forecast (1984-2007)			
	1Q ahead	2Q ahead	3Q ahead	4Q ahead	1Q ahead	2Q ahead	3Q ahead	4Q ahead
Disagreement	0.56	0.52	0.55	0.60	0.60	0.30	0.44	0.58
Model 0	0.36	0.49	0.56	0.52	0.32	0.24	0.39	0.57
Model 1	0.56	0.51	0.67	0.44	0.63	0.33	0.46	0.42
Model 2	0.59	0.51	0.67	0.53	0.62	0.32	0.42	0.47
Model 3	0.62	0.54	0.64	0.51	0.56	0.24	0.60	0.70
Model 4	0.61	0.54	0.61	0.53	0.61	0.31	0.62	0.71
Model 5	0.57	0.52	0.66	0.49	0.58	0.20	0.37	0.50
Model 6	0.57	0.53	0.67	0.54	0.63	0.34	0.39	0.33
Model 7	0.62	0.56	0.61	0.52	0.58	0.37	0.62	0.69
Model 8	0.63	0.56	0.61	0.51	0.61	0.34	0.61	0.71

Note: This table presents the correlations between survey and alternative measures of uncertainty. Alternative measures of uncertainty are generated by the sum of the variance of aggregate shocks from Models 0 to 8 and the disagreement from the survey. In particular, in Model 0, the squared error in the mean forecasts is used as a proxy for the variance of aggregate shocks. In Models 1 through 8, the variance of aggregate shocks is generated from the following models:

Model 1: GARCH (1, 1) with normal distribution;

Model 2: GARCH (1, 1) with t-distribution (6 degree of freedom);

Model 3: GARCH (0, 1) with the average of mean squared errors (MSE) over the last 10 years and normal distribution;

Model 4: GARCH (0, 1) with the average of mean squared errors (MSE) over the last 10 years and t-distribution (6 degree of freedom);

Model 5: Power GARCH (1, 1) with normal distribution;

Model 6: Power GARCH (1, 1) with t-distribution (6 degree of freedom);

Model 7: Power GARCH (0, 1) with the average of root mean squared errors (RMSE) over the last 10 years and normal distribution;

Model 8: Power GARCH (0, 1) with the average of root mean squared errors (RMSE) over the last 10 years and t-distribution (6 degree of freedom).

Interestingly, Reifschneider and Tulip (2007) have recently suggested a similar measure of past forecast uncertainty using squared individual forecast errors of a number of private and government forecasters averaged over 1986-2006. The purpose is to use this average historical uncertainty based on past predictive accuracy as a benchmark against which FOMC participants can assess their present uncertainty. In order to generate this benchmark for a “typical” uncertainty to be associated with the individual forecasts, they first calculate the individual root mean squared error (RMSE) over the period and then average across forecasters of the individual RMSEs to obtain:

$$RMSE_1 = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{T} \sum_{t=1}^T (A_t - F_{it})^2} .$$

Note that the above measure is different from the one suggested by our analysis. Instead, we suggest using

$$RMSE_2 = \sqrt{\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (A_t - F_{it})^2}$$

to estimate the typical uncertainty of a randomly drawn forecaster from the sample. It is clear that the Reifschneider-Tulip measure, will have the disagreement and the squared consensus forecast error as components of uncertainty. Also, because of the averaging of squared consensus forecast errors over the last twenty years, it may not be very sensitive to occasional large forecast errors, and thus, may be a reasonable approximation for the average variance of *ex ante* aggregate shocks over the period.

However, according to Jensen’s inequality, we can easily show that in general $RMSE_1 \leq RMSE_2$, the latter having been justified in our previous analysis as the appropriate measure of benchmark *ex post* uncertainty. In order to gauge the extent of underestimation in our sample, we calculated the two measures using our data during 1986-2006. We found that the Reifschneider-Tulip measure underestimates the benchmark uncertainty, by 4% to 8% for inflation forecasts. The degree of underestimation is even more pronounced for GDP forecasts ranging from 8% to 13%. Also, we find that this benchmark measure of historical uncertainty based on average *ex post* predictive accuracy can be sensitive to occasional large forecast errors. For instance, the one-quarter ahead GDP forecast for 1995 is associated with an unusually large error due to the sudden slowdown of the U.S. economy. If we take out this large forecast error from our calculations for GDP forecasts, the R-T measures decrease from 0.47 to 0.40 and from 0.54 to 0.47, respectively.

Measures of uncertainty based on forecast errors averaged over 1986-2006

	SPF inflation forecast				SPF GDP forecast			
	1Q ahead	2Q ahead	3Q ahead	4Q ahead	1Q ahead	2Q ahead	3Q ahead	4Q ahead
RMSE ₁	0.49	0.52	0.56	0.64	0.47	0.52	0.62	0.97
RMSE ₂	0.51	0.57	0.60	0.67	0.54	0.59	0.70	1.06

Note: $RMSE_1 = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{T} \sum_{t=1}^T (A_t - F_{it})^2}$ and $RMSE_2 = \sqrt{\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (A_t - F_{it})^2}$.