Ifo Institute for Economic Research at the University of Munich

ifo Working Papers

Measuring Forecast Uncertainty by Disagreement: The Missing Link

Kajal Lahiri
Xuguang Sheng

Ifo Working Paper No. 60

August 2008

Measuring Forecast Uncertainty by Disagreement: The Missing Link*

Abstract

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, we find direct evidence in support of our hypothesis. Our results support the use of GARCH-type models, rather than the ex post squared error in consensus forecasts, to estimate the ex ante variability of aggregate shocks as a component of aggregate uncertainty.

JEL Code: E17, E37.
Keywords: Aggregate shocks, public information, forecast disagreement, forecast horizon, forecast uncertainty, panel data, private information.

Kajal Lahiri**
Department of Economics
University at Albany
SUNY
Albany, NY 12222
USA
Phone: +1(0)518/442-4758
klahiri@albany.edu

Xuguang Sheng
Department of Economics
University at Albany
SUNY
Albany, NY 12222
USA
Phone: +1(0)518/442-4830
xs7410@albany.edu

* We thank Rob Engle and Victor Zarnowitz for many helpful discussions and encouragement. However, we alone are responsible for any remaining errors and shortcomings.

** Corresponding Author.

1. INTRODUCTION

Forecast uncertainty is playing an increasingly important role in macroeconomics and monetary policy making. For instance, effective November 2008, the U.S. Federal Open Market Committee (FOMC) will publish information about uncertainty associated with their economic outlooks. Since the mid-90s the Bank of England has been reporting fan charts that show subjective confidence bands surrounding official forecasts. Since forecast uncertainty is unobservable, economists have experimented with alternative proxies for it. One of the more popular measures has been forecast disagreement, simply calculated as the dispersion in alternative point forecasts. When disagreement is taken to indicate uncertainty, the underlying assumption is that this inter-personal dispersion measure is an acceptable proxy for the average dispersion of intra-personal predictive probabilities held by individual experts. The validity of this assumption can by no means be taken for granted. Since the seminal work of Zarnowitz and Lambros (1987), economists have studied but disagreed on whether disagreement is a good proxy for uncertainty.[1] As pointed out by Bomberger (1996) and Giordani and Söderlind (2003), disagreement remains to be theoretically an unfounded measure of uncertainty. 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.[2]

In this paper, we establish a simple relationship connecting forecast uncertainty to disagreement. Using a standard decomposition of forecast errors into common and

[1] See, for instance, Bomberger (1996, 1999), Rich and Butler (1998), Giordani and Söderlind (2003), Lahiri and Liu (2005), and Boreo, Smith and Wallis (2007).

[2] See Zhang (2006) and references therein. Barry and Jennings (1992), Abarbanell et al. (1995), Barron et al. (1998), Diether et al. (2002) and Johnson (2004) have argued that disagreement alone is not sufficient to approximate uncertainty.

idiosyncratic components, we show that forecast uncertainty equals disagreement plus the variance of future aggregate shocks that accumulate over the horizons. This finding has important implications for the empirical studies using disagreement as a proxy for uncertainty. It suggests that the robustness of the proxy depends on the variance of aggregate shocks over time and across horizons. It also simplifies the multi-dimensional covariance matrix of individual forecast errors in Barry and Jennings (1992) in terms of the variance of aggregate shocks, which can be easily interpreted as the uncertainty shared by all forecasters due to their exposure to future common shocks.

Using a panel of density forecasts from Survey of Professional Forecasters over 1969-2007, we find direct evidence in support of our hypothesized time and horizon effects. As for the time effect, disagreement is found to be a reliable measure for uncertainty in a stable period. In periods with large volatility of aggregate shocks, however, disagreement becomes less reliable a proxy. As for the horizon effect, we find that the longer the forecast horizon, the larger is the difference between disagreement and uncertainty.

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.

The reminder of the paper is organized as follows. In section 2, we develop the theoretical model and derive the relationship between disagreement and uncertainty. Section 3 tests empirically whether disagreement is a reliable proxy for uncertainty and suggests a method to construct the *ex ante* measure of uncertainty. Section 4 concludes.

2. 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

$$e_{ith} = A_t - F_{ith}\,. \tag{1}$$

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}):

$$e_{ith} = \lambda_{th} + \varepsilon_{ith}\,, \tag{2}$$

$$\lambda_{th} = \sum_{j=1}^{h} u_{tj}\,. \tag{3}$$

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 (3) 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:

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

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

Assumption 2:

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

Assumption 3:

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

Thus, aggregate shocks are assumed to be uncorrelated over time and horizons (assumption 1). The idiosyncratic errors (ε_{ith}) capture forecaster heterogeneity due to differences in information acquisition and processing, interpretation, judgment, forecasting models, etc., and are taken to be mutually uncorrelated at all leads and lags (assumption 2). In addition, the common component and idiosyncratic disturbances are assumed to be uncorrelated at all leads and lags (assumption 3), which is a standard assumption in the literature. Taken together, assumptions 1 to 3 imply that the individual forecast error is a zero-mean stationary process for any *h* and has the factor model interpretation.

The observed disagreement (d_{th}) among forecasters is the variance of their point forecasts which, given (1) and (2), can be expressed as:

$$d_{th} \equiv \frac{1}{N-1}\sum_{i=1}^{N}(F_{ith} - F_{.th})^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\varepsilon_{ith} - \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{ith})^2, \qquad (4)$$

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}:

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

$$= \frac{1}{N-1}\sum_{i=1}^{N}(\sigma^2_{\varepsilon|ith} + \frac{1}{N^2}\sum_{i=1}^{N}\sigma^2_{\varepsilon|ith} - \frac{2}{N}\sigma^2_{\varepsilon|ith})$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sigma^2_{\varepsilon|ith}. \tag{5}$$

Thus, not surprisingly, we find that D_{th} is determined by the average variance of idiosyncratic errors.[3]

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

$$U_{ith} \equiv Var(A_t - F_{ith}) = Var(\lambda_{th} + \varepsilon_{ith}) = \sigma^2_{\lambda|th} + \sigma^2_{\varepsilon|ith}. \tag{6}$$

Individual forecast uncertainty in (6) is comprised of two components: perceived uncertainty associated with forthcoming common shocks, $\sigma^2_{\lambda|th}$ and idiosyncratic shocks, $\sigma^2_{\varepsilon|ith}$. 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.[4] Given our model, U_{th} can be expressed as a function of the model parameters as:

$$U_{th} = \sigma^2_{\lambda|th} + \frac{1}{N}\sum_{i=1}^{N}\sigma^2_{\varepsilon|ith}. \tag{7}$$

After substituting (5) into (7), we get

$$U_{th} = \sigma^2_{\lambda|th} + D_{th}. \tag{8}$$

[3] The number of forecasters in the survey changes over both *t* and *h*. For simplicity, we suppress the subscripts *t* and *h* of *N* in equation (4) and thereafter.
[4] See also Lahiri et al. (1988), Bomberger (1996), Giordani and Söderlind (2003), and Boero et al. (2007).

Given the model assumptions, forecast uncertainty, disagreement and the variance of forthcoming aggregate shocks are expected to be related in the sample as in (8) – 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. In periods where the volatility of aggregate shocks is high, disagreement can become a tenuous proxy for uncertainty.

In the context of equation (8), one can understand the efforts of Bomberger (1996) who 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. Our model assumptions, though admittedly simple, rule out any feedback from perceived future variability of common shocks to current idiosyncratic individual variances. However, it is possible that enhanced future uncertainty about common shocks affects current individual $\sigma^2_{\varepsilon|ith}$ and co-vary with $\sigma^2_{\lambda|th}$. This is how Bomberger's (1996) econometric exercise can be justified. On the other hand, as Zarnowitz and Lambros (1987) have pointed out, there may be periods where all forecasters agree on relatively high macroeconomic uncertainty in the immediate future, and hence disagreement between forecasters will be low even though uncertainty is high. The opposite is also possible where forecasters disagree a lot about their mean forecasts, but they are confident about their individual predictions. This situation will arise when forecasters disagree on otherwise precise models and scenarios that should be used to depict the movement of the economy over the forecasting horizon. Thus, lacking any theoretical basis, the strength and the stability of the relationship between disagreement and overall forecast uncertainty (not merely $\sigma^2_{\lambda|th}$ or consensus uncertainty) becomes an empirical issue. But our result clearly suggests that the relationship will crucially depend on the sample period, the target variable, and length of the forecast horizon. Our analysis also helps to reconcile the divergent findings in previous empirical studies examining the appropriateness of disagreement as a proxy for forecast uncertainty. 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

$$l_{th} = A_t + \eta_{th},\ \eta_{th} \sim N(0, 1/\sigma^2_{\eta|th}), \tag{9}$$

and one private signal, represented by

$$s_{ith} = A_t + \zeta_{ith},\ \zeta_{ith} \sim N(0, 1/\sigma^2_{\zeta|ith}). \tag{10}$$

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

$$F_{ith} \equiv E(A_t | l_{th}, s_{ith}) = (\sigma^2_{\eta|th} l_{th} + \sigma^2_{\zeta|ith} s_{ith}) \big/ (\sigma^2_{\eta|th} + \sigma^2_{\zeta|ith}), \tag{11}$$ [5]

and the conditional variance of A_t as

$$U_{ith} \equiv Var(A_t | l_{th}, s_{ith}) = 1 \big/ (\sigma^2_{\eta|th} + \sigma^2_{\zeta|ith}). \tag{12}$$

The individual forecast uncertainty defined in (12) reflects the uncertainty in both the public and private information, which is similar to (6) 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_{th}) as the average of the individual uncertainties $U_{th} \equiv \frac{1}{N}\sum_{i=1}^{N} U_{ith}$. Given the Bayesian learning model, U_{th} can be expressed as:

$$U_{th} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(\sigma^2_{\eta|th} + \sigma^2_{\zeta|ith})}. \tag{13}$$

[5] Under the assumption that $\sigma^2_{\zeta|ith} = \sigma^2_{\zeta|th}$ for all *i*, the individual forecast error can be written as $e_{ith} = [-\sigma^2_{\eta|th} \big/ (\sigma^2_{\eta|th} + \sigma^2_{\zeta|th})]\eta_{th} + [-\sigma^2_{\zeta|th} \big/ (\sigma^2_{\eta|th} + \sigma^2_{\zeta|th})]\zeta_{ith}$, where the first and second term on the right-hand side correspond to λ_{th} and ε_{ith} in (2), respectively.

Note that overall forecast uncertainty in (13), derived in the context of Bayesian learning framework, provides the justification that the aggregate uncertainty can be defined as the simple average of individual uncertainties as in (7). It is a combined uncertainty in the context of forecast combination literature with equal weights.[6]

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 $\sigma^2_{\eta|th}$ and $\sigma^2_{\zeta|ith}$, the pre-posterior variance is the variance based on the distribution of the signals l_{th} and s_{ith} for $i = 1,2,...,N$. The disagreement among forecasters can then be measured as

$$D_{th} \equiv E\left[\frac{1}{N-1}\sum_{i=1}^{N}(F_{ith} - \frac{1}{N}\sum_{i=1}^{N}F_{ith})^2\right]$$

$$= \frac{1}{N}E\left[\sum_{i=1}^{N}F_{ith}^2\right] - \frac{1}{N(N-1)}E\left[\sum_{i=1}^{N}\sum_{j\neq i}^{N}F_{ith}F_{jth}\right]. \qquad (14)$$

After substituting for F_{ith} from (11), we get

$$D_{th} = \left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{(\sigma^2_{\eta|th} + \sigma^2_{\zeta|ith})}\right] - \left[\frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{\sigma^2_{\eta|th}}{(\sigma^2_{\eta|th} + \sigma^2_{\zeta|ith})(\sigma^2_{\eta|th} + \sigma^2_{\zeta|jth})}\right]. \qquad (15)$$

Note that the first term on the right-hand side of (15) is forecast uncertainty, U_{th} and the second term is the average covariance among forecast errors, C_{th}, where

$$C_{th} = \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}^{N}Cov(A_t - F_{ith}, A_t - F_{jth}). \qquad (16)$$

[6] Our measure of uncertainty is different from the "combined uncertainty" as defined by the variance of aggregate density forecast in Wallis (2005), which includes both our measure of uncertainty and the disagreement as its components.

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 in (16) as

$$C_{th} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i}^{N} E[(\lambda_{th} + \varepsilon_{ith})(\lambda_{th} + \varepsilon_{jth})] = \sigma^2_{\lambda|th}, \tag{17}$$

which can be easily interpreted as the uncertainty shared by all forecasters due to their exposure to common shocks. Thus, (17) greatly simplifies the results in Barry and Jennings (1992) and Barron et al. (1998), and gives the relationship (8).

3. EMPIRICAL TEST OF THE RELATIONSHIP BETWEEN UNCERTAINTY AND DISAGREEMENT

This section begins with a short description of data on density forecasts used in this study. In subsequent sections, we present empirical evidence in support of our hypothesized relationship between disagreement and uncertainty over time and horizons. We then evaluate the appropriateness of using squared error of mean forecasts as a proxy for the variance of aggregate shocks that has been extensively used in recent accounting literature. Our suggestion to construct a robust measure of *ex ante* uncertainty given a panel of forecasts is presented at the end.

3.1 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. The historical time series of forecasts in this survey is quite lengthy (since the fourth quarter of 1968), and there are a number of changes in the surveys that make the data challenging to work with. 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 3½, 2½, 1½, and ½ 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.[7] 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.[8]

To test the hypothesized relationships, we also need the actual values of inflation and output growth. As is well known, the NIPA data often go through serious revisions. Obviously, the most recent revision is not a good choice, since it involves adjustment of definitions and classifications. Consistent with the findings in Harvey and Newbold (2003) that the unrevised data approximates the forecasters' objective better, we choose the first release of annual inflation and output growth to compute the actual values. These are the real-time data available from the Federal Reserve Bank of Philadelphia.[9]

[7] The Philadelphia Fed is uncertain about the target years referred to in the surveys made in the first quarter of 1985 and 1986. We deleted those forecasters who were obviously misled by the wrong wording of the question and used the rest of the responses.

[8] See Giordani and Söderlind (2003) and Lahiri and Liu (2005) for a detailed discussion on the specification and construction of the analytical sample, and hence not repeated here.

[9] All calculations reported in this paper were also repeated with the so called "first final" (i.e., the third monthly revision) and July revisions. The main results and conclusions were unchanged.

3.2 Test of the relationship between uncertainty and disagreement

Note that the variance of forecast error in (6) can be interpreted as the variance of random variable A_t as perceived by individual *i*, given information available at time $t-h$, which is conceptually the same as the variance of the density forecast reported by individual *i*. Taking the average of the variances of individual densities yields estimates of forecast uncertainty as defined in (7).

To get appropriate measures of forecast disagreement, we need to control for any possible individual bias in the sample. Following Davies and Lahiri (1995), the individual forecast error has a 3-dimensional nested structure in the presence of individual bias ϕ_{ih}:

$$e_{ith} \equiv A_t - F_{ith} = \phi_{ih} + \lambda_{th} + \varepsilon_{ith}. \tag{18}$$

The systematic individual bias, $\hat{\phi}_{ih}$, can be estimated as

$$\hat{\phi}_{ih} = \frac{1}{T}\sum_{t=1}^{T}(A_t - F_{ith}). \tag{19}$$

By adding these individual biases to the forecasts, we get unbiased forecasts and forecast disagreement.[10]

We should note that equation (8) 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_{th}) at time $t-h$, but before the actual value A_t was realized.

[10] Since they were estimated to be relatively small, individual biases did not affect forecast disagreement by any significant amount.

Following Engle (1983), note that one can always decompose the average squared individual forecast errors as:

$$\frac{1}{N}\sum_{i=1}^{N}(A_t - F_{ith})^2 = (A_t - F_{.th})^2 + (1-\frac{1}{N})d_{th}\,. \tag{20}$$

Taking expectations on both sides given all available information at time *t* including F_{ith} and d_{th}, we get the following conditional relationship between aggregate uncertainty, the variance of consensus forecast errors and observed disagreement:

$$U_{th} = E(A_t - F_{.th})^2 + d_{th}. \tag{21}$$

Now focusing on the first term on the right-hand side of (21), it can be alternatively written as

$$E(A_t - F_{.th})^2 = \frac{1}{N^2}E\left[\sum_{i=1}^{N}(A_t - F_{ith})^2\right] + \frac{1}{N^2}E\left[\sum_{i=1}^{N}\sum_{j\neq i}^{N}(A_t - F_{ith})(A_t - F_{jth})\right]. \tag{22}$$

In the context of forecast combination, Batchelor and Dua (1995) had a similar decomposition. Given our framework, (22) can be expressed as

$$E(A_t - F_{.th})^2 = \sigma^2_{\lambda|th} + \frac{1}{N^2}\sum_{i=1}^{N}\sigma^2_{\varepsilon|ith}\,. \tag{23}$$

We should point out that the uncertainty about the consensus forecast in (23) defined by Bomberger (1996) is different from our measure of forecast uncertainty in (7). The uncertainty about the consensus forecast is less than the average of the individual uncertainties due to the fact that combining individual forecasts implicitly pools the diverse idiosyncratic errors. Note that, as the number of forecasters goes to infinity, the uncertainty about the consensus forecast will reflect only the uncertainty in the common information.

Substituting (23) in (21), we obtain

$$U_{th} = \sigma^2_{\lambda|th} + \frac{1}{N^2}\sum_{i=1}^{N}\sigma^2_{\varepsilon|ith} + d_{th}. \tag{24}$$

For typical values of N and $\sigma^2_{\varepsilon|ith}$ in our context, the second term on the right-hand side of (24) will be very close to zero and can be ignored.[11] 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, disagreement and their difference, which is an estimate of the variance of *ex ante* aggregate shock, are plotted in Figures 1 to 4. Their average values are given in Table 1. Several points are worth noting. 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. Note also that the estimated variances of aggregate shocks are systematically much bigger for GDP growth than inflation at all horizons. This finding implies that it is more difficult to forecast real GDP growth than inflation, and is consistent with most recent studies on forecast evaluation that report significantly higher RMSE for real GDP than for inflation forecasts.[12]

[11] In our sample, the average values of $\frac{1}{N^2}\sum_{i=1}^{N}\sigma^2_{\varepsilon|ith}$ lie between 0.01 and 0.02 for both inflation and output growth forecasts.

[12] See, for instance, Öller and Barot (2000), Banerjee and Marcellino (2006), and Reifschneider and Tulip (2007).

Second, somewhat unexpectedly, in some quarters disagreement exceeds uncertainty, especially for inflation. Certainly, one reason is the imprecision in the estimation of uncertainty and disagreement based on a finite sample of survey respondents. After all, relationships (8) and (24) are expected to hold only on the average. However, there are other possibilities that should be pointed out. It could be that survey measure of uncertainty does not represent the "true" or objective uncertainty correctly. Diebold et al. (1999) concluded that survey uncertainty overestimates the true values. However, Giordani and Söderlind (2003) reached an opposite conclusion. Following the latter approach, in Table 2 we report the average percentage times the 90% predictive interval covers the actual outcomes after fitting a uniform distribution over the bins during 1969-2007. We find that survey measures of uncertainty are well calibrated for all horizons except 4-quarter ahead forecasts. For the 4-quarter ahead forecasts, the survey measure underestimates the objective uncertainty by 13% for inflation and 17% for output growth forecasts. This possible underestimation of the true uncertainty by survey densities can rule out a few of the negative estimates of the variance of aggregate shocks.[13] Also, if we believe that, for a particular horizon, the extent of under or over-estimation is time invariant, the survey uncertainty will continue to be a meaningful indicator for true forecast uncertainty. Even if adjusted for the degree of underestimation by 13%, the uncertainty is still far less than disagreement at 4-quarter ahead inflation forecast for 1980. This can be a sign of the occurrence of structural break. Recall that for our decomposition of forecast errors into common and idiosyncratic components, the

[13] Following Giordani and Söderlind (2003), we also fitted normal distributions over histograms and repeated the same comparison exercise. As expected, the normal approximation suggested even more underestimation. Many recent studies have, however, avoided the practice of fitting normal distribution to the individual density forecasts because the majority of the respondents seldom assign probabilities to more than 3 intervals, see Engelberg et al. (2006).

individual forecast errors were assumed to be a stationary process. As is well known, inflation rose sharply and unexpectedly during 1979-1981, and is characterized by a structural break in the inflation process. Thus, the stationary assumption is violated and accordingly (24) may not hold during periods of structural breaks. If the economic system is temporarily non-stationary due to structural breaks and regime change, there will typically be many different beliefs about the future course of the economy. This leads to forecasters adopting disparate forecasting functions and as a result, their predictions will generate extraordinary disagreement. Uncertainty, on the other hand, is seen to be very sticky in terms of its high autocorrelation and low volatility and as a result, responds slowly to even rapid changes in the economic environment.[14] Thus, whereas the relatively large negative variance of aggregate shocks may suggest periods of structural breaks and regime change, the smaller ones can be attributed to imprecision in small sample estimation.

Third, Figures 1-4 suggest that the volatility of aggregate shocks declined sharply after 1991 for both inflation and output growth. This finding contributes to our understanding of the factors behind *Great Moderation* - the well-documented decline in macroeconomic volatility in the United States since 1984. Our result suggests that the decline in macroeconomic volatility during 1984-1991 can not be attributed to "good luck", since the economy was hit by unforeseen large shocks during this period (cf. Campbell, 2007), and instead must be explained by other factors, such as structural changes (cf. McConnell and Perez-Quiros, 2000) or improved monetary policy (cf.

[14] This is also true for time series measures of uncertainty. Giordani and Söderlind (2003) and Lahiri and Liu (2005) show that the GARCH measure of uncertainty fails to capture the increase in inflation uncertainty around the second oil price shock.

Mishkin, 2007). After 1991, the shocks hitting the economy became smaller and more stable, and thus played a large role in the reduction of macroeconomic volatility.

Now we can test formally the implications of (24) that the relationship between uncertainty and disagreement depends on the variance of aggregate shocks over time and across horizons. By plotting actual inflation rate, we find its average value during 1969-1983 to be at least 2.5 times than that during 1984-2007, consistent with the stylized fact documented in the literature, cf. Stock and Watson (2007). As is well known, higher rates of inflation are generally associated with higher variability of inflation and presumably greater uncertainty about future rates. 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:

$$U_{th} = \beta D_{th} + \rho_1 H_1 + \rho_2 H_2 + \rho_3 H_3 + \rho_4 H_4 + \varepsilon_{th}, \tag{25}$$

where $H_i = 1$ if the forecast is made at horizon *i* for $i = 1,2,3,4,$ and 0 otherwise.

Table 3 shows the estimation results. 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 3, 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 (see Figure 1). 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.[15]

3.3 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) extended the model in Barry and Jennings (1992) and 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. Yet, without direct information on uncertainty, the validity of their suggestion in finite samples can never be established. Our analysis below addresses this question.

[15] Two exceptions are the recent papers by Lahiri and Sheng (2007) and Patton and Timmermann (2007), who explicitly modeled the evolution of survey forecasts over horizons.

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

$$\hat{U}_{th} = (A_t - F_{.th})^2 + (1 - \frac{1}{N})d_{th}. \qquad (26)$$

Because forecast errors are known to respondents only after the announcement of actual values, (26) 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. Figures 5 and 6 plot these two measures of uncertainty in inflation and output growth forecasts during 1984-2007. The general message is that, compared to survey measure of uncertainty, *ex post* uncertainty from (26) 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. This is unfortunate because, being unanticipated, these errors should not have affected the forecast uncertainty that predates the observed forecast error. The regression results in Table 4 reinforce some of the features from these graphs. For inflation forecasts, the estimated coefficient of *ex post* uncertainty is almost zero during the unstable period 1969-83. Even in the stable period, the coefficients are estimated to be very small for both inflation and output growth forecasts. Comparing $\overline{R}^2$ in Tables 3 and 4, we see that disagreement alone is a reasonable proxy for uncertainty. However, adding the squared error in the consensus forecast to disagreement turns out to be a significantly worse proxy for uncertainty than the disagreement alone. $\overline{R}^2$ falls from 0.34 to 0.09 during 1969-83 and from 0.39 to 0.30 during 1984-2007 for inflation, implying that the squared forecast errors contribute negatively to explaining survey uncertainty. For real GDP, the squared forecast errors have practically no additional explanatory power, as $\overline{R}^2$ increases from 0.53 to 0.54.

To understand this puzzle, note the decomposition in (20). Comparing (20) with (26), it immediately follows that *ex post* uncertainty is nothing but the average squared individual forecast errors.[16] Clearly, forecast uncertainty constructed according to Barron et al. (1998) depends on the realization of individual forecast errors. But forecast error is necessarily an *ex post* quantity, which reflects unexpected shocks after the forecast is made, and thus should not affect uncertainty at the time a forecast is issued. 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. Their measure has been used to study the impact of special events, such as Regulation Fair Disclosure, on the forecasting environment of financial analysts, see, for example, Mohanram and Sunder (2006) and references therein. 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 (see Figures 5 and 6).

Engle (1983) demonstrated that the average squared individual forecast errors do not show patterns similar to ARCH measures of uncertainty.[17] Our findings here, together

[16] During our sample period, the squared error of the mean forecast accounts for 40% to 70% of *ex post* uncertainty in output growth forecast and from 30% to 60% in inflation forecast, as the horizon gets longer from 1- to 4-quarter ahead. The remainder is attributable to disagreement.
[17] As shown in Table 2 of Engle (1983), the average squared individual forecast errors are 31.78 (1947/12-1952/6), 1.35 (1962/6-1966/12) and 13.01 (1971/6-1975/12), but the corresponding ARCH uncertainty is 19.22, 2.57 and 3.37, respectively.

with the empirical evidence presented in Engle (1983), strongly caution against using the squared error in the mean forecast as a component of overall forecast uncertainty. We show that forecast disagreement by itself, without the *ex post* mean squared error, correlates better with the observed survey uncertainty.

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:

$$RMSE_1 = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{1}{T}\sum_{t=1}^{T}(A_t - F_{it})^2}\ . \tag{27}$$

Note that the above measure is different from the one suggested by our analysis. Instead, according to (20), one should use

$$RMSE_2 = \sqrt{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(A_t - F_{it})^2} \tag{28}$$

to estimate the typical uncertainty of a randomly drawn forecaster from the sample. It is clear that the Reifschneider-Tulip measure (27), like (28), 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, (27) 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. As can be seen in Table 5, the Reifschneider-Tulip measure underestimates the benchmark uncertainty (28), 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 measures based on (27) and (28) decrease from 0.47 to 0.40 and from 0.54 to 0.47, respectively (cf. Table 5).

3.4 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_{th} and the variance of aggregate shocks generated conditionally by GARCH-type models, $\hat{\sigma}^2_{\lambda|th}$ (cf. Engle, 1982; Bollerslev, 1986) to estimate U_{th}:

$$\hat{U}_{th} = \hat{\sigma}^2_{\lambda|th} + d_{th}. \tag{29}$$

The justification is as follows. Uncertainty comes from two sources: the error components in common information and in private information. $\hat{\sigma}^2_{\lambda|th}$ captures the imprecision in common information, and d_{th} reflects the same in forecasters' idiosyncratic information and diversity in forecasting models. The measure of uncertainty in (29) avoids the drawback of the inability to capture the heterogeneity of forecasting models in using GARCH measure of uncertainty alone. Our suggestion is supported by the findings in Batchelor and Dua (1993) and Bomberger (1996); in a comparison of ARCH and survey measures of uncertainty, these two studies concluded that the former tends to be lower than the latter, and more importantly the former is less variable over time than the latter. Thus, if one accepts survey measures as valid, ARCH measure alone underestimates the level and the variation in uncertainty over time.

In order to generate GARCH-type estimates of the variability of aggregate shocks, we first filter the mean forecast errors for possible autocorrelation, see Harvey and Newbold (2003). 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 $\sigma^2_{\lambda|t}$ 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:

$$e_t \sim N(0, \sigma^2_{\lambda|t}),\ \sigma^2_{\lambda|t} = \alpha_0 + \alpha_1 e^2_{t-1} + \alpha_2 \sigma^2_{\lambda|t-1}, \tag{30}$$

where e_t is the serially uncorrelated mean forecast error. Equation (30) has been estimated using the quasi-maximum likelihood (cf. Bollerslev and Wooldridge, 1992) for the 1984-2007 subsample and for each horizon. Consistent with many earlier studies, in Model 2 we estimated (30) using the *t*-distribution with 6 degrees of freedom. As an alternative specification, we replaced the lagged mean squared forecast error in (30) with the average of mean squared errors over the last ten years.[18] In Model 3, we estimated $\sigma^2_{\lambda|t}$ using the following model specification:

$$e_t \sim N(0, \sigma^2_{\lambda|t}), \ \sigma^2_{\lambda|t} = \beta_0 + \beta_1 (\sum_{s=1}^{10} MSE_{t-s}/10) + \beta_2 \sigma^2_{\lambda|t-1}. \qquad (31)$$

Model 4 estimated (31) 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. The estimation results, not reported here, show that the lagged variance of aggregate shocks was significant at the 5% level in the majority of the cases, but the lagged mean forecast errors, as well as the average of mean squared errors over the last ten years, are only significant in some cases, depending on the horizons and variables under study.[19]

According to (29), forecast uncertainty is generated by the sum of the estimated variance of aggregate shocks $\hat{\sigma}^2_{\lambda|th}$ from GARCH-type models and the disagreement from the survey. Table 6 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 in Table 6), 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

[18] During 1974-1981, SPF did not ask for the annual average forecast. We matched the reported quarterly point forecasts with the real time data to derive the implied annual forecasts for the current year.
[19] Following Bomberger (1996), we also added disagreement in the variance equation of the GARCH models and found that disagreement never became significant at the 5% level. This is consistent with the findings in Rich and Butler (1998).

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 – a point that we have established before in section 3.3.

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.[20]

We plot the evolution of uncertainty generated from the best models in inflation and output growth forecasts over time in Figures 5 and 6. 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

[20] We also estimated Models 1, 2, 5 and 6 during 1969-2007. We find that the generated uncertainty according to these four models cannot beat the disagreement alone to match the survey measure of uncertainty when we include the unstable period.

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.

4. CONCLUDING REMARKS

Due to the ready availability of point forecasts, disagreement among forecasters has been widely used as a proxy for aggregate uncertainty in the economics, accounting and finance literature. Lacking theoretical basis, empirical evidence has been mixed as to whether the disagreement is a reliable measure for the uncertainty. Using a standard decomposition of forecast errors into common and idiosyncratic shocks in a panel data setting, our paper demonstrates that under certain regularity conditions, the difference between uncertainty and disagreement is the perceived variance of future aggregate shocks that accumulate over horizons. This result has important implications. It implies that the robustness of the relationship between uncertainty and disagreement depends on the variance of aggregate shocks over time and across horizons. Using the SPF density forecasts for inflation and output growth, we find direct evidence in support of our hypothesized time and horizon effects. As for the time effect, disagreement is found to be a reliable measure for uncertainty in a stable period. In periods with large volatility of aggregate shocks, however, disagreement becomes less useful a proxy. As for the horizon effect, we find that the longer the forecast horizon, the larger is the difference between disagreement and uncertainty. Though disagreement alone tends to understate the level of uncertainty, our empirical results suggest that one can safely use disagreement as a proxy for uncertainty in a regression context, provided the forecast environment is relatively stable. By subtracting observed disagreement from uncertainty using density forecasts, we obtain a truly *ex ante* measure of aggregate shocks that befell on the economy. These aggregate shocks are available to a policy maker before the actual values are realized, and show remarkable reduction in the volatility after 1991.

Our results do not support the use of squared mean forecast errors to construct *ex ante* uncertainty, as often practiced in recent accounting and finance research. Since forecast error is an *ex post* measure reflecting unexpected shocks after the forecast is made, it should not affect uncertainty at the time of forecast. In order to construct an *ex ante* measure of forecast uncertainty, one should use the sum of the observed disagreement from the survey and the projected variance of aggregate shocks generated by a suitably specified GARCH model. We find that this approach performs much better than the use of squared forecast errors in matching the survey measure of uncertainty, and is less sensitive to occasional large forecast surprises.

Reference:

Abarbanell JS, Lanen WN, Verrecchia RE. 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. *Journal of Accounting and Economics* 20: 31-60.

Banerjee A, Marcellino M. 2006. Are there any reliable leading indicators for US inflation and GDP growth? *International Journal of Forecasting* 22: 137-151.

Barron OE, Kim O, Lim SC, Stevens DE. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73: 421-433.

Barry CB, Jennings RH. 1992. Information and diversity of analyst opinion. *Journal of Financial and Quantitative Analysis* 27: 169-183.

Batchelor R, Dua P. 1993. Survey vs. ARCH measures of inflation uncertainty. *Oxford Bulletin of Economics and Statistics* 55: 341-353.

Batchelor R, Dua P. 1995. Forecast diversity and the benefits of combining forecasts. *Management Science* 41: 68-75.

Boero G, Smith J, Wallis KF. 2007. Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters. Forthcoming in *Economic Journal.*

Bollerslev T. 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31: 307-327.

Bollerslev T, Wooldridge JM. 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews* 11: 143-172.

Bomberger WA. 1996. Disagreement as a measure of uncertainty. *Journal of Money, Credit and Banking* 28: 381-392.

Bomberger WA. 1999. Disagreement and uncertainty: a reply to Rich and Butler. *Journal of Money, Credit and Banking* 31: 273-276.

Campbell SD. 2007. Macroeconomic volatility, predictability, and uncertainty in the great moderation: evidence from the Survey of Professional Forecasters. *Journal of Business and Economic Statistics* 25: 191-200.

Davies A, Lahiri K. 1995. A new framework for analyzing survey forecasts using three-dimensional panel data. *Journal of Econometrics* 68: 205-227.

Diebold, FX, Tay, AS, Wallis, KF. 1999. Evaluating density forecasts of inflation: The survey of professional forecasters. In: Engle, R.F., White, H. (Eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger. Oxford University Press, Oxford.

Diether KB, Malloy CJ, Scherbian A. 2002. Differences of opinion and the cross section of stock returns. *The Journal of Finance* 57: 2113-2141.

Engle RF. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987-1008.

Engle RF. 1983. Estimates of the variance of U.S. inflation based upon the ARCH model. *Journal of Money, Credit and Banking* 15: 286-301.

Engelberg J, Manski CF, and Williams J. 2006. Comparing the point predictions and subjective probability distributions of professional forecasters, Forthcoming in *Journal of Business and Economic Statistics*.

Giordani P, Söderlind P. 2003. Inflation forecast uncertainty. *European Economic Review* 47: 1037-1059.

Harvey D, Newbold P. 2003. The non-normality of some macroeconomic forecast errors. *International Journal of Forecasting* 19: 635-653.

Johnson TC. 2004. Forecast dispersion and the cross section of expected returns. *The Journal of Finance* 59: 1957-1978.

Lahiri K, Teigland C, Zaporowski M. 1988. Interest rates and the subjective probability distribution of inflation forecasts. *Journal of Money*, *Credit and Banking* 20: 233-248.

Lahiri K, Liu F. 2005. ARCH models for multi-period forecast uncertainty - a reality check using a panel of density forecasts. In *Advances in Econometrics*, *volume 20: Econometric Analysis of Financial Time Series*, edited by T. Fomby and D. Terrell, pp. 321-363. Elsevier, JAI Press.

Lahiri K, Liu F. 2006. Modeling multi-period inflation uncertainty using a panel of density forecasts. *Journal of Applied Econometrics* 21: 1199-1220.

Lahiri K, Sheng X. 2007. Evolution of forecast disagreement in a Bayesian learning model. Forthcoming in *Journal of Econometrics*.

McConnell MM, Perez-Quiros G. 2003. Output fluctuations in the United States: what has changed since the early 1980's? *American Economic Review* 90: 1464-1476.

Mishkin FS. 2007. Inflation dynamics. Speech delivered at the Annual Macro Conference, Federal Reserve Bank of San Francisco, California (March 23).

Mohanram PS, Sunder SV. 2006. How has regulation FD affected the operations of financial analysts? *Contemporary Accounting Research* 23: 491-525.

Öller LE, Barot B. 2000. The accuracy of European growth and inflation forecasts. *International Journal of Forecasting* 16: 293-315.

Reifschneider D, Tulip P. 2007. Gauging the uncertainty of the economic outlook from historical forecasting errors. Federal Reserve Board, Finance and Economics Discussion Series: 60.

Rich RW, Butler JS. 1998. Disagreement as a measure of uncertainty: a comment on Bomberger. *Journal of Money, Credit and Banking* 30: 411-419.

Stock JH, Watson MW. 2007. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* 39: 3-33.

Timmermann A, Patton A. 2007. Learning in real time: theory and empirical evidence from the term structure of survey forecasts. working paper, UCSD.

Wallis KF. 2005. Combining density and interval forecasts: a modest proposal. *Oxford Bulletin of Economics and Statistics* 67: 983-994.

Zarnowitz V, Lambros LA. 1987. Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95: 591-621.

Zhang XF. 2006. Information uncertainty and stock returns. *The Journal of Finance* 61: 105-137.

Table 1. 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

Table 2. Comparison of 90% predictive interval with actual outcomes

Horizon	4Q ahead	3Q ahead	2Q ahead	1Q ahead
SPF Inflation	78.48	85.46	89.34	88.22
SPF Output growth	74.35	87.17	85.95	83.33

Note: This table shows the percentage of times that the 90% predictive interval covers the actual outcomes. Predictive intervals are constructed from SPF individual density forecasts during 1969-2007 by fitting uniform distribution over histograms.

Table 3. Regression of survey measure of uncertainty on disagreement over time

	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

Note: Standard errors are in parentheses. One asterisk denotes that the estimated values are significant at the 5% critical level.

Table 4. Regression of survey measure of uncertainty on *ex post* uncertainty

	SPF inflation forecast		SPF GDP forecast
	1969-1983	1984-2007	1984-2007
Ex post uncertainty	0.02	0.27*	0.25*
	(0.02)	(0.07)	(0.05)
H1	0.46*	0.24*	0.30*
	(0.22)	(0.02)	(0.01)
H2	0.49*	0.40*	0.62*
	(0.05)	(0.02)	(0.02)
H3	0.55*	0.49*	0.84*
	(0.06)	(0.03)	(0.03)
H4	0.75*	0.48*	0.93*
	(0.06)	(0.05)	(0.06)
Adj. R^2	0.09	0.30	0.54

Note: Standard errors are in parentheses. One asterisk denotes that the estimated values are significant at the 5% critical level.

Table 5. 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 $RMSE_2 = \sqrt{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(A_t - F_{it})^2}$.

Table 6. 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).

Figure 1. Uncertainty (solid line) and disagreement (dotted line) in inflation forecasts

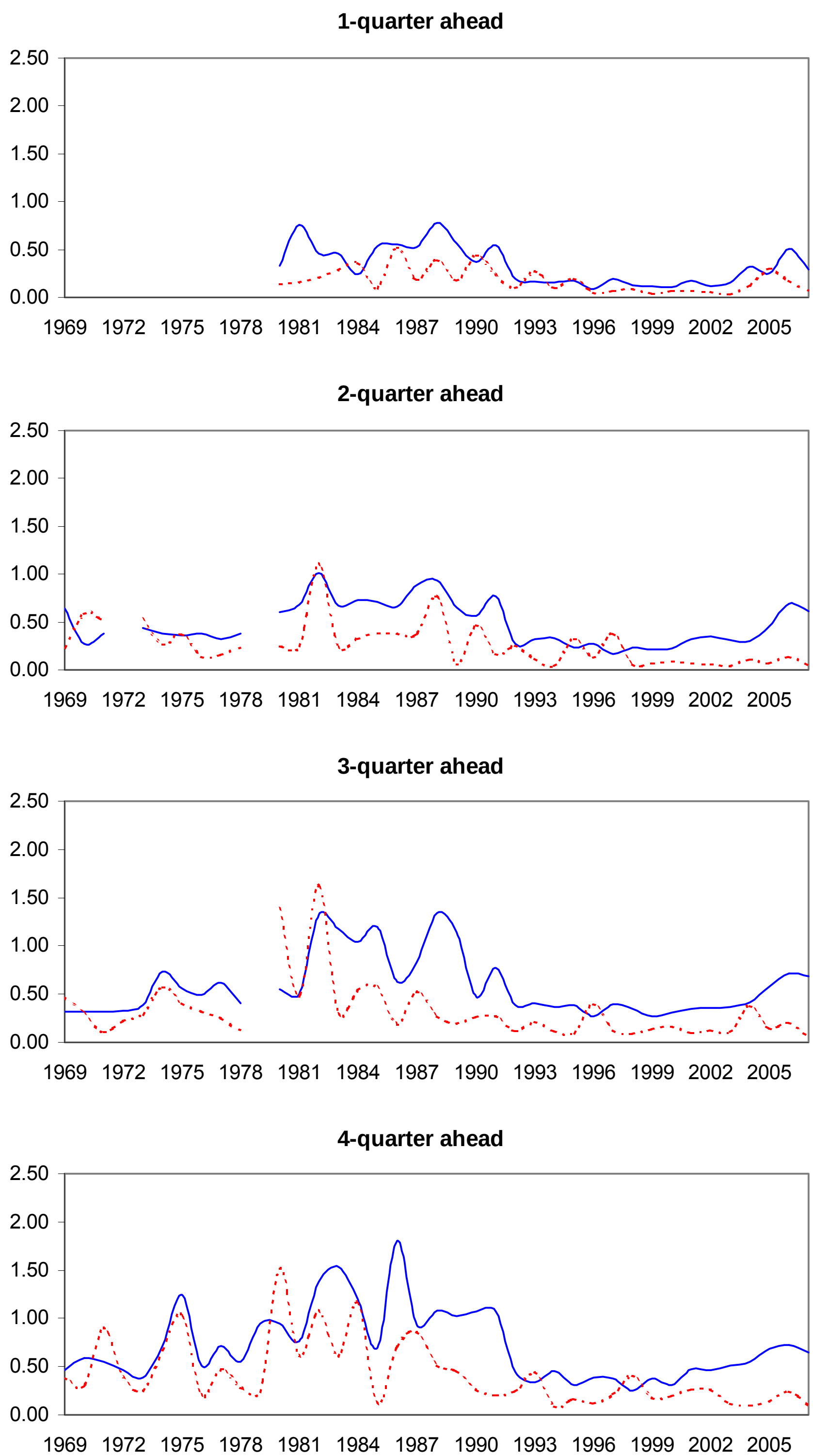


Figure 2. Uncertainty (solid line) and disagreement (dotted line) in real GDP forecasts

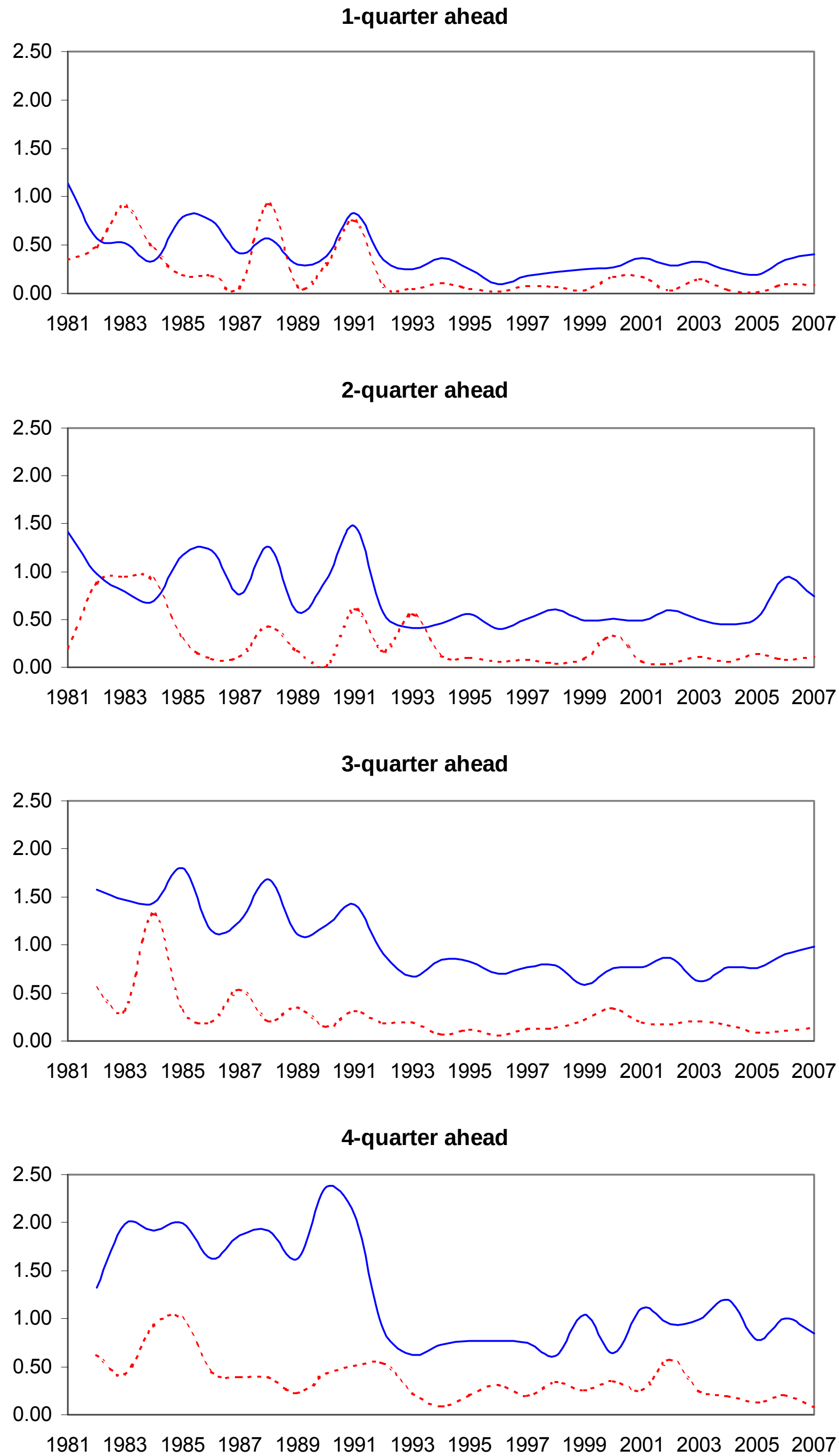


Figure 3. Estimated variance of aggregate shocks in inflation forecasts

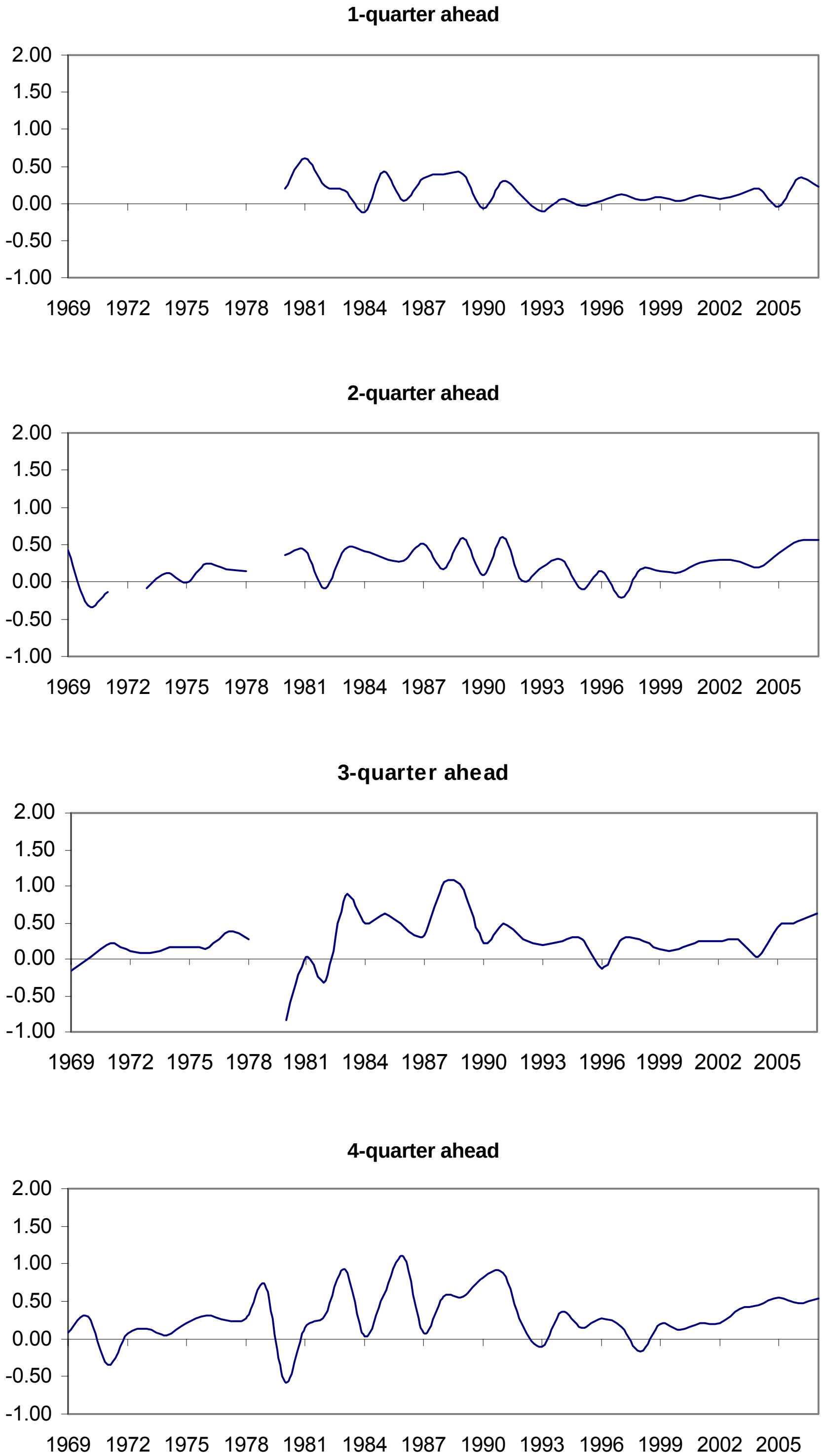


Figure 4. Estimated variance of aggregate shocks in output growth forecasts

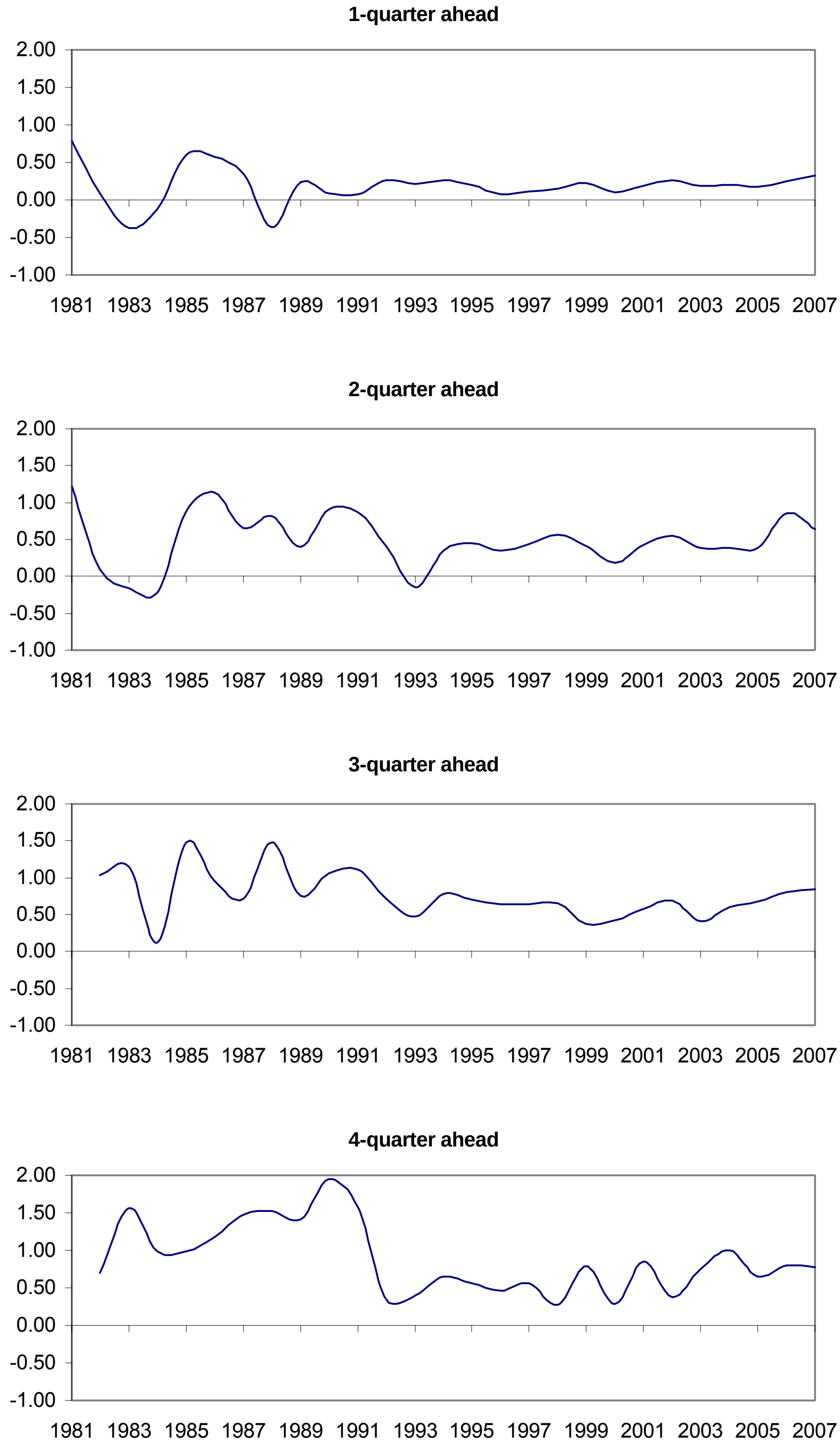


Figure 5. Measures of uncertainty in inflation forecasts:
Survey measure of uncertainty (solid line)
Uncertainty using squared error of mean forecast (dotted line)
Uncertainty from GARCH-type model (line with diamond)

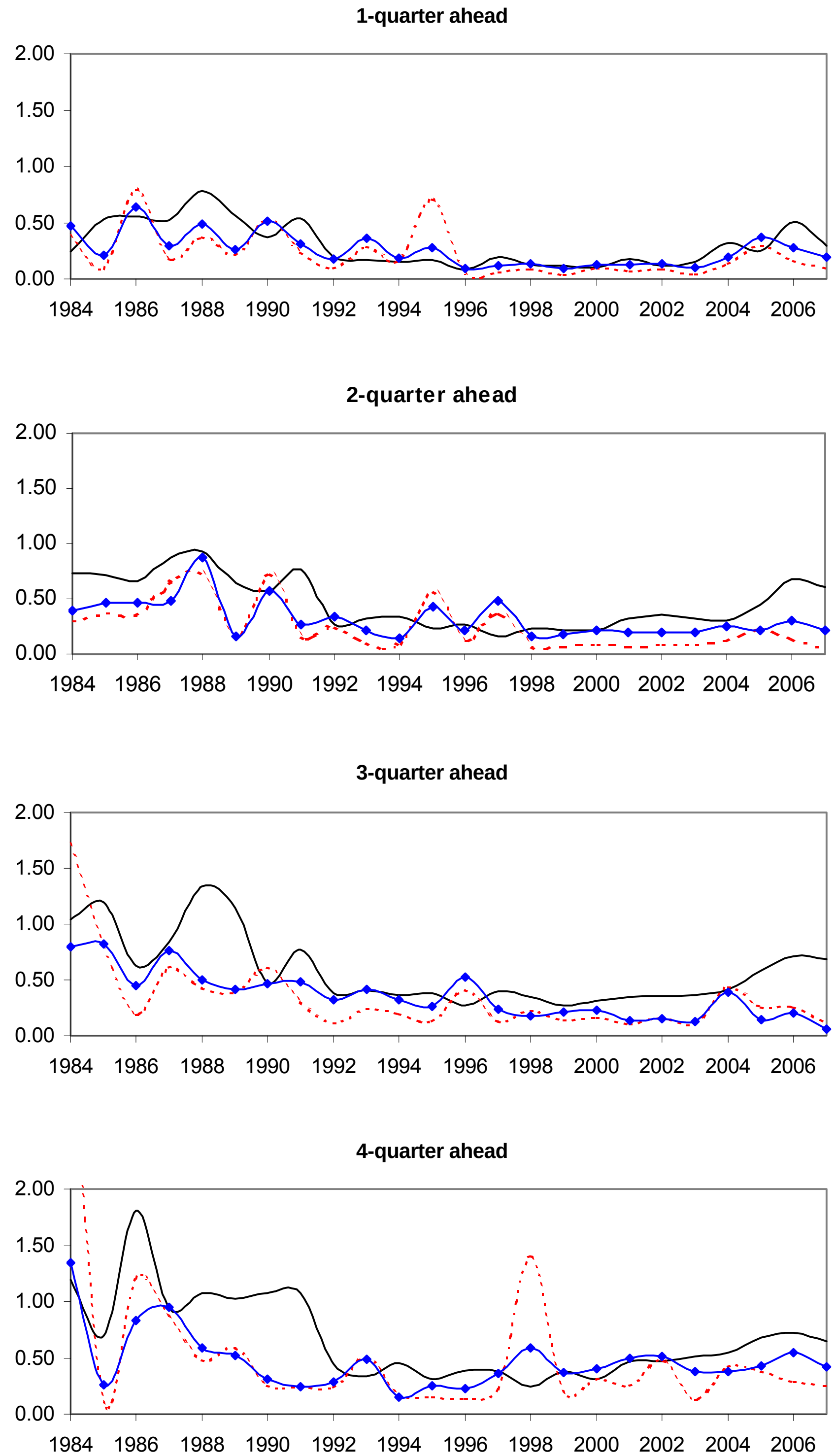
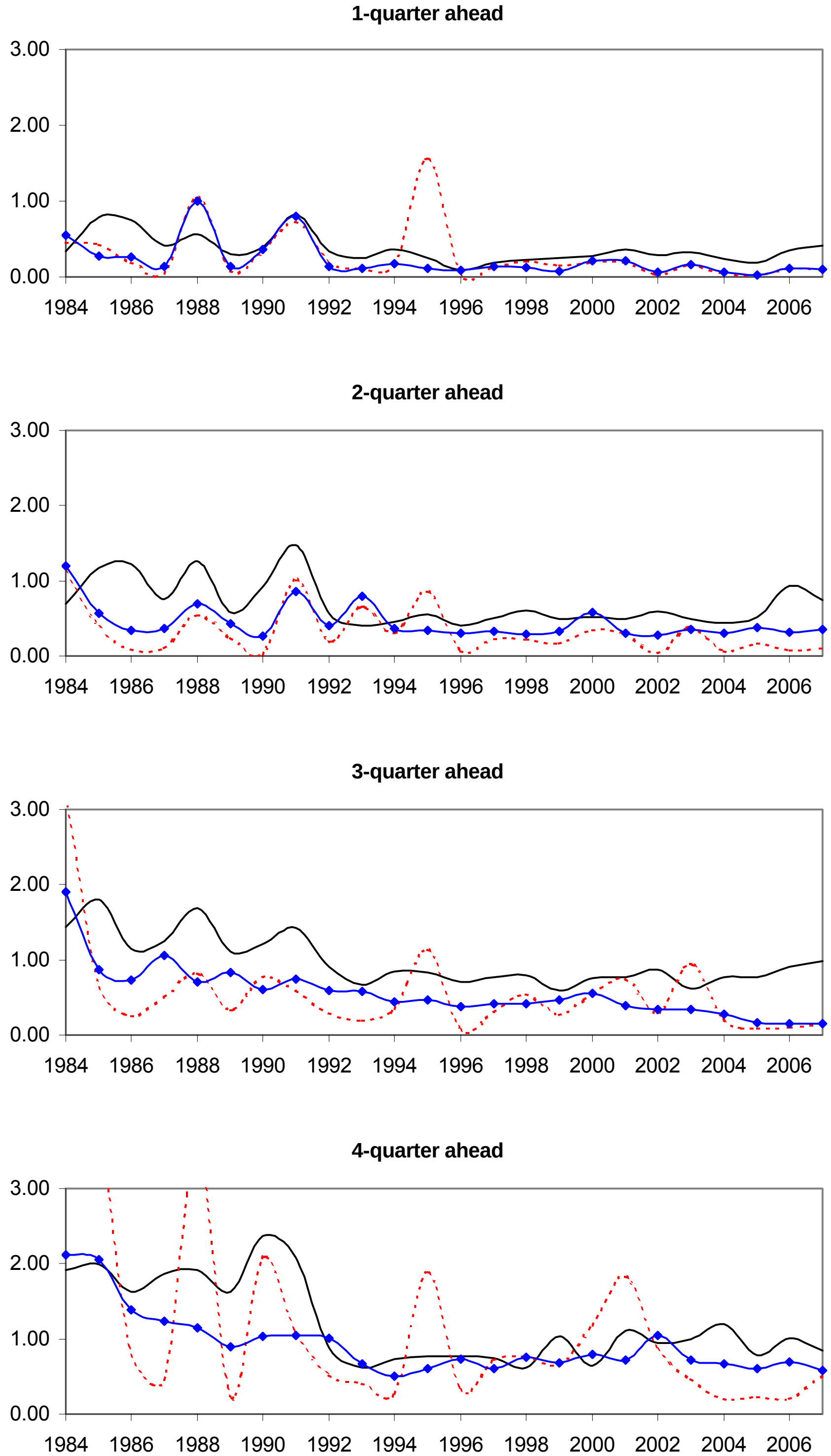


Figure 6. Measures of uncertainty in output growth forecasts:
Survey measure of uncertainty (solid line)
Uncertainty using squared error of mean forecast (dotted line)
Uncertainty from GARCH-type model (line with diamond)



Ifo Working Papers

No. 59 Overesch, M. and G. Wamser, Who Cares about Corporate Taxation? Asymmetric Tax Effects on Outbound FDI, April 2008.

No. 58 Eicher, T.S: and T. Strobel, Germany's Continued Productivity Slump: An Industry Analysis, March 2008.

No. 57 Robinzonov, N. and K. Wohlrabe, Freedom of Choice in Macroeconomic Forecasting: An Illustration with German Industrial Production and Linear Models, March 2008.

No. 56 Grundig, B., Why is the share of women willing to work in East Germany larger than in West Germany? A logit model of extensive labour supply decision, February 2008.

No. 55 Henzel, S., Learning Trend Inflation – Can Signal Extraction Explain Survey Forecasts?, February 2008.

No. 54 Sinn, H.-W., Das grüne Paradoxon: Warum man das Angebot bei der Klimapolitik nicht vergessen darf, Januar 2008.

No. 53 Schwerdt, G. and J. Turunen, Changes in Human Capital: Implications for Productivity Growth in the Euro Area, December 2007.

No. 52 Berlemann, M. und G. Vogt, Kurzfristige Wachstumseffekte von Naturkatastrophen – Eine empirische Analyse der Flutkatastrophe vom August 2002 in Sachsen, November 2007.

No. 51 Huck, S. and G.K. Lünser, Group Reputations – An Experimental Foray, November 2007.

No. 50 Meier, V. and G. Schütz, The Economics of Tracking and Non-Tracking, October 2007.

No. 49 Buettner, T. and A. Ebertz, Quality of Life in the Regions – Results for German Counties, September 2007.

No. 48 Mayr, J. and D. Ulbricht, VAR Model Averaging for Multi-Step Forecasting, August 2007.

No. 47 Becker, S.O. and K. Wohlrabe, Micro Data at the Ifo Institute for Economic Research – The "Ifo Business Survey", Usage and Access, August 2007.

No. 46 Hülsewig, O., J. Mayr and S. Sorbe, Assessing the Forecast Properties of the CESifo World Economic Climate Indicator: Evidence for the Euro Area, May 2007.

No. 45 Buettner, T., Reform der Gemeindefinanzen, April 2007.

No. 44 Abberger, K., S.O. Becker, B. Hofmann und K. Wohlrabe, Mikrodaten im ifo Institut – Bestand, Verwendung und Zugang, März 2007.

No. 43 Jäckle, R., Health and Wages. Panel data estimates considering selection and endogeneity, March 2007.

No. 42 Mayr, J. and D. Ulbricht, Log versus Level in VAR Forecasting: 16 Million Empirical Answers – Expect the Unexpected, February 2007.

No. 41 Oberndorfer, U., D. Ulbricht and J. Ketterer, Lost in Transmission? Stock Market Impacts of the 2006 European Gas Crisis, February 2007.

No. 40 Abberger, K., Forecasting Quarter-on-Quarter Changes of German GDP with Monthly Business Tendency Survey Results, January 2007.

No. 39 Batchelor, R., Forecaster Behaviour and Bias in Macroeconomic Forecasts, January 2007.

No. 38 Sülzle, K., Innovation and Adoption of Electronic Business Technologies, December 2006.

No. 37 Overesch, M. and G. Wamser, German Inbound Investment, Corporate Tax Planning, and Thin-Capitalization Rules – A Difference-in-Differences Approach, December 2006.

No. 36 Kempkes, G. and C. Pohl, The Efficiency of German Universities – Some Evidence from Non-Parametric and Parametric Methods, October 2006.

No. 35 Kuhlmann, A., German Productivity – A Reassessment via the New Ifo Productivity Database, October 2006.

No. 34 Kuhlmann, A., What is the X-Factor in the German Electricity Industry?, September 2006.

No. 33 Temple, J. and L. Wößmann, Dualism and cross-country growth regressions, August 2006.