



Evolution of forecast disagreement in a Bayesian learning model[☆]

Kajal Lahiri^{*}, Xuguang Sheng

Department of Economics, University at Albany - SUNY, Albany, NY 12222, United States

ARTICLE INFO

Article history:

Received 31 July 2007

Received in revised form

30 January 2008

Accepted 8 February 2008

Available online 30 April 2008

JEL classification:

E17

E37

Keywords:

Forecast disagreement

Forecast horizon

Forecast efficiency

Consensus forecast

Differential interpretation

Public information

Bayesian learning

Panel data

ABSTRACT

We estimate a Bayesian learning model with heterogeneity aimed at explaining expert forecast disagreement and its evolution over horizons. Disagreement is postulated to have three components due to differences in: (i) the initial prior beliefs, (ii) the weights attached on priors, and (iii) interpreting public information. The fixed-target, multi-horizon, cross-country feature of the panel data allows us to estimate the relative importance of each component precisely. The first component explains nearly all to 30% of forecast disagreement as the horizon decreases from 24 months to 1 month. This finding firmly establishes the role of initial prior beliefs in generating expectation stickiness. We find the second component to have barely any effect on the evolution of forecast disagreement among experts. The importance of the third component increases from almost nothing to 70% as the horizon gets shorter via its interaction with the quality of the incoming news. We propose a new test of forecast efficiency in the context of Bayesian information processing and find significant heterogeneity in the nature of inefficiency across horizons and countries.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Until very recently disagreement among economic agents had very little role to play in standard economic theory. The presumption was that agents will merge to a common “prior” because of learning from the same infinite sequence of signals.¹ However, various survey data on expectations from many countries over last fifty years have produced mounting evidence on substantial interpersonal heterogeneity in how people perceive the current and form inferences about the future economic conditions.² Beginning with Lucas (1973), the observed disagreement was attributed to

individuals having been exposed to different histories of observations. Currently there is a growing feeling that agents not only have differential information sets, but may also interpret the same public information differentially. Varian (1989), Harris and Raviv (1993), Kandel and Pearson (1995), among others, have explicitly rejected a core tenet of the efficient market hypothesis that agents start off with a rational common prior and develop heterogeneous beliefs only because they receive differential information. In certain learning environments, rational agents may not converge to a common belief even when exposed to the same information. Heterogeneity in the processing of information and lack of common beliefs is a central theme that has emerged in many areas of economics.

Disagreement has assumed an important role in macroeconomics due to the sticky-information macro model of Mankiw and Reis (2002) and Mankiw et al. (2003). In this model, the households only occasionally pay attention to news, and this inattention generates disagreement in aggregate expectations endogenously. However, as noted by Sims (2003) and Williams (2003), the sticky-information model may not characterize professional forecasters well. In a competitive market for professional forecasts, the primary reason why these forecasters disagree is certainly not due to lags in gathering relevant information. Nonetheless, the quality of forecasts is likely to form an upper bound for that of laymen who mostly get information from the commentaries and reports by experts in the news media, cf. Carroll (2003). Thus, it is natural to

[☆] Earlier versions of this paper were presented at New York Camp Econometrics (2006) and the Far Eastern Meeting of the Econometric Society (2006) in Beijing. We thank Jushan Bai, Badi Baltagi, John Jones, Peter Schmidt, Herman Stekler, Victor Zarnowitz, Arnold Zellner, two anonymous referees, and participants in above conferences for helpful comments and suggestions. We alone are responsible for any remaining errors and shortcomings.

^{*} Corresponding author. Tel.: +1 518 442 4758; fax: +1 518 442 4736.

E-mail addresses: klahiri@albany.edu (K. Lahiri), xs7410@albany.edu (X. Sheng).

¹ See Savage (1972) and Aumann (1976) for theoretical foundations for the hypothesis.

² See Jonung (1981), Zarnowitz and Lambros (1987), Lahiri et al. (1988), Maddala (1991), Ivanova and Lahiri (1998), Souleles (2004), Patton and Timmermann (2007), among many others.

look into such questions as: do experts hold divergent prior beliefs about macroeconomic events even when they have access to a wealth of public information? Do they interpret the same public information differently? How does forecast disagreement changes as information of varying quality arrives? We address these questions using a Bayesian learning model based on monthly GDP forecasts for G7 countries made by 285 professional forecasters during 1990–2004.

Our paper makes the following three contributions to the literature on macroeconomic forecasting and belief formation. First, we provide comprehensive evidence on certain patterns in multi-horizon forecasts and extend the literature on learning and disagreement in a number of ways. Based on corporate earnings and inflation forecasts, respectively, Kandel and Pearson (1995) and Kandel and Zilberfarb (1999) found evidence in favor of the hypothesis that economic agents interpret publicly available information differently. Using fixed-target, multi-horizon, cross-country panel data on real GDP forecasts, we find significant heterogeneity in how experts interpret public information. We find that, as the horizons get shorter, the “interpretation bias” decreases but not uniformly; it depends on the timing of public information arrival. Since the uncertainty associated with new information over different forecast horizons changes in a predictable manner, our evidence highlights the hypothesis recently put forward by Acemoglu et al. (2006) that the uncertainty about the informativeness of news is a key to expert disagreement.

Second, in recent years various models have been proposed in the macroeconomic literature aimed at explaining the evolution of forecast disagreement *over time*. None of these models, however, investigate the issue that we specifically address here, that is, the evolution of forecast disagreement *over forecast horizons*.³ By studying the pattern of forecast disagreement over horizons with fixed targets, we can learn a great deal about the importance of initial prior beliefs, the timing of information arrival and the efficiency in the use of public information. We find experts to start off with widely divergent initial prior beliefs at very long horizons. Their initial beliefs propagate forward onto the whole series of forecasts, generating significant amount of inertia in expectations formation. This result has important implications for the ongoing discussion on expectation stickiness and the efficient market hypothesis (Mankiw and Reis, 2006; Morris and Shin, 2006). The striking finding from our forecast revision analysis is that the first major public information seems to arrive when the forecast horizon is around 15 months for all seven sampled countries. Afterwards, as the horizon decreases and the information on target year GDP gets more certain, forecasters not only give more weight to new information but also interpret them more similarly, resulting in rapidly converging forecasts.

Third, without depending on certain *ad hoc* proxies for public information, we propose a test of forecast efficiency based on the property that forecast errors should be uncorrelated with the most recent forecast revisions, thus generalizing the weak efficiency test of Nordhaus (1987). Remarkably, we find that forecasters do not always underweight public information; at very long horizons they tend to revise forecasts unnecessarily. For other horizons, the actual weights attached to new information appear to be significantly less than the efficient weights implied by the optimal Bayesian information processing rule.

The rest of our paper is organized as follows. Section 2 presents some stylized facts based on the cross-country forecast data.

Section 3 proposes the Bayesian learning model that is consistent with the stylized facts. Section 4 presents the estimation of the model parameters using panel data, and Section 5 explores the forecast efficiency in utilizing public information. Section 6 concludes.

2. Data

2.1. Data source

The data for this study are taken from “*Consensus Forecasts: A Digest of International Economic Forecasts*”, published by Consensus Economics Inc. We study the panel forecasts of annual GDP growth. Survey respondents make their first forecasts when there are 24 months to the end of the target year; that is, they start forecasting in January of the previous year, and their last forecast is reported at the beginning of December of the target year. So for each country and for each target year we have 24 forecasts of varying horizons. Our analysis covers the forecasts made for 14 target years from 1991 to 2004. The countries we study are seven major industrialized countries – Canada, France, Germany, Italy, Japan, the United Kingdom and the United States.⁴ The number of panelists ranges from 30 to 60 for G7 countries. Altogether we have more than 50,000 forecasts for GDP growth. For annual GDP growth rate, we use an early revision 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* are regularly sold in a wide variety of markets and hence one would expect these professional forecasts to satisfy a certain level of accuracy in contrast to laymen expectations. Second, the names of the respondents are published next to their forecasts. This lends further credibility to the individual forecasts as poor forecasts damage the respondents’ reputation. Third, the cross-country nature of the survey data provides us a unique opportunity to compare professional forecasts across countries. Fourth, since the importance of private information in forecasting real GDP is expected to be very small compared to corporate earnings and even inflation, we can identify the news to GDP growth as mostly public. Finally, forecasts for fairly long horizons (currently from 24- to 1-month ahead) are available. This fixed-event scheme enables us to study carefully the effect of information arrival and uncertainty on measures of forecast disagreement.

2.2. Twin empirical findings

Following the literature, we measure forecast disagreement as the variance of forecasts *across* professional forecasters. We calculate the disagreement for each horizon, each year, and for each country. We find forecast disagreement to vary dramatically over time and forecast horizons. The results imply that occasionally the economies go through substantial changes that are very difficult to forecast by virtually any of the techniques currently used to forecast economic variables. We merely observe that heterogeneity of expectations is the norm in the market place.

Since we have a rich panel data at our disposal, decomposition of the total sum of squares in forecasts into between- and within-agent variations can be revealing. Let $F_{it,h}$ be the forecast of GDP growth made by agent i , for the target year t , and h months ahead of

³ The importance of forecast horizon has recently been recognized by Yang and Shintani (2006), Granger (2007) and Lee and Nelson (2007).

⁴ Note that the target for GDP in Germany changes in our data sample due to unification. We used forecasts for West Germany made for the target years 1991–1995, and for unified Germany for the target years 1996–2004.

the end of the target year. In this context, we introduce three new measures – within-agent variation (S_h^w), between-agent variation (S_h^b), and total variation (S_h^f) in forecasts at each horizon, adjusted for the number of forecasters,

$$S_h^w = \sum_{i=1}^{N_h} \sum_{t=1}^{T_{ih}} (F_{it} - \bar{F}_{i,h})^2 / \sum_{i=1}^{N_h} T_{ih} \quad (1)$$

$$S_h^b = \sum_{i=1}^{N_h} T_{ih} (\bar{F}_{i,h} - \bar{\bar{F}}_{..h})^2 / \sum_{i=1}^{N_h} T_{ih} \quad (2)$$

$$S_h^f = \sum_{i=1}^{N_h} \sum_{t=1}^{T_{ih}} (F_{it} - \bar{\bar{F}}_{..h})^2 / \sum_{i=1}^{N_h} T_{ih} \quad (3)$$

where $\bar{F}_{i,h} = \sum_{t=1}^{T_{ih}} F_{it} / T_{ih}$ is the mean forecasts over time and $\bar{\bar{F}}_{..h} = \sum_{i=1}^{N_h} \bar{F}_{i,h} / N_h$ the overall mean forecasts over time and across agents. Since not all forecasters respond in all surveys, we dealt with the unbalanced panel by taking appropriate summations over different survey dates and horizons, see Baltagi et al. (2001). Forecasters who responded less than 10% of the times were deleted from our analysis such that our results are not sensitive to a few extreme observations. By construction, the total variation in forecasts is the sum of within-agent and between-agent variation, that is,

$$S_h^f = S_h^w + S_h^b. \quad (4)$$

Fig. 1 plots the within-agent variation in forecasts (dotted line). In order to study the general pattern of forecast disagreement over horizons, we also plot the average disagreement over 14 target years for each horizon (solid line) on the same graph.⁵ Plots start when the horizon is 24 months and end when the horizon is 1 month. Although the magnitudes vary a lot across countries, their evolution over horizons share certain common features.

First, forecast disagreement at the 24-month horizon is high and stays almost unchanged till about the 16-month horizon. It starts to decline sharply at 15-month horizon and keeps declining as the horizon gets shorter. The same pattern is observed when we looked at inter-quartile or inter-decile ranges, cf. Bowles et al. (2007). Second, the within-agent variation in forecasts, however, starts with a relatively very low value at the 24-month horizon and remains almost the same at horizons 24 to 16 months. It then increases sharply as the forecast horizon decreases from 15 months to 1 month.

Some parts of the first finding have been documented in the literature. Based on a subset of our database, Harvey et al. (2001) and Gallo et al. (2002) find that the cross-section standard deviation decreases as forecast horizon shortens. We extend the above studies to include 7 countries and 14 target years, and thus provide more comprehensive evidence concerning the nature of the dynamics of disagreement among professional forecasters. Using the Blue Chip Economic Indicators data, Gregory et al. (2001) and Gregory and Yetman (2004) tested whether forecast differences decay to zero at the same point in time. Their results suggest that forecasters move toward consensus when they are making repeated GDP forecasts. Our study, by providing a formal learning model, can be regarded as providing an explicit microeconomic foundation for their decay function. However, since their data include forecasts made from July of the previous year to May of the current year, the disagreement measures for the long horizons (24 to 18 months ahead) were not analyzed in these papers. The persistently high level of forecast disagreement in real

GDP forecasts at these long horizons has not been documented in the literature. As explained in Section 3, this finding, coupled with the low within-agent variation in forecasts for the same horizons, has very important bearing on the role of initial prior beliefs in the evolution of the shorter term forecasts.

2.3. Variation in forecast revisions

With fixed-target forecasts, an analysis of forecast revisions reveals important information about when major public signals arrive and whether experts interpret them differently. Forecast revision (R_{it}) is defined as the difference between two successive forecasts for the same individual i and the same target year t , i.e. $R_{it} = F_{it} - F_{it+1}$. We can divide total variation in forecast revisions ($\sum_{i=1}^{N_h} \sum_{t=1}^{T_{ih}} (R_{it} - \bar{R}_{..h})^2 / \sum_{i=1}^{N_h} T_{ih}$) into within-agent ($\sum_{i=1}^{N_h} \sum_{t=1}^{T_{ih}} (R_{it} - \bar{R}_{i,h})^2 / \sum_{i=1}^{N_h} T_{ih}$) and between-agent variation ($\sum_{i=1}^{N_h} T_{ih} (\bar{R}_{i,h} - \bar{R}_{..h})^2 / \sum_{i=1}^{N_h} T_{ih}$) at each horizon.

Table 1 shows the ratio of between-agent variation to total variation in forecast revisions. With a few exceptions, between-agent variation explains 6% to 25% of the total variation in forecast revisions, depending on the country and horizon. We take this as the first evidence that professional forecasters on some occasions interpret the same public signal quite differently.

The between-agent variation, however, is relatively small, since total variation in forecast revision is mainly driven by within-agent variation. This is not unexpected because our forecasters are experts, and the target is a widely discussed macroeconomic entity. However, because of its relative size, we should also be interested in the evolution of within-agent variation, which shows how professional forecasters revise their forecasts intrapersonally after observing new public signals. Fig. 2 plots the total variation in GDP forecast revisions and its components over horizons. Whenever we see a big jump in within-agent variation at certain horizons, it means that professional forecasters make major revisions at that specific horizon. The first big spike in within-agent variation is observed at horizon 15 months for all sampled countries, which simply means that professional forecasters observe the first relevant public signal and revise their forecasts at the beginning of October of the previous year.⁶ One may wonder about the sources of information that is incorporated in forecasts at this horizon. It seems that forecasters extract information from relevant monthly indicators such as employment, industrial production, manufacturing index, etc. that have predictive power for next year's GDP growth. In addition, various leading indicators (e.g. interest rate spreads, stock market index, unemployment insurance claims, etc.) have predictive power beyond a year and are available promptly as valuable information sources. Also, there may be pressure from the users of forecasts for definitive forecasts for next year GDP growth around this horizon.⁷ For the US, the within-agent variation gets a boost again at horizon 11 months after the base-year GDP growth is known, which, as expected, affects the forecasts for year-over-year growth rate. For other G7 countries, this spike is discernable at horizons 9 and 10 months, depending on the timing of their base-year GDP announcements. As forecast horizon declines, within-agent variation gets a boost whenever the first release of GDP growth for the previous quarter becomes available.

⁶ Note the exact timing of information arrival. The first substantive public signal arrives during the month of September of the previous year. The professional forecasters, however, make their forecasts at the beginning of each month and as a result, the monthly signals get reflected in forecasts made at the beginning of October of the previous year.

⁷ It is interesting to note that prior to 1980, Blue Chip Economic Indicators produced monthly GDP growth forecasts for the next year beginning October of the previous year. Thus, extending the horizon beyond 16 months after 1980 could have been a demand-driven development.

⁵ Using the framework of Davies and Lahiri (1995), we can easily show that in a balanced panel, N times the average disagreement plus $N(T-1)$ times the variance of the systematic forecaster-specific biases is equal to the between-agent variation S_h^b . Thus, if forecasts are unbiased, N times the average disagreement is the same as the between-agent variation.

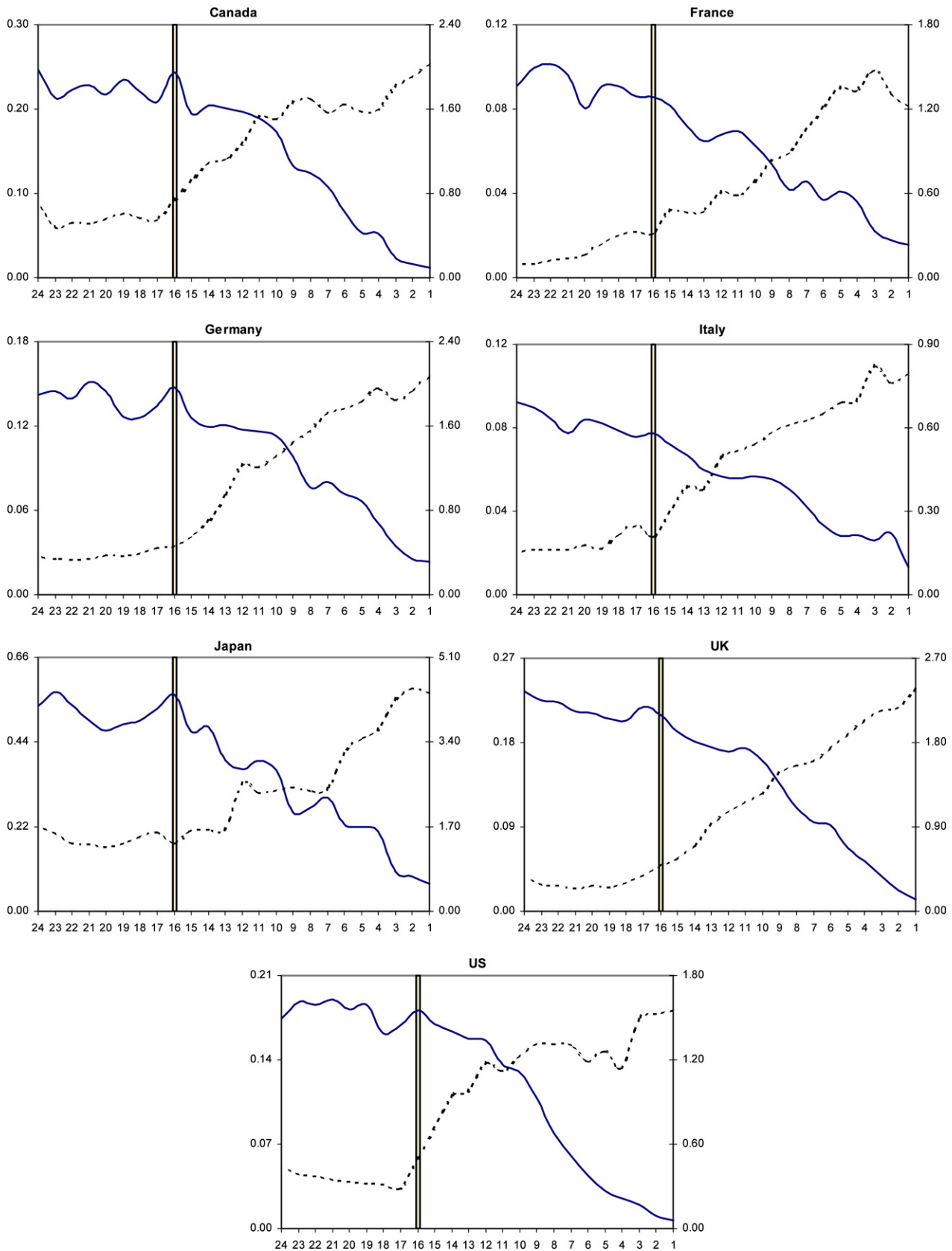


Fig. 1. Disagreement (solid line, left scale) and within-agent variation (dotted line, right scale) in forecasts.

For example, for the United States, within-agent variation jumps at horizon 8 months as a result of the advance release of first quarter GDP growth at the end of April. Then it gets a boost again

5 months as a result of the advance release of second quarter of GDP growth at the end of July. The synchronization of the movements in within-agent (and total) variation in forecast revisions with the

Table 1
Ratio of between-agent variation to total variation in forecast revisions

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	0.14	0.10	0.08	0.06	0.19	0.14	0.09
2	0.24	0.17	0.22	0.21	0.08	0.16	0.18
3	0.14	0.16	0.10	0.11	0.13	0.11	0.14
4	0.15	0.16	0.10	0.15	0.14	0.11	0.19
5	0.20	0.18	0.05	0.21	0.18	0.16	0.12
6	0.23	0.18	0.15	0.11	0.11	0.20	0.22
7	0.12	0.09	0.07	0.12	0.27	0.16	0.12
8	0.17	0.09	0.11	0.15	0.15	0.16	0.15
9	0.13	0.06	0.15	0.16	0.17	0.16	0.14
10	0.15	0.12	0.14	0.08	0.14	0.15	0.14
11	0.13	0.13	0.08	0.09	0.19	0.15	0.12
12	0.14	0.11	0.10	0.12	0.16	0.15	0.18
13	0.20	0.10	0.10	0.09	0.26	0.14	0.15
14	0.17	0.24	0.10	0.19	0.14	0.16	0.24
15	0.07	0.10	0.15	0.13	0.14	0.14	0.14
16	0.13	0.12	0.16	0.22	0.21	0.10	0.19
17	0.23	0.19	0.10	0.17	0.15	0.18	0.18
18	0.09	0.31	0.08	0.17	0.22	0.15	0.25
19	0.13	0.12	0.10	0.07	0.23	0.16	0.17
20	0.20	0.19	0.10	0.11	0.27	0.10	0.16
21	0.16	0.13	0.11	0.19	0.17	0.13	0.14
22	0.17	0.28	0.18	0.10	0.19	0.18	0.23
23	0.33	0.27	0.13	0.31	0.23	0.17	0.19

timing of the advance release of GDP growth figures holds for other countries in our sample as well.⁸

3. Bayesian learning model

In this section, we develop a Bayesian learning model to explain the evolution of forecast disagreement over horizons. The model explicitly recognizes that professional forecasters begin forecasting with specific prior beliefs about the GDP growth for the next year and that they progressively learn over horizons to modify their initial beliefs.

As shown in Fig. 1, the within-agent variation in forecasts starts with a very low value at the 24-month horizon and stays almost the same for another 8 months. We interpret this as evidence that each agent forms a prior belief about GDP growth for the next year and stays intrapersonally stable for several months. Batchelor (2007) also noted that forecasters seem to start their forecasting round by adopting a relatively optimistic or pessimistic view of growth, and those who start optimistic one year also start off being optimistic in other years. In addition, we found that the initial forecast disagreement is relatively high, and thus agents hold interpersonally divergent initial prior beliefs. It may be rational to go to extremes when ignorant, cf. Arrow and Hurwitz (1972). Consistent with the empirical findings above, we make the following specific assumption about forecasters' initial prior beliefs.

Assumption 1. The initial prior belief of GDP growth for the target year t , held by the forecaster i , at the 24-month horizon is represented by the normal density with the mean F_{it24} and the precision (i.e. the reciprocal of the variance) a_{it24} for $i = 1, \dots, N$, $t = 1, \dots, T$.⁹

⁸ For Canada, France and Germany, within-agent variation jumps at horizons 7 and 4 months as a result of the release of first quarter GDP growth at the end of May and second quarter GDP growth at the end of August, respectively. For Italy and Japan, within-agent variation peaks at horizons 6 and 3 months, since their first and second quarter GDP growth figures are released at the end of June and September, respectively. For the United Kingdom, within-agent variation peaks at forecast horizon 6 months since the complete estimate of first quarter GDP announcement is not available until the end of June, which is 3 months after the end of first quarter.

⁹ The precisions of initial prior beliefs are allowed to be different across forecasters. This assumption is confirmed by recent studies using density forecasts that document the heterogeneity in forecast uncertainty (cf. Lahiri and Liu (2006)).

In thinking about why professional forecasters disagree at these relatively long horizons, note that for many the problem they face is not necessarily to estimate a finite number of parameters that appear in a parametric representation of the economy, but rather to pick a model of the economy from the set of all possible alternatives that might characterize the world at that time. Heterogeneity of forecasting models is the norm in the market place (cf. Kurz (1994) and Kandel and Pearson (1995)). Many of the forecasts are based on expert judgments also.¹⁰ A wealth of historical data relevant for forecasting real GDP growth is publicly available to all forecasters. Thus, it is not the availability of relevant data but the models and methods used to interpret them that are different from one forecaster to another. Due to the length of the forecasting horizon, forecasters face very high uncertainty in interpreting available information based on whatever model or judgment they are using, and hence disagree a lot about GDP growth in the next year. At these long horizons, the usable signals from observable indicators get lost in the unforecastable noise of the chosen predictive or judgmental model (Granger, 1996). In addition, misspecification of the econometric models and possible non-stationarity in the data generating process add to the discord.

As is also clear from Fig. 1, both the disagreement and within-agent variation in forecasts remain relatively stable for 24- to 16-month ahead forecasts. Together, these two empirical regularities imply that experts stick to their initial prior beliefs and their forecasts do not show any sign of convergence during this period. Between 24 to 16 months before the end of the target year, information about GDP growth in the next year is only indirect and mixed with noise; also pertinent information may not be forthcoming. It may require potentially infinite investment to acquire and process useful information. So professional forecasters face an environment where learning is very difficult at these long horizons. As long as there is no substantial evidence that would dramatically surprise them, forecasters will not revise their prior beliefs in big ways because of the so-called predisposition effects (cf. Brock and Hommes (1997) and Branch (2004)). A similar point

¹⁰ Batchelor and Dua (1991) reported that the single most important forecasting technique used by the Blue Chip panel was judgment (51%), followed by econometric modeling (28%), and time series analysis (21%).

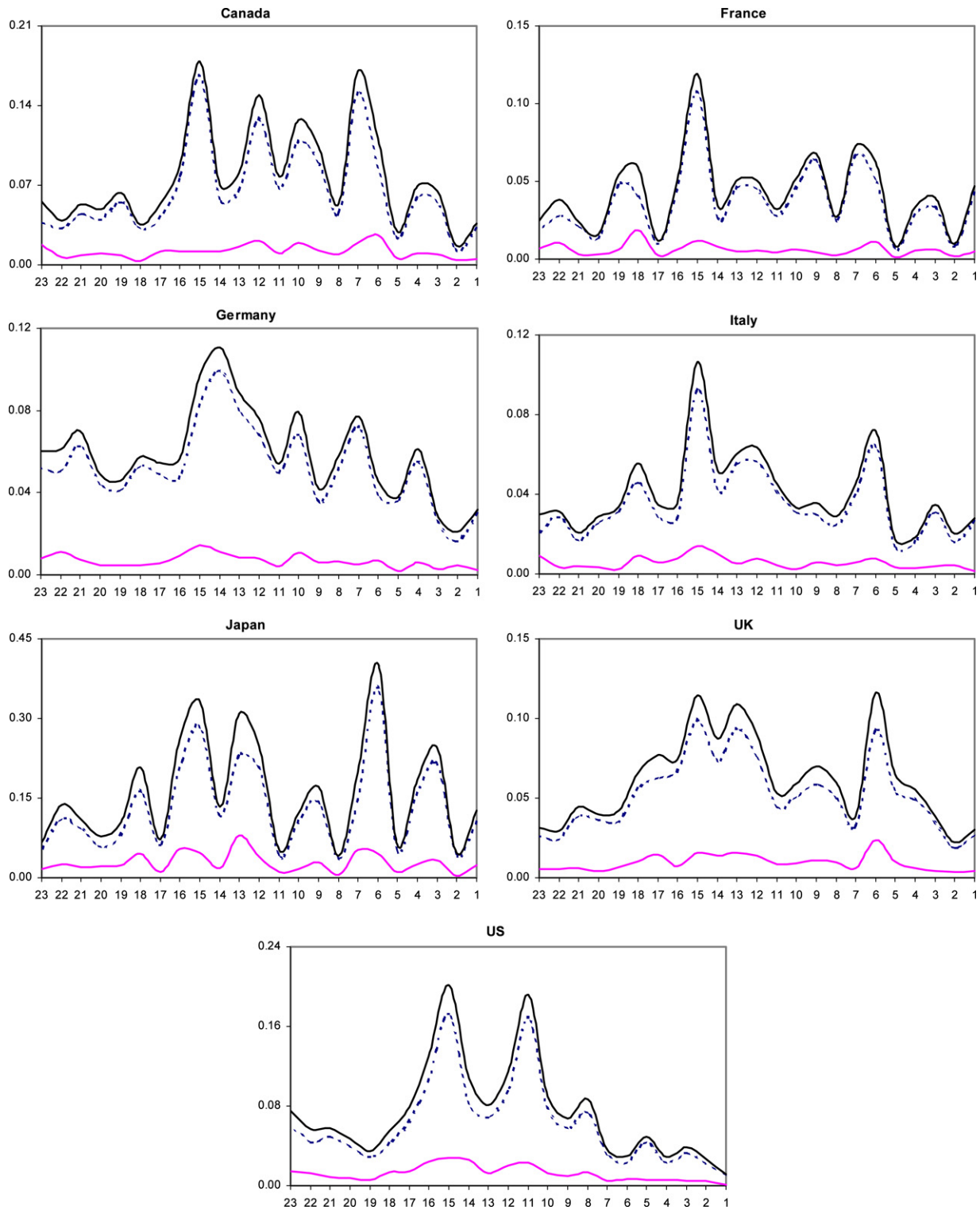


Fig. 2. Variation in forecast revision: between-agent variation (bottom line), within-agent variation (middle line), total variation (top line).

has also been made by Clements (1997) while analyzing UK macro forecasts for up to 12 quarters ahead. In a time period during which not much relevant information arrives, it seems reasonable for forecasters to take an initial view and not to change it in subsequent forecasts, except for random adjustments.

An important implication of the above findings is that there will typically be an “agreement to disagree” between professional forecasters. In other words, given their divergent initial beliefs, forecasters will agree that, after seeing the same public signals

year after year, they will continue to disagree in their longer-run outlooks. This evidence is consistent with the theoretical prediction in Acemoglu et al. (2006) who showed that, starting with heterogeneous prior beliefs, agents will not converge to a consensus even after observing the same infinite sequence of signals when there is a lot of uncertainty in the public information. Thus, we have “equilibrium disagreement” among professional forecasters at 24- to approximately 16-month horizons. This is interesting because the common prior assumption, typically made

in the learning literature, leads to the well-known no “agreement to disagree” result (cf. Aumann (1987)).

We now turn to analyzing the evolution of forecast disagreement at shorter horizons. Consistent with our broad empirical findings, we make further assumptions concerning information arrival.

Assumption 2. At forecast horizon h months, forecasters receive public signal L_{th} concerning the target-year GDP growth but may not interpret it identically. In particular, individual i 's estimate of GDP growth of the target year, Y_{ith} , conditional only on the new public signal that is observed at forecast horizon h months, can be written as

$$Y_{ith} = L_{th} - \varepsilon_{ith}, \quad \varepsilon_{ith} \sim N(\mu_{ith}, b_{ith}). \tag{5}$$

Assumption 2 allows for the possibility that agents can interpret the same public signal differently, which is captured by μ_{ith} with associated uncertainty b_{ith} . At each month, all agents observe new public signal but disagree on its effect on GDP growth for the target year. One expert can interpret the signal more optimistically or pessimistically than another. The difference in the mean beliefs, conditional on the public signal, is simply $\mu_{ith} - \mu_{jth}$ for any $i \neq j$, t and h . 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. Assumption 2 is in line with the empirical evidence presented above about significant between-agent variation in forecast revisions, and also with a larger finance literature that equally informed agents can interpret the same information differently (cf. Varian (1989), Harris and Raviv (1993), Kandel and Pearson (1995), Veronesi (2000) and Dominitz and Manski (2005)).

Bayes rule implies that under the normality assumptions, the posterior mean is the weighted average of the prior mean and the likelihood:

$$F_{ith} = \lambda_{ith}F_{ith+1} + (1 - \lambda_{ith})(L_{th} - \mu_{ith}), \tag{6}$$

where $\lambda_{ith} = a_{ith+1}/(a_{ith+1} + b_{ith})$ is the weight attached on prior beliefs. For convenience, the following population parameters are defined across professional forecasters for target year t at horizon h months: F_{th} and $\sigma_{F_{th}}^2$ are the mean and variance of forecasts of GDP growth; λ_{th} and $\sigma_{\lambda_{th}}^2$ are the mean and variance of the population distribution of the weights attached on prior beliefs; and μ_{th} and $\sigma_{\mu_{th}}^2$ are the mean and variance of the population distribution of interpretation of public information. From (6), the variance of the posterior forecasts can be defined as:

$$\begin{aligned} \text{Var}(F_{ith}) &= \text{Var}(\lambda_{ith}F_{ith+1}) + \text{Var}[(1 - \lambda_{ith})(L_{th} - \mu_{ith})] \\ &\quad + 2\text{Cov}[\lambda_{ith}F_{ith+1}, (1 - \lambda_{ith})(L_{th} - \mu_{ith})]. \end{aligned} \tag{7}$$

Following Kandel and Zilberfarb (1999), we make the following assumption, which simplifies the algebra.

Assumption 3. F_{ith+1} , λ_{ith} and μ_{ith} are mutually independent of each other for any t and h .

Recall that the formula $\text{Var}(XY) = \text{Var}(X)\text{Var}(Y) + \text{Var}(X)[E(Y)]^2 + \text{Var}(Y)[E(X)]^2$ holds if the random variables X and Y are independent of each other. Then under Assumption 3, we have

$$\text{Var}(\lambda_{ith}F_{ith+1}) = \sigma_{\lambda_{th}}^2 \sigma_{F_{th+1}}^2 + \sigma_{\lambda_{th}}^2 F_{th+1}^2 + \sigma_{F_{th+1}}^2 \lambda_{th}^2, \tag{8}$$

$$\begin{aligned} \text{Var}[(1 - \lambda_{ith})(L_{th} - \mu_{ith})] &= \sigma_{\lambda_{th}}^2 \sigma_{\mu_{th}}^2 + \sigma_{\lambda_{th}}^2 (L_{th} - \mu_{th})^2 \\ &\quad + \sigma_{\mu_{th}}^2 (1 - \lambda_{th})^2, \end{aligned} \tag{9}$$

and

$$\text{Cov}[\lambda_{ith}F_{ith+1}, (1 - \lambda_{ith})(L_{th} - \mu_{ith})] = -(L_{th} - \mu_{th})F_{th+1}\sigma_{\lambda_{th}}^2. \tag{10}$$

Substituting (8)–(10) into (7), we get

$$\begin{aligned} \sigma_{F_{ith}}^2 &= \sigma_{F_{th+1}}^2 (\sigma_{\lambda_{th}}^2 + \lambda_{th}^2) + \sigma_{\mu_{th}}^2 [\sigma_{\lambda_{th}}^2 + (1 - \lambda_{th})^2] \\ &\quad + \sigma_{\lambda_{th}}^2 (L_{th} - \mu_{th} - F_{th+1})^2. \end{aligned} \tag{11}$$

From (6), one can define forecast revision as

$$\Delta F_{ith} \equiv F_{ith} - F_{ith+1} = (1 - \lambda_{ith})(L_{th} - \mu_{ith} - F_{ith+1}). \tag{12}$$

Under Assumption 3, the mean forecast revision, ΔF_{th} , can be derived as

$$\Delta F_{th} \equiv E(\Delta F_{ith}) = (1 - \lambda_{th})(L_{th} - \mu_{th} - F_{th+1}). \tag{13}$$

Substituting $(L_{th} - \mu_{th} - F_{th+1}) = \Delta F_{th}/(1 - \lambda_{th})$ from (13) into (11), we get

$$\begin{aligned} \sigma_{F_{ith}}^2 &= \sigma_{F_{th+1}}^2 (\sigma_{\lambda_{th}}^2 + \lambda_{th}^2) + \sigma_{\mu_{th}}^2 [\sigma_{\lambda_{th}}^2 + (1 - \lambda_{th})^2] \\ &\quad + \sigma_{\lambda_{th}}^2 [\Delta F_{th}/(1 - \lambda_{th})]^2. \end{aligned} \tag{14}$$

In (14) forecast disagreement is posited to have three components due to inter-agent differences in (i) the prior beliefs, $\sigma_{F_{th+1}}^2$; (ii) the weights attached on priors, $\sigma_{\lambda_{th}}^2$ and (iii) the interpretation of the public signal, $\sigma_{\mu_{th}}^2$, cf. Kandel and Zilberfarb (1999). It encompasses a number of special cases. In one case where all agents attach the same weight on their prior beliefs (i.e. $\sigma_{\lambda_{th}}^2 = 0$ for any t and h), (14) becomes

$$\sigma_{F_{ith}}^2 = \lambda_{th}^2 \sigma_{F_{th+1}}^2 + (1 - \lambda_{th})^2 \sigma_{\mu_{th}}^2. \tag{15}$$

In another case where all agents interpret the public signal identically (i.e. $\sigma_{\mu_{th}}^2 = 0$ for any t and h), (14) is simplified to

$$\sigma_{F_{ith}}^2 = \lambda_{th}^2 \sigma_{F_{th+1}}^2 + \sigma_{\lambda_{th}}^2 [\Delta F_{th}/(1 - \lambda_{th})]^2 + \sigma_{F_{th+1}}^2 \sigma_{\lambda_{th}}^2. \tag{16}$$

Our model suggests the following scenario of expectation formation by professional forecasters. At 24 months ahead, agents hold divergent initial prior beliefs about GDP growth for the next year. Agents stick to their initial prior beliefs during horizons 24 to 16 months due to predisposition effect and lack of dependable public information. Starting from horizon 15 months, agents observe the same public signal but interpret it differently from one another. Each agent then forms new posterior distributions by combining his prior and perception of the public information using the Bayes rule. At horizon 14 months, each agent combines his updated prior belief, which is the posterior formed last month, and his perception of new information, which is observed during the current month, to derive the new posterior. The process continues until the end of forecasting exercise for that target year. This cycle of prior to posterior is actually a very solid way of characterizing the forecasting process: forecasters take what knowledge they have in hand and update it with the arrival of new information (cf. El-Gamal and Grether (1995)).

This section is closely related to the literature on learning in expectation formations. Notable works in this area go back to Cyert and DeGroot (1974), Townsend (1983), Haltiwanger and Waldman (1985) and Bray and Savin (1986). These papers all allow for heterogeneity in priors and with Bayesian learning, the heterogeneity disappears in the limit. More recently, Acemoglu et al. (2006) show that asymptotic agreement is not assured when there is uncertainty in the interpretation of public signals.

It is interesting to look at two limiting cases in our framework. In one case, public information is so uncertain that its precision b_{ith} approaches zero in the limit. As a result, forecasters will attach 100% weight to their prior beliefs and the disagreement will reflect only the diversity in their prior beliefs, that is, $\sigma_{F_{ith}}^2 = \sigma_{F_{th+1}}^2$. Thus, even after observing a sequence of public signals, forecasters will continue to sustain the original disagreement due to its high uncertainty. This seems to characterize our longer horizon forecasts. In the other polar case, the public information

arrives with precision b_{ith} approaching infinity in the limit. In this case, forecasters attach 100% weight to the new signal, and the disagreement will depend only on the extent of experts' differential interpretation of information, that is, $\sigma_{F_{ith}}^2 = \sigma_{\mu_{ith}}^2$. Since it is expected that experts will interpret highly certain information similarly, forecast disagreement will disappear in the limit. This seems to characterize our shorter horizon forecasts inside the target year.

4. Estimation results

In this section, we show how to estimate the parameters in our learning model using the panel data. We then proceed to give explanations for the empirical results.

4.1. Estimation of λ_h

Let us first focus on estimating the weight attached on prior belief (λ_h). Eq. (6) shows that

$$F_{ith} = \lambda_{ith}F_{ith+1} + \varepsilon_{ith}, \tag{17}$$

where $\varepsilon_{ith} = (1 - \lambda_{ith})(L_{th} - \mu_{ith})$ is the error term. Eq. (17) corresponds to the most general specification of the panel data regression model. However, as pointed out by Hsiao and Pesaran (2008), this general formulation is not estimable, since the number of parameters to be estimated exceeds the number of observations. Therefore, it is essential to impose some restrictions on its parameters and error structures. In order to reduce the number of parameters in (17), we assume

$$\lambda_{ith} = \lambda_{ih} = \lambda_h + v_{ih}, \tag{18}$$

where v_{ih} is assumed to be mean zero, mutually independent of each other, and 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 estimable version of (17) becomes

$$\Delta F_{ith} = \beta_h F_{ith+1} + u_{ith}, \tag{19}$$

where $\beta_h = \lambda_h - 1$ and $u_{ith} = \varepsilon_{ith} + v_{ih}F_{ith+1}$. Note unlike standard random coefficient dynamic panel data models, u_{ith} is uncorrelated with F_{ith+1} in (19). To see this, by model construction, F_{ith+1} is independent of λ_{ith} and μ_{ith} , and thus $E(\varepsilon_{ith}F_{ith+1}) = 0$. Furthermore, through continuous substitutions, $F_{ith+1} = \sum_{j=0}^{\infty} \left[\prod_{s=0}^{j-1} (\lambda_{h+1+s} + v_{ih+1+s}) \right] \varepsilon_{ith+1+j}$, and thus $E(v_{ih}F_{ith+1}) = 0$. However, u_{ith} might be cross-sectionally correlated due to common aggregate shocks. Following the line of reasoning in Pesaran (2006), we proxied the common effects in the residuals by \bar{F}_{th+1} and $\Delta \bar{F}_{th}$ with the assumption that u_{ith} is serially uncorrelated in this augmented specification.¹¹

It may seem desirable to estimate the panel data model in (19) with all three dimensions by imposing some smooth functional form for β_h over horizons, cf. Gregory and Yetman (2004). However, as shown later, the estimated β varies unevenly over horizons and across countries, depending on the timing of public information arrival. Thus, we estimated (19) separately for each horizon, augmented by a few additional regressors to

¹¹ Initially we included both ΔF_{ith+1} and $\Delta \bar{F}_{th+1}$ in the regressions to control for serial correlation in the residuals. It turns out that the coefficients on these two terms were statistically insignificant at the 5% critical value for most of horizons and countries. As a result, we could safely drop these two terms to maintain a more parsimonious model.

asymptotically filter out the effects of the cross-section correlation in the residuals¹²:

$$\Delta F_{ith} = \beta_h F_{ith+1} + c_h z_{ith} + \eta_{ith}, \tag{20}$$

where $z_{ith} = (1, \bar{F}_{th+1}, \Delta \bar{F}_{th})$ and $E(\eta_{ith} \eta_{i't'h'}) = \sigma_{\eta(t)}^2$, for $i = i'$, $t = t'$, $h = h'$ and 0, otherwise.

Table 2 shows the estimates of the weights attached to public information $1 - \hat{\lambda}_h$ (i.e., $1 - \hat{\lambda}_h = -\hat{\beta}_h$). On the average, 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 it becomes more precise and certain. In addition, there are significant boosts in estimated weights at certain horizons, whenever new GDP growth estimates become available. For example, the estimated weight for the United States gets a boost at horizon 8 months as a result of the advance release of the first quarter GDP growth at the end of April. It gets a boost again at horizons 5 and 2 months as a result of the advance release of the second and third quarter GDP growth, respectively. The matching in the pattern of horizon-specific weights attached to public information with the timing of the GDP growth announcements also holds for other countries except for France and Germany. In these two countries forecasters may not be using new public information efficiently at certain horizons. This possibility is explored formally in Section 5.

4.2. Estimation of $\sigma_{\lambda_{ih}}^2$ and $\sigma_{\mu_{ih}}^2$

Now we turn our attention to estimating the variance of the weights attached to prior beliefs ($\sigma_{\lambda_{ih}}^2$) and the variance of interpretation of public information ($\sigma_{\mu_{ih}}^2$). Substituting estimated values of $\hat{\lambda}_h$ into (14) and rearranging it, we have

$$\sigma_{F_{ith}}^2 - \hat{\lambda}_h^2 \sigma_{F_{ith+1}}^2 = \sigma_{\lambda_{ith}}^2 \sigma_{\mu_{ith}}^2 + \sigma_{\lambda_{ith}}^2 [\sigma_{F_{ith+1}}^2 + \Delta F_{th}^2 / (1 - \hat{\lambda}_h)^2] + \sigma_{\mu_{ith}}^2 (1 - \hat{\lambda}_h)^2. \tag{21}$$

Several econometric issues arise in the estimation of (21). First, certain restrictions have to be imposed on its parameters. Since we are interested in horizon-specific effects, we assume that both $\sigma_{\lambda_{ith}}^2$ and $\sigma_{\mu_{ith}}^2$ are invariant to target years. Second, with only 14 observations for each horizon and each country, the estimates may be unreliable. To increase the number of observations, we pooled all seven countries. As a forecaster-specific variable, $\sigma_{\lambda_{ih}}^2$ is expected to be similar across countries, and was accordingly constrained to be same.¹³ In the pooled data set, the residuals are likely to be cross-country correlated because of common shocks in G7 economies. We estimated (21) by OLS with standard errors of parameters calculated using panel-corrected standard errors (PCSE), which corrects for contemporaneous correlation in the residuals.¹⁴ Finally, whenever $\hat{\lambda}_h$ is close to 1, which is occasionally the case at very long horizons, the second term on the right-hand side in (21) (i.e. $\sigma_{F_{ith+1}}^2 + \Delta F_{th}^2 / (1 - \hat{\lambda}_h)^2$) will blow up and the third term on the right-hand side in (21) (i.e. $(1 - \hat{\lambda}_h)^2$) will approach zero. To deal with this issue, we arbitrarily put an upper bound of 0.90 on $\hat{\lambda}_h$.

Table 3 presents the estimated parameters. As reported in the first column, the estimates of $\hat{\sigma}_{\lambda_{ih}}^2$ are almost zero for all

¹² This estimator is referred to as common correlated effects (CCE) estimator, which seems to have satisfactory small sample properties in a moderate sample size like ours where $N = 30$ and $T = 14$ (Pesaran, 2006). For an application of this estimator in the context of panel unit root tests, see Pesaran (2007).

¹³ As a robustness check, we estimated $\sigma_{\lambda_{ih}}^2$ separately for each country. The F -test could not reject the hypothesis that they are same across countries at the 5% significance level.

¹⁴ With sample size $N = 7$ and $T = 14$, Beck and Katz (1995) show that generalized least squares underestimates the standard errors quite severely.

Table 2

Estimates of the weight attached to public information, $1 - \hat{\lambda}_h$

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	0.65 (0.06)	0.61 (0.06)	0.38 (0.04)	0.64 (0.05)	0.48 (0.05)	0.60 (0.03)	0.50 (0.03)
2	0.51 (0.05)	0.28 (0.04)	0.35 (0.03)	0.22 (0.05)	0.25 (0.04)	0.43 (0.03)	0.67 (0.03)
3	0.65 (0.04)	0.61 (0.04)	0.39 (0.03)	0.34 (0.06)	0.54 (0.04)	0.38 (0.03)	0.44 (0.04)
4	0.35 (0.05)	0.27 (0.04)	0.42 (0.03)	0.17 (0.06)	0.22 (0.04)	0.36 (0.03)	0.38 (0.04)
5	0.31 (0.04)	0.10 (0.03)	0.22 (0.03)	0.25 (0.05)	0.15 (0.03)	0.36 (0.03)	0.45 (0.04)
6	0.49 (0.05)	0.62 (0.04)	0.28 (0.03)	0.56 (0.05)	0.37 (0.04)	0.31 (0.04)	0.33 (0.03)
7	0.38 (0.05)	0.23 (0.05)	0.30 (0.04)	0.35 (0.05)	0.13 (0.04)	0.19 (0.03)	0.30 (0.03)
8	0.19 (0.04)	0.19 (0.03)	0.29 (0.03)	0.14 (0.05)	0.11 (0.02)	0.26 (0.03)	0.38 (0.03)
9	0.30 (0.04)	0.36 (0.04)	0.19 (0.03)	0.18 (0.05)	0.24 (0.03)	0.23 (0.03)	0.26 (0.03)
10	0.23 (0.04)	0.16 (0.04)	0.29 (0.04)	0.23 (0.05)	0.12 (0.03)	0.16 (0.03)	0.19 (0.03)
11	0.20 (0.04)	0.15 (0.03)	0.21 (0.03)	0.38 (0.04)	0.01 (0.02)	0.10 (0.02)	0.33 (0.03)
12	0.24 (0.05)	0.22 (0.04)	0.24 (0.03)	0.29 (0.06)	0.29 (0.03)	0.22 (0.03)	0.23 (0.03)
13	0.20 (0.04)	0.15 (0.04)	0.25 (0.04)	0.32 (0.06)	0.25 (0.03)	0.20 (0.03)	0.19 (0.03)
14	0.17 (0.04)	0.12 (0.04)	0.38 (0.03)	0.29 (0.05)	0.08 (0.03)	0.11 (0.03)	0.20 (0.03)
15	0.25 (0.04)	0.36 (0.05)	0.19 (0.03)	0.30 (0.07)	0.20 (0.04)	0.17 (0.03)	0.20 (0.04)
16	0.05 (0.04)	0.23 (0.04)	0.13 (0.03)	0.14 (0.06)	0.03 (0.04)	0.13 (0.03)	0.13 (0.03)
17	0.18 (0.03)	0.08 (0.03)	0.13 (0.03)	0.29 (0.06)	0.07 (0.02)	0.08 (0.03)	0.17 (0.03)
18	0.13 (0.02)	0.31 (0.05)	0.21 (0.03)	0.37 (0.06)	0.07 (0.04)	0.13 (0.02)	0.21 (0.03)
19	0.04 (0.03)	0.15 (0.04)	0.21 (0.03)	0.29 (0.05)	0.10 (0.04)	0.13 (0.02)	0.09 (0.02)
20	0.15 (0.03)	0.12 (0.03)	0.16 (0.03)	0.15 (0.05)	0.13 (0.03)	0.07 (0.02)	0.12 (0.03)
21	0.09 (0.03)	0.16 (0.03)	0.23 (0.03)	0.14 (0.04)	0.12 (0.04)	0.13 (0.02)	0.11 (0.03)
22	0.05 (0.03)	0.16 (0.05)	0.18 (0.03)	0.17 (0.05)	0.23 (0.04)	0.06 (0.02)	0.20 (0.03)
23	0.14 (0.04)	0.08 (0.05)	0.25 (0.03)	0.20 (0.05)	0.02 (0.03)	0.07 (0.02)	0.20 (0.04)

Note: Standard errors are in parenthesis.

horizons, which simply means that professional forecasters place very similar weights on their updated prior beliefs. It is interesting to compare our results with Kandel and Zilberfarb (1999). In their paper, $\sigma_{\lambda|ih}^2$ was not estimated directly. Instead, they assumed that λ_{ith} is distributed Beta, that is, $\lambda_{ith} \sim Beta(a, b_{ih})$ and accordingly $\sigma_{\lambda|ih}^2$ was expressed as

$$\sigma_{\lambda|ih}^2 = \lambda_h^2(1 - \lambda_h)/(a + \lambda_h). \tag{22}$$

Substituting $a = 3$ assumed in their calculations and the estimated values of λ_h from our Table 2 in (22) yields approximate values of $\sigma_{\lambda|ih}^2$ for different horizons. We find $\sigma_{\lambda|ih}^2$ to range from 0.016 to 0.038 implied by their assumption $a = 3$. In contrast, the cross-country feature of our data allows us to estimate $\sigma_{\lambda|ih}^2$ directly and our estimates show that it ranges from 0.001 to 0.018 for almost all horizons. The difference in these estimates, however, turned out to be inconsequential for our conclusions in that this channel was found to be unimportant in generating experts' disagreement using either set of $\sigma_{\lambda|ih}^2$ estimates.

The estimated values of $\hat{\sigma}_{\mu|ih}^2$ are presented in the last seven columns of Table 3. Three points bear emphasis. First, for majority of the horizons, the estimated variance of the interpretation of public information is statistically significant at the 5%

level for all countries.¹⁵ This provides direct evidence in support of the hypothesis that equally informed agents can sometimes interpret the same public information differently. To the best of our knowledge, this parameter has never been estimated directly, though Kandel and Pearson (1995) and Kandel and Zilberfarb (1999) pioneered testing such a hypothesis indirectly. Our finding has important implications for belief formation and learning. It says that a multitude of competing models can simultaneously exist to interpret new information. Differential interpretation of public information can be a challenge for establishing the credibility of monetary policies. Also, heterogeneity in analysts' expectations and thus trading in the asset markets can arise from model diversity. Second, on average, $\hat{\sigma}_{\mu|ih}^2$ becomes smaller as forecast horizon declines. Using forecasts made only for

¹⁵ Note that $\hat{\sigma}_{\mu|ih}^2$ is considerably higher for Japan than other countries. This may suggest that public information is very differently perceived by Japanese forecasters. However, these estimates also came with considerably larger standard errors. This is because forecast disagreement in Japan is almost three times more than in other countries, see Fig. 1. During our sample period, the Japanese economy has behaved differently in many other ways. See, for example, Stock and Watson (2005) and references therein.

Table 3
Estimates of $\sigma_{\lambda|h}^2$ and $\sigma_{\mu|h}^2$

Horizon	$\sigma_{\lambda h}^2$	$\sigma_{\mu h}^2$						
		Canada	France	Germany	Italy	Japan	UK	US
1	0.047 (0.005)	0.02 (0.01)	0.02 (0.00)	0.05 (0.03)	0.02 (0.01)	0.15 (0.04)	0.02 (0.00)	0.01 (0.01)
2	0.000 (0.000)	0.05 (0.01)	0.12 (0.05)	0.12 (0.06)	0.32 (0.14)	1.09 (0.55)	0.06 (0.02)	0.02 (0.01)
3	0.004 (0.002)	0.04 (0.01)	0.05 (0.01)	0.10 (0.02)	0.20 (0.09)	0.24 (0.08)	0.11 (0.02)	0.07 (0.02)
4	0.000 (0.000)	0.23 (0.07)	0.22 (0.05)	0.19 (0.05)	0.65 (0.28)	2.57 (0.97)	0.26 (0.05)	0.10 (0.02)
5	0.012 (0.003)	0.12 (0.08)	0.56 (0.23)	0.42 (0.10)	0.13 (0.11)	2.19 (1.22)	0.20 (0.10)	0.08 (0.02)
6	0.008 (0.002)	0.23 (0.10)	0.08 (0.02)	0.43 (0.13)	0.09 (0.02)	0.45 (0.36)	0.49 (0.13)	0.14 (0.03)
7	0.003 (0.001)	0.49 (0.12)	0.47 (0.18)	0.62 (0.26)	0.31 (0.14)	6.28 (2.60)	0.60 (0.26)	0.23 (0.09)
8	0.017 (0.003)	0.33 (0.69)	0.19 (0.15)	0.06 (0.30)	1.13 (1.15)	2.97 (1.81)	0.36 (0.23)	0.19 (0.07)
9	0.001 (0.001)	0.50 (0.41)	0.31 (0.13)	0.74 (0.27)	0.87 (0.31)	1.39 (1.14)	0.90 (0.19)	0.47 (0.17)
10	0.006 (0.001)	1.01 (0.31)	0.78 (0.37)	0.81 (0.25)	0.47 (0.11)	3.85 (1.53)	1.30 (0.44)	0.85 (0.31)
11	0.004 (0.001)	1.75 (1.38)	1.18 (0.47)	1.24 (0.44)	0.26 (0.22)	7.71 (2.21)	2.40 (0.71)	0.68 (0.25)
12	0.002 (0.001)	2.33 (1.19)	1.21 (0.43)	1.18 (0.50)	0.76 (0.45)	1.76 (0.94)	0.66 (0.29)	1.10 (0.31)
13	0.006 (0.001)	1.89 (0.68)	0.86 (0.32)	0.74 (0.22)	0.50 (0.15)	2.43 (1.11)	1.99 (0.57)	1.28 (0.72)
14	0.005 (0.002)	2.89 (1.18)	1.18 (0.63)	0.45 (0.17)	0.66 (0.24)	10.47 (2.63)	2.94 (1.34)	1.38 (0.76)
15	0.007 (0.000)	0.44 (0.56)	0.48 (0.10)	0.61 (0.38)	0.51 (0.10)	0.65 (2.84)	1.65 (0.76)	0.91 (0.55)
16	0.004 (0.001)	4.27 (0.97)	0.53 (0.08)	2.65 (0.74)	0.41 (0.55)	14.85 (4.73)	2.19 (0.65)	2.85 (1.33)
17	0.013 (0.002)	1.34 (1.47)	0.85 (0.44)	1.83 (0.52)	0.84 (0.45)	4.79 (1.64)	2.66 (0.86)	1.55 (0.82)
18	0.012 (0.004)	0.57 (0.93)	0.75 (0.51)	1.29 (0.46)	0.27 (0.07)	6.42 (4.26)	2.06 (0.66)	0.99 (0.42)
19	0.007 (0.002)	5.38 (1.37)	1.08 (0.28)	0.65 (0.28)	0.42 (0.13)	12.76 (5.36)	2.54 (0.72)	2.63 (0.76)
20	0.011 (0.004)	2.48 (0.91)	0.32 (0.23)	1.20 (0.65)	1.31 (0.47)	4.28 (1.87)	2.62 (0.91)	1.96 (0.35)
21	0.000 (0.000)	4.06 (0.91)	0.94 (0.45)	1.23 (0.34)	0.92 (0.47)	5.52 (1.89)	2.24 (0.40)	3.11 (1.09)
22	0.018 (0.003)	2.47 (0.90)	0.93 (0.67)	0.94 (0.57)	0.46 (0.49)	4.14 (4.26)	2.44 (0.71)	1.95 (0.56)
23	0.020 (0.017)	3.04 (2.06)	1.63 (0.81)	1.01 (0.54)	0.67 (0.25)	6.71 (3.66)	2.75 (0.75)	0.90 (0.81)

Note: Standard errors are in parenthesis.

3- and 9-month horizons, Kandel and Zilberfarb (1999) hypothesized that the distribution of “interpretation bias” for longer horizons should exhibit higher variance. Our estimation result, based on forecasts for horizons from 24 months to 1 month, provides comprehensive evidence in support of their conjecture. Third, the difference in the interpretation of public signal depends on the degree of uncertainty in new information. We observe a sharp decline in $\hat{\sigma}_{\mu|h}^2$ for almost all sampled countries around horizon 15 months, when the first relevant public information arrives. Then depending on the timing of the release of GDP growth, we observe subsequent declines in the estimated values of $\hat{\sigma}_{\mu|h}^2$. For example, there is a sudden drop in $\hat{\sigma}_{\mu|h}^2$ for the United States at horizon 8 months, when the first quarter GDP growth for the target year becomes available. Note that the informativeness and nature of the public signals at $h = 15$ and $h = 8$ are quite different. At horizon 15 months, public signals are in the form of monthly indicators and leading indexes that contain only indirect and limited information regarding the target year GDP growth. At horizons 8 months and less, public information comes in the form of actual GDP estimates that becomes

part of target year GDP growth.¹⁶ Forecasters interpret these more certain information more homogeneously, leading to a sharp decline in forecast disagreement at $h = 8$ compared to $h = 15$. Thus, we find evidence for the hypothesis that asymptotic agreement is achievable when the uncertainty about interpretation of public signals disappears.

With the estimates of parameters in hand, we can compare the predicted disagreement by our model to the disagreement observed in the survey data over all target years in our sample. Substituting the parameter estimates in (14) and dynamically generating the fitted values yield the predicted disagreement by our model. We find that on the average our estimated model explains about 60% of the total variation in observed disagreement over the target years. Considering the fact that forecast disagreement varies a lot from year to year for any specific horizon, our model does a good job in explaining the evolution of forecast disagreement over target years and horizons.

¹⁶ Interestingly, Patton and Timmermann (2007) modeled this shift in the informativeness of signals at $h = 12$.

Table 4
Contribution of three channels in explaining forecast disagreement

Horizon	Forecast disagreement due to differences in		
	Initial prior beliefs	Weight attached on priors	Interpretation of public information
1	0.35	0.05	0.60
2	0.48	0.00	0.52
3	0.37	0.03	0.60
4	0.18	0.02	0.80
5	0.26	0.00	0.74
6	0.31	0.08	0.62
7	0.40	0.00	0.60
8	0.49	0.11	0.40
9	0.44	0.00	0.56
10	0.47	0.00	0.53
11	0.53	0.02	0.45
12	0.52	0.02	0.46
13	0.41	0.00	0.59
14	0.49	0.00	0.51
15	0.34	0.18	0.48
16	0.42	0.02	0.56
17	0.49	0.03	0.48
18	0.49	0.03	0.48
19	0.66	0.05	0.29
20	0.88	0.07	0.05
21	0.98	0.00	0.02
22	0.93	0.04	0.03
23	1.00	0.00	0.00

Note: The sum of contributions of all three channels is normalized to 1 for each horizon.

Recall that forecast disagreement is postulated to have three components due to differences in: (i) the initial prior beliefs, (ii) the weights attached on updated priors, and (iii) the interpretation of public information. To gauge the relative importance of each of three channels, we perform the following experiment. With the second channel being closed, we calculate the predicted disagreement by our model in (14). We then regress the observed disagreement on predicted one and get R^2 . The difference in R^2 with all three channels on and with the second channel closed shows approximately the contribution of the second channel to forecast disagreement. By a similar argument, we separate out the contributions of each of the remaining two channels. Table 4 presents the relative importance of each channel over horizons. As expected, the importance of the first channel - differences in the initial prior beliefs - steadily declines, as forecast horizons get shorter. The second channel - differences in the weights attached to updated priors - accounts only for 3% of forecast disagreement on average, and thus has barely any effect on the evolution of forecast disagreement over horizons. The third channel - differences in the interpretation of public information - explains about 50%, and thus becomes a major source of forecast disagreement for horizon 18 months and less. Note that $\sigma_{\mu_{i|h}}^2$, the heterogeneity parameter in the interpretation of public signals, affects forecast disagreement only through its interaction with the weight attached to public information, which depend on the quality of public information.

5. Efficiency of using public information

As is well known, Bayes' theorem implies that under the normality assumption, the posterior mean of the target variable is a weighted average of the prior mean and the likelihood. Correspondingly, forecaster i 's conditional estimate of target-year GDP growth Y_t , given $F_{i,t+h}$ and $L_{i,t}$, is formed as

$$E(Y_t | F_{i,t+h}, L_{i,t}) = \lambda_{i,t} F_{i,t+h} + (1 - \lambda_{i,t})(L_{i,t} - \mu_{i,t}). \tag{23}$$

Zellner (1988) has shown that the above Bayesian information updating rule is 100% efficient, since no information is lost or added when (23) is employed. Thus the weight attached on prior belief,

$\lambda_{i,t} = a_{i,t+h} / (a_{i,t+h} + b_{i,t})$, is the efficient weight under which $F_{i,t} = E(Y_t | F_{i,t+h}, L_{i,t})$ as in (6). However, forecasters may not apply the efficient weight $\lambda_{i,t}$ in making forecasts. For simplicity, the forecast made by agent i for the target year t and h months ahead of the end of target year is assumed also to have the form

$$F_{i,t} = \delta_{i,t} F_{i,t+h} + (1 - \delta_{i,t})(L_{i,t} - \mu_{i,t}), \tag{24}$$

where $\delta_{i,t}$ is the actual weight forecaster i places on his prior belief. We observe that forecaster i overweights his prior belief (or underweights public information) if $\delta_{i,t} > \lambda_{i,t}$.

Interestingly, Chen and Jiang (2006) derived another form of test for forecast efficiency by combining (23) and (24). Their test is based on the relationship between an analyst's forecast error and the deviation of his forecast from the public information, that is, $E(Y_t - F_{i,t} | F_{i,t+h}, L_{i,t}) = \frac{\lambda_{i,t} - \delta_{i,t}}{\delta_{i,t}} (F_{i,t} - L_{i,t})$ by assuming identical interpretation of public information (i.e. $\mu_{i,t} = 0$ for any i, t and h). In their empirical analysis, public information is approximated by a weighted average of previous period's forecasts.¹⁷ Instead, without depending on certain *ad hoc* proxy for public information, our test is based on the observable properties of forecasts.

Our alternative formulation of the efficiency test builds on the relation between forecast error and forecast revision. Given the actual forecast in (24), we can express the expected forecast error as:

$$E(Y_t - F_{i,t} | F_{i,t+h}, L_{i,t}) = \frac{\delta_{i,t} - \lambda_{i,t}}{1 - \delta_{i,t}} (F_{i,t} - F_{i,t+h}), \tag{25}$$

which yields a testable relation between forecast error ($E_{i,t}$) and forecast revision ($R_{i,t}$)¹⁸:

$$E_{i,t} = \theta_{i,t} R_{i,t} + \varepsilon_{i,t}, \tag{26}$$

where $E_{i,t} = Y_t - F_{i,t}$, $R_{i,t} = F_{i,t} - F_{i,t+h}$ and $\theta_{i,t} = (\delta_{i,t} - \lambda_{i,t}) / (1 - \delta_{i,t})$. Under the null hypothesis that forecasters use efficient weights (i.e., $\delta_{i,t} = \lambda_{i,t}$), the coefficient estimate $\hat{\theta}_{i,t}$ should be zero. Since

¹⁷ Alternatively, Batchelor and Dua (1992) approximated public information with the actual value plus some random error.

¹⁸ It is interesting to note that this relationship is one of the two joint conditions used by Ehrbeck and Waldmann (1996) to test their theory of rational strategic bias.

δ_{it} lies between 0 and 1, a positive (negative) $\hat{\theta}_{it}$ suggests underweighting (overweighting) of public information. The intuition behind the relationship is straightforward. Whereas forecast revision can be taken as a measure of how forecasters interpret the importance of public information in real time, forecast error is the *ex post* “prize” they get as a result of revising their forecasts. Suppose that forecasters make large revisions at horizon h months but the performance of the forecasts do not improve much at that horizon; then one may conjecture that forecasters overweight new public information.

Although our test looks similar to the weak efficiency test in Nordhaus (1987),¹⁹ our formulation is different in an important way. While weak efficiency requires that current forecast errors be uncorrelated with *past* forecast revisions (cf. Proposition 1 in Nordhaus (1987)), our test requires that current forecast errors be uncorrelated with *current* forecast revisions. Since current forecast revisions reflect the forecaster’s perception of the latest public information concerning the target variable, our test is expected to have more power than the weak efficiency test. On the other hand, our formulation is not the *strong* form of efficiency test in the sense of Muth (1961) that requires expectations to reflect *all* relevant available information immediately. Thus we test a *semi-strong* form of efficiency in the spirit of Fama (1970).

Note that the weak efficiency test of Nordhaus (1987) also requires that forecast revisions be uncorrelated with past forecast revisions (cf. Proposition 2 in Nordhaus (1987)). As pointed out by Clements (1997), however, the maintained assumption for this proposition is that forecasters’ behavior in incorporating new information is similar at all horizons. Due to possible offsetting news in the future and the associated additional uncertainty, forecasts at longer horizons may be stickier than those at shorter horizons. More importantly, since the flow and quality of information in our case is uneven and lumpy over successive months, forecasters may absorb information differently at different horizons. For these reasons we do not use the Nordhaus formulation here.²⁰

Rather than pooling over all horizons, as is the case with most studies on fixed-event forecasts, we test forecast efficiency for each horizon. To perform the test, we check if $\theta = 0$ in the following regression for any specific horizon h :

$$E_{it} = \alpha + \theta R_{it} + \varepsilon_{it}. \quad (27)$$

If θ is found to be significantly positive (negative), the conclusion is that forecasters underweight (overweight) public information. Unfortunately, test based on OLS regression in (27) yields inconsistent standard error estimates due to the fact that the residuals may be serially correlated at longer horizons, and also cross-sectionally correlated due to common shocks. The serial correlations arise because the realized values of Y_{t-2} and Y_{t-1} are not known when some of the F_{it} ’s are reported. Hence the corresponding forecast errors E_{it-2} and E_{it-1} are not observable at longer horizons. Depending on the timing of GDP announcements, the structure of the serial correlation is slightly different at different horizons for different countries in our sample. To be safe, we allowed for the maximum order of serial correlation under efficiency in each case. Thus, we specified a MA (2) process for horizons 19 to 22 months, a MA (1) process for horizons 10 to 18 months, and no serial correlation for the remaining short horizons. We estimate the coefficients in (27) using GMM after controlling for both cross-section and serial correlation in the residuals using the appropriate weighting matrix. Estimation results are shown in Table 5. Although many estimates are not close to zero, given

the standard errors of the estimates, the data do not reject the null hypothesis of forecast efficiency for most of horizons and countries. However, evidence also indicates certain degree of forecast inefficiency for some countries and some horizons. More specifically we note the following:

First, forecasters seem to put more than efficient weight on new public information at very long horizons, as displayed by many statistically significant and negative coefficient estimates at longer horizons. Since we have found that public signals concerning next year GDP growth are not very informative during the initial eight monthly rounds of forecasting, experts are found to make unnecessary revisions during this period. While analyzing British fixed-target forecasts with horizons pooled up to 12 quarters, Clements (1995, 1997) also found negative autocorrelations in forecast revisions and interpreted them as evidence of absence of significant news over the period. Note that depending on the quality of public information, certain amount of random adjustment in forecasts can be rationalized in our Bayesian learning model, even if the signals are not informative. Our forecasters may also believe incorrectly that they have superior private information with respect to next year GDP growth. Moreover, the pressure from their clients may push forecasters to revise their forecasts often, since issuing the same forecast continuously for a couple of months gives the impression of their inactiveness. Not surprisingly, in terms of a standard MSE skill score, the forecasts at these long horizons were found to be no better than the initial 24-month ahead forecasts.²¹

Second, we find that forecasters underweight public information in the middle horizons, as indicated by the statistically significant and positive estimates. As the forecast horizon gets shorter, the base-year GDP growth numbers become available with increasing certainty, which plays a large role in forecasting year-over-year growth rate. Furthermore, as we approach the end of target year, current-year GDP announcements and data revisions become part of target-year GDP growth. As a result, forecasters should put a high weight on the newly arrived public information. The degree of underweighting public information, however, is largest for France and Germany. This finding, based on individual forecasts, complements the recent empirical evidence presented by Isiklar et al. (2006).

How large is the forecasters’ underweighting of public information? To investigate this question, we calculated the efficient weights that should be attached to public information when the actual weights are adjusted for the degree of inefficiency.²² Fig. 3 plots both the efficient and actual weights attached to public information over the horizons. Efficient weights appear jagged because they incorporate new public information quickly. Actual weights appear smoother and staggered, for they are seen to let new public information seep in slowly. On the average, forecasters underweight new public information by 5% to 35%, depending on the horizons and countries under study. At horizons 15 months or less, while the degree of underweighting of new public information for United States, Italy and Japan is below 10%, the underweighting for Canada, France, Germany and the United Kingdom ranges between 25% and 35%.

Another striking finding in Fig. 3 is that the efficient weights on public information increase from almost nothing to 70%, as forecast horizon gets shorter. Put differently, forecasters attach nearly 100%

¹⁹ For the recent applications of Nordhaus’ weak efficiency test, see, Clements and Taylor (2001) and Isiklar et al. (2006).

²⁰ Note that in a data set like ours, over half of the forecasts are not adjusted at all at any particular horizon, with more such cases in the long horizons. This creates a serious limited dependent variable problem while regressing forecast revisions on their lagged values under the Nordhaus weak efficiency test.

²¹ Diebold and Kilian (2001) have proposed using $p_{s,24} = 1 - MSE_s/MSE_{24}$ to measure forecast improvement made at horizon s months over 24-month ahead forecasts for $s = 1, \dots, 23$. Using the mean GDP forecasts, Isiklar and Lahiri (2007) find this measure to remain almost same over horizons 24–18 months.

²² Note that the estimated weights in Table 2 are *actual* rather than *efficient* weights attached to public information by professional forecasters.

Table 5
Test of efficiency in using public information

Horizon	Canada	France	Germany	Italy	Japan	UK	US
1	-0.25 ^a (-3.81)	0.10 (1.46)	-0.14 ^a (-4.14)	-0.08 (-0.61)	0.26 (1.87)	-0.01 (-0.16)	0.67 ^a (7.49)
2	0.44 ^a (2.54)	0.86 ^a (3.77)	0.06 (0.51)	0.06 (0.42)	-0.17 (-1.53)	-0.25 ^a (-2.26)	-0.40 ^a (-2.88)
3	0.10 (1.19)	0.30 ^a (2.85)	0.32 ^a (3.64)	0.26 (1.27)	0.00 (0.08)	-0.07 (-0.78)	-0.24 ^a (-2.21)
4	0.31 ^a (2.18)	0.32 ^a (1.98)	0.48 ^a (6.45)	0.22 (0.59)	0.00 (0.01)	0.30 ^a (3.57)	-0.18 (-0.95)
5	0.31 (1.23)	-0.07 (-0.25)	0.60 ^a (6.59)	0.09 (0.40)	-0.17 (-0.79)	0.27 ^a (3.40)	-0.22 (-1.82)
6	0.08 (0.79)	0.00 (0.00)	0.45 ^a (3.38)	0.25 (1.61)	0.16 ^a (2.80)	0.22 (1.61)	0.04 (0.20)
7	0.43 ^a (4.20)	0.28 ^a (1.97)	0.41 ^a (2.89)	0.39 (1.85)	-0.37 ^a (-3.85)	0.98 ^a (4.42)	0.44 ^a (3.99)
8	0.52 ^a (2.08)	1.30 ^a (7.00)	0.41 ^a (2.37)	-0.23 (-0.46)	0.62 (1.83)	0.50 ^a (3.55)	0.29 ^a (4.39)
9	0.56 ^a (4.23)	0.81 ^a (3.04)	0.31 (1.30)	-0.13 (-0.34)	0.91 ^a (4.62)	0.88 ^a (6.04)	0.19 (1.23)
10	1.00 ^a (5.57)	1.38 ^a (4.85)	0.13 (0.93)	-0.14 (-0.36)	0.37 ^a (2.04)	0.86 ^a (4.87)	-0.01 (-0.11)
11	1.65 ^a (7.33)	1.37 ^a (4.92)	0.60 (1.77)	0.47 (1.55)	0.43 ^a (2.05)	1.34 ^a (5.35)	0.53 ^a (7.20)
12	0.81 ^a (2.35)	0.81 ^a (2.50)	0.38 ^a (2.26)	0.17 (0.77)	0.39 (1.79)	1.00 ^a (5.09)	0.40 ^a (2.38)
13	0.96 ^a (2.02)	0.68 ^a (2.44)	0.09 (0.50)	0.10 (0.53)	0.07 (0.35)	0.51 ^a (2.68)	1.08 ^a (3.26)
14	-0.16 (-0.36)	0.92 (1.89)	0.58 ^a (3.70)	0.00 (0.01)	-0.01 (-0.04)	0.11 (0.45)	0.19 (0.79)
15	-0.75 ^a (-3.00)	0.44 ^a (2.14)	0.96 ^a (4.48)	0.44 ^a (3.12)	-0.18 (-0.92)	0.74 ^a (3.32)	-0.32 ^a (-2.01)
16	-1.18 ^a (-2.11)	0.70 ^a (2.51)	0.68 ^a (2.58)	0.17 (0.56)	0.10 (0.40)	0.54 (1.74)	0.50 ^a (2.90)
17	0.38 (0.97)	1.71 ^a (2.24)	0.70 (1.88)	1.07 (1.89)	-0.06 (-0.13)	-0.34 (-1.02)	0.11 (0.42)
18	0.75 (1.40)	-0.99 (-1.84)	-0.14 (-0.44)	0.20 (0.70)	0.00 (0.01)	0.30 (0.86)	0.69 (1.93)
19	-0.79 (-1.50)	-2.03 ^a (-6.84)	-0.51 (-1.32)	0.89 ^a (2.72)	-1.31 ^a (-3.74)	-0.10 (-0.18)	0.58 (1.14)
20	1.41 ^a (2.96)	-1.54 ^a (-2.31)	-0.97 ^a (-2.67)	-0.40 (-0.65)	-1.18 ^a (-2.38)	0.25 (0.54)	0.00 (0.01)
21	0.05 (0.10)	1.04 ^a (2.75)	0.22 (1.10)	0.90 (1.76)	-0.46 (-1.00)	0.16 (0.36)	-0.13 (-0.42)
22	-0.18 (-0.25)	-0.65 (-1.75)	-0.26 (-0.54)	-1.92 ^a (-4.06)	1.00 ^a (3.31)	0.20 (0.25)	-0.57 (-1.45)
23	-0.70 (-1.04)	-0.03 (-0.07)	-0.25 (-0.47)	-0.88 (-1.07)	0.81 ^a (2.89)	-0.71 (-1.42)	-0.75 ^a (-2.40)

Note: *t*-statistics are shown in parenthesis.

^a Denotes that the estimated values are statistically significant at the 5% level.

to 30% on their updated priors, as the horizon decreases from 24 months to 1 month. Even at the 1-month forecast horizon, there is still certain degree of disagreement among experts, as is shown in Fig. 1. Apart from the uncertainty associated with the remaining month, the future data revisions and measurement errors have to be a significant factor for the residual disagreement at the end of the forecasting period.²³ This finding establishes the role of prior beliefs in generating expectation stickiness. To see this formally, iterating (23) backward, we get

$$E(Y_t | F_{it+h+1}, L_{th}) = \prod_{j=h}^{23} \lambda_{itj} F_{24} + (1 - \lambda_{itth})(L_{th} - \mu_{itth}) + \sum_{j=h+1}^{23} \left(\prod_{s=h}^{j-1} \lambda_{its} \right) (1 - \lambda_{itj})(L_{tj} - \mu_{itj}). \quad (28)$$

²³ Note that certain degree of disagreement can be rationalized if forecasters' loss functions are asymmetric and heterogeneous; see Capistrán and Timmermann (2006). Using density forecasts, Lahiri and Liu (2007) find evidence of mild departure from symmetric loss functions that become less as forecast horizon gets short.

In (28) the optimal forecast made at horizon *h* is a weighted average of three components: the initial prior beliefs, current public information and all past public information. The initial prior belief causes forecast stickiness in two ways. First, it enters into the current forecast directly and is propagated forward onto the whole series of forecasts for the target year, though its importance declines over horizons. Second, it allows for all past public information to affect current forecast in a staggered way. Without the role of prior beliefs (i.e. $\lambda_{itth} = 0$ for all *h*), current forecast reflects only the latest information about the target variable. As aptly noted by Zellner (2002), this “anchoring” like effect, much emphasized in the psychological literature, is a result of optimal Bayesian information processing in the presence of an initial prior, see Tversky and Kahneman (1974). Thus, stickiness of expectations in itself does not necessarily contradict the forecast efficiency hypothesis. Instead the Bayesian learning model allows for 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 and Reis (2006) and Morris and Shin (2006)).

Our analysis, however, shows that there is more stickiness in the recorded forecasts than the optimal Bayesian learning model would allow. What are the potential sources of this

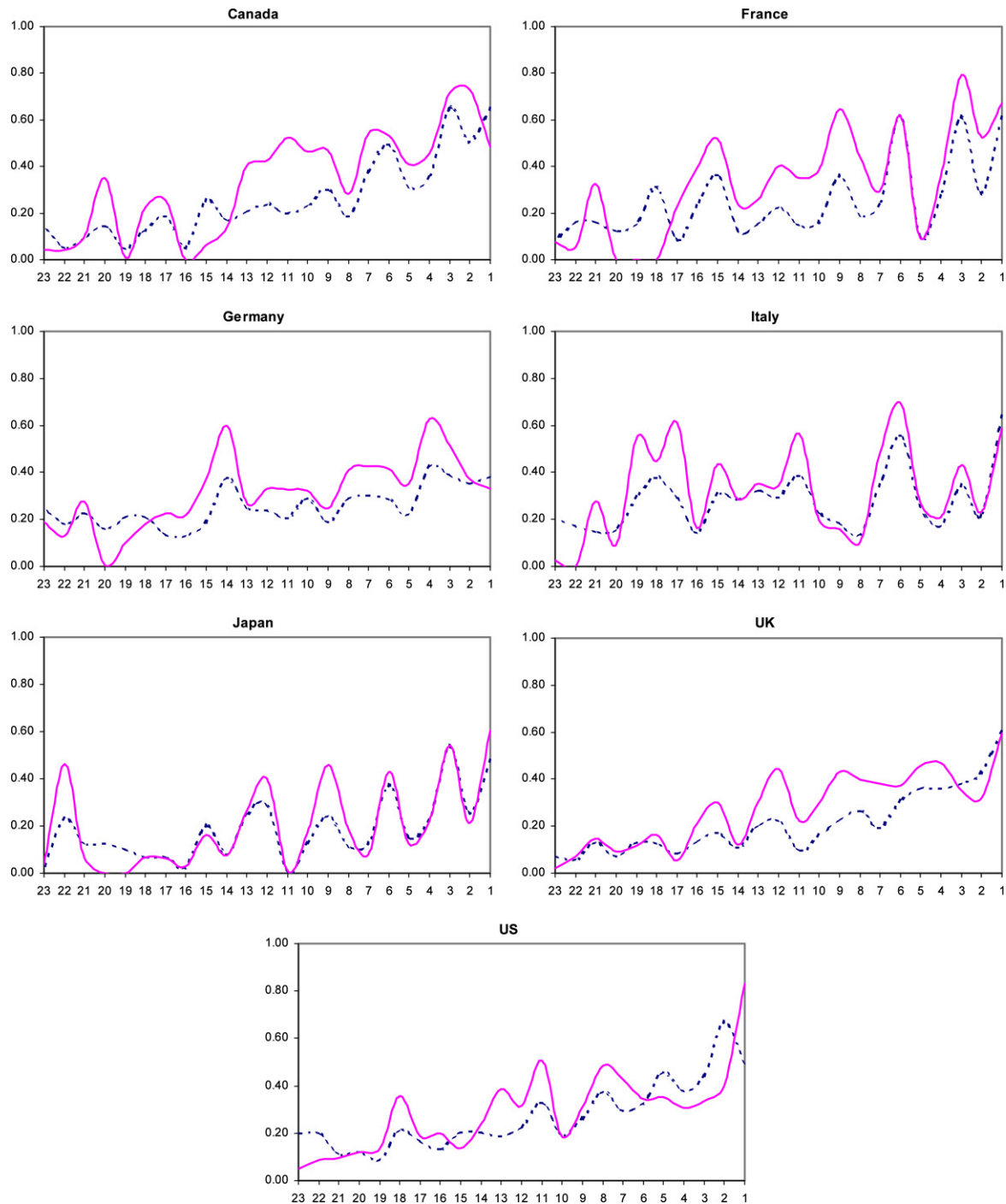


Fig. 3. Efficient weight (solid line) and actual weight (dotted line) attached to public information.

forecast inefficiency? Since the forecasters in the survey are not anonymous, the possibility exists that at least part of the inefficiency can be explained by strategic behavior among them. Ehrbeck and Waldmann (1996) and Laster et al. (1999) have explored this possibility.

Ehrbeck and Waldmann (1996) proposed two models of rational bias when the less able forecasters mimic the pattern of predictions associated with abler forecasters in order to look good before the outcome is observed. They derived two testable implications of their theory – one involving a relationship between forecast errors and forecast revisions, as in our Eq. (27), and the other specifying a relationship between mean squared prediction errors and mean squared forecast revisions across forecasters.

Their model of “rational stubbornness”, where less able forecasters are conservative in adjusting their forecasts, makes both the relationships positive. Their second model of “rational jumpiness”, where less able forecasters overshoot their forecasts, makes the two relationships negative. Using limited forecast data on T-bill rate they found no evidence of rational bias. Based on our comprehensive forecast data on GDP growth, we also found no support for their hypothesis for any of the horizons.

To reconcile forecaster behavior with strategic rationality, Laster et al. (1999) built an economic model in which forecasters’ rewards are based on two criteria: accuracy and ability to generate publicity for their firms. Their model implies that the more an industry seeks publicity, the more it will tend

to produce unconventional forecasts. Following their criterion, we divided our forecasters into six industry categories.²⁴ Our regression procedure, which is exactly the same as theirs, shows that the mean absolute deviations from the consensus forecasts are relatively large for securities firms, small for industrial corporations and banks, and somewhere in between for independent forecasters, econometric modelers and forecasters in the residual category. These results are not consistent with the Laster et al. (1999) hypothesis that independent forecasters tend to deviate most from the consensus. Taken together, our empirical results suggest that forecast inefficiency cannot be explained by the types of strategic rationality suggested in these two papers.²⁵ However, it is certainly possible that some other unexplored behavioral models can explain the observed inefficiency.

6. Concluding remarks

We have established four broad empirical patterns in real GDP growth forecasts for G7 countries in a multi-horizon context. First, expert disagreement in 24-month ahead forecasts is relatively high for all sampled countries, and remains almost the same for another eight monthly rounds of forecasting. The disagreement then steadily falls as the horizon gets shorter from 15 months to 1 month. Second, the within-agent variation in forecasts, however, starts with a low value at 24-month ahead and remains almost constant from horizons 24 to 16 months. It then increases sharply as horizon decreases from 15 months to 1 month. Third, it is within- and not between-agent variation in the data that dominates experts' forecast revision process. Fourth, the within-agent variation in the data reveals that experts observe and incorporate the first major public signal around horizon 15 months to the end of target year for all G7 countries.

Any model aimed at explaining the formation of expectation by experts over horizons must then capture the major elements of above findings. The Bayesian model in our paper explicitly recognizes that professional forecasters begin forecasting with specific prior beliefs about the GDP growth for the next year and that they learn to modify their initial beliefs over horizons with the arrival of public information. We postulate forecast disagreement to have three components due to differences in: (i) the initial prior beliefs, (ii) the weights attached on updated priors, and (iii) the interpretation of the public signal. The fixed-target, multi-horizon and cross-country feature of the panel data allows us to estimate and gauge the relative importance of each component precisely.

The first component (i.e., the diversity in initial prior beliefs) explains nearly 100% to 30% of forecast disagreement as the horizon decreases from 24 months to 1 month. This finding establishes the role of initial prior beliefs in generating expectation stickiness in two ways. On the one hand, it enters into current forecast directly, though its importance declines over horizons. On the other, it allows for all past public information to affect current forecast in a staggered way. This "anchoring" like effect, much emphasized in the psychological literature, is the result of Bayesian optimal information processing rule.

We find the second component (i.e., the variation in weights attached by experts to priors) to barely have any effect on forecast disagreement, since professional forecasters place very similar weights on their updated prior beliefs. This channel, however, might account for a large part of the disagreement in laymen's expectations.

Depending on the timing and quality of new public information, the significance of the third component increases from almost nothing to 70% as the horizon gets shorter. Since the importance of private information ought to be small in real GDP forecasts relative to that in earnings or even inflation forecasts, our empirical result provides direct evidence in support of Kandel and Pearson (1995) and Kandel and Zilberfarb (1999) on the role of differential interpretation of public signals in generating expert disagreement. This finding has important implications on belief formation and learning in many areas of economics. It says that there can simultaneously exist two or more competing models used to interpret public information, which can be a great challenge for establishing the credibility of macroeconomic policies. Also, heterogeneity in analysts' expectations and thus trading in the asset markets can arise from model uncertainty.

Without relying on certain *ad hoc* proxies for public information as in the conventional literature, we propose a new test of forecast efficiency based on the observable properties of fixed-target forecasts. Our test requires that current forecast errors be uncorrelated with current forecast revisions and is expected to have more power to identify forecast inefficiency for each horizon. We find forecasters to overweight public information at very long horizons, but significantly underweight in the middle horizons. On average, forecasters underweight new public information by 5% to 35%. Especially for forecasters in France and Germany, the degree of underweighting public information is the highest among G7 countries. After adjusting for the degree of inefficiency, we recover the efficient weights attached on public information. Efficient weights appear jagged because they incorporate new public information quickly. Actual weights appear smoother and staggered, for they let new public information seep in slowly.

Our findings have several important implications. First, they sharpen our understanding of the role and importance of forecast horizons in forecasting, and in multi-period decision-making problems. Many studies have focused on the evolution of forecast disagreement over time, without specifying the underlying forecast horizons. The fixed-target scheme, one of the unique features of our data, enables us to study carefully the effect of information arrival and its uncertainty on forecast disagreement. Second, our findings help distinguish between alternative explanations for the heterogeneous expectations held by professional forecasters. The sticky-information model may not be suitable as a description of the behavior of professional forecasters. The Bayesian learning model, which permits differential interpretation of public information and the role of priors, seems to explain a major part of the variation in the expectations reported by professional forecasters, and is consistent with the salient features of the multi-country forecast data set. Third, the efficiency test we have proposed in this paper can be utilized to study the degree of inefficiency in the use of public information in many other areas of economics and finance.

References

- Acemoglu, D., Chernozhukov, V., Wernz, M., 2006. Learning and disagreement in an uncertain world. NBER working paper No. 12648.
- Arrow, K.J., Hurwitz, L., 1972. An optimality criterion for decision making under ignorance. In: C.F. Carter and J.L. Ford (Eds.): *Uncertainty and Expectations in Economics: Essays in Honor of G.L.S. Shackle*. A.M. Kelly, New Jersey.
- Aumann, R., 1976. Agreeing to disagree. *The Annals of Statistics* 4, 1236–1239.
- Aumann, R., 1987. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55, 1–18.
- Baltagi, B.H., Song, S.H., Jung, B.C., 2001. The unbalanced nested error component regression model. *Journal of Econometrics* 101, 357–381.
- Batchelor, R., 2007. Bias in macroeconomic forecasts. *International Journal of Forecasting* 23, 189–203.
- Batchelor, R., Dua, P., 1991. Blue Chip rationality tests. *Journal of Money, Credit and Banking* 23, 692–705.
- Batchelor, R., Dua, P., 1992. Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting* 11, 169–181.

²⁴ We thank Michael Jerison, Massimiliano Marcellino and editors of *Consensus Forecasts* for helping us to identify the affiliations of some of the European forecasters.

²⁵ The regression results are available on request.

- Beck, N., Katz, J.N., 1995. What to do (and not to do) with time-series cross-section data. *The American Political Science Review* 89, 634–647.
- 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. ECB Occasional Study No. 59.
- Branch, W.A., 2004. The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal* 114, 592–621.
- Bray, M.M., Savin, N.E., 1986. Rational expectation equilibrium, learning, and model specification. *Econometrica* 54, 1129–1160.
- Brock, W.A., Hommes, C.H., 1997. A rational route to randomness. *Econometrica* 65, 1059–1160.
- Capistrán, C., Timmermann, A., 2006. Disagreement and biases in inflation expectations. Working paper, Department of Economics, UCSD.
- Carroll, C., 2003. Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118, 269–298.
- Chen, Q., Jiang, W., 2006. Analysts' weighting of private and public information. *Review of Financial Studies* 19, 319–355.
- Clements, M.P., 1995. Rationality and the role of judgment in macroeconomic forecasting. *Economic Journal* 105, 410–420.
- Clements, M.P., 1997. Evaluating the rationality of fixed-event forecasts. *Journal of Forecasting* 16, 225–239.
- Clements, M.P., Taylor, N., 2001. Robust evaluation of fixed-event forecast rationality. *Journal of Forecasting* 20, 285–295.
- Cyert, R.M., DeGroot, M.H., 1974. Rational expectations and Bayesian analysis. *Journal of Political Economy* 82, 521–536.
- Davies, A., Lahiri, K., 1995. A new framework for analyzing survey forecasts using three-dimensional panel data. *Journal of Econometrics* 68, 205–227.
- 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.
- Ehrbeck, T., Waldmann, R., 1996. Why are professional forecasters biased? Agency versus behavioral explanations. *Quarterly Journal of Economics* 61, 21–40.
- El-Gamal, M., Grether, D.M., 1995. Are people Bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association* 90, 1137–1145.
- Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25, 383–417.
- Gallo, G.M., Granger, C.W.J., Jeon, Y., 2002. Copycats and common swings: The impact of the use of forecasts in information sets. IMF Staff Papers 49.
- Granger, C.W.J., 1996. Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics* 11, 455–473.
- Granger, C.W.J., 2007. Forecasting – looking back and forward. *Journal of Econometrics* 138, 3–13.
- Gregory, A.W., Smith, G.W., Yetman, J., 2001. Testing for forecast consensus. *Journal of Business and Economic Statistics* 19, 34–43.
- Gregory, A.W., Yetman, J., 2004. The evolution of consensus in macroeconomic forecasting. *International Journal of Forecasting* 20, 461–473.
- Haltiwanger, J., Waldman, M., 1985. Rational expectations and the limits of rationality: An analysis of heterogeneity. *American Economic Review* 75, 326–340.
- Harris, M., Raviv, A., 1993. Differences of opinion make a horse race. *Review of Financial Studies* 6, 473–506.
- Harvey, D., Leybourne, S.J., Newbold, P., 2001. Analysis of a panel of UK macroeconomic forecasts. *Econometrics Journal* 4, 37–55.
- Hsiao, C., Pesaran, M.H., 2008. Random coefficient panel data models. In: Matyas, L., Sevestre, P. (Eds.), *The Econometrics of Panel Data*, 3rd ed. Springer, Boston, pp. 185–213.
- 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.
- Ivanova, D., Lahiri, K., 1998. A time series and cross sectional analysis of consumer sentiment and its components. In: Oppenländer, K.H., Poser, G. (Eds.), *Social Structural Change - Consequences for Business Cycle Surveys*. Aldershot: Ashgate Publishing, UK, pp. 337–362.
- Jonung, L., 1981. Perceived and expected rates of inflation in Sweden. *American Economic Review* 71, 961–968.
- 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.
- Kurz, M., 1994. On the structure and diversity of rational beliefs. *Economic Theory* 4, 877–900.
- Lahiri, K., Teigland, C., Zaporowsky, M., 1988. Interest rates and the subjective probability distribution of price forecasts. *Journal of Money, Credit and Banking* 20, 233–247.
- 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., 2007. On the estimation of forecasters' loss functions using density forecasts. Paper presented at the 27th International Symposium on Forecasting, NY.
- Laster, D., Bennett, P., Geoum, I.S., 1999. Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics* 114, 293–318.
- Lee, J., Nelson, C.R., 2007. Expectation horizon and the Phillips curve: The solution to an empirical puzzle. *Journal of Applied Econometrics* 22, 161–178.
- Lucas, R.E., 1973. Some international evidence on inflation-output tradeoffs. *American Economic Review* 63, 326–334.
- Maddala, G.S., 1991. Survey data on expectations: What have we learnt? In: Nerlove, M. (Ed.), *Issues in Contemporary Economics*, vol. 2. Macmillan I.E.A., pp. 319–344.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295–1328.
- 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. NBER Macroeconomics Annual 2003.
- Morris, S., Shin, H.S., 2006. Inertia of forward-looking expectations. *American Economic Review* 96, 152–157.
- Muth, J.F., 1961. Rational expectations and the theory of price movements. *Econometrica* 29, 315–335.
- Nordhaus, W., 1987. Forecasting efficiency: Concepts and applications. *Review of Economics and Statistics* 69, 667–674.
- Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74, 967–1012.
- Pesaran, M.H., 2007. A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics* 22, 265–312.
- 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.
- Savage, L.J., 1972. *The Foundations of Statistics*. Dover reprint, New York, 1954.
- Sims, C.A., 2003. Implications of rational inattention. *Journal of Monetary Economics* 50, 665–690.
- Souleles, N., 2004. Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the Michigan consumer sentiment surveys. *Journal of Money, Credit and Banking* 36, 39–72.
- Stock, J.H., Watson, M.W., 2005. Understanding changes in international business cycle dynamics. *Journal of the European Economic Association* 3, 966–1006.
- Townsend, R.M., 1983. Forecasting the forecasts of others. *Journal of Political Economy* 91, 546–588.
- Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185, 1124–1131.
- Varian, H., 1989. Differences of opinion in financial markets. In: Stone, C. (Ed.), *Financial Risk: Theory, Evidence and Implications*. Kluwer, Boston, pp. 3–37.
- Veronesi, P., 2000. How does information quality affect stock returns? *Journal of Finance* 55, 807–837.
- Williams, J.C., 2003. Discussion of disagreement about inflation expectations. NBER Macroeconomics Annual 2003.
- Yang, K., Shintani, M., 2006. Does the prediction horizon matter for the forward premium anomaly? Evidence from panel data. *Economics Letters* 93, 255–260.
- Zarnowitz, V., Lambros, L.A., 1987. Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95, 591–621.
- Zellner, A., 1988. Optimal information processing and Bayes' theorem. *The American Statistician* 42, 278–284 with invited discussion.
- Zellner, A., 2002. Information processing and Bayesian analysis. *Journal of Econometrics* 107, 41–50.