Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 26 (2010) 265–292
international journal of forecasting
www.elsevier.com/locate/ijforecast

Learning and heterogeneity in GDP and inflation forecasts

Kajal Lahiri[a,*], Xuguang Sheng[b]

[a] Department of Economics, University at Albany - SUNY, Albany, NY 12222, United States
[b] Department of Economics, SUNY Fredonia, Fredonia, NY 14063, United States

Abstract

Using a Bayesian learning model with heterogeneity across agents, our study aims to identify the relative importance of alternative pathways through which professional forecasters disagree and reach consensus on the term structure of inflation and real GDP forecasts, resulting in different patterns of forecast accuracy. There are two primary sources of forecast disagreement in our model: differences in prior beliefs, and differences in the interpretation of new public information. Estimated model parameters, together with two separate case studies on (i) the dynamics of forecast disagreement in the aftermath of the 9/11 terrorist attack in the US, and (ii) the successful inflation targeting experience of Italy after 1997, firmly establish the importance of these two pathways to expert disagreement, and help to explain the relative forecasting accuracy of these two macroeconomic variables.

Keywords: Bayesian learning; Public information; Panel data; Forecast disagreement; Forecast horizon; Forecast efficiency; GDP; Inflation targeting

1. Introduction

An analysis of forecast revisions and their cross-sectional dispersion can reveal important information on how efficiently and uniformly forecasters react to new information. Using monthly fixed-target survey forecasts for real GDP, Lahiri and Sheng (2008) estimated a Bayesian learning model aimed at explaining the role of priors in forecast disagreement and its evolution over various horizons. In this paper we extend our analysis to both real GDP and inflation forecasts using more recent data, and explain certain important differences in the ways professional forecasters treat these two variables for producing multi-period forecasts. We find that when predicting inflation, professional forecasters (i) make smaller forecast errors; (ii) disagree to a lesser extent; and (iii) start revising their forecasts much earlier, compared to predicting real GDP. Even though the first of these results has been implicit in most studies of forecast

* Corresponding author. Tel.: +1 518 442 4758; fax: +1 518 442 4736.
E-mail addresses: klahiri@albany.edu (K. Lahiri), sheng@fredonia.edu (X. Sheng).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2009.12.009

evaluation,[1] none of these empirical results are well articulated in the forecasting literature.

At least part of the explanation for the superior forecasting record of some variables has to lie in the nature of their data generating processes. In reality, however, the predictability can be improved by incorporating additional information from diverse sources and using more complicated models. In real time, the forecasters face both additional uncertainty due to data revisions and the possibility of breaks due to unstable data generating processes. Also, one could reasonably ask why the data generating processes differ between variables. To understand these issues more comprehensively, we also need to explore the underlying expectation formation processes and the role of individual heterogeneity in incorporating new information. Using a Bayesian information processing framework, our study aims to identify the relative importance of the alternative pathways through which professional forecasters adapt to new information and determine the term structure of forecasts, resulting in different patterns of forecast accuracy.

We find that experts start off with widely divergent prior beliefs at very long horizons. Their initial beliefs propagate forward to the whole series of forecasts, generating a significant amount of inertia in expectations formation. This "anchoring"-type effect, which has been much emphasized in the psychological literature, is a result of optimal Bayesian information processing that efficiently combines priors with new information (see Zellner, 2002). However, our analysis shows that there is more pervasive stickiness in the recorded real GDP forecasts than in the inflation forecasts, due to the inefficient use of new information.

The rest of the paper is organized as follows. In the next section, we present some stylized facts based on the cross-country forecast data. In Section 3, we explore the data generating processes of the target variables. In Section 4, we estimate the Bayesian learning model and present empirical evidence on the alternative pathways for generating disagreement. Section 4 also presents two case studies on (i) the dynamics of forecast disagreement after the 9/11 terrorist attack in the US, and (ii) the inflation targeting experience of Italy after 1997. We investigate forecast efficiency in utilizing public information for both real GDP and inflation in Section 5, and Section 6 concludes.

2. Some stylized facts

This section starts with a brief introduction to the data used in our analysis. We then highlight a few stylized facts concerning the evolution of consensus forecasts, forecast accuracy, forecast disagreement and forecast revisions in real GDP and inflation. We find some important differences in the ways professional forecasters treat these two macroeconomic variables.

2.1. Data

The data used in this study are taken from "*Consensus Forecasts: A Digest of International Economic Forecasts*", published by Consensus Economics Inc. We study a panel of forecasts of annual real GDP growth and inflation. The survey respondents start forecasting in January of the previous year, and their last forecast is reported at the beginning of December of the target year. Thus, for each country and target year, we have 24 forecasts at various horizons. Our data start with the January 1990 forecasts and end with the December 2007 forecasts, giving predictions for 17 target years 1991–2007 and seven major industrialized (G7) countries — Canada, France, Germany, Italy, Japan, the United Kingdom and the United States.[2] Inflation is measured by the annual percentage change in the consumer price index for all G7 countries except the United Kingdom.[3] The forecasting institutions, numbering between 20 and 40, are typically banks, securities firms, econometric modelers, industrial corporations and independent forecasters. Thus, they are all professional private market forecasters. Since most of the institutions are located in the countries for which they are forecasting, country-specific expertise is guaranteed. Altogether, we have more than

[1] See, for example, Banerjee and Marcellino (2006), Öller and Barot (2000), Stock and Watson (2003), and Zarnowitz and Braun (1993) over various sample periods and countries.

[2] Note that the targets for GDP and inflation in Germany change over our data sample due to unification. We use forecasts for West Germany made for the target years 1991–1995, and for unified Germany for the target years 1996–2007.

[3] For the UK, the inflation rate is based on the Retail Price Index (RPI). However, from April 1997 onward, forecasts are solicited for the RPI excluding mortgage interest costs.

115,000 forecasts for real GDP and inflation. In the following analysis, we use an early announcement as the actual value, which is published in the May issues of *Consensus Forecasts* immediately following the target year.

For the current study, this data set has many advantages over some other more commonly used surveys. First, *Consensus Forecasts* is regularly sold in a wide variety of markets, and the names of the respondents are published next to their forecasts. Hence, one would expect these professional forecasts to achieve a higher level of accuracy compared to laymen's expectations, as poor forecasts will damage the forecaster's reputation. Second, since private information is expected to be relatively unimportant in forecasting GDP and inflation compared to such variables as corporate earnings, company stock prices, etc., we can identify the news which is relevant to GDP and inflation as being mostly public information. Finally, forecasts for fairly long horizons, currently from 24 months to 1 month ahead, are available. This fixed-event scheme enables us to study the role of heterogeneity in priors and their effects on expert disagreement for a sequence of 24 forecasts for 17 target years.

2.2. Evolution of forecast accuracy over horizons

In Fig. 1 we plot the average root mean squared forecast errors (RMSE) for real GDP and inflation using individual data over the period 1991–2007. Two findings stand out. First, compared to inflation, professional forecasters make larger forecast errors in real GDP at all horizons observed. This finding is consistent with most previous studies of forecast evaluation. Second, the RMSE for real GDP at horizons of 24–18 months often stays relatively flat, but for inflation it declines steadily right from the beginning (i.e., $h = 24$). This latter finding has not yet been explored in the forecasting literature.

2.3. Evolution of consensus forecasts

Next, we examine the plots of consensus (i.e., mean) forecasts and the realized actual values of real GDP and inflation over the period 1991–2007. These plots start when the forecast horizon is 24, which is in January of the previous year, and end when the forecast horizon is 0, which is the actual realization.[4]

First, as can be seen from these plots, for the first few rounds of forecasting (for horizons 24 to 18 months) and for the majority of years and countries, the consensus forecasts do not seem to change very much – more so for real GDP than for inflation. This empirical observation leads us to believe that, over these horizons, forecasters do not receive enough dependable information to enable them to revise their forecasts systematically. Second, the initial 24-month-ahead forecasts for all countries seem to start from a relatively narrow band, and then, as information accumulates, they tend to diverge from these initial starting points and move toward their final destinations. Thus, the variability of the mean forecasts over the target years is very small at the longer horizons, and increases rapidly as the forecast horizon gets shorter. Note also that the initial inflation forecasts seem to be bunched together more than the initial GDP forecasts, and have become less variable toward the latter part of the sample.

All in all, a close look at these graphs reveals certain regularities in the way in which the fixed-target consensus forecasts evolve over time. We now proceed to examine more rigorously the underlying dynamics in forecaster disagreement around these consensus forecasts and the timing of the arrival of important information when forecasters break away from their initial estimates. The dynamics of forecast disagreement contain important information about the sources of forecast inefficiency.

2.4. Evolution of forecast disagreement

Following the literature, we measure forecast disagreement as the variance of the forecasts across professional forecasters.[5] In order to study the general pattern of forecast disagreement over horizons, in Fig. 2 we plot the average disagreement over 17 target years at each of the 24 forecast horizons for

[4] To save space, these graphs are not reported here, but can be found in our working paper at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1266219.

[5] In our study sample, where only more frequent respondents were included, the inter-quartile range and the variance of the individual forecasts were found to be very similar, cf. Döpke and Fritsche (2006).

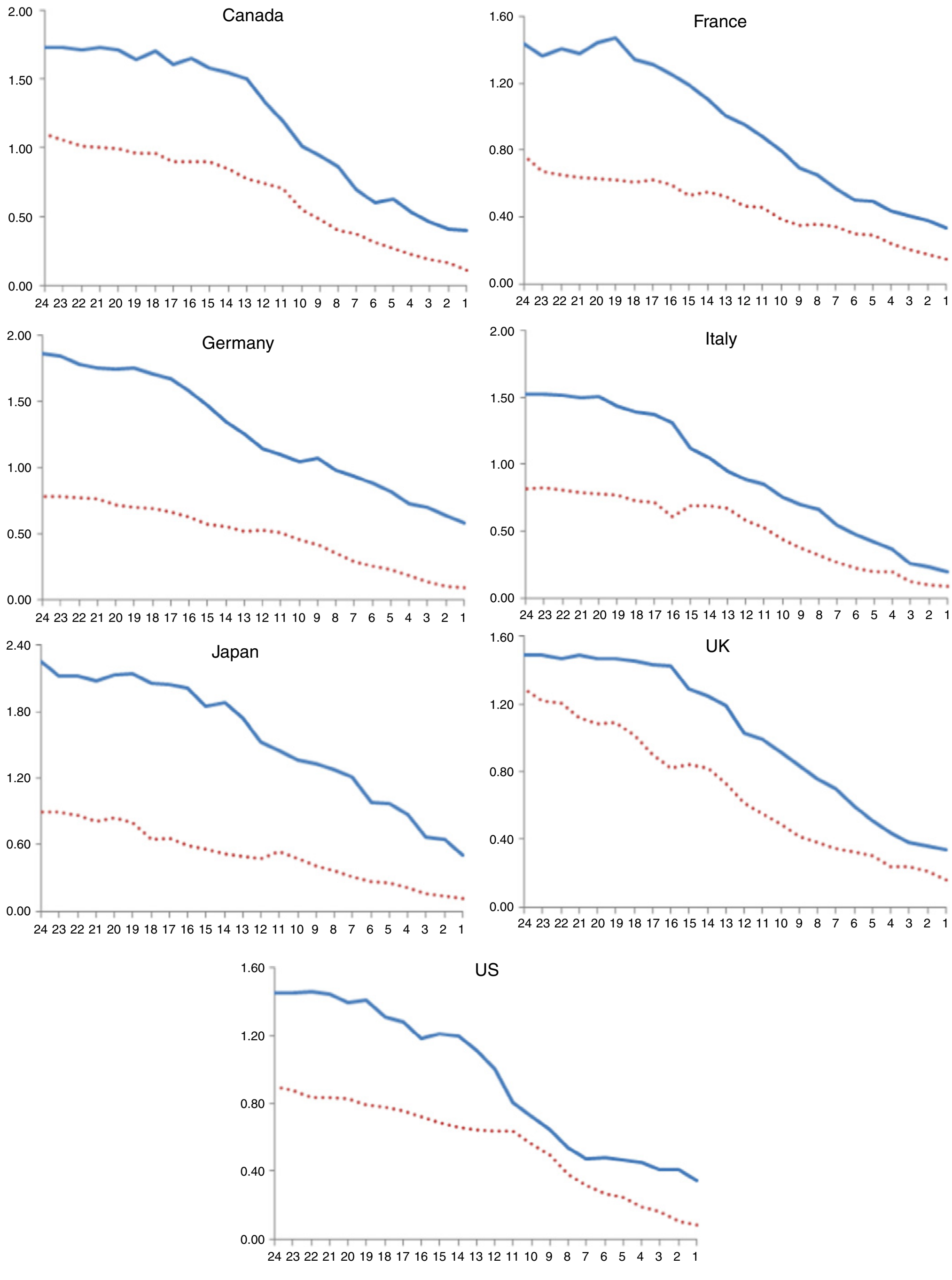


Fig. 1. RMSEs of the real GDP (solid lines) and inflation (dotted lines) forecasts.

the GDP and inflation forecasts. Although the magnitude of the disagreement varies a lot across countries (France, Italy and Germany have comparatively low disagreements), the extent of the disagreement among

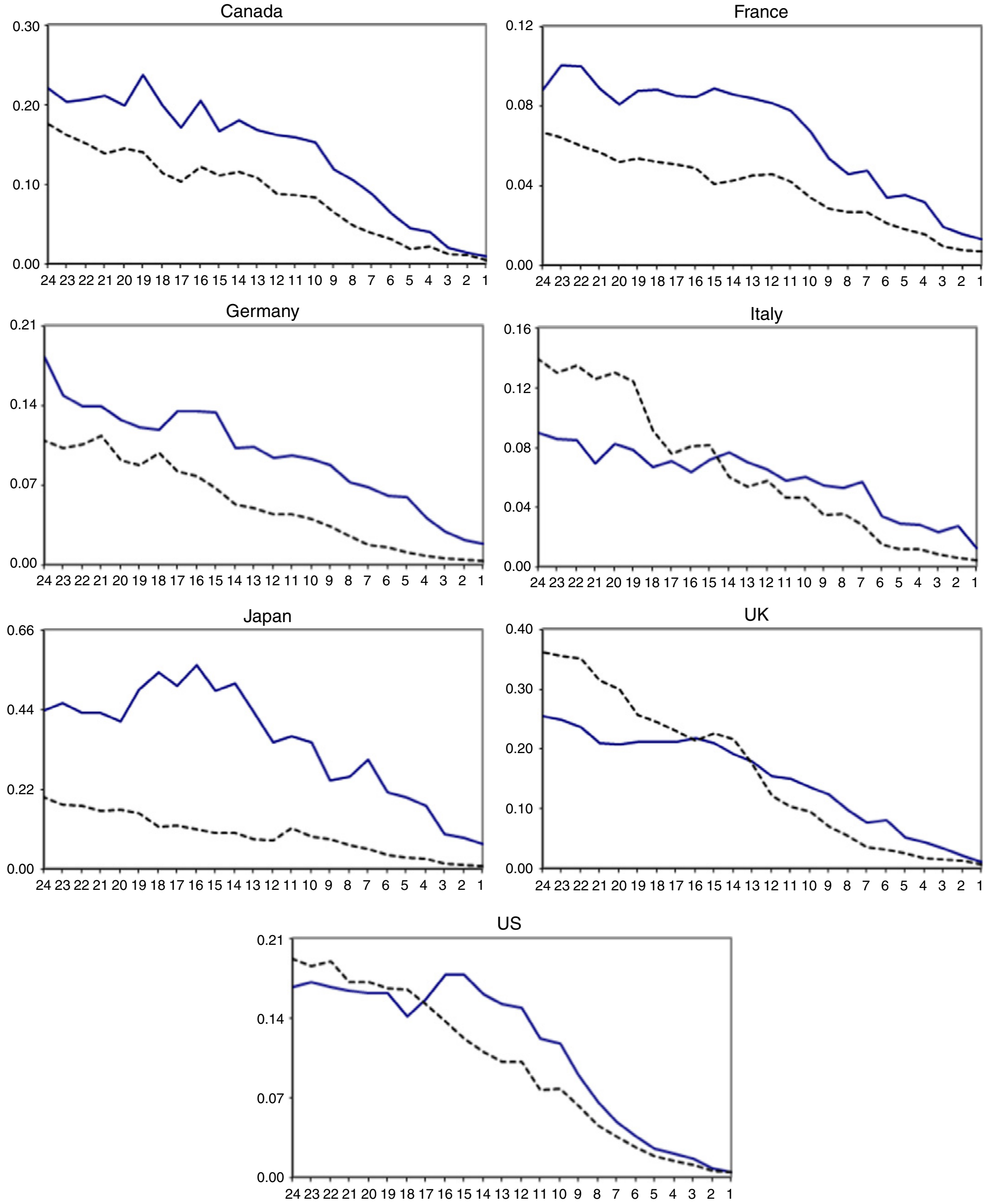


Fig. 2. Forecast disagreement in GDP (solid lines) and inflation (dotted lines).

professional forecasters is less when predicting inflation than GDP, on average.[6] For GDP forecasts, the disagreement is very high at the 24-month horizon and stays almost unchanged or declines very slightly until about the 16-month horizon, after which it starts to decrease sharply at the 15-month horizon and keeps

[6] Note that for the UK, if we ignore the initial two years and the period 1994–96, during which the definition of the price variable was changed, the disagreement at $h = 24$ will be much smaller for inflation than for real GDP.

declining as the horizon gets shorter. For inflation forecasts the disagreement is also relatively high at the beginning, but unlike forecasts of real GDP, it declines monotonically as the horizon gets shorter from 24 months to 1 month.

2.5. Evolution of forecast revisions over horizons

With fixed-target forecasts, an analysis of forecast revisions gives us critical information about when major public information arrives, when it is processed, and the extent to which experts interpret the information differently. Let F_{ith} be the forecast of the target variable made by agent i, for the target year t, h months before the end of the target year. Forecast revision is defined as the difference between two successive forecasts for the same individual i and the same target year t, i.e. $R_{ith} = F_{ith} - F_{ith+1}$. The decomposition of the total sum of squares of forecast revisions into between- and within-agent variations reveals important characteristics of the forecasts in the aggregate. In this context, we introduce three measures, the within-agent variation (S_h^w), between-agent variation (S_h^b) and total variation (S_h^t) in revisions, adjusted for the number of respondents in each month:

$$S_h^w = \sum_{i=1}^{N_h} \sum_{t=1}^{T_{ih}} (R_{ith} - \overline{R}_{i.h})^2 \Big/ \sum_{i=1}^{N_h} T_{ih} \tag{1}$$

$$S_h^b = \sum_{i=1}^{N_h} T_{ih} (\overline{R}_{i.h} - \overline{\overline{R}}_{..h})^2 \Big/ \sum_{i=1}^{N_h} T_{ih} \tag{2}$$

$$S_h^t = \sum_{i=1}^{N_h} \sum_{t=1}^{T_{ih}} (R_{ith} - \overline{\overline{R}}_{..h})^2 \Big/ \sum_{i=1}^{N_h} T_{ih}, \tag{3}$$

where $\overline{R}_{i.h} = \sum_{t=1}^{T_{ih}} R_{ith}/T_{ih}$ is the mean forecast revision over time at horizon h for agent i and $\overline{\overline{R}}_{..h} = \sum_{i=1}^{N_h} \overline{R}_{i.h}/N_h$ is the overall mean revision over time and across agents. Forecasters who responded less than 10% of the time are excluded from our forecast revision analysis, in order to ensure that our results are not dominated by a very few extreme observations. By construction, the total variation in forecast revisions is the sum of the within-agent and between-agent variation; that is, $S_h^t = S_h^w + S_h^b$.

Depending on the horizon, the country and the target variable, we find that the between-agent variation explains 4% to 34% of the total variation in forecast revisions.[7] Over all horizons, the between-agent variation accounts for 10%–15% and 12%–17% of the total variation in GDP and inflation forecasts on average, respectively. This variation across agents can be attributed to different prior beliefs and to a differential interpretation of the same public information. The between-agent variation, however, is relatively small, with the total variation in the forecast revision being driven mainly by within-agent variation. This is to be expected, since our forecasters are professional experts and the targets are widely discussed macroeconomic entities.

Because of its relative size, we are also interested in the evolution of within-agent variation over different forecast horizons. Note that the within-agent variation is the average across-time variation of forecast revisions at each monthly horizon. Figs. 3a and 3b plot the total variation and its components in real GDP and inflation forecast revisions, respectively. Whenever we see a big jump in the within-agent variation at a certain horizon, it indicates that professional forecasters make major revisions at that specific horizon.

For GDP forecasts, the first big spike is observed at horizon 15 for all countries sampled, which simply suggests that professional forecasters observe the first relevant public signal and revise their forecasts at the beginning of the October of the previous year. Depending on the timing of their base-year GDP announcements, within-agent variation gets another boost at horizons of 11 to 9 months, which, as expected, affects the forecasts of the year-over-year growth rate. As the forecast horizon declines, the within-agent variation gets a boost whenever the first release of GDP growth for the previous quarter becomes available.

As for inflation forecasts, we also observe a big spike in within-agent variation around horizon 15 for all sampled countries. It is remarkable that a substantial forecast revision takes place at this horizon for both inflation and real GDP.[8] However, for inflation, professional forecasters start making major

[7] The detailed tabulations can be found at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1266219.

[8] One might have thought that the most important revision would take place at around $h = 11$, when the last year's actual value becomes known, cf. Patton and Timmermann (2007).

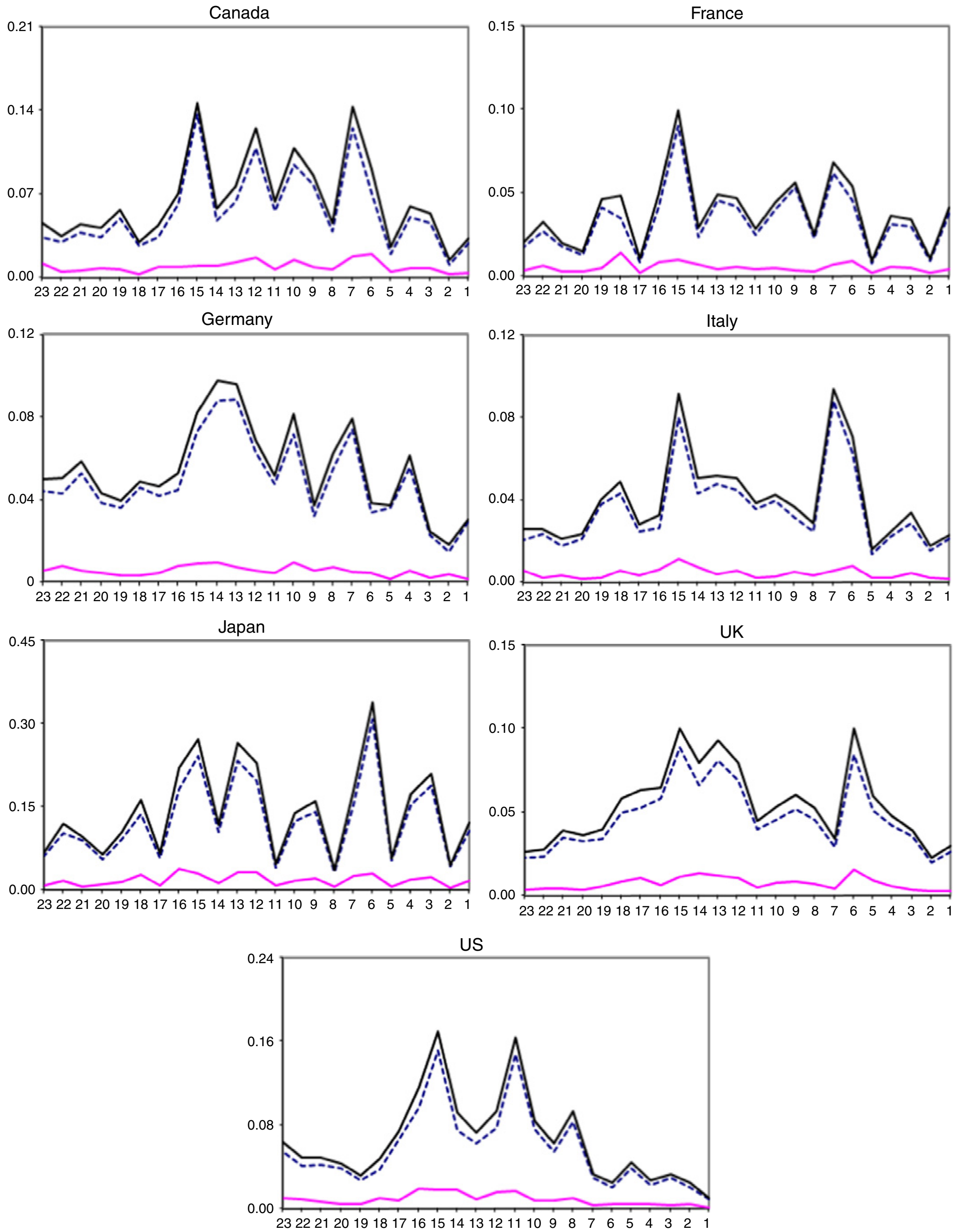


Fig. 3a. GDP forecast revision: between-agent variation (lower solid line), within-agent variation (dashed line), and total variation (upper solid line).

forecast revisions much earlier: they are discernible at horizon 22 for Canada, Germany and the UK and horizon 18 for all of the other G7 countries except Italy, which has its highest peak at the 15-month horizon. Recent research by Banerjee and Marcellino (2006), Gürkaynak, Levin, Marder, and

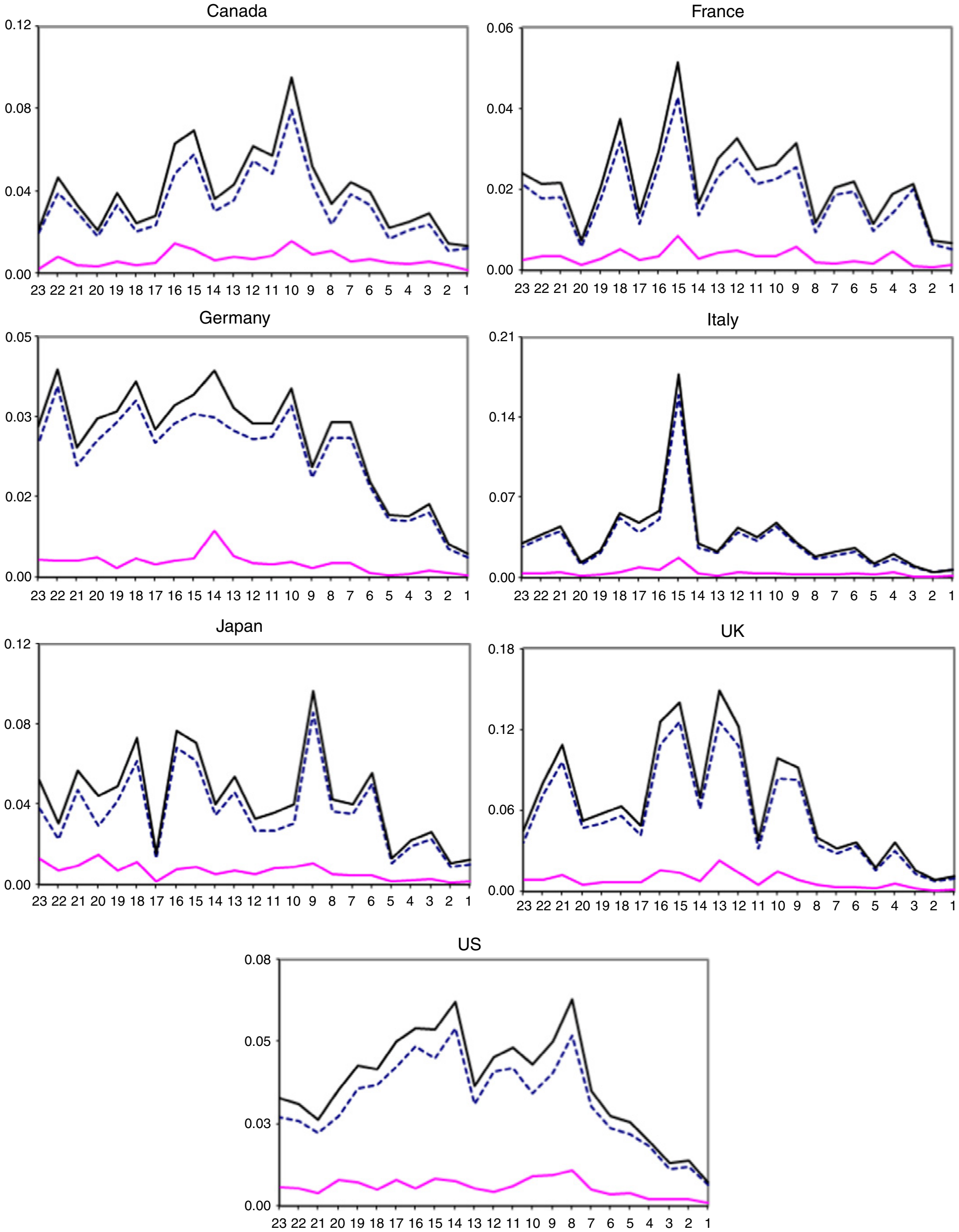


Fig. 3b. Inflation forecast revision: between-agent variation (lower solid line), within-agent variation (dashed line), and total variation (upper solid line).

Swanson (2007) and others has shown that numerous monthly indicators are regularly utilized by market forecasters for gauging future expectations of these macro variables. As the horizon gets shorter, the within-agent variation gets a boost whenever some relevant information concerning inflation for the target year becomes available, including quarterly IPD announcements, various monthly variables and leading indicators. It is interesting to note that even for inflation, forecasters do not seem to revise their forecasts uniformly every month. This gives some credence to the hypothesis put forward by Mankiw, Reis, and Wolfers (2003) that forecasters do not update their information on a continuous basis.

3. Exploring the data generating processes

We find that, historically, real GDP has been a much more difficult variable to predict than inflation.[9] One might think that this could be attributed to the variability of the underlying series. However, it is not the variability, but rather the predictability of the target variable that is one of the important factors in the analysis. This is the focus of this section.

Following Galbraith (2003) and Galbraith and Tkacz (2007), we calculate the forecast content and content horizons for the quarterly GDP and monthly inflation rates for all seven countries in our sample over the period 1990–2007. The forecast content is defined as the proportionate gain in the mean squared forecast error (MSE) from the best fitting autoregressive model over the unconditional mean of the series as the benchmark. The forecast content horizon is defined as the horizon beyond which the forecast content is close to zero. Galbraith (2003) has characterized the content function of AR(p) models analytically, taking into account the uncertainty associated with parameter estimation. We allow p to be no greater than 4 for quarterly GDP data, and 8 for monthly inflation data. The value of p is chosen using the Schwarz information criterion, with an upper bound. The benchmark values were the unconditional means of the individual series during the period 1990–2007. All of the data used in this section are downloaded from *DataStream*.

The results of the estimation of forecast content functions are presented in Figs. 4a and 4b for GDP and inflation, respectively. For annual GDP growth using quarterly data, the forecast content becomes less than 0.05 when the horizon exceeds six quarters. However, for annual inflation using monthly data, the corresponding forecast content horizons are much longer. For Germany and Italy, the content horizon extends beyond 36 months; for the other G7 countries, it is around 24 months. These findings are consistent with the results reported by Galbraith (2003), who looked at the predictability of GDP and inflation for Canada and the United States.

We should point out that our forecast content functions are based purely on linear autoregressive models of the target variables. In reality, the forecast content and predictability can be (and possibly are) improved upon by incorporating additional information and using more complicated models.[10] In addition, the forecast content functions are typically estimated using the currently available revised data. For variables like real GDP which go through a substantial number of data revisions, their predictability in real time can be quite different. Since the variance of the early revisions of a variable is necessarily less than that of the revised series, the predictability of a series may seem to be lower than one could get using real time data. In that sense, the forecast content from the simple AR model provides an overall *lower* bound on the true predictability of a series. For real GDP, Croushore (2006) reports mixed evidence on the effect of data revisions on predictability, depending on the sample period. Since the data revisions are relatively small for inflation, they have very little effect on predictability. In our analysis, the relative ranking of different countries in terms of RMSE does not match their relative ranking in terms of forecast content horizons as obtained from Galbraith's method for either of the variables. The ranking can also depend on the specific benchmark used in the analysis. Thus, it is necessary to study the predictability of real GDP and inflation by professional forecasters in real time with respect to a more natural benchmark.

Following Diebold and Kilian (2001), we define a skill score $p_{s,24}$ as the proportionate MSE gain in the s-month-ahead forecast over the initial forecast made 24 months ahead as the naïve benchmark, i.e., $p_{s,24} = 1 - (MSE_s/MSE_{24})$, where MSE_s is the

[9] This is despite the fact that the trend component in inflation has become less predictable in recent years (see Stock & Watson, 2007). See also Mishkin (2007).

[10] However, Galbraith and Tkacz (2007) found that forecast content horizons do not improve even when dynamic factor models with many predictors are used in place of simple univariate autoregressive models.

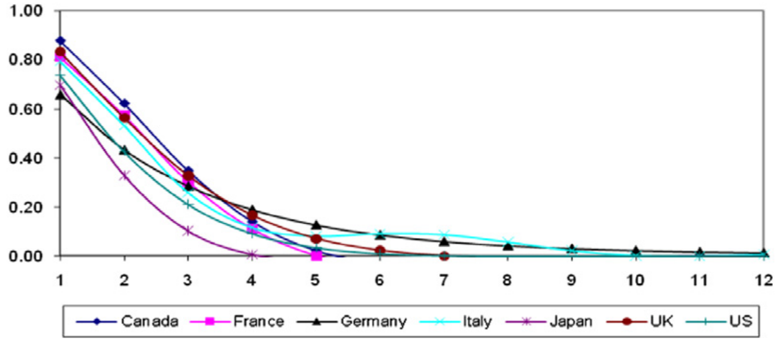


Fig. 4a. Real GDP predictability based on AR(p) models (quarterly horizons).

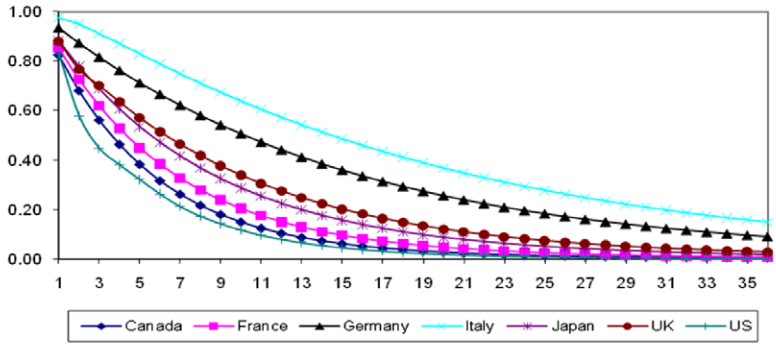


Fig. 4b. Inflation predictability based on AR(p) models (monthly horizons).

mean squared error for horizon $s = 1, 2, \ldots, 23$.[11] The measure of predictability $p_{s,24}$ indicates the improvement in the forecasts of a target variable at the forecast horizon s with respect to its predictability at the 24-month horizon as the horizon decreases. Thus, variables with different variances and forecast difficulties can be compared naturally using $p_{s,24}$. Large values of $p_{s,24}$ imply that forecasts made at horizon s improve significantly over the 24-month-ahead benchmark forecast.

Fig. 5 plots the statistic $p_{s,24}$ for GDP and inflation forecasts for all seven countries. It is clear that, for most countries, the inflation content function dominates that for real GDP, meaning that as the horizon shortens, useful information is absorbed in inflation forecasts more promptly. The dominance of inflation forecasts is especially noteworthy for Canada, France, Japan and the UK. We also find that the wedge is larger at the longer horizons, echoing earlier evidence that real GDP forecasts do not add any value during the first 6–8 rounds of forecasting. For inflation forecasts, however, each additional month increases the information content of the forecasts over that of the previous month, even at longer horizons. This provides additional evidence in support of the conclusion that real GDP is inherently more difficult to forecast than inflation, and shows that our professional forecasters have been more successful in processing the relevant information for predicting inflation than that for real GDP. As the horizon falls from $h = 24$ to $h = 1$, the mean squared error of the inflation forecasts decreases substantially, causing the inflation skill score $p_{s,24}$ to approach 100% at a faster rate than that of real GDP.

4. Understanding the individual forecasts

To understand more fully the relative forecasting records of real GDP and inflation, one has to explore their underlying expectations generating processes and recognize that the ability and willingness of individual forecasters to absorb new information at

[11] A similar measure was used by Öller and Teterukovsky (2007).

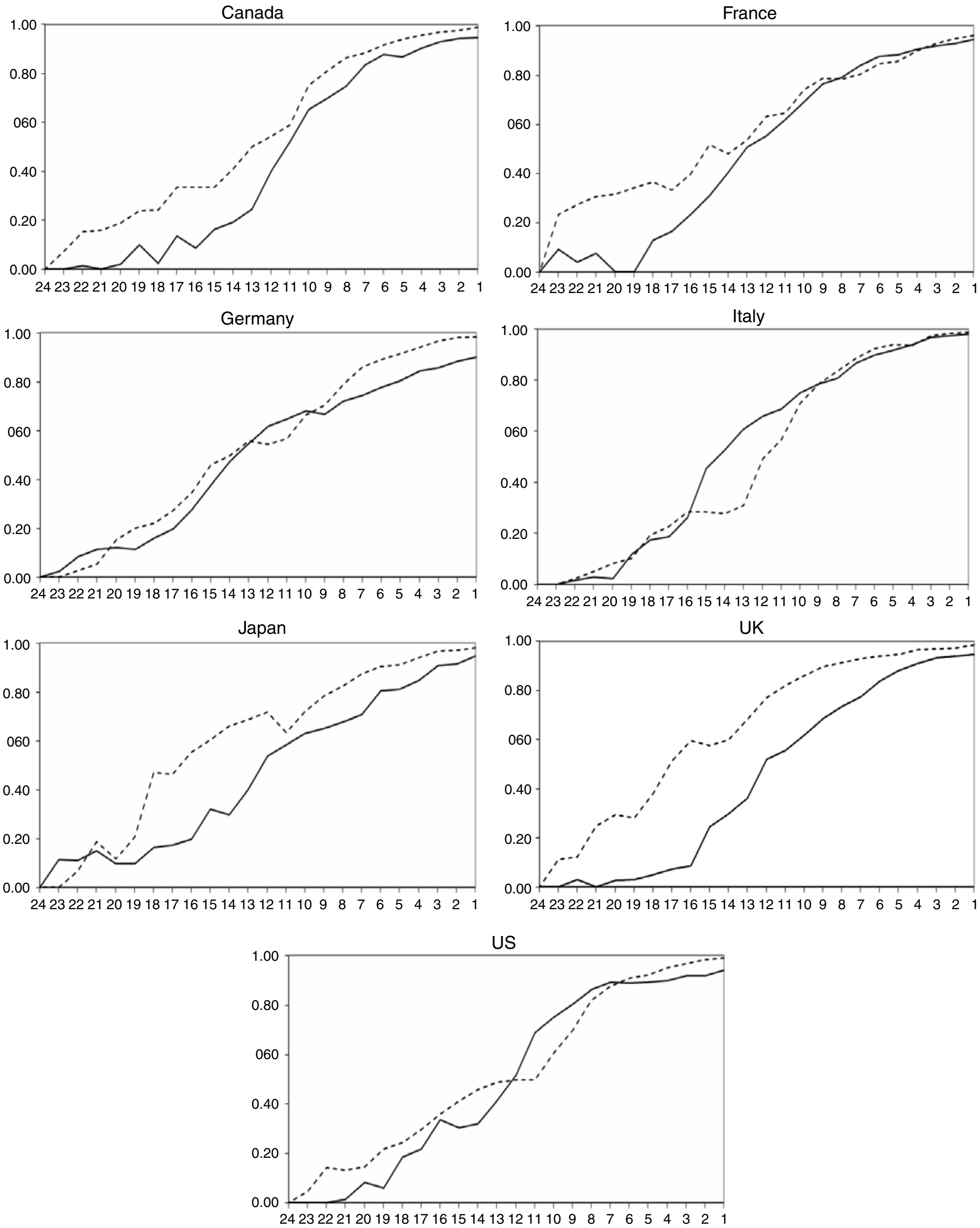


Fig. 5. Predictability of GDP (solid line) and inflation forecasts (dotted line) based on a real time information set.

different forecast horizons differ depending on the nature of the target variable. These possibilities are formally explored in this section, where we develop a simple Bayesian learning model aimed at identifying the relative importance of alternative pathways through which professional forecasters predict the term structure of forecasts, resulting in certain patterns of forecast accuracy.

4.1. The model

We have seen that professional forecasters do not seem to adjust their initial forecasts much in predicting real GDP, and that they sustain their initial disagreement during the initial rounds of forecasting.[12] We have also noted that at long horizons, consensus forecasts vary very little over time, since the idiosyncratic components in each year's forecasts cancel out while averaging over forecasters. Our Bayesian model explicitly recognizes these twin facts, and hypothesizes that professional forecasters begin forecasting with specific prior beliefs at the 24-month horizon.[13]

In thinking about why professional forecasters disagree regarding their long-run forecasts, note that a wealth of historical information on GDP and inflation are publicly available to all forecasters for estimating the long-run unconditional values of the series. Thus, it is not the availability of relevant data but the models, methods and philosophies used to interpret them that differ between forecasters. This is consistent with the finding of Döpke and Fritsche (2006) that forecasters do not share a common belief about what is an adequate model of the economy. Due to the length of the forecasting horizon, experts face very high levels of uncertainty in interpreting the available information based on whatever model or judgment they are using, and hence they disagree a lot about GDP and inflation in the long- or medium-run, see Zarnowitz and Lambros (1987).

Accordingly, we assume that the prior belief of the target variable for the year t, held by the forecaster i at the 24-month horizon, $\hat{F}_{it24}$, is represented by $\hat{F}_{it24} \sim N(F_{it24}, a_{it24}^{-1})$ for $i = 1, \ldots, N$, $t = 1, \ldots, T$, where F_{it24} and a_{it24} are the mean and the precision of agent i's prior belief, respectively.[14]

With the arrival of new public information, experts learn progressively to modify their initial beliefs over horizons. Consistent with our broad empirical findings on the fixed-target forecasts, we assume that at horizon h, forecasters receive a public signal L_{th} concerning the target variable, but that they may not all interpret it identically. In particular, individual i's estimate, Y_{ith}, of the target variable, conditional only on the new public signal that is observed at forecast horizon h, can be written as $Y_{ith} \sim N(L_{th} - \mu_{ith}, b_{ith}^{-1})$. Note that Y_{ith} is not observed.

This assumption allows for the possibility that agents can interpret the same public signal differently, which is captured by μ_{ith} with an associated uncertainty b_{ith}. Each month, all agents observe a new public signal, but they disagree on its effect on the target year. One expert may interpret the signal more optimistically or pessimistically than another. The precision of public information b_{ith} allows individual forecasters some latitude in interpreting public signals, and is a key parameter in generating expert disagreement and also forecast accuracy; see Acemoglu, Chernozhukov, and Yildiz (2006). This is in line with the empirical evidence presented above about a significant amount of between-agent variation in forecast revisions, and also with the large body of finance literature showing that equally informed agents can interpret the same information differently (cf. Dominitz & Manski, 2005; Kandel & Zilberfarb, 1999).

The Bayes rule implies that under the normality assumption, agent i's posterior mean is the weighted average of his prior mean and his estimate of the target variable, conditional only on the new public signal:

$$F_{ith} = \lambda_{ith} F_{ith+1} + (1 - \lambda_{ith})(L_{th} - \mu_{ith}), \tag{4}$$

with his posterior precision $a_{ith} = a_{ith+1} + b_{ith}$, where $\lambda_{ith} = a_{ith+1}/(a_{ith+1} + b_{ith})$ is the weight attached to prior beliefs.

For convenience, the following population parameters are defined across professional forecasters for the

[12] In the case of inflation, meaningful updating seems to begin at horizons slightly longer than 24 months.

[13] Even though 24-month-ahead forecasts are strictly medium-run forecasts, there is some evidence suggesting that these forecasts are in fact very close to being long-run forecasts. In Figs. 6a and 6b we have plotted 10-year forecasts of real GDP and inflation for the US, obtained from the Survey of Professional Forecasters (SPF) of the Philadelphia Federal Reserve Bank, against 24-month-ahead forecasts from our *Consensus Forecasts* database over the period 1992–2006. The corresponding disagreements are also reported. To match the timing of the forecasts, we compare the February forecasts of *Consensus Forecasts* with the first quarter forecasts from SPF (which are reported in the middle of the quarter). Even though, as expected, the variations in the 10-year long-run expectations are slightly muted, the two series are remarkably similar in terms of the mean values of the forecasts and the disagreement measures.

[14] The precisions of prior beliefs are allowed to differ across forecasters. This assumption is corroborated by recent studies using density forecasts that document the heterogeneity in forecast uncertainty, see for example Boero, Smith, and Wallis (2008), Bowles et al. (2007), and Lahiri and Liu (2006).

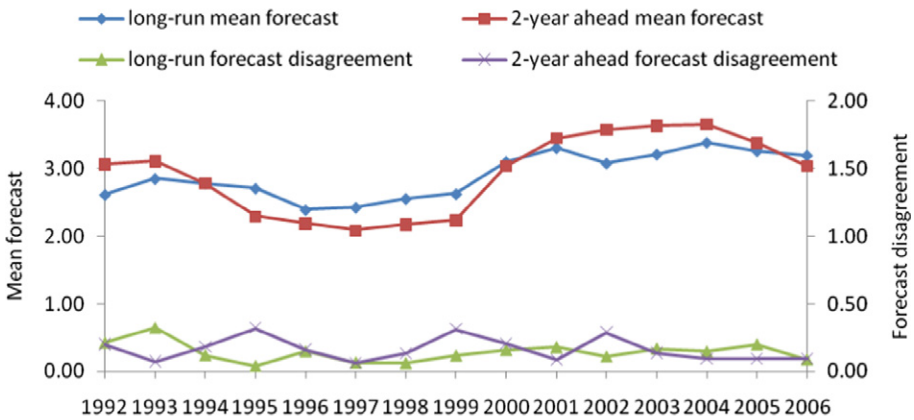


Fig. 6a. Evolution of the mean and disagreement in real GDP forecasts.

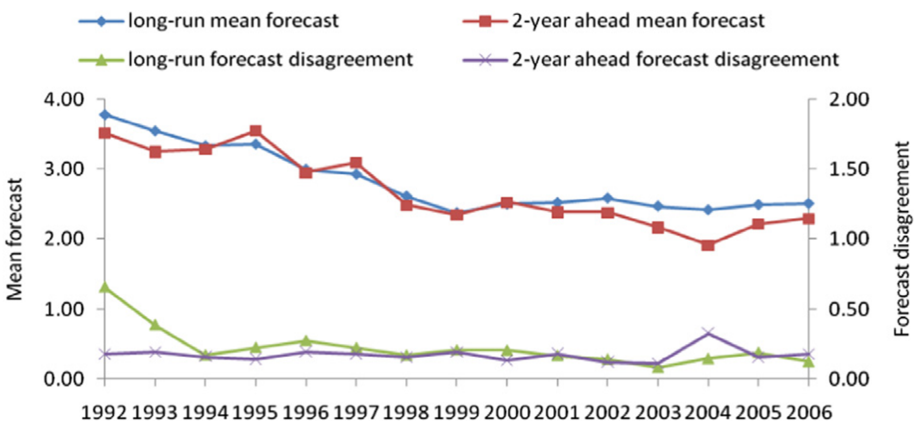


Fig. 6b. Evolution of the mean and disagreement in inflation forecasts.

target year t at horizon h:

$$E_i(F_{ith}) = F_{th}, \qquad \text{var}_i(F_{ith}) = \sigma^2_{F|th}; \tag{5}$$
$$E_i(\lambda_{ith}) = \lambda_{th}, \qquad \text{var}_i(\lambda_{ith}) = \sigma^2_{\lambda|th};$$
$$E_i(\mu_{ith}) = \mu_{th}, \qquad \text{var}_i(\mu_{ith}) = \sigma^2_{\mu|th}.$$

Since we expect that the prior mean F_{ith+1} will be independent of the prior precision a_{ith+1}, we can safely assume that F_{ith+1}, λ_{ith} and μ_{ith} are mutually independent of each other for any t and h. Lahiri and Sheng (2008) derived the following relationship between the disagreements in two consecutive rounds of fixed-target forecasting:

$$\begin{aligned}\sigma^2_{F|th} &= \sigma^2_{F|th+1}(\sigma^2_{\lambda|th} + \lambda^2_{th}) \\ &\quad + \sigma^2_{\mu|th}[\sigma^2_{\lambda|th} + (1-\lambda_{th})^2] \\ &\quad + \sigma^2_{\lambda|th}[\Delta F_{th}/(1-\lambda_{th})]^2,\end{aligned} \tag{6}$$

where $\Delta F_{th} = F_{th} - F_{th+1}$. In Eq. (6) the dynamics of the forecast disagreement over forecast horizons are seen to be governed by three parameters of the model, representing across-forecaster differences in (i) prior beliefs, $\sigma^2_{F|th+1}$; (ii) the weights attached to priors, $\sigma^2_{\lambda|th}$; and (iii) the interpretation of public signals, $\sigma^2_{\mu|th}$. The equation encompasses a number of special cases. In the case where all agents attach the same weight to their prior beliefs relative to the reliability of the sample information (i.e. $\sigma^2_{\lambda|th} = 0$ for any t and h), Eq. (6) becomes

$$\sigma^2_{F|th} = \lambda^2_{th}\sigma^2_{F|th+1} + (1-\lambda_{th})^2\sigma^2_{\mu|th}. \tag{7}$$

4.2. Estimation

Let us first focus on estimating one of our structural parameters — the weight attached to a prior belief relative to the reliability of the sample information. Eq. (4) shows that

$$F_{ith} = \lambda_{ith} F_{ith+1} + \varepsilon_{ith}, \tag{8}$$

where $\varepsilon_{ith} = (1 - \lambda_{ith})(L_{th} - \mu_{ith})$ is the error term. By construction, ε_{ith} and F_{ith+1} are independent for any t and h. In estimating the above equation, several econometric issues arise.

First, note that Eq. (8) cannot be estimated, since the number of parameters to be estimated exceeds the number of observations. We assume that

$$\lambda_{ith} = \lambda_{ih} = \lambda_h + v_{ih}, \tag{9}$$

where v_{ih} have a mean of zero, are mutually independent of each other, and are independent over forecast horizons. We regress the forecast revision (ΔF_{ith}) on the lagged forecast (F_{ith+1}) to circumvent the possible problem of spurious regression. Thus, the version of Eq. (8) to be estimated becomes

$$\Delta F_{ith} = \beta_h F_{ith+1} + u_{ith}, \tag{10}$$

where $\beta_h = \lambda_h - 1$ and $u_{ith} = \varepsilon_{ith} + v_{ih} F_{ith+1}$.

Second, u_{ith} will be correlated across forecasters, because, conditional on t and h, it has a nonzero mean that depends on L_{th}. To solve this problem, we rewrite Eq. (10) as

$$\varepsilon_{ith} = \Delta F_{ith} - \beta_h F_{ith+1} - v_{ih} F_{ith+1}. \tag{11}$$

Taking the expectations of Eq. (11) over i conditional on t and h, we get

$$E_i(\varepsilon_{ith}) = \Delta F_{th} - \beta_h F_{th+1}. \tag{12}$$

Subtracting Eq. (12) from Eq. (11), we obtain

$$\Delta F_{ith} - \Delta F_{th} = \beta_h (F_{ith+1} - F_{th+1}) + w_{ith}, \tag{13}$$

where $w_{ith} = \varepsilon_{ith} - E_i(\varepsilon_{ith}) + v_{ih} F_{ith+1}$. In contrast to u_{ith} in Eq. (10), the error w_{ith} has a zero mean.

Third, it may seem desirable to estimate the panel data model in Eq. (13) with all three dimensions by imposing a smooth functional form for β_h over horizons, as was done by Gregory and Yetman (2004). However, as is shown later, the estimated β_h varies unevenly over horizons, depending on the lumpiness and timing of the public information arrival.

Finally, w_{ith} might be serially correlated. Let w_{ith} follow the AR(1) process $w_{ith} = \rho_h w_{ith+1} + \eta_{ith}$; then Eq. (13) can be rewritten as

$$\begin{aligned} \Delta F_{ith} - \Delta F_{th} &= \beta_h (F_{ith+1} - F_{th+1}) \\ &+ \rho_h (\Delta F_{ith+1} - \Delta F_{th+1}) \\ &- \rho_h \beta_h (F_{ith+2} - F_{th+2}) + \eta_{ith}, \end{aligned} \tag{14}$$

where $E(\eta_{ith}\eta_{i't'h'}) = \sigma^2_{\eta(i)}$ for $i = i', t = t', h = h'$ and 0 otherwise. Using nonlinear least squares, we estimate Eq. (14) for each horizon after controlling for the heterogeneity in the error term.

Tables 1a and 1b present the estimated weights attached to public information for the GDP and inflation forecasts, respectively. In predicting both real GDP and inflation, forecasters give a lower weight to public information at longer horizons because of its low perceived quality, and a higher weight at short horizons as the information becomes more precise. At longer horizons, the priors are relatively more important. Another important observation is that, on average over all horizons, professional forecasters attach a higher weight to public information in predicting inflation than in predicting GDP. Recall from Eq. (4) that the relative weight attached to public information is a function of the precisions of new information and priors, i.e. $1 - \lambda_{ith} = b_{ith}/(a_{ith+1} + b_{ith}) = (b_{ith}/a_{ith+1})/(1 + b_{ith}/a_{ith+1})$. It therefore follows that the ratio of the precision of new information to the precision of prior belief, b_{ith}/a_{ith+1}, is higher, and thus public information is perceived to be more precise and certain in predicting inflation than GDP. This finding could possibly be explained by the fact that initial GDP announcements are revised more heavily than price indexes, are observed only quarterly, and involve substantial measurement errors. The repeated arrival of substantial real GDP revisions in real time makes all "news" related to GDP less precise, and hence GDP forecasts tilt away from the proper use of current information. On the other hand, the retail price index for the UK is never revised after its initial release, and hence it will have no such bias. Also, the more frequent communication of the latest inflationary developments to the general public and the commitment to long-run price stability by central banks may make adjustments to inflationary expectations dependent more on current news, and less on priors.

Table 1a
Estimated weights attached to public information in GDP forecasts.

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	0.52	0.73	0.35	0.52	0.46	0.49	0.38
	(0.06)	(0.32)	(0.04)	(0.06)	(0.08)	(0.03)	(0.03)
2	0.46	0.24	0.30	0.18	0.24	0.38	0.41
	(0.06)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
3	0.39	0.45	0.34	0.24	0.46	0.31	0.41
	(0.04)	(0.05)	(0.04)	(0.05)	(0.06)	(0.03)	(0.05)
4	0.40	0.20	0.30	0.18	0.18	0.26	0.28
	(0.08)	(0.04)	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)
5	0.24	0.11	0.22	0.23	0.14	0.23	0.37
	(0.03)	(0.03)	(0.04)	(0.05)	(0.03)	(0.02)	(0.03)
6	0.32	0.54	0.18	0.38	0.23	0.19	0.23
	(0.04)	(0.08)	(0.03)	(0.05)	(0.03)	(0.03)	(0.02)
7	0.31	0.32	0.21	0.28	0.13	0.19	0.21
	(0.05)	(0.09)	(0.03)	(0.06)	(0.04)	(0.02)	(0.02)
8	0.17	0.12	0.23	0.16	0.09	0.19	0.34
	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)	(0.02)	(0.03)
9	0.16	0.25	0.15	0.15	0.19	0.13	0.21
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.02)	(0.03)
10	0.22	0.14	0.22	0.20	0.10	0.15	0.15
	(0.05)	(0.04)	(0.03)	(0.05)	(0.03)	(0.02)	(0.03)
11	0.14	0.13	0.14	0.21	0.01	0.07	0.24
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
12	0.17	0.13	0.16	0.20	0.23	0.15	0.19
	(0.04)	(0.03)	(0.03)	(0.05)	(0.03)	(0.02)	(0.03)
13	0.23	0.14	0.23	0.26	0.20	0.17	0.17
	(0.05)	(0.03)	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)
14	0.12	0.05	0.28	0.24	0.03	0.10	0.12
	(0.03)	(0.03)	(0.03)	(0.06)	(0.02)	(0.02)	(0.03)
15	0.22	0.21	0.14	0.27	0.16	0.13	0.12
	(0.05)	(0.04)	(0.03)	(0.07)	(0.04)	(0.03)	(0.03)
16	0.05	0.20	0.13	0.15	0.04	0.08	0.13
	(0.04)	(0.04)	(0.03)	(0.05)	(0.03)	(0.02)	(0.03)
17	0.17	0.07	0.09	0.21	0.05	0.04	0.09
	(0.03)	(0.02)	(0.03)	(0.05)	(0.03)	(0.02)	(0.02)
18	0.09	0.14	0.15	0.23	0.04	0.08	0.16
	(0.02)	(0.04)	(0.03)	(0.05)	(0.03)	(0.02)	(0.03)
19	0.08	0.17	0.15	0.21	0.10	0.06	0.05
	(0.03)	(0.05)	(0.02)	(0.05)	(0.03)	(0.02)	(0.02)
20	0.15	0.12	0.13	0.12	0.07	0.05	0.10
	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)
21	0.07	0.12	0.19	0.13	0.09	0.10	0.03
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)
22	0.08	0.08	0.08	0.09	0.17	0.05	0.13
	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)	(0.03)

Note: Standard errors are in parentheses.

4.3. The role of prior beliefs

Recall that forecast disagreement is posited to have three components (see Eq. (6)). Lahiri and Sheng (2008) find the second component, i.e. differences in the weights attached by different experts to their prior beliefs, to have barely any effect on GDP

Table 1b
Estimated weight attached to public information in inflation forecasts.

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	0.58	0.36	0.42	0.55	0.41	0.53	0.49
	(0.08)	(0.05)	(0.05)	(0.09)	(0.06)	(0.04)	(0.05)
2	0.37	0.41	0.52	0.37	0.39	0.27	0.57
	(0.07)	(0.05)	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)
3	0.74	0.65	0.45	0.44	0.52	0.40	0.33
	(0.05)	(0.08)	(0.04)	(0.05)	(0.06)	(0.04)	(0.04)
4	0.42	0.31	0.44	0.27	0.33	0.65	0.45
	(0.07)	(0.05)	(0.04)	(0.05)	(0.05)	(0.17)	(0.04)
5	0.50	0.24	0.42	0.26	0.24	0.34	0.47
	(0.07)	(0.06)	(0.05)	(0.06)	(0.03)	(0.06)	(0.05)
6	0.37	0.31	0.40	0.33	0.43	0.33	0.34
	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)
7	0.34	0.31	0.36	0.18	0.29	0.35	0.31
	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)	(0.03)	(0.03)
8	0.33	0.23	0.33	0.15	0.28	0.31	0.39
	(0.05)	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)
9	0.23	0.48	0.26	0.31	0.42	0.27	0.27
	(0.04)	(0.08)	(0.04)	(0.05)	(0.07)	(0.02)	(0.03)
10	0.26	0.29	0.27	0.29	0.12	0.34	0.19
	(0.06)	(0.04)	(0.04)	(0.06)	(0.03)	(0.07)	(0.03)
11	0.22	0.19	0.22	0.24	0.14	0.14	0.24
	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.02)	(0.03)
12	0.30	0.25	0.19	0.34	0.22	0.27	0.18
	(0.05)	(0.05)	(0.03)	(0.07)	(0.04)	(0.02)	(0.03)
13	0.10	0.24	0.24	0.23	0.28	0.31	0.13
	(0.04)	(0.06)	(0.03)	(0.04)	(0.05)	(0.05)	(0.03)
14	0.14	0.13	0.25	0.25	0.19	0.07	0.24
	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)	(0.02)	(0.03)
15	0.12	0.31	0.23	0.53	0.25	0.10	0.15
	(0.03)	(0.05)	(0.03)	(0.22)	(0.05)	(0.02)	(0.03)
16	0.14	0.19	0.16	0.17	0.15	0.16	0.14
	(0.05)	(0.04)	(0.02)	(0.06)	(0.03)	(0.02)	(0.02)
17	0.16	0.12	0.16	0.20	0.05	0.12	0.14
	(0.03)	(0.03)	(0.03)	(0.04)	(0.02)	(0.02)	(0.03)
18	0.15	0.22	0.11	0.33	0.19	0.08	0.11
	(0.02)	(0.05)	(0.03)	(0.06)	(0.05)	(0.02)	(0.02)
19	0.17	0.21	0.16	0.06	0.13	0.12	0.12
	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)	(0.02)	(0.02)
20	0.04	0.08	0.14	0.05	0.07	0.07	0.08
	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)
21	0.10	0.24	0.07	0.09	0.22	0.11	0.11
	(0.03)	(0.05)	(0.02)	(0.03)	(0.07)	(0.02)	(0.02)
22	0.08	0.09	0.13	0.12	0.07	0.05	0.06
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)	(0.02)

Note: Standard errors are in parentheses.

forecast disagreement, since professional forecasters place very similar weights on their prior beliefs.[15] We thus maintain a more parsimonious model in which forecast disagreement arises from two possible sources: differences in forecaster's prior beliefs, and differences in their interpretation of public information, as in Eq. (7).

15 This component, however, might account for a large part of the disagreement in laymen's expectations.

Substituting $\hat{\lambda}_h$ into Eq. (7), we get estimates of the heterogeneity parameter in the interpretation of public signals, $\sigma^2_{\mu|h}$, as the sample average of $(\sigma^2_{F|th} - \hat{\lambda}^2_h \sigma^2_{F|th+1})/(1-\hat{\lambda}_h)^2$ over the target years. Note that differences in the interpretations of public information only affect forecast disagreement through its interaction with the weight attached to public information.

With the parameter estimates in hand, we can check how well the disagreement predicted by our model matches the disagreement observed in the survey data. Substituting the parameter estimates of λ_h and $\sigma^2_{\mu|h}$ into

$$\sigma^2_{F|th} = \prod_{j=h}^{23} \lambda^2_j \sigma^2_{F|t24} + (1-\lambda_h)^2 \sigma^2_{\mu|th} + \sum_{j=h+1}^{23} \left(\prod_{s=h}^{j-1} \lambda^2_s \right) \left(1-\lambda_j\right)^2 \sigma^2_{\mu|j}, \qquad (15)$$

we get the dynamically generated forecast disagreement at each horizon that is predicted by our model.

We find that, depending on the country, our estimated model explains between about 20% and 56% of the total variation in the observed GDP forecast disagreement over all target years and horizons. The corresponding figures for the inflation forecasts are much higher, ranging from 40% to 74%.[16] It is interesting to note that, using a dynamic structural time series model with measurement errors and assuming forecast efficiency, Patton and Timmermann (2007) could successfully mimic the dispersion in the term structure of US real GDP forecasts, but could not mimic that of inflation at short horizons. We only face a similar problem for Italy's inflation forecast dispersion at very short horizons. As is found in Section 5 below, this can be explained by the fact that forecasters in Italy overweight public information at these short horizons. Considering the fact that the forecast disagreement varies a lot from year to year for any specific horizon due to various exogenous factors (e.g., recessions, 9/11, Katrina, policy actions, etc.), and that our theoretical model is meant to explain only the term structure of forecasts, the estimated model does a good job of explaining the evolution of the disagreement over target years and horizons — though admittedly more so for the inflation forecasts than for the GDP forecasts.

The contribution of the heterogeneity in (updated) prior beliefs to explaining the disagreement in GDP and inflation forecasts is presented in Fig. 7. With a few exceptions, the diversity in their priors plays a larger role in explaining expert disagreement in forecasting GDP than in forecasting inflation. As expected, the importance of the prior beliefs declines steadily as the forecast horizons get shorter. However, even at the end of the forecasting rounds, 1 month ahead, the diversity in the updated priors still explains about 14%–47% of GDP and 25%–38% of inflation forecast disagreement. This finding firmly establishes the role of heterogeneity in prior beliefs in generating inter-personal differences in individual forecasts over the whole term structure. Patton and Timmermann (2007) also established the role of priors in their study of disagreement in US GDP and inflation forecasts, but found little effect of differential information.

To formally see the role of the priors, we iterate Eq. (4) backwards to get

$$E(Y_t \mid F_{ith+1}, L_{th}) = \prod_{j=h}^{23} \lambda_{itj} F_{24} + (1-\lambda_{ith})(L_{th} - \mu_{ith}) + \sum_{j=h+1}^{23} \left(\prod_{s=h}^{j-1} \lambda_{its} \right) (1-\lambda_{itj})(L_{tj} - \mu_{itj}). \qquad (16)$$

In Eq. (16) the optimal forecast made at horizon h is a weighted average of three components: the prior beliefs, current public information, and all past public information. The prior belief causes expectation stickiness in two ways. First, it enters directly into the current forecast and is propagated forward into the whole series of forecasts for the target year, though its importance declines over forecast horizons. This is consistent with the findings of Batchelor (2007) that biases due to optimism or pessimism in the priors persist throughout the forecasting cycle. Second, it allows all past public information to affect the current forecast in a staggered way. Without the role of prior

[16] For GDP forecasts, our model explains about 56%, 29%, 50%, 20%, 27%, 54% and 50% of the total variation in observed forecast disagreement over the target years and horizons for Canada, France, Germany, Italy, Japan, the UK and the US, respectively. The corresponding figures for the inflation forecasts are 59%, 40%, 68%, 52%, 56%, 74% and 72%, respectively.

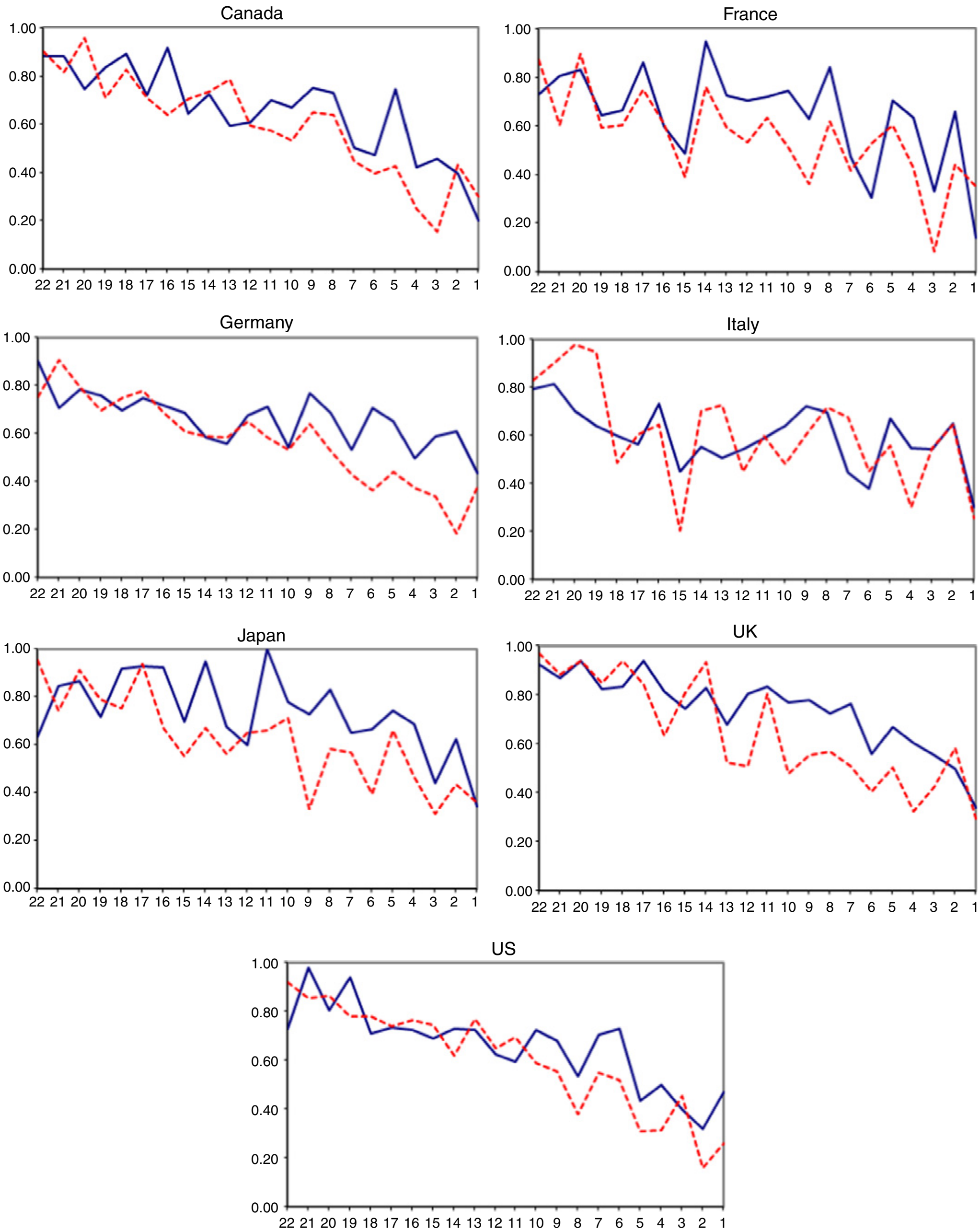


Fig. 7. Contribution of differences in prior beliefs to explaining GDP (solid line) and inflation (dotted line) forecast disagreement.

beliefs (i.e. $\lambda_{ith} = 0$ for all h), the current forecast reflects only the latest information about the target variable. Thus, a stickiness of expectations in itself does not necessarily contradict the forecast efficiency hypothesis. Instead, the Bayesian learning model allows for a certain amount of inertia in expectations, and thus offers an additional cue to the ongoing discussion on the micro foundation of expectation stickiness (cf. Mankiw & Reis, 2006; Morris & Shin, 2006).

4.4. The role of heterogeneity in the interpretation of new public information

Apart from the diversity in prior beliefs, a second factor that explains the forecast disagreement in our model is the heterogeneity in the interpretation of new public information by experts. As Tables 1a and 1b reveal, the latter pathway becomes increasingly more important at shorter horizons, and provides evidence in support of the hypothesis that equally informed agents can sometimes interpret the same public information differently. In Section 4.4.1, we present a case study of the 9/11 terrorist attack on the US that will firmly establish the role of this channel in generating expert disagreement. In Section 4.4.2, we present another interesting case study on the Italian inflation targeting regime, where the monetary authority successfully reduced inflation forecast disagreement by first anchoring the long-term expectations within a very narrow range and then limiting the heterogeneity in the interpretation of incoming news over the term structure of forecasts.

4.4.1. The impact of the 9/11 terrorist attack on forecast disagreement: A case study

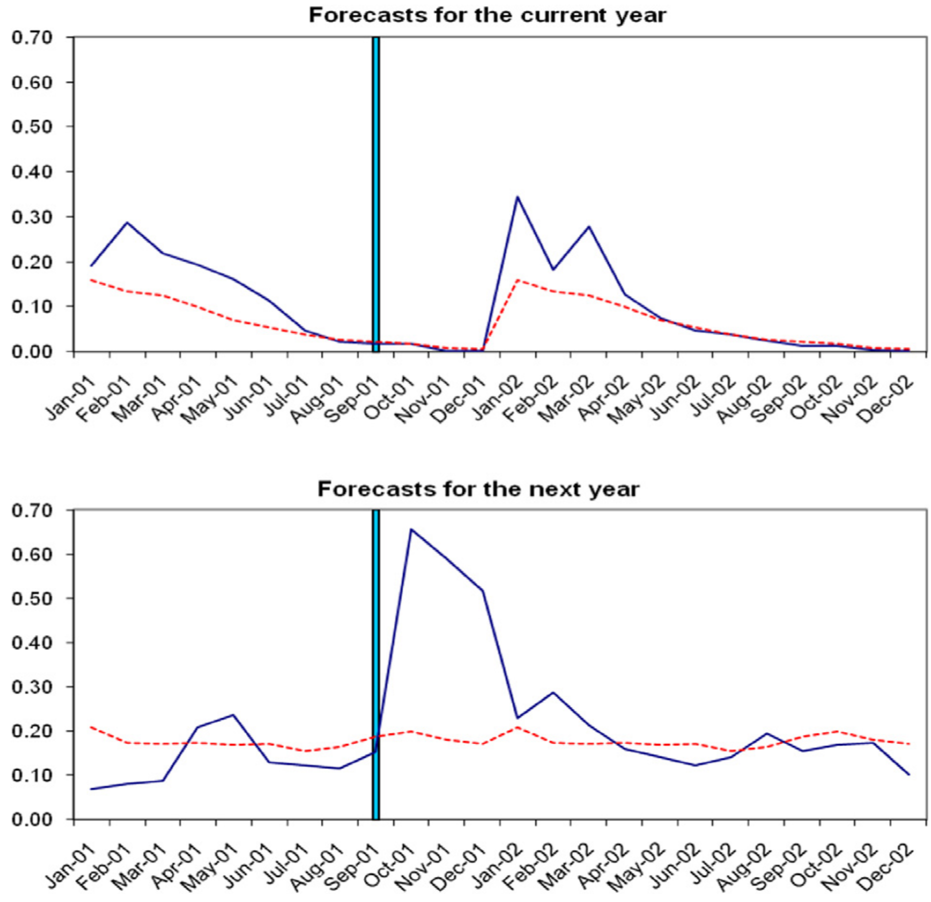
As was mentioned in Section 2, we expect private information about GDP and inflation to be of limited importance relative to public information. However, we cannot rule out the possibility of the simultaneous arrival of public and private information. In this section, we study the evolution of forecast disagreement in the aftermath of the September 11, 2001, terrorist attack on the US. This event provides an opportunity to establish the importance of the differential interpretation of public information in generating forecast disagreement, where any confounding role of either prior beliefs or private information can be ruled out. Patton and Timmermann (2007) also looked at the evolution of consensus forecasts around 9/11, but with a different purpose.

Fig. 8 plots the effect of 9/11 on the evolution of GDP forecast disagreement. The horizontal axis shows the month/year when the forecasts were made. The upper and lower panels trace (solid lines) the disagreement in experts' forecasts made for the current year and the next year, respectively, in different months from January 2001 to December 2002. Since disagreement, *ceteris paribus*, is higher for longer horizon forecasts, we have also plotted the average disagreement (dotted lines) over the period 1991–2000 for each monthly horizon, for the purpose of a benchmark comparison. Thus, the effect of the 9/11 attack on disagreement will be the vertical difference between the solid and dotted lines.

Let us focus first on forecast disagreement in predicting the current year's GDP growth for 2001 (upper panel). Prior to 9/11, expert disagreement was a little higher than the ten-year historical average, possibly due to the recession that started in March 2001. After 9/11, however, the disagreement did not increase during the October–December 2001 forecasts. There are two obvious reasons for this. First, since we are considering the current-year GDP growth, with only three months of the year remaining, even a big shock can only have a limited effect on the current year's growth. Second, the total impact of a shock is sure to be distributed over time, and three months is too short a period to capture the total impact. Thus, when the horizon is very short, the impact of an unexpected shock on forecast disagreement will be similarly small. Turning to the forecast disagreement in predicting the current year's GDP growth for 2002, however, we find additional disagreement during the period January–May 2002, relative to the historical values of these months. The disagreement doubled to 0.34 in January 2002, and it then took 4 more months for the disagreement to get back to its historical level. Note that during this period, the consensus real GDP forecast increased from 0.9% in January 2002 to 2.8% in May 2002, as the economy was recovering from the recession.

The lower panel is more interesting, and plots the forecast disagreement in predicting the next year's GDP growth rates for 2002 and 2003. The disagreement was remarkably close to the historical average prior to 9/11, despite the recession. The disagreement then more than tripled to 0.60 in October 2001, and stayed high relative to the historical average until January 2002.[17] However, after another few months the disagreement quickly fell back to the historical level, suggesting that the impact of a shock on forecast disagreement is also small when the

[17] It is interesting to note that we did not find any significant impact of 9/11 on the evolution of the consensus forecast and disagreement on inflation forecasts.



Note: The horizontal axis shows the month/year when the forecasts were made. The graph plots forecast disagreement for the target year (solid line) against a benchmark disagreement averaged over the period 1991–2000 (dotted line).

Fig. 8. Effect of 9/11 on the evolution of the disagreement in US GDP forecasts.

horizon is very long. The revisions to the next year's growth forecasts were just the opposite of that for the disagreement during the period October 2001–May 2002 — the growth forecasts were downgraded as the disagreement rose, and vice versa. Our results suggest that an unanticipated shock tends to have the maximum impact on the yearly GDP forecasts and dispersions if it comes during the middle horizons when there are 10 to 14 months remaining until the end of the target year. The extra disagreement then takes about 4–5 months to dissipate to its historical levels. Isiklar and Lahiri (2007) and Isiklar, Lahiri, and Loungani (2006) found very similar results on the response pattern of mean forecasts to shocks using a VAR analysis.

There are two antecedents to the present case study. Mankiw et al. (2003) studied the evolution of the forecast distribution as a part of household learning after a regime change due to the Volker disinflation policy during the period 1979–82. In another classic paper, Kandel and Pearson (1995) established the importance of heterogeneity in the interpretation of public information by looking at analysts' forecasts before and after earnings announcements. However, they could not rule out the possibility of the simultaneous arrival of private and public information about the value of the announcement. We circumvent this problem by looking at a completely unanticipated but universally observed common shock. The only reason for experts disagreeing in this case is that they used different models and methods, and each interpreted the effect of this event on the economy differently. A differential interpretation of public information can be a great challenge in establishing the credibility and effectiveness of monetary policies — an issue that we examine more carefully in the next section.

4.4.2. Italy under inflation targeting: Another case study

In 1998, the Governing Council of the ECB interpreted the Maastricht Treaty as a mandate to maintain price inflation close to 2% over the medium term. In recent years, a number of studies have concluded that, due to the official inflation targeting policies of the central banks in Europe and Canada, long-run inflation expectations have become more anchored in these countries than in the United States. Beechey, Johannsen, and Levin (2007) used survey data from the ECB and the Philadelphia Federal Reserve Bank to show that over the period 2000–2006, the disagreement in long-run inflation expectations was lower in the euro area than in the US. Note that our model implies that the effect of inflation targeting on prior beliefs will be transmitted to expert disagreement over the whole term structure of forecasts via the Bayesian updating process. Since Italy's performance in achieving price stability in recent years has been particularly noteworthy, we estimated the Bayesian learning model parameters before and after the successful implementation of inflation targeting using Italian forecasts.[18] Fig. 9a clearly shows that since 1997, there has been a sharp and permanent decline in the 24-month-ahead inflation forecasts and disagreement in Italy. Thus, we split the sample into the periods 1991–1997 and 1998–2007, and estimate the parameters using the pre- and post-inflation target regimes.

As expected, Fig. 9b shows that forecast disagreement is markedly lower at all horizons after 1997. The estimates of the relative weights attached to incoming news are given in Fig. 9c, where we find that, on average, agents attach more importance to current news than to priors under inflation targeting. Thus, the enhanced communication strategy of the ECB under inflation targeting, combined with the ECB's credibility, has made new information more dependable in Italy. However, note that at certain horizons the updated prior becomes temporarily more important. This is because, as Fig. 5b shows, forecasters do not update the new information every month by the same amount, and during the months of relative inactivity in forecast revisions, the prior becomes relatively more important. Finally, Fig. 9d shows how inflation targeting affects another parameter of our model — the difference across forecasters in the interpretation of new information. Clearly, the disagreement due to heterogeneity in this parameter has been reduced significantly at all horizons in the post-1997 period. Thus, this case study of Italy shows how inflation targeting not only reduces the variability of long-run expectations, but also limits the heterogeneity in how experts interpret new information, resulting in reduced disagreement and aggregate forecast errors throughout the term structure of the forecasts. The existing literature on inflation targeting has only considered the effect of targeting on the reduced volatility of long-term mean expectations around the desired value.[19] Johnson (2002), who at least recognized the possibility of a differential interpretation of public information being a source of forecast disagreement, is an exception.

5. Forecast efficiency

We find that professional forecasters make smaller forecast errors in predicting inflation than real GDP, which can partly be explained by the underlying data

[18] As one referee pointed out, the introduction of inflation targeting was a big policy change, but other exogenous changes in Italy pre- and post-1997 cannot be ruled out as being factors behind the price stability.

[19] Beechey et al. (2007) and Gürkaynak et al. (2007) studied the sensitivity to new information of the inflation compensation embedded in inflation swaps and in the yield spread between long maturity nominal and inflation-indexed government bonds. They found that macroeconomic surprises affect inflation compensation in the US but not in the inflation targeting countries. For our purpose, note that these studies report a substantial variability of inflation compensation, much of which remains unexplained in both inflation-targeting and non-targeting countries. This unexplained variability in the implied long-run inflation expectations can be justified in terms of forecasters having prior distributions on long-run expectations with low levels of precision.

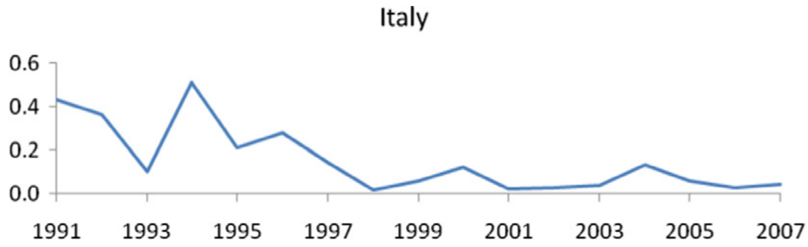


Fig. 9a. Disagreement in 24-month-ahead inflation forecasts over time.

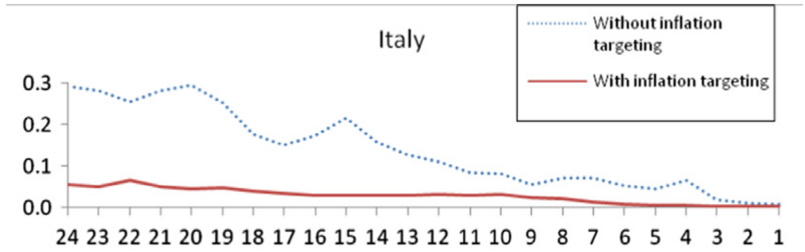


Fig. 9b. Disagreement in inflation forecasts over horizons.

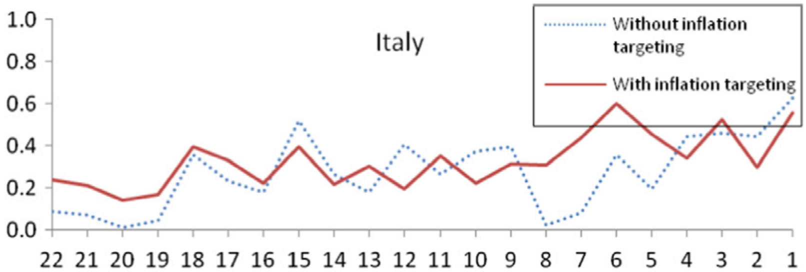


Fig. 9c. Weights attached to public information.

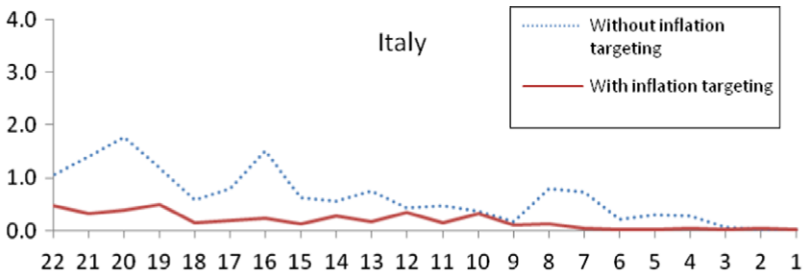


Fig. 9d. Variance across forecasters in interpreting public information.

generating processes. Our analysis also reveals that they put more weight on new information when forecasting inflation than when forecasting real GDP. The question is, are these weights efficient? The weights will be inefficient if the forecasters perceive the relative preciseness of new information incorrectly, coupled by an undue confidence in their own prior beliefs. One possible explanation for the relative superiority of the inflation forecasts is that forecasters use public information more efficiently when forecasting inflation than when forecasting real GDP. This possibility is explored formally in this section.

Following Eq. (4), forecaster i's estimate of the target variable Y_t, conditional on F_{ith+1} and L_{th}, is given by

$$E(Y_t \mid F_{ith+1}, L_{th}) = \lambda_{ith} F_{ith+1} + (1 - \lambda_{ith})(L_{th} - \mu_{ith}). \tag{17}$$

Zellner (1988) has shown that the above Bayesian information updating rule is 100% efficient, since no information is either lost or added when Eq. (17) is employed. Thus, the weight attached to the prior belief, $\lambda_{ith} = a_{ith+1}/(a_{ith+1} + b_{ith})$, is the efficient weight. However, for various reasons, forecasters may not be able to perceive the relative precision of the incoming information relative to the prior, and may fail to apply the efficient weight λ_{ith} when making forecasts. For simplicity, the forecast made by agent i for the target year t, h months before the end of the target year, is assumed to also have the form

$$F_{ith} = \delta_{ith} F_{ith+1} + (1 - \delta_{ith})(L_{th} - \mu_{ith}), \tag{18}$$

where δ_{ith} is the actual weight that forecaster i attached to his prior belief. We observe that forecaster i underweights public information if $\delta_{ith} > \lambda_{ith}$.

Combining Eqs. (17) and (18), Lahiri and Sheng (2008) derived a new test for forecast efficiency under the Bayesian learning framework. Their formulation of the efficiency test builds on the relationship between forecast error and forecast revision as:

$$E(Y_t - F_{ith} \mid F_{ith+1}, L_{th}) = \theta_{ith}(F_{ith} - F_{ith+1}), \tag{19}$$

where $\theta_{ith} = (\delta_{ith} - \lambda_{ith})/(1 - \delta_{ith})$. Under the null hypothesis that forecasters use efficient weights (i.e., $\delta_{ith} = \lambda_{ith}$), θ_{ith} should be zero. Since δ_{ith} lies between 0 and 1, a positive (negative) value of θ_{ith} suggests that public information is being underweighted (overweighted). The intuition behind the relationship is straightforward. Whereas forecast revisions can be taken as a measure of the way in which forecasters interpret the importance of public information in real time, forecast errors are the *ex post* "prize" that they get as a result of revising their forecasts. Suppose that forecasters make large revisions at horizon h but that the performances of the forecasts do not improve much at that horizon; one may then conjecture that the forecasters are overweighting new public information.

To perform the test, we check whether $\theta_h = 0$ in the regression for any specific horizon h:

$$Y_t - F_{ith} = \alpha_h + \theta_h(F_{ith} - F_{ith+1}) + \varepsilon_{ith}. \tag{20}$$

Following the method of Lahiri and Sheng (2008), we estimate the coefficients in Eq. (20) using GMM, controlling for both cross-sectional correlation and serial correlation in the residuals, and using the appropriate weighting matrix. The estimation results are shown in Tables 2a and 2b for the GDP and inflation forecasts, respectively. Although many of the estimates are not close to zero (particularly for real GDP), given the standard errors of the estimates, the test fails to reject the null hypothesis of forecast efficiency for more than half of the horizons and countries. However, the evidence also indicates that there is significant forecast inefficiency for some countries and horizons. We address this issue below.

For the GDP forecasts, we note the following. First, forecasters seem to put more than the efficient weight on new public information at very long horizons, given the many statistically significant and negative coefficient estimates. Since we find that the public signals concerning next year's GDP growth are not very informative during the initial eight monthly rounds of forecasting, experts are found to make unnecessary but mostly small revisions during this period.[20] Second, we find that forecasters underweight public information fairly substantially at the middle horizons. Since, as was shown by Isiklar and Lahiri (2007), a unit shock has the maximum effect on a target variable at the middle horizons, this under-weighting of public information turns out to be very significant. As the horizon gets shorter, the base-year GDP growth numbers become available with increasing certainty. Furthermore, as we approach the end of the target year, current-year GDP announcements and data revisions become part of the target-year GDP growth. As a result, forecasters should put a higher weight on the newly arrived public information. The degree of underweighting of public information, however, is largest for Canada, France and Germany. This finding, based on individual forecasts, complements the recent empirical evidence presented by Isiklar et al. (2006).

[20] While analyzing British fixed-target forecasts with horizons pooled up to 12 quarters, Clements (1995) also found negative autocorrelations in forecast revisions and interpreted them as evidence of an absence of significant news over the period.

Table 2a
Test of efficiency in the use of public information in GDP forecasts.

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	−0.17	0.01	0.11	−0.10	0.21*	−0.18*	0.61*
	(−1.37)	(0.11)	(0.76)	(−1.04)	(2.29)	(−2.09)	(3.61)
2	0.67*	0.39	1.03*	0.02	−0.23	−0.33*	−0.30*
	(3.44)	(1.72)	(5.01)	(0.14)	(−1.25)	(−3.08)	(−2.41)
3	0.14	0.19	0.63*	0.28*	0.02	−0.14	−0.16
	(1.27)	(1.46)	(3.33)	(2.62)	(0.25)	(−1.62)	(−1.37)
4	0.29*	0.09	0.72*	0.41*	0.10	0.16	−0.14
	(2.38)	(0.57)	(6.03)	(2.20)	(0.86)	(1.83)	(−1.07)
5	0.13	0.18	1.00*	0.48	−0.18	0.24*	−0.30*
	(0.54)	(0.53)	(5.61)	(1.80)	(−0.82)	(2.68)	(−2.72)
6	0.00	0.10	1.16*	0.08	0.06	0.10	−0.13
	(0.02)	(0.77)	(6.15)	(0.64)	(0.62)	(1.16)	(−0.86)
7	0.47*	0.30*	0.38*	−0.01	−0.25	0.96*	0.34*
	(4.12)	(2.29)	(2.59)	(−0.09)	(−1.45)	(5.79)	(2.56)
8	0.68*	1.06*	0.43*	0.26	0.41	0.48*	0.21*
	(2.63)	(4.38)	(2.28)	(1.07)	(1.06)	(3.12)	(2.27)
9	0.60*	0.87*	0.32	0.03	0.77*	0.49*	0.27*
	(3.03)	(5.12)	(1.23)	(0.12)	(4.27)	(3.24)	(2.04)
10	1.01*	1.38*	0.09	0.16	−0.15	0.40*	0.09
	(5.27)	(6.59)	(0.52)	(0.69)	(−0.66)	(2.20)	(0.70)
11	1.74*	1.25*	0.82*	0.62*	0.18	1.12*	0.50*
	(6.01)	(4.03)	(3.63)	(2.53)	(0.45)	(5.43)	(5.23)
12	0.89*	0.84*	0.95*	0.29	0.23	0.85*	0.45*
	(3.74)	(3.22)	(4.56)	(1.20)	(1.19)	(5.22)	(2.82)
13	0.78*	0.74*	0.84*	0.34	0.01	0.09	0.96*
	(2.33)	(2.76)	(4.49)	(1.36)	(0.03)	(0.51)	(4.99)
14	−0.23	1.05*	0.99*	0.56*	0.10	−0.12	0.25
	(−0.56)	(2.85)	(4.81)	(2.08)	(0.34)	(−0.61)	(1.29)
15	−0.68*	0.37	1.21*	0.63*	−0.09	0.51*	−0.39*
	(−2.64)	(1.78)	(5.02)	(2.94)	(−0.44)	(2.82)	(−2.63)
16	−0.90*	0.72*	0.92*	0.01	0.12	0.38	0.59*
	(−2.42)	(2.14)	(2.85)	(0.02)	(0.49)	(1.55)	(3.31)
17	0.48	1.92*	1.03*	1.25*	0.15	−0.38	0.06
	(0.98)	(2.45)	(2.92)	(2.44)	(0.33)	(−1.52)	(0.23)
18	1.02	−0.98*	0.03	−0.13	0.14	0.19	0.70*
	(1.75)	(−2.78)	(0.09)	(−0.38)	(0.43)	(0.74)	(2.19)
19	−0.80	−2.07*	−0.15	−0.22	−0.74	−0.11	0.73
	(−1.92)	(−6.42)	(−0.38)	(−0.57)	(−1.65)	(−0.36)	(1.81)
20	1.56*	−1.23*	−0.47	−0.46	−0.99	0.00	−0.02
	(3.25)	(−2.11)	(−1.16)	(−0.89)	(−1.76)	(0.01)	(−0.06)
21	0.08	0.61	0.51	0.40	−0.18	0.20	−0.13
	(0.17)	(1.12)	(1.47)	(0.72)	(−0.40)	(0.63)	(−0.39)
22	−0.28	−0.78	0.01	−1.73*	0.89*	−0.01	−0.52
	(−0.49)	(−1.73)	(0.04)	(−3.43)	(2.07)	(−0.01)	(−1.49)
23	−0.69	−0.29	−0.27	−0.70	0.28	−0.39	−0.73*
	(−1.65)	(−0.49)	(−0.69)	(−1.34)	(0.42)	(−1.12)	(−2.57)

Note: t-statistics are shown in parentheses.
* Indicates that the estimated values are statistically significant at the 5% level.

Considering inflation forecasts, the picture is better and shows much less inefficiency, both quantitatively and based on the number of statistically significant parameters. For example, in the US, forecasts are

Table 2b
Test of efficiency in the use of public information in inflation forecasts.

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	−0.09	−0.05	−0.15*	−0.43*	−0.28*	0.22*	−0.17*
	(−1.39)	(−0.54)	(−2.22)	(−7.06)	(−4.07)	(3.01)	(−3.39)
2	−0.27*	0.06	−0.01	−0.10	−0.31*	−0.12	0.00
	(−3.24)	(0.59)	(−0.08)	(−1.16)	(−3.83)	(−1.04)	(−0.05)
3	−0.14*	0.02	0.21*	−0.09	−0.09	−0.31*	−0.17*
	(−2.02)	(0.26)	(3.84)	(−1.14)	(−1.49)	(−3.43)	(−2.39)
4	0.00	0.26*	0.18*	−0.53*	−0.11	0.02	0.05
	(0.01)	(2.44)	(2.13)	(−5.34)	(−1.25)	(0.26)	(0.67)
5	0.22	−0.15	0.04	−0.38*	−0.22	−0.08	−0.24*
	(1.91)	(−0.89)	(0.41)	(−2.71)	(−1.61)	(−0.79)	(−3.21)
6	−0.14	−0.23	−0.03	0.09	−0.23*	−0.42*	−0.01
	(−1.34)	(−1.96)	(−0.30)	(0.96)	(−3.22)	(−5.42)	(−0.12)
7	−0.05	−0.12	0.05	0.37*	0.01	0.02	0.15
	(−0.46)	(−0.95)	(0.68)	(2.88)	(0.07)	(0.27)	(1.76)
8	0.39*	0.02	0.22*	0.55*	−0.09	−0.19*	0.19*
	(2.74)	(0.09)	(2.18)	(3.10)	(−0.81)	(−2.08)	(2.50)
9	0.26*	−0.02	0.07	0.70*	−0.34*	−0.02	0.38*
	(2.01)	(−0.12)	(0.49)	(4.68)	(−4.20)	(−0.33)	(3.55)
10	0.17	0.38*	0.34*	0.54*	−0.25	−0.36*	0.16
	(1.51)	(2.63)	(3.08)	(3.60)	(−1.79)	(−4.73)	(1.16)
11	−0.14	−0.02	0.07	0.59*	−0.47*	0.25	−0.13
	(−0.76)	(−0.14)	(0.53)	(2.86)	(−3.29)	(1.88)	(−0.89)
12	−0.09	0.13	−0.58*	0.28	−0.32*	0.10	−0.45*
	(−0.49)	(0.85)	(−4.00)	(1.33)	(−2.12)	(1.16)	(−2.96)
13	−0.21	−0.07	0.05	0.93*	−0.01	−0.29*	−0.16
	(−0.88)	(−0.36)	(0.33)	(2.83)	(−0.08)	(−3.29)	(−0.91)
14	0.55*	−0.11	−0.06	0.66*	−0.50*	−0.06	−0.35*
	(2.03)	(−0.45)	(−0.46)	(2.25)	(−3.33)	(−0.45)	(−2.55)
15	0.01	0.02	0.27	−0.77*	−0.13	−0.74*	−0.17
	(0.03)	(0.16)	(1.73)	(−6.78)	(−1.07)	(−7.54)	(−1.12)
16	−0.16	0.29	0.25	0.31	0.06	0.08	−0.06
	(−0.73)	(1.35)	(1.45)	(1.43)	(0.48)	(0.76)	(−0.35)
17	0.76*	0.12	−0.21	0.13	−0.33	0.93*	0.05
	(2.19)	(0.37)	(−1.05)	(0.46)	(−1.09)	(5.27)	(0.31)
18	0.43	−0.47*	−0.37*	0.30	0.03	0.05	−0.10
	(1.16)	(−2.45)	(−2.11)	(1.22)	(0.18)	(0.27)	(−0.49)
19	−0.23	−0.34	0.12	0.64	0.24	0.16	−0.07
	(−0.79)	(−1.31)	(0.65)	(1.65)	(1.07)	(0.84)	(−0.37)
20	0.05	0.03	0.28	0.86	−0.33	0.25	0.15
	(0.12)	(0.06)	(1.31)	(1.61)	(−1.33)	(1.20)	(0.64)
21	−0.21	−0.13	−0.31	−0.28	−0.03	0.17	−0.28
	(−0.63)	(−0.44)	(−1.24)	(−0.91)	(−0.13)	(1.18)	(−1.00)
22	−0.25	0.05	−0.14	0.03	0.00	−0.11	−0.22
	(−0.84)	(0.16)	(−0.70)	(0.10)	(0.00)	(−0.58)	(−0.82)
23	0.75	0.08	−0.16	−0.40	−0.36	0.29	−0.35
	(1.58)	(0.23)	(−0.61)	(−0.98)	(−1.46)	(1.11)	(−1.24)

Note: t-statistics are shown in parentheses.
* Indicates that the estimated values are statistically significant at the 5% level.

inefficient only for 7 of the 24 horizons for inflation, but for 13 horizons for the GDP forecasts. The numbers are very similar for the other six countries. Forecasters seem to put more than the efficient weight on new public information at very short horizons, if at all. For the middle horizons, the evidence is mixed. While forecasters underweight public information for Canada and Italy, they overweight it for Japan and the UK in predicting inflation. Note that our tests in Eq. (20) were conducted using real time actual data. Since data revisions are a lot more formidable in real GDP than in inflation, the use of revised actual data would have revealed relatively more inefficiency in real GDP forecasts than in inflation forecasts.

In summary, our analysis shows that, given the Bayesian learning model, there is more pervasive stickiness and inefficiency in the recorded real GDP forecasts than in the inflation forecasts.

6. Concluding remarks

Based on data from several industrialized countries, we establish that when predicting inflation, professional forecasters (i) make smaller forecast errors; (ii) disagree to a lesser extent; and (iii) begin revising their forecasts much earlier, relative to real GDP. Even though the first of these results has been implicit in most studies of forecast evaluations, none of these empirical results are well articulated in the forecasting literature, even though, as Granger (1996) noted, in order to increase the perceived quality of macro forecasts, we should be cognizant of the variables that are relatively easy to forecast.

To understand these interesting differences between real GDP and inflation forecasts, we first explore the underlying data generating processes for the target variables. Using a standard autoregressive model, we find that the real GDP should indeed be more difficult to forecast; in particular, we find that real GDP forecasts do not have any predictive value over naïve benchmarks beyond the 18-month horizon, while for inflation the content horizon is around 24 months, and even up to 36 months for some countries.

In reality, the predictability can be improved upon by incorporating additional information and using nonlinear models. In that sense, the forecast content from the simple autoregressive model provides an overall *lower* bound on the true predictability of a series. To better understand the relative forecasting records of these two macro variables, we develop a simple Bayesian learning model aimed at identifying the relative importance of alternative pathways through which professional forecasters adapt to new information, resulting in different patterns of forecast accuracy over the term structure. In our model, forecast disagreement arises from two sources: differences in prior beliefs, and differences in forecasters' interpretations of public information. The importance of the second pathway is identified by analyzing the evolution of forecast disagreement over horizons, and is a key factor in explaining the differential forecast accuracies of the two target variables.

We find that diversity in the prior beliefs of forecasters explains between 40% and nearly 100% of the disagreement in the GDP forecasts, as the horizon decreases from 24 months to 1 month. However, the corresponding numbers are much lower for inflation. The rest of the explained forecast disagreement is driven by heterogeneity in the interpretation of new information. This empirical finding, together with two case studies on (i) forecast disagreement around the 9/11 terrorist attack, and (ii) the inflation targeting experience of Italy after 1997, provides strong support for the role of differential interpretations of public information in generating expert disagreement in macroeconomic forecasts, and also provides an additional explanation of the relative superiority of inflation forecasts over real GDP forecasts. We find that a reduction in expert disagreement translates to a corresponding reduction in aggregate forecast error.

We explore the possibility that if forecasters do not use information efficiently because they place sub-optimal weights on new information, the forecast content from the simple autoregressive model might provide an overall *upper* bound on the true predictability of a series. Following a test for forecast efficiency developed by Lahiri and Sheng (2008) in a Bayesian learning framework, we find that, on the whole, inflation forecasts are more efficient than real GDP forecasts. There is overwhelming evidence that professional forecasters significantly underweight public information in the middle horizons when predicting real GDP. For inflation we find very little systematic misuse of new information. This evidence can be rationalized by the fact that the initial GDP announcements are revised more heavily than price indexes, are observed only quarterly, and involve substantial measurement errors. Our case study on the Italian inflation experience in the 1990s shows how the inflation targeting policy of its Central Bank reduced the forecast disagreement and aggregate forecast errors throughout the term structure of forecasts. Thus, a more frequent communication of the latest inflationary developments to the general public and the commitment to long-run price stability by the monetary authorities in our sample countries may make the adjustments to inflationary expectations more dependent on current news, resulting in superior forecasts.

Finally, we should point out that the relative superiority of inflation forecasts compared to those of real GDP can also be determined by the demand side of the forecasting market; i.e., the professional forecasters may devote more effort to generating better forecasts if that is what their clients demand. Indeed, Sinclair, Gamber, Stekler, and Reed (2009) have shown that, in the context of a forward-looking Taylor rule as the yardstick for Fed's monetary policy, inflation forecast errors are implicitly considered to be three times more costly than those of real GDP. Also, Capistrán and Timmermann (2008) have argued that a certain degree of inefficiency in real GDP forecasts compared to inflation can be rationalized if clients' loss functions are asymmetric, and differ between real GDP and inflation. Using density forecast data, Lahiri and Liu (2009) find some evidence to this effect, particularly at longer horizons.

It is interesting that as part of the Fed's major changes in its communication strategies, effective September 2007, the horizon of the projections for GDP growth and inflation made by all FOMC members has been extended from two years to three.[21] We have found that at present the real GDP forecasts do not seem to have any value beyond the 18-month horizon. If the demand side of the forecasting market has any effect on forecast quality, we may expect that the content horizon for these forecasts will lengthen in the future as a result of this change in FOMC policy. However, only time will show whether this happens.

Acknowledgements

An earlier version of this paper was presented at the 28th International Symposium on Forecasting, at the 9th IWH-CIREQ Macroeconometric Workshop, and in a seminar at the City University of Hong Kong. We thank Antony Davies, Frank Diebold, Jean-Marie Dufour, John Galbraith, David Hendry, Vikas Kakkar, Lars-Erik Öller, Barbara Rossi, and Mark Watson for many helpful comments and suggestions. We have also benefited greatly from the suggestions and advice of the guest editor Gael Martin and three anonymous referees. However, we alone are responsible for any remaining errors and shortcomings.

[21] See Bernanke (2007).

References

Acemoglu, D., Chernozhukov, V., & Yildiz, M. (2006). *Learning and disagreement in an uncertain world*. NBER working paper No. 12648.

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

Batchelor, R. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting*, *23*, 189–203.

Beechey, M. J., Johannsen, B. K., & Levin, A. T. (2007). *Are long-run inflation expectations anchored more firmly in the Euro area than in the United States*? Board of Governors of the Federal Reserve System, CEPR Discussion Paper No. DP6536.

Bernanke, B. S. (2007). Federal Reserve communications. *Speech delivered at the Cato Institute 25th annual monetary conference*.

Boero, G., Smith, J., & Wallis, K. F. (2008). Uncertainty and disagreement in economic prediction: The Bank of England survey of external forecasters. *Economic Journal*, *118*, 1107–1127.

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., & Rautanen, T. (2007). *The ECB Survey of Professional Forecasters (SPF): A review after eight years' experience*. European Central Bank occasional paper series, No. 59.

Capistrán, C., & Timmermann, A. (2008). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking*, *41*, 365–396.

Clements, M. P. (1995). Rationality and the role of judgment in macroeconomic forecasting. *Economic Journal*, *105*, 410–420.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting* (pp. 961–982). Amsterdam: North-Holland.

Diebold, F. X., & Kilian, L. (2001). Measuring predictability: Theory and macroeconomic applications. *Journal of Applied Econometrics*, *16*, 657–669.

Dominitz, J., & Manski, C. F. (2005). *Measuring and interpreting expectations of equity returns*. NBER working paper No. 11313.

Döpke, J., & Fritsche, U. (2006). When do forecasters disagree? An assessment of German growth and inflation forecast dispersion. *International Journal of Forecasting*, *22*, 125–135.

Galbraith, J. W. (2003). Content horizons for univariate time-series forecasts. *International Journal of Forecasting*, *19*, 43–55.

Galbraith, J. W., & Tkacz, G. (2007). Forecast content and content horizons for some important macroeconomic time series. *Canadian Journal of Economics*, *40*, 935–953.

Granger, C. W. J. (1996). Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics*, *11*, 455–473.

Gregory, A. W., & Yetman, J. (2004). The evolution of consensus in macroeconomic forecasting. *International Journal of Forecasting*, *20*, 461–473.

Gürkaynak, R., Levin, A. T., Marder, A. N., & Swanson, E. T. (2007). Inflation targeting and the anchoring of inflation expectations in the Western Hemisphere. In *Economic review, Federal Reserve Bank of San Francisco* (pp. 25–47).

Isiklar, G., & Lahiri, K. (2007). How far ahead can we forecast? Evidence from cross-country surveys. *International Journal of Forecasting*, *23*, 167–187.

Isiklar, G., Lahiri, K., & Loungani, P. (2006). How quickly do forecasters incorporate news? Evidence from cross-country surveys. *Journal of Applied Econometrics*, *21*, 703–725.

Johnson, D. J. (2002). The effect of inflation targeting on the behavior of expected inflation: Evidence from a 11 country panel. *Journal of Monetary Economics*, *49*, 1521–1538.

Kandel, E., & Pearson, N. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, *103*, 831–872.

Kandel, E., & Zilberfarb, B. Z. (1999). Differential interpretation of information in inflation forecasts. *Review of Economics and Statistics*, *81*, 217–226.

Lahiri, K., & Liu, F. (2006). Modeling multi-period inflation uncertainty using a panel of density forecasts. *Journal of Applied Econometrics*, *21*, 1199–1219.

Lahiri, K., & Liu, F. (2009). On the estimation of forecasters' loss function using density forecasts. *Presented at the Joint Statistical Meetings*.

Lahiri, K., & Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, *144*, 325–340.

Mankiw, N. G., & Reis, R. (2006). Pervasive stickiness. *American Economic Review*, *96*, 164–169.

Mankiw, N. G., Reis, R., & Wolfers, J. (2003). Disagreement about inflation expectations. In *NBER macroeconomics annual 2003*.

Mishkin, F. S. (2007). *Inflation dynamics*. NBER working paper No. 13147.

Morris, S., & Shin, H. S. (2006). Inertia of forward-looking expectations. *American Economic Review*, *96*, 152–157.

Öller, L.-E., & Barot, B. (2000). The accuracy of European growth and inflation forecasts. *International Journal of Forecasting*, *16*, 293–315.

Öller, L.-E., & Teterukovsky, A. (2007). Quantifying the quality of macroeconomic variables. *International Journal of Forecasting*, *23*, 205–217.

Patton, A. J., & Timmermann, A. (2007). *Learning in real time: Theory and empirical evidence from the term structure of survey forecasts*. Working paper, Department of Economics, UCSD.

Sinclair, T. M., Gamber, E. N., Stekler, H. O., & Reed, E. (2009). Jointly evaluating GDP and inflation forecasts in the context of the Taylor rule. *Presented at the first macroeconomic conference, IFO/INSEE/SAE*.

Stock, J. H., & Watson, M. W. (2003). Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature*, *41*, 788–829.

Stock, J. H., & Watson, M. W. (2007). Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, *39*, 3–33.

Zarnowitz, V., & Lambros, L. A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, *95*, 591–621.

Zarnowitz, V., & Braun, P. (1993). Twenty-two years of NBER-ASA quarterly economic outlook surveys: Aspects and comparisons of forecasting performance. In J. H. Stock, & M. W. Watson (Eds.), *Business cycles, indicators and forecasting*. University of Chicago Press and NBER.

Zellner, A. (1988). Optimal information processing and Bayes' theorem. *American Statistician*, *42*, 278–284.

Zellner, A. (2002). Information processing and Bayesian analysis. *Journal of Econometrics*, *107*, 41–50.

Kajal Lahiri is Distinguished Professor of Economics and Health Policy Management at the University at Albany – SUNY, and Honorary Fellow of the International Institute of Forecasters. His research interests include econometric methodology, business cycle forecasting, forecast evaluation, and health economics. Dr. Lahiri is on the editorial boards of the *Journal of Econometrics*, *International Journal of Forecasting*, *Empirical Economics*, and the *Journal of Business Cycle Measurement and Analysis*.

Xuguang Sheng is Assistant Professor of Economics at SUNY Fredonia. His research interests include panel unit roots and cointegration, Bayesian methodology, and economic forecasting. He has published in the *Journal of Econometrics* and the *Journal of Applied Econometrics*.