

# How Quickly is Macroeconomic Uncertainty Resolved?

## Theory and Empirical Evidence from the Term Structure of Forecast Errors\*

Andrew J. Patton                      Allan Timmermann  
London School of Economics      University of California San Diego

26 June 2007

PRELIMINARY. COMMENTS WELCOME.

### Abstract

How quickly is uncertainty about macroeconomic variables such as output growth and inflation resolved over time? To answer this question, we develop a theoretical framework for understanding the resolution of uncertainty about economic variables with a partially predictable component. Our model incorporates the effect of measurement errors and heterogeneity in individual forecasters' prior beliefs and their information signals. It also accounts for updates to agents' beliefs about past, current and future variables. We use the model to develop insights into the complete term structure of forecasts and forecast errors as a function of the length of the forecast horizon, and test its implications on a data set comprising forecasts of annual GDP growth and inflation with horizons ranging from 1 to 24 months. Consistent with the model, the steepness of the relation between the forecast horizon and the forecast precision changes around the 12-month horizon and both the forecast error variance and the dispersion among forecasters declines along a concave path as the distance to the event draws closer. The rate at which this decline occurs is shown to contain information not only about the rate at which uncertainty is resolved but also about the size of a persistent component in GDP growth and inflation.

---

\*We thank Roy Batchelor, Mike McCracken, Hashem Pesaran, Michela Verardo, Mark Watson and seminar participants at the Board of Governors of the Federal Reserve, Cambridge, City University London, Duke, London School of Economics, Universite Libre Bruxelles (ECARES), Tilburg and Stanford for helpful comments and suggestions.

# 1 Introduction

Uncertainty about macroeconomic variables such as output growth and inflation influences the actions taken by policy makers, businesses and individuals. News on such variables is therefore keenly monitored through time and is quickly reflected in agents' beliefs about future events. The rate at which uncertainty about the economy is resolved is economically important because of the irreversibility and lags in many economic decisions ('time to build') as emphasized in dynamic macroeconomic analysis (Kydland and Prescott (1982) and Dixit and Pindyck (1994)). The level and resolution of uncertainty may also have welfare implications: Ramey and Ramey (1995) link output growth to the degree of uncertainty surrounding it, arguing that firms scale back planned output during periods with high levels of uncertainty. Finally, macroeconomic uncertainty has also been shown to be an important determinant of asset prices and volatility in financial markets.<sup>1</sup>

While uncertainty about the state of the economy and its resolution through time is of great importance to economic agents, surprisingly little is known about it. In part this reflects the difficulty in modeling how information about macroeconomic variables evolves through time and across agents. A formal modeling approach requires simultaneously making assumptions about the structure of the economy and about the forecasting models and predictor variables used by agents. As such, this approach is complicated by the existence of literally hundreds of economic state variables that could be adopted in such models, (Stock and Watson (2005)), and the lack of information about which models agents actually use.

This paper proposes a very different approach for extracting information about how agents learn about the state of the economy. Using a new panel data set containing survey forecasts of GDP growth and inflation across many different horizons ( $T$  different time periods and  $H$  different horizons, say), we are able to estimate the rate at which economic forecasters learn about these macroeconomic variables. We do so by modeling agents' learning problem and then matching the full term structure of forecast errors at different forecast horizons with the moments implied by this model.

A key contribution of this paper is to develop a framework for studying panels of forecasts

---

<sup>1</sup>Ederington and Lee (1996) and Andersen et al (2003) find that macroeconomic announcements have a significant effect on T-bill futures and exchange rates, respectively, whose mean often jumps following such news, while Beber and Brandt (2006) find that implied volatilities and trading volumes in options markets for stocks and bonds are closely related to macroeconomic uncertainty.

containing numerous different forecast horizons (“large  $H$ ”), and present empirical results that shed new light on the speed at which the uncertainty surrounding realizations of macroeconomic variables is resolved. The “large  $H$ ” nature of our panel enables us to answer a number of interesting questions that are intractable with forecasts of just one or two different horizons. For example, we are able to separately identify the relative importance of the predictable and unpredictable components of a given variable, as well as the degree of persistence in the persistent component. With some further structure, we are also able to estimate the extent of measurement error in the forecasters’ estimates of the current and lagged value of the variable of interest. This use of survey forecast data has received relatively little attention in the literature.<sup>2</sup>

Expectations and the degree of uncertainty about macroeconomic variables can vary considerably through time, even when the date of the variable in question remains fixed. Consequently, much can be learned by studying how agents update their beliefs about the same “event”. Figure 1 provides an illustration of this by showing Consensus forecasts of US GDP growth for 2002 as this evolved each month from January 2001 (corresponding to a 24-month horizon) to December 2002 (a one-month horizon).<sup>3</sup> The participants in these surveys are professional forecasters such as investment banks, think tanks or quasi-public research institutions. A comparison of the initial and final forecasts—at 3.5% and 2.5%, respectively—shows a fairly sizeable reduction in the projected growth, but fails to incorporate the full picture of the dramatic revisions that occurred in the interim. At the beginning of September 2001, the growth forecast for 2002 was 2.7%. Following the events of 9/11, the October 2001 forecast fell to 1.2%, i.e. by a full 1.5%—the largest single-month forecast revision observed in more than a decade. It declined even further to 0.8% in November 2001 before stabilizing. Expectations of 2002 GDP growth then increased by 1.7% from January through April of 2002, from which point the subsequent forecasts were within 0.5% of the actual growth figure, which came in just below 2.5%.

The cross-sectional dispersion saw similarly dramatic changes: Prior to 9/11 the dispersion

---

<sup>2</sup>Exceptions include Davies and Lahiri (1995), Clements (1997), Swidler and Ketcher (1990) and Chernov and Müller (2007). The first three of these studies are, however, concerned only with rationality testing and do not address the question of extracting information about agents’ updating process, nor do they address the dispersion among forecasters. Chernov and Müller, on the other hand, construct a parametric model of the term structure of U.S. inflation forecasts from four different surveys, in order to combine them with information from Treasury yields and macroeconomic variables.

<sup>3</sup>The data is described in further detail in Section 3.1.

across forecasters was close to 0.7%. Dispersion then more than doubled to around 1.6% in October through December of 2001, before falling back to its normal level once again.

While this is only one episode in our data, it illustrates several of the features of professional forecasters' updating processes that survey data can shed light on: rapid adjustment to news (reflected both in the consensus and in the cross-sectional dispersion), decreasing forecast errors and declining cross-sectional dispersion as the forecast horizon is reduced.

To match these features of our data, we extend our baseline model to incorporate heterogeneity in agents' prior beliefs and their information signals and derive the implications for how the cross-sectional dispersion evolves as a function of the forecast horizon<sup>4</sup>. Our empirical analysis strongly suggests the need for modeling heterogeneity across forecasters both in terms of the signals they receive and in terms of their prior beliefs about the long-run values of the variables of interest.

We find many interesting empirical results. First, perhaps unsurprisingly we find that uncertainty about output growth and inflation declines as the forecast horizon gets shorter. Second, consistent with a simple model containing a persistent component in the predicted variable, uncertainty falls at a slower rate in the "next-year" forecasts ( $h \geq 12$ ) than in the "current-year" forecasts ( $h < 12$ ), following a concave pattern. Third, we find that the cross-sectional dispersion is systematically linked to the forecast horizon and shrinks as the horizons is reduced.

The plan of the paper is as follows. Section 2 presents a simple framework for understanding how uncertainty is resolved through time and how the consensus forecast is updated as the forecast horizon is reduced. Section 3 follows up with empirical results on the consensus forecasts. Section 4 extends the model to cover cross-sectional dispersion among forecasters by allowing for heterogeneity in agents' information and their prior beliefs. Section 5 provides empirical results for this extended model. Section 6 concludes. Technical derivations are provided in an appendix.

## 2 A Model for the Term Structure of Forecast Errors

This section develops a simple benchmark model for how agents update their beliefs about macroeconomic variables such as output growth and inflation rates. Our analysis exploits the rich information available by studying how forecasts of a variable measured at a low frequency (e.g., annual

---

<sup>4</sup>Lahiri and Sheng (2006) undertake a similar study of forecast dispersions, though their econometric approach is quite different to the one we employ.

GDP growth) are updated at a higher frequency (monthly, in our case). Moreover, since we shall be concerned with flow variables that agents gradually learn about as new information arrives prior to and during the period of their measurement, the fact that part of the outcome may be known prior to the end of the measurement period (the “event date”) introduces complications. It also means that the timing of the forecasts has to be carefully considered.

Our analysis assumes that forecasters choose their forecasts to minimize the expected value of a loss function  $L$  that depends on the forecast error,  $e_{t,t-h} = z_t - \hat{z}_{t,t-h}$ , where  $z_t$  is the predicted variable,  $\hat{z}_{t,t-h}$  is the forecast computed at time  $t - h$ ,  $t$  is the event date and  $h$  is the forecast horizon. Assuming squared loss,  $L(e) = e^2$ , the optimal  $h$ -period forecast is simply the conditional expectation of  $z_t$  given information at time  $t - h$ ,  $\mathcal{F}_{t-h}$ :<sup>5</sup>

$$\hat{z}_{t,t-h}^* = E[z_t | \mathcal{F}_{t-h}]. \quad (1)$$

Survey data on expectations has been the subject of many studies—see Pesaran and Weale (2006) for a recent review. The focus of this literature has, however, mainly been on testing the rationality of survey expectations as opposed to understanding how the precision of the forecasts evolves over time. This is related to the fact that survey data usually takes the form of rolling event forecasts of objects measured at different points in time (using a fixed forecast horizon but a varying date) such as a sequence of year-ahead forecasts of growth in GDP. While it may be of economic interest to ask if the standard deviation of the forecast error is the same across different subsamples, forecast efficiency implies no particular ranking of the error variances across different subsamples since the variance of the predicted variable need not be constant. For example, the forecast error associated with US GDP growth may have declined over time (Kim and Nelson (1999) and McConnell and Perez-Quiros (2000)), but this need not imply that forecasters are getting better if, as is widely believed, the volatility of US output growth has also come down.

To study agents’ learning process we keep the event date,  $t$ , fixed and vary the forecast horizon,  $h$ . As illustrated in Figure 1, this allows us to track how agents update their beliefs through time. As pointed out by Nordhaus (1987) and Clements (1997), such fixed-event forecasts are a largely unexplored resource compared with rolling-event forecasts which vary the date of the forecast while holding the horizon constant.

---

<sup>5</sup>The assumption that forecasters make efficient use of the most recent information is most appropriate for professional forecasters such as those we shall consider in our empirical analysis but is less likely to hold for households which may only update their views less frequently, see Carroll (2003).

## 2.1 Benchmark Model

We first propose a simple model that ignores heterogeneity among agents along with measurement errors in the predicted variable. This model is sufficiently simple and tractable that it allows us to establish intuition for the factors determining the full term structure of forecast errors. Derivations get complicated very quickly as additional features are added.

Since the predicted variable in our application is measured less frequently than the forecasts are revised, it is convenient to describe the target variable as a rolling sum of a higher-frequency variable. To this end, let  $y_t$  denote the single-period variable (e.g., log-first differences of GDP or a price index tracking inflation), while the rolling sum of the 12 most recent single-period observations of  $y$  is denoted  $z_t$  :

$$z_t = \sum_{j=0}^{11} y_{t-j}. \quad (2)$$

Our benchmark model is based on a decomposition of  $y_t$  into a persistent (and thus predictable) first-order autoregressive component,  $x_t$ , and a temporary component,  $u_t$ :

$$y_t = x_t + u_t \quad (3)$$

$$x_t = \phi x_{t-1} + \varepsilon_t, \quad -1 < \phi < 1$$

$$u_t \sim iid(0, \sigma_u^2),$$

$$\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$$

$$E[u_t \varepsilon_s] = 0 \quad \forall t, s.$$

$\phi$  measures the persistence of  $x_t$ , while  $u_t$  and  $\varepsilon_t$  are innovations that are both serially uncorrelated and mutually uncorrelated. Without loss of generality, we assume that the unconditional mean of  $x_t$ , and thus  $y_t$  and  $z_t$ , is zero. Assuming that both  $x_t$  and  $y_t$  are observed at time  $t$ , the forecaster's information set at time  $t$  is  $\mathcal{F}_t = \sigma([x_{t-j}, y_{t-j}], j = 0, 1, 2, \dots)$ .

The assumption that the predicted variable contains a first-order autoregressive component, while clearly an approximation, is likely to capture well the presence of a persistent component in most macroeconomic data. For example, much of the dynamics in the common factors extracted from large cross-sections of macroeconomic variables by Stock and Watson (2002) is captured by low-order autoregressive terms.

This simple model allows a complete characterization of how the mean squared forecast error (MSE) evolves as a function of the forecast horizon ( $h$ ):

**Proposition 1** Suppose that  $y_t$  can be decomposed into a persistent component ( $x_t$ ) and a temporary component ( $u_t$ ) satisfying (3) and forecasters minimize the squared loss.

(1) The optimal forecast of  $z_t = \sum_{j=0}^{11} y_{t-j}$  given information at time  $t-h$ ,  $\hat{z}_{t,t-h}^*$ , is given by

$$\hat{z}_{t,t-h}^* = \begin{cases} \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi} x_{t-h}, & \text{for } h \geq 12 \\ \frac{\phi(1-\phi^h)}{1-\phi} x_{t-h} + \sum_{j=h}^{11} y_{t-j}, & \text{for } h < 12 \end{cases}.$$

(2) The mean squared forecast error as a function of the forecast horizon is given by

$$E[e_{t,t-h}^2] = \begin{cases} 12\sigma_u^2 + \frac{1}{(1-\phi)^2} \left( 12 - 2\frac{\phi(1-\phi^{12})}{1-\phi} + \frac{\phi^2(1-\phi^{24})}{1-\phi^2} \right) \sigma_\varepsilon^2 + \frac{\phi^2(1-\phi^{12})^2(1-\phi^{2h-24})}{(1-\phi)^3(1+\phi)} \sigma_\varepsilon^2 & \text{for } h \geq 12 \\ h\sigma_u^2 + \frac{1}{(1-\phi)^2} \left( h - 2\frac{\phi(1-\phi^h)}{1-\phi} + \frac{\phi^2(1-\phi^{2h})}{1-\phi^2} \right) \sigma_\varepsilon^2 & \text{for } h < 12 \end{cases}$$

(3) As the forecast horizon is reduced from  $h$  to  $h-1$  periods, the decline in the mean squared forecast error is given by

$$\Delta MSE_h \equiv E[e_{t,t-h}^2] - E[e_{t,t-h+1}^2] = \begin{cases} \sigma_\varepsilon^2 \frac{(1-\phi^{12})^2}{(1-\phi)^2} \phi^{2h-24}, & \text{for } h \geq 12 \\ \sigma_u^2 + \sigma_\varepsilon^2 \frac{(1-\phi^h)^2}{(1-\phi)^2}, & \text{for } h < 12 \end{cases}.$$

The proof of Proposition 1 is in the Appendix. Proposition 1 is simple to interpret: At each point in time an optimal forecast makes efficient use of the most recent information. Forecasts computed prior to the measurement period (i.e., those with  $h \geq 12$ ) make use of the most recent value of  $x$  since this is the only predictable component of  $y$ . During the measurement period (when  $h < 12$ ), those values of  $y$  that are already observed are used directly in the forecast, which is the second term in the expression for  $\hat{z}_{t,t-h}^*$  for  $h < 12$ .

Turning to part 2 of Proposition 1, the first term in the expression for the MSE captures the unpredictable component,  $u_t$ . The second term captures uncertainty about shocks to the remaining values of the persistent component,  $x_t$ , over the measurement period. The additional term in the expression for  $h \geq 12$  comes from having to predict  $x_{t-11}$ , the initial value of the persistent component at the beginning of the measurement period.

As  $h \rightarrow \infty$ , the optimal forecast converges to the unconditional mean of  $z_t$  (normalized to zero in our model). This forecast generates the upper bound for the MSE of an optimal forecast, which is the unconditional variance of  $z_t$ :

$$\lim_{h \rightarrow \infty} E[e_{t,t-h}^2] = 12\sigma_u^2 + \frac{\sigma_\varepsilon^2}{(1-\phi)^2} \left( 12 - 2\frac{\phi(1-\phi^{12})}{1-\phi} + \frac{\phi^2(1-\phi^{24}) + \phi^2(1-\phi^{12})^2}{1-\phi^2} \right). \quad (4)$$

The third part of Proposition 1 shows the effect of reducing the forecast horizon by a single period, from  $h$  to  $h-1$ , so the forecasters' information set expands from  $\mathcal{F}_{t-h}$  to  $\mathcal{F}_{t-h+1}$ . Uncertainty, as measured by the MSE-value, should of course (weakly) decline and part 3 provides the magnitude of this decline. The speed of uncertainty resolution rises as the forecast horizon is reduced. The reason is twofold. First, prior to the measurement period ( $h \geq 12$ ) the only new information relevant to forecasting  $z$  is that which helps predict the value of  $x$  at the start of the measurement period. The further back in time the forecast is produced, the less valuable the information is. In the limit as  $h \rightarrow \infty$ ,  $\Delta MSE \rightarrow 0$ . Second, during the measurement period the forecaster observes part of the actual variable and so uncertainty about the current value of  $y$  gets completely removed by adding one more observation<sup>6</sup>.

To illustrate Proposition 1, Figure 2 plots the root mean squared error (RMSE) for  $h = 1, 2, \dots, 24$  using parameters similar to those we obtain in the empirical analysis for U.S. GDP growth. Holding the unconditional variance of annual GDP growth and the degree of predictability of GDP growth fixed, we show the impact of varying the persistence parameter,  $\phi$ . The figure shows the large impact that this parameter has on the shape of the MSE function. The variance of the forecast error grows linearly as a function of the length of the forecast horizon if  $y$  has no persistent component ( $\phi = 0$ ). Conversely, the persistent component gives rise to a more gradual decline in the forecast error variance as the horizon is reduced. In effect uncertainty is resolved more gradually the higher the value of  $\phi$ . Notice also how the change in RMSE gets smaller at the longest horizons, irrespective of the value of  $\phi$ .

The benchmark model (3) is helpful in establishing intuition for the drivers of how macroeconomic uncertainty gets resolved through time. However, it also has some significant shortcomings. Most obviously, it assumes that forecasters observe the predicted variable without error, and so uncertainty vanishes completely as  $h \rightarrow 0$ . Our empirical work, described in detail in Section 3, indicates that this assumption is in conflict with the data. To account for this, we next extend the model to allow for measurement errors.

---

<sup>6</sup>Isiklar and Lahiri (2007) use observed values of  $\Delta MSE_h$  across 18 countries to estimate the longest horizons at which survey forecasts provide useful information.



## 2.2 Measurement Errors

Macroeconomic variables are, to varying degrees, subject to measurement errors as reflected in data revisions and changes in benchmark weights. Such errors are less important for survey-based inflation measures such as the consumer price index (CPI). Revisions are, however, very common for measures of GNP which are generally calculated once a quarter (e.g., Croushore and Stark (2001), Mahadeva and Muscatelli (2005) and Croushore (2006)). Measurement errors make the forecasters' signal extraction problem more difficult: the greater the measurement error, the noisier are past observations of  $y$  and hence the less precise the forecasters' readings of the state of the economy. They also mean that forecasters cannot simply "plug in" observed values of past  $y$ 's during the measurement period ( $h < 12$ ).

To account for these effects, we use a Kalman filter-based approach. While this captures the signal extraction problem that forecasters have to deal with, such an approach does not lend itself to easily interpretable formulas for the term structure of forecast errors, however. To gain intuition we therefore first consider a simplified model which captures the spirit of the measurement error problem, where the persistent component,  $x_t$ , is perfectly observable while the predicted variable,  $y_t$ , is observable only with noise:

$$\tilde{y}_t = y_t + \eta_t, \quad \eta_t \sim iid(0, \sigma_\eta^2). \quad (5)$$

Further assume that the measurement error is mean zero and uncorrelated with all other innovations, i.e. for all  $t, s$ ,  $E[\eta_t] = E[\varepsilon_s \eta_t] = E[u_s \eta_t] = 0$ . This model nests the "no noise" model for  $\sigma_\eta^2 = 0$ . Measurement errors clearly affect variables such as the GNP. In addition, the persistent component is likely to be surrounded by considerable noise, particularly if it is extracted not just from the underlying variable itself but, as seems more likely, also is based on additional information sources. We discuss this further below.

Proposition 2 establishes the optimal forecast along with the variance of the forecast error for this model assuming that the forecasters information set is given by  $\tilde{\mathcal{F}}_t = \sigma([x_{t-j}, \tilde{y}_{t-j}], j = 0, 1, 2, \dots)$ .

**Proposition 2** *Suppose that the predicted variable,  $y_t$ , follows the process (3) but is subject to measurement error (5) so the forecasters' information set is  $\tilde{\mathcal{F}}_t = \sigma([x_{t-j}, \tilde{y}_{t-j}], j = 0, 1, 2, \dots)$ ,*

(1) *The optimal estimate of  $y_t$  conditional on  $\tilde{\mathcal{F}}_{t+j}$  ( $j \geq 0$ ) takes the form*

$$E[y_t | \tilde{\mathcal{F}}_{t+j}] = \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} \cdot x_t + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \cdot \tilde{y}_t.$$

(2) The optimal forecast of  $z_t = \sum_{j=0}^{11} y_{t-j}$  is given by

$$\hat{z}_{t,t-h}^* = \begin{cases} \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi} x_{t-h}, & \text{for } h \geq 12 \\ \frac{\phi(1-\phi^h)}{1-\phi} x_{t-h} + \sum_{j=h}^{11} \left( \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} x_{t-j} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \tilde{y}_{t-j} \right), & \text{for } h < 12 \end{cases}.$$

(3) The mean squared forecast error is given by

$$E [e_{t,t-h}^2] = \begin{cases} 12\sigma_u^2 + \frac{1}{(1-\phi)^2} \left( 12 - 2\frac{\phi(1-\phi^{12})}{1-\phi} + \frac{\phi^2(1-\phi^{24})}{1-\phi^2} \right) \sigma_\varepsilon^2 + \frac{\phi^2(1-\phi^{12})^2(1-\phi^{2j-24})}{(1-\phi)^3(1+\phi)} \sigma_\varepsilon^2, & \text{for } h \geq 12 \\ h\sigma_u^2 + \frac{1}{(1-\phi)^2} \left( h - 2\frac{\phi(1-\phi^h)}{1-\phi} + \frac{\phi^2(1-\phi^{2h})}{1-\phi^2} \right) \sigma_\varepsilon^2 + (12-h) \frac{\sigma_u^2 \sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2}, & \text{for } h < 12 \end{cases}$$

Allowing for measurement error in the reported value of  $y_t$ , the forecaster has two imperfect estimates of the true value of  $y_t$ , namely the persistent component,  $x_t$ , and the value  $\tilde{y}_t$ , which is measured with noise. The first part of Proposition 2 shows that an optimal estimate of the true value of  $y_t$  combines information from these two sources according to their relative accuracy. If the measurement error is very small ( $\sigma_\eta^2 \rightarrow 0$ ), then the weight given to the measured variable  $\tilde{y}_t$  goes to one and the weight on  $x_t$  goes to zero. Conversely, if the measured value of  $y_t$  is very noisy, so  $\sigma_\eta^2$  is large, then the weight attached to  $\tilde{y}_t$  goes to zero and the forecaster just uses the predictable component,  $x_t$ , to proxy for  $y_t$ . Finally, if  $\sigma_u^2 \rightarrow 0$  then the predictable component dominates  $y_t$ , making  $x_t$  a good proxy for  $y_t$  and so again the weight attached to  $\tilde{y}_t$  goes to zero.

The second part of the proposition reiterates that the measurement error does not affect  $V(e_{t,t-h})$  for  $h \geq 12$ . Only the initial value of  $x$  at the start of the measurement period,  $x_{t-11}$ , matters to these forecasts and  $x$  is assumed to be known. However, such errors give rise to an additional term in the variance of the forecast error during the measurement period ( $h < 12$ ) because the realized values of  $y_t$  no longer are fully observed. Once again, if  $\sigma_\eta^2 \rightarrow 0$  or  $\sigma_u^2 \rightarrow 0$ , then this term vanishes.

Agents' updating processes allow us to characterize the precision of their information signals—or conversely quantify the size of the measurement error in the underlying variable. Figure 3 illustrates the impact of measurement error on the structure of MSE. The degree of measurement error is described as  $\sigma_\eta^2 = k^2 \sigma_u^2$  so  $k$  measures the size of the measurement error in terms of the innovation variance for  $y$ . The greater is  $k$ , the larger the measurement error. In the absence of measurement errors the MSE will converge to zero as  $h \rightarrow 0$ , whereas in the presence of measurement error the MSE will converge to some positive quantity. However as the horizon,  $h$ , shrinks towards zero, the relative importance of measurement errors grows. Moreover, the slope of the term structure gets

flatter as the size of the measurement error increases. In contrast to Figure 2, however, measurement error plays no part for long-horizon forecasts, since its impact on overall uncertainty is small relative to other sources of uncertainty. This also shows that the persistence ( $\phi$ ) and measurement error ( $\sigma_\eta^2$ ) parameters are separately identified by jointly considering short and long ends of the term structure of MSE-values.

### 3 Empirical Results: The Term Structure of Consensus Forecasts

We next present empirical results for the consensus forecasts. After describing the data source, we first present forecast efficiency tests in order to see if our assumption that forecasters employ information efficiently and have squared loss can be maintained as a working hypothesis. We next present estimation results both for the simple model that ignores measurement errors and for an extended model that accounts for such errors. Finally, we use our estimates to discuss the speed with which macroeconomic uncertainty is resolved over time.

#### 3.1 Data and Tests of Forecast Rationality

Our data is taken from the Consensus Economics Inc. forecasts which comprise polls of more than 600 private sector forecasters. Each month participants are asked about their views of a range of variables for the major economies and the consensus (average) forecast as well as the dispersion in views across survey participants are recorded. Our analysis focuses on US real GDP growth and inflation for the current and subsequent year. This gives us 24 monthly next-year and current-year forecasts over the period 1991-2004 or a total of  $24 \times 14 = 336$  monthly observations. Naturally these observations are not independent draws but are subject to a set of tight restrictions across horizons, as revealed by the term structure analysis in the previous section. The typical number of participants in surveys for the GDP and inflation rate is on the order of 20-50 forecasters. To measure the realized value of the target variable (GDP growth or inflation), we use second release data published during the following year's October issue of the IMF's World Economic Outlook.<sup>7</sup>

---

<sup>7</sup>Results are very similar when the first release (available in the April issue of the subsequent year's World Economic Outlook) is used instead.

### 3.2 Bias and Efficiency Tests

As a prelude to our analysis of the term structure of forecast errors, we initially undertake statistical tests that check for biases and serial correlation in the forecast errors. It follows from (1) that the conditional and unconditional bias in the (optimal) forecast error,  $e_{t,t-h}^* = z_t - \hat{z}_{t,t-h}^*$ , should be zero:

$$E[e_{t,t-h}^*] = E[e_{t,t-h}^* | \mathcal{F}_{t-h}] = 0. \quad (6)$$

To see if this holds, we bootstrapped  $p$ -values for the null of a zero mean forecast error. The bootstrap was implemented by sampling from the empirical distribution function, using the stationary bootstrap of Politis and Romano (1994), while imposing the null hypothesis that recenters the forecast error distribution on zero. There was no evidence of significant biases in the forecasts of GDP growth at any of the 24 horizons. For the inflation series, the bootstrap  $p$ -values revealed evidence of a small, positive bias at some horizons.

We also regressed the realized value on an intercept and the predicted value (a so-called ‘‘Mincer-Zarnowitz’’ regression) and tested whether the coefficients on these regressors are zero and one, respectively, as implied by the hypothesis of unbiased forecasts. These tests revealed no evidence of biases in the GDP growth or inflation forecasts at any of the horizons. Similarly, consistent with (6) tests did not reveal any evidence of serial correlation in the forecast errors. See Table 1 for the complete set of results.

Turning to the efficiency tests and defining the MSE-value associated with an  $h$ -period horizon,  $MSE_h \equiv E[e_{t,t-h}^2]$ , the null hypothesis implied by the simple assumption that uncertainty is resolved over time can be stated as:

$$H_0 : MSE_1 \leq MSE_2 \leq \dots \leq MSE_{24}.$$

The alternative hypothesis is that there exists at least one horizon for which these inequalities fail to hold:

$$H_a : MSE_i > MSE_j \text{ for some } i < j.$$

Testing that this holds is complicated because of the partial overlap among forecast errors which induces dependence between the MSE-values at different horizons. Furthermore, under the null, 23 dependent inequalities should hold simultaneously which again poses a non-standard hypothesis testing problem.

To deal with these complications, we employ the bootstrap “reality check” test of White (2000), which is ideally suited to testing many inequality constraints when the covariance matrix of the variables of interest (the squared forecast errors, in our case) is not available. In our application the covariance *is* available, but it is not full rank due to the fact that we have more horizons than time periods. Recalling that  $\Delta MSE_{h+1} \equiv MSE_{h+1} - MSE_h$ , the null hypothesis can be re-written as follows, making the link to White’s test more apparent:

$$H_0 : \min_{h=1,2,\dots,23} \Delta MSE_{h+1} \geq 0$$

vs.  $H_a : \min_{h=1,2,\dots,23} \Delta MSE_{h+1} < 0$

The test is conducted by generating 10,000 block vector bootstrap samples of the original squared forecast errors, and using these to construct the bootstrap distribution of the test statistic. We use the stationary bootstrap of Politis and Romano (1994) to generate the block bootstrap samples, with the average block length set equal to 24 months so as to account for the maximum overlap in the forecast errors. We also report the p-values obtained from Hansen’s (2005) refinement of the reality check<sup>8</sup>.

Empirical results from applying this test to GDP growth yielded  $p$ -values of 0.447 (White) and 0.042 (Hansen), respectively, while the corresponding values for inflation were 0.424 (White) and 0.237 (Hansen), respectively.

We conclude that few of the efficiency regressions or tests for serial correlation suggest lack of optimality for our forecasts under squared loss. Hence, we shall proceed to estimate the parameters of our model under the assumption that forecasters use information efficiently.

### 3.3 Parameter Estimates and Tests

The simple benchmark model contains just three free parameters, namely the variance of the innovations in the temporary ( $\sigma_u^2$ ) and persistent ( $\sigma_\varepsilon^2$ ) components, and the persistence parameter,  $\phi$ , for the predictable component. The expressions for the MSE as a function of  $h$ , stated in Proposition 1 for the benchmark model and in the appendix for the learning-based model that uses a Kalman filter, enable us to use GMM to estimate the unknown parameters given a panel of forecast errors measured at various horizons. Access to multi-horizon forecasts is crucial to our

---

<sup>8</sup> Alternative tests of forecast rationality for this form of data are presented in Davies and Lahiri (1995), Clements (1997) and Isiklar, *et al.* (2006).

analysis: These parameters are not separately identifiable if forecasts for a single horizon are all that is available. In contrast, since the variance of the  $h$ -period forecast error grows linearly in  $\sigma_u^2$  while  $\sigma_\varepsilon^2$  and  $\phi$  generally affect the MSE in a non-linear fashion, these parameters can be identified from a sequence of MSE-values corresponding to different forecast horizons,  $h$ , provided at least three different horizons are available.

We estimate the parameters using the moment conditions obtained by matching the sample MSE at various forecast horizons to the population mean squared errors implied by our model:

$$\hat{\theta}_T \equiv \arg \min_{\theta \in \Theta} g_T(\theta)' W_T g_T(\theta) \quad (7)$$

$$g_T(\theta) \equiv \frac{1}{T} \sum_{t=1}^T \begin{bmatrix} e_{t,t-1}^2 - MSE_1(\theta) \\ e_{t,t-2}^2 - MSE_2(\theta) \\ \vdots \\ e_{t,t-24}^2 - MSE_{24}(\theta) \end{bmatrix} \quad (8)$$

where  $\theta \equiv [\sigma_u^2, \sigma_\varepsilon^2, \phi]'$  and  $MSE_h(\theta)$  is obtained using Proposition 1 and the Appendix.

Clearly, we have over-identifying restrictions available, and so the choice of weighting matrix,  $W_T$ , in the GMM estimation is important. We use the identity matrix as the weighting matrix so that all horizons get equal weight in the estimation procedure; this is not fully efficient, but is justified by our focus on modeling the entire term structure of forecast errors. Nevertheless, we still require the covariance matrix of the sample moments to compute standard errors and a test of the over-identifying conditions. Given that our sample is only 14 years long it is not feasible to estimate this matrix directly from the data since this would require controlling for the correlation between the sample moments induced by overlaps across the 24 horizons. Fortunately, given the simple structure of our model, for any given parameter value we can compute a full-rank model-implied covariance matrix of the sample moments despite the fact that our time series is shorter than the number of horizons. Under the assumption that the model is correctly specified, this matrix captures the correlation between sample moments induced by overlaps and serial persistence.<sup>9</sup>

Figure 4 plots the sample root mean squared forecast error (RMSE) for output growth and inflation at the 24 different horizons. In the case of output growth the RMSE shrinks from about

---

<sup>9</sup>While it is possible to derive analytical expressions characterising the covariance matrix, these expressions are extremely long and tedious. We instead simulated 10,000 non-overlapping “years” of data from the model to compute the covariance matrix of the sample moments. Details on this are provided in the appendix.

1.8% at the 24-month horizon to 1% at the 12-month horizon and 0.5% at the 1-month horizon. For inflation it ranges from 0.8% at the two-year horizon to 0.4% at the 12-month horizon and less than 0.1 at the 1-month horizon. Forecast precision improves systematically as the forecast horizon is reduced, as expected. Moreover, consistent with Proposition 1, the rate at which the RMSE declines is smaller in the next-year forecasts at horizons ( $h \geq 12$ ) than in current-year forecasts ( $h < 12$ ).<sup>10</sup>

The fitted values from the models with and without measurement error, also shown in Figure 4, clearly illustrate the limitation of the specification with no measurement error. This model assumes that forecasters get a very precise reading of the outcome towards the end of the current year and hence forces the fitted estimate of the RMSE to decline sharply as the forecast horizon shrinks. This property is clearly at odds with the GDP growth data and means that the benchmark model without measurement error does not succeed in capturing the behavior of the RMSE at both the short and long horizons. For inflation forecasts the assumption of zero measurement error appears consistent with the data.

### 3.4 Introducing Measurement Errors

Although the model used in Proposition 2 is useful for understanding how measurement error impacts the term structure of MSE-values, it is unrealistic in its treatment of the two components of the predicted variable: it allows for a measurement error in the unpredictable component, while assuming that the predictable component,  $x_t$ , is perfectly observable. In reality,  $x_t$  must be extracted from data and thus it is likely measured with substantial error. A more realistic approach allows both  $x_t$  and  $y_t$  to be measured with error and is best-handled by writing the model in state-space form and estimating it using the Kalman filter. Using this framework leaves the state equation

---

<sup>10</sup>Note that Figure 4 reveals no “lumps” in the term structure of forecast errors. One might have expected that around the time of quarterly releases of macroeconomic data the RMSE plot would drop sharply downwards. This is not the case for either GDP growth or inflation, which is consistent with the work of Giannone, *et al.* (2007) who consider how macroeconomic forecasts smoothly incorporate news about the macroeconomy between formal announcement dates.

unchanged:

$$\begin{aligned} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & \phi \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ \varepsilon_t \end{bmatrix} \\ \begin{bmatrix} u_t \\ \varepsilon_t \end{bmatrix} &\sim iid \left( 0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} \right), \end{aligned} \tag{9}$$

while the measurement equation becomes:

$$\begin{aligned} \begin{bmatrix} \tilde{y}_t \\ \tilde{x}_t \end{bmatrix} &= \begin{bmatrix} y_t \\ x_t \end{bmatrix} + \begin{bmatrix} \eta_t \\ \psi_t \end{bmatrix} \\ \begin{bmatrix} \eta_t \\ \psi_t \end{bmatrix} &\sim iid \left( 0, \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\psi^2 \end{bmatrix} \right) \end{aligned} \tag{10}$$

This is a very simple state-space system<sup>11</sup>. Unfortunately, however, it does not yield a formula for the term structure of MSEs that is readily interpretable and so we relegate it and other details to the technical appendix. The key difficulty that arises is best illustrated by considering current-year forecasts. When producing a current-year forecast at time  $t-h$ , economic agents must use past and current information to backcast realizations dated at time  $t-11, \dots, t-h-1$ ; they must also produce a “nowcast” for the current month and, finally, must predict future realizations,  $y_{t-h+1}, \dots, y_t$ . When the persistent component,  $x_t$ , is not observable, the resulting forecast errors will generally be serially correlated even after conditioning on all information that is available to the agents. For example, a large positive realization of  $\eta_{t-h}$  will not only lead to overly optimistic projections for current and future values of  $y$ , but will increase the entire sequence of backcast values. Handling this problem is difficult and requires expressing the backcast, nowcast and forecast errors in terms of the primitive shocks,  $u_t, \varepsilon_t, \eta_t$  and  $\psi_t$ , which are serially uncorrelated. We show how to accomplish this in the appendix.

This extended model introduces two further parameters which reflect the magnitude of measurement errors ( $\sigma_\eta^2$  and  $\sigma_\psi^2$ ). To reduce the set of unknown parameters to a tractable number, we do

---

<sup>11</sup>Faust, *et al.* (2005) find that revisions to U.S. GDP figures are essentially unpredictable, motivating the simple *iid* noise structure used above. These authors find revisions to GDP figures in some other G7 countries are significantly predictable, and our model would thus need to be extended to apply to those countries. As noted in the previous section, measurement error does not appear important for our inflation data and so the structure of the measurement equation is less important for this variable.



two things. First, as a normalization we set  $\sigma_\psi^2 \rightarrow \infty$ , effectively removing  $\tilde{x}_t$  from the measurement equation. This choice reflects that, in practice, it is the predicted variable that is observed with noise, i.e.  $\tilde{y}_t$ , whereas the persistent component is a latent variable that must be extracted from the observations of the predicted variable. Furthermore, even though both  $\sigma_\eta^2$  and  $\sigma_u^2$  are well-identified in theory, in practice they are difficult to estimate separately. We therefore set  $\sigma_\eta$  to be proportional to  $\sigma_u$  :  $\sigma_\eta = k \cdot \sigma_u$  and estimated the model for  $k = \{0.01, 0.25, 0.5, 1, 2, 3, 4, 5, 10\}$ . The goodness-of-fit of the model (as measured by Hansen’s (1982)  $J$ -test of over-identifying restrictions) was generally robust for  $1 \leq k \leq 4$ . We set  $k = 2$  in the estimation.

Figure 4 shows that this specification does a much better job at matching the decay pattern in observed RMSE for US output growth as the forecast horizon,  $h$ , shrinks to zero. In the case of US inflation, there is little to distinguish between the models with and without measurement error. This is consistent with Croushore and Stark (2001) who report that revisions in reported GDP figures tend to be larger than those in reported inflation figures.

Table 2 presents parameter estimates for the model with measurement errors fitted to the Consensus forecasts. The predictable component in inflation appears to be slightly more persistent than that in output growth<sup>12</sup>. Moreover, the model passes the specification tests for both variables and thus there is little statistical evidence against our simple specification, once measurement errors are considered.

The parameter estimates in Table 2 can be difficult to interpret since the ‘explained’ part of the predicted variable also depends on  $\phi$ . Plots of the predictive  $R^2$  as a function of the forecast horizon,  $h$ , are easier to interpret. We show these in Figure 5 for US GDP growth and inflation.<sup>13</sup>

<sup>12</sup>The implied first-order autocorrelation coefficients for quarterly (annual) GDP growth and inflation are 0.60 and 0.69 (0.80 and 0.88) respectively.

<sup>13</sup>For the computation of the  $R^2 = 1 - MSFE/V[Z]$ , we use the model-implied MSFE in the numerator, and the sample variance of the variables in the denominator. This is done because the sample variance is not one of the moments matched in our estimation (it corresponds to matching the  $h \rightarrow \infty$  forecast horizon) and so the model-implied unconditional variance is a poor estimate of the true variance. Also note that we use data from 1991-2004 to estimate the variance of inflation, and data from 1971-2004 to estimate the variance of GDP growth. We use a longer sample for GDP growth because we obtain *negative*  $R^2$  values using sample starting in 1991. (This is true using sample MSFEs as well as model-implied MSFEs.) Presumably this is because the sample variance of GDP growth since 1991 is very low relative to historical data, see McConnell and Perez-Quiros (2000). Using any starting point after about 1980 led to negative  $R^2$  values, while using almost any starting point before 1980 lead to very similar  $R^2$  values.

In the case of GDP growth, the  $R^2$  rises from 0.3 at the 24-month horizon to values of 0.6 and 0.9 at the 12-month and 1-month horizons, respectively. These values suggest great uncertainty about output growth two-years ahead of time, but also show that uncertainty is greatly reduced over the following months.

Turning to the inflation plots in the lower panels of Figure 5, these suggest that more is known about inflation than about GDP growth two years ahead of time. In part this reflects the higher degree of persistence of inflation compared with output growth. The  $R^2$  associated with the 24, 12 and 1-month forecast horizons are 0.8, 0.9 and close to 1, respectively. At least during the sample period covered here, a great deal appears to have been known about inflation even two years ahead of time. While additional news about inflation emerge during the course of the year, there is clearly less incremental information to be gained about inflation than about output growth during the year in question.

### 3.5 The estimated components of GDP growth and inflation

Our model for the term structure of consensus forecast errors is based on the decomposition of the target variable, GDP growth or inflation, into a persistent component,  $x_t$ , and an unpredictable component,  $u_t$ . Our GMM estimation procedure does not require the estimation of the sample paths for  $x_t$  and  $u_t$ , unlike MLE, however with the estimated parameter vector and the panel of forecasts we are able to infer the forecaster’s estimated values of these variables. We use the long-horizon forecasts ( $h \geq 12$ ) to infer the forecaster’s estimate of the persistent component, and the short-horizon ( $h < 12$ ) forecasts to infer the forecaster’s estimate of the unpredictable component. Intuitively, one can think of our estimates of these two components as an alternative representation of the two forecasts the forecaster makes at each point in time (the “next year”,  $h \geq 12$ , and the “current year”,  $h < 12$ , forecasts). We can obtain both of these components without needing to make any further identifying assumptions, and without needing to employ any data other than the panel of forecasts. Details are presented in the appendix.

In Figure 6 we present the estimated persistent components of GDP growth and inflation, as implied by the observed consensus forecasts and the parameters of our model, for each month in our sample. For reference we plot both the “filtered” estimates, which are estimates of  $E \left[ x_t | \tilde{\mathcal{F}}_t \right]$ , and the “smoothed” estimates, which are estimates of  $E \left[ x_t | \tilde{\mathcal{F}}_T \right]$ . The estimates for GDP growth reveal that the panel of professional forecasters estimated the level of GDP growth in the early 1990s

quite well, but were consistently surprised by the strong GDP growth in the mid to late 1990s: the estimated persistent component of GDP growth hovered around 1.5% annualized, whereas the actual GDP growth in that period was closer to 4%. Since the 2001 recession the persistent component has consistently been above the realized values of GDP growth.

Similarly, our panel of forecasters was consistently surprised by the declining inflation of the 1990s, with our estimated persistent component generally lying above the realized values of inflation. In the latter part of the sample the estimated persistent component is more in line with realized inflation, consistent with the view that forecasters took some time to adjust their views on long-run inflation in the US.

## 4 Dispersion Among Forecasters

So far our analysis concentrated on explaining properties of the evolution in the consensus forecasts and forecast errors and we ignored heterogeneity among forecasters. In actuality, as indicated by Figure 1, there is often considerable disagreement among forecasters. We shall model disagreement as arising from two possible sources: differences in the information obtained by each individual forecaster, or differences in their prior beliefs. We define the cross-sectional dispersion among forecasters as

$$d_{t,t-h}^2 \equiv \frac{1}{N_{t,t-h}} \sum_{i=1}^{N_{t,t-h}} (\hat{z}_{i,t,t-h} - \bar{z}_{t,t-h})^2 \quad (11)$$

where  $\bar{z}_{t,t-h} \equiv \frac{1}{N_{t,t-h}} \sum_{i=1}^{N_{t,t-h}} \hat{z}_{i,t,t-h}$  is the consensus forecast of  $z_t$ , computed at time  $t-h$ ,  $\hat{z}_{i,t,t-h}$  is forecaster  $i$ 's prediction of  $z_t$  at time  $t-h$  and  $N_{t,t-h}$  is the number of forecasts available at time  $t$  for forecast horizon  $h$ . Our data set does not contain information on  $N_{t,t-h}$ , although it is known to be around 20 to 50. In our computations we set  $N_{t,t-h} = 30$  for all  $t, h$ .<sup>14</sup>

To capture heterogeneity in the forecasters' information, we assume that each forecaster observes

---

<sup>14</sup>We could theoretically estimate this, imposing for example that  $N_{t,t-h} = N_h$ , as a substitute for allowing for a more standard residual term in our model. We elect to employ the latter modelling approach for simplicity; the details of our residual term are described below.

a different signal of the current value of  $y_t$ , denoted  $\tilde{y}_{i,t}$ :

$$\begin{aligned}\tilde{y}_{i,t} &= y_t + \eta_t + \nu_{i,t} \\ \eta_t &\sim iid(0, \sigma_\nu^2) \quad \forall t \\ \nu_{i,t} &\sim iid(0, \sigma_\nu^2) \quad \forall t, i \\ E[\nu_{i,t}\eta_s] &= 0 \quad \forall t, s, i\end{aligned}\tag{12}$$

Individual forecasters' measurements of  $y_t$  are contaminated with a common source of noise, denoted  $\eta_t$  as in the model for consensus MSE, and independent idiosyncratic noise, denoted  $\nu_{i,t}$ . The participants in the survey we use are not formally able to observe each others' forecasts for the current period but they do observe previous survey forecasts<sup>15</sup>. For this reason, we include a second measurement variable,  $\tilde{y}_{t-1}$ , which is the measured value of  $y_{t-1}$  contaminated with only the common noise:

$$\tilde{y}_{t-1} = y_{t-1} + \eta_{t-1}\tag{13}$$

From this, the individual forecaster is able to compute the optimal forecast from the variables observable to him:

$$\hat{z}_{i,t,t-h}^* \equiv E[z_t | \tilde{\mathcal{F}}_{i,t-h}], \quad \tilde{\mathcal{F}}_{i,t-h} = \{\tilde{y}_{i,t-h-j}, \tilde{y}_{t-h-1-j}\}_{j=0}^{t-h}.\tag{14}$$

Differences in signals about the predicted variable alone are unlikely to explain the observed degree of dispersion in the forecasts. The simplest way to verify this is to consider dispersion for very long horizons: as  $h \rightarrow \infty$  the optimal forecasts converge towards the unconditional mean of the predicted variable. Since we assume that all forecasters have the same (true) model this implies that dispersion should asymptote to zero as  $h \rightarrow \infty$ . As we shall see in the empirical analysis, this implication is in stark contrast with our data, which suggests instead that the cross-sectional dispersion converges to a constant but non-zero level as the forecast horizon grows. Thus there must be a source of dispersion beyond that deriving from differences in signals.

We therefore consider a second source of dispersion by assuming that each forecaster comes with prior beliefs about the unconditional average of  $z_t$ , denoted  $\mu_i$ . We assume that forecaster  $i$  shrinks the optimal forecast based on his information set  $\tilde{\mathcal{F}}_{i,t-h}$  towards his prior belief about

---

<sup>15</sup>As the participants in this survey are professional forecasters they may be able to observe each others' current forecasts through published versions of their forecasts, for example: investment bank newsletters or recommendations. If this is possible, then we would expect to find  $\sigma_\nu$  close to zero.

the unconditional mean of  $z_t$ . The degree of shrinkage is governed by a parameter  $\kappa^2 \geq 0$ , with low values of  $\kappa^2$  implying a small weight on the data-based forecast  $\hat{z}_{i,t,t-h}^*$  (i.e., a large degree of shrinkage towards the prior belief) and large values of  $\kappa^2$  implying a high weight on  $\hat{z}_{i,t,t-h}^*$ . As  $\kappa^2 \rightarrow 0$  the forecaster places all weight ( $\omega_h$ ) on his prior beliefs and none on the data; as  $\kappa^2 \rightarrow \infty$  the forecaster places no weight on his prior beliefs.

$$\begin{aligned}\hat{z}_{i,t-h,t} &= \omega_h \mu_i + (1 - \omega_h) E[z_t | \mathcal{F}_{i,t-h}], \\ \omega_h &= \frac{E[e_{i,t,t-h}^2]}{\kappa^2 + E[e_{i,t,t-h}^2]} \\ e_{i,t,t-h} &\equiv z_t - E[z_t | \mathcal{F}_{i,t-h}].\end{aligned}\tag{15}$$

We allow the weights placed on the prior and the optimal expectation  $E[z_t | \tilde{\mathcal{F}}_{i,t-h}]$  to vary across the forecast horizons in a manner consistent with standard forecast combinations: as  $\hat{z}_{i,t,t-h}^* \equiv E[z_t | \tilde{\mathcal{F}}_{i,t-h}]$  becomes more accurate (i.e., as  $E[e_{i,t,t-h}^2]$  decreases) the weight attached to that forecast increases. Thus for short horizons the weight put on the prior is reduced, while for long horizons the weight attached to the prior grows<sup>16</sup>. Furthermore, note that

$$\omega_h \rightarrow \frac{V[z_t]}{\kappa^2 + V[z_t]} \text{ as } h \rightarrow \infty.$$

For analytical tractability, and for better finite sample identification of  $\kappa^2$ , we impose that  $\kappa^2$  is constant across all forecasters<sup>17</sup>.

The additional term  $\mu_i$  could arise even in a classical setting if we consider that the forecasters may use different models for long-run growth or inflation (for example, models with or without cointegrating relationships imposed) or if forecasters choose to use different sample periods for the computation of their forecasts. In both cases, the  $\mu_i$  term would generally be time-varying, but we leave that possibility aside for now.

Alternatively, the dispersion of forecasts at long horizons could be the outcome of a game played between individual forecasters, with only limited connection to the statistical properties of

---

<sup>16</sup>Lahiri and Sheng (2006) also propose a parametric model for the cross-sectional dispersion of macroeconomic forecasts as a function of the forecast horizon. They model the dispersion term structure directly, rather than through a combined model of the data generating process and the individual forecasters' prediction process as above.

<sup>17</sup>As a normalization we assume that  $N^{-1} \sum_{i=1}^N \mu_i = 0$  since we cannot separately identify  $N^{-1} \sum_{i=1}^N \mu_i$  and  $\sigma_\mu^2 \equiv N^{-1} \sum_{i=1}^N \mu_i^2$  from our data on forecast dispersions. This normalization is reasonable if we think that the number of "optimistic" forecasters is approximately equal to the number of "pessimistic" forecasters.

the underlying data. Laster, *et al.* (1999), for example, show that under certain conditions the equilibrium distribution of individual forecasts is proportional to the conditional distribution of the target variable, which has some similarities with our specification above. Lamont (2002) and Ottaviani and Sørensen (2006) consider other strategic environments that can generate increased cross-sectional dispersion in individual forecasts.

#### 4.1 Empirical Results: Cross-Sectional Dispersion

Equations (12) to (15) allow us to characterize the mean of the cross-sectional dispersion,  $\delta_h^2 \equiv E \left[ d_{t,t-h}^2 \right]$ . Given our model for the mean of the term structure of dispersion in beliefs, all that remains is to specify a residual term for the model. Since the dispersion is measured by the cross-sectional variance, it is sensible to allow the innovation term to be heteroskedastic, with variance related to the level of the dispersion. This form of heteroskedasticity, where the cross-sectional dispersion increases with the level of the predicted variable, has been documented empirically for inflation data by e.g. Grier and Perry (1998). We use the following model:

$$\begin{aligned} d_{t,t-h}^2 &= \delta_h^2 \cdot \lambda_{t,t-h} \\ E[\lambda_{t,t-h}] &= 1 \\ V[\lambda_{t,t-h}] &= \sigma_\lambda^2, \end{aligned} \tag{16}$$

where  $d_{t,t-h}^2$  is the observed value of the cross-sectional dispersion. In particular, we assume that the parameter capturing time variation in the cross-sectional dispersion,  $\lambda_{t,t-h}$ , is log-normally distributed with unit mean:

$$\lambda_{t,t-h} \sim iid \log N \left( -\frac{1}{2}\sigma_\lambda^2, \sigma_\lambda^2 \right).$$

Thus, our model for dispersion introduces four additional parameters relative to the model based only on the consensus forecast. Three parameters,  $\sigma_\mu^2$ ,  $\sigma_\nu^2$  and  $\kappa^2$ , relate to the term structure of forecast dispersions—i.e. how the dispersion changes as a function of the forecast horizon,  $h$ —while the fourth parameter,  $\sigma_\lambda^2$ , relates to the variance of forecast dispersions through time.

To estimate  $\sigma_\lambda^2$  we need to incorporate information from the degree of variability in dispersions. In addition to the term structures of consensus MSE-values and cross-sectional dispersion (each yielding up to 24 moment conditions) we also include moments implied by the term structure of dispersion variances to estimate the parameters of our full model. In total this model for dispersion

uses up to 72 moment conditions to estimate 8 parameters:

Data generating process parameters:  $\sigma_u^2, \sigma_\varepsilon^2, \phi$

Measurement error parameter:  $\sigma_\eta^2$

Forecaster dispersion parameters:  $\sigma_\nu^2, \sigma_\mu^2, \kappa^2$

Dispersion error parameter:  $\sigma_\lambda^2$ .

As in the analysis of the consensus data, to reduce the number of parameters we fix  $\sigma_\eta = k \cdot \sigma_u$  with  $k = 2$  and we estimate the remaining parameters by GMM with identity weight matrix, see the appendix for details.

$$\hat{\theta}_T \equiv \arg \min_{\theta \in \Theta} g_T(\theta)' g_T(\theta) \quad (17)$$

$$g_T(\theta) \equiv \frac{1}{T} \sum_{t=1}^T \begin{bmatrix} e_{t,t-1}^2 - MSE_1(\theta) \\ \vdots \\ e_{t,t-24}^2 - MSE_{24}(\theta) \\ d_{t,t-1}^2 - \delta_1^2(\theta) \\ \vdots \\ d_{t,t-24}^2 - \delta_{24}^2(\theta) \\ (d_{t,t-1}^2 - \delta_1^2(\theta))^2 - \delta_1^4(\theta) (\exp(\sigma_\lambda^2) - 1) \\ \vdots \\ (d_{t,t-24}^2 - \delta_{24}^2(\theta))^2 - \delta_{24}^4(\theta) (\exp(\sigma_\lambda^2) - 1) \end{bmatrix} \quad (18)$$

Panel A of Table 3 reports parameter estimates for this model. Compared with Table 2, the estimates of the parameters defining the dynamics of the target variables are essentially unchanged. The estimates of  $\kappa$  and  $\sigma_\mu$  suggest considerable heterogeneity across forecasters in our panel, whereas the estimates of  $\sigma_\nu$  indicate that differences in individual signals may not be important, consistent with the possibility that the individual forecasters in our panel are able to observe each others' contemporaneous forecasts, rather than with a one-period lag.

Figure 7 shows the cross-sectional dispersion (in standard deviation format) in output growth and inflation forecasts as a function of the forecast horizon. The cross-sectional dispersion of output growth declines only slowly for horizons in excess of 12 months, but declines rapidly for  $h < 12$  months from a level near 0.4 at the 12-month horizon to around 0.1 at the 1-month horizon. For

inflation, again there is a systematic reduction in the dispersion as the forecast horizon shrinks. The cross-sectional dispersion declines from around 0.45 at the 24-month horizon to 0.3 at the 12-month horizon and 0.1 at the 1-month horizon.

Our tests of the over-identifying restrictions for each model indicate that the model provides a good fit to the GDP growth consensus forecast and forecast dispersion, with the  $p$ -value for that test being 0.86. Moreover, the top panel of Figure 7 confirms that the model provides a close fit to the empirical term structure of forecast dispersions. This panel also shows that the model with  $\sigma_\nu$  set to zero provides almost as good a fit as the model with this parameter freely estimated. This indicates that differences in individual information about GDP growth, modelled by  $\nu_{it}$ , are not important for explaining forecast dispersion; the most important features are the differences in prior beliefs about long-run GDP growth and the accuracy of Kalman filter-based forecasts (as they affect the weight given to the prior relative to the Kalman filter forecast).

The model for inflation forecasts and dispersions is rejected by the test of over-identifying restrictions. The model fits dispersion well for horizons greater than 12 months, but for horizons less than 9 months the observed dispersion is systematically above what is predicted by our model. Given the functional form specified for the weight attached to the prior belief about long-run inflation versus the Kalman filter-based forecast, the model predicts that each forecaster will place 95.0% and 99.1% weight on his/her Kalman filter-based forecast for  $h = 3$  and 1. The Kalman filter forecasts are very similar across forecasters at short horizons and thus our model predicts that dispersion will be low.

In contrast, the observed dispersion is relatively high, particularly when compared with the observed forecast errors: observed dispersion (in standard deviations) for horizons 3 and 1 are 0.11 and 0.07, compared with the RMSE of the consensus forecast at these horizons of 0.08 and 0.05. Compare this with the corresponding figures for the GDP forecasts, with dispersions of 0.14 and 0.08 and RMSE of 0.61 and 0.56. Thus, the dispersion of inflation forecasts is around 25% greater than the RMSE of the consensus forecast for short horizons, whereas the dispersion of GDP growth forecasts is around 75% *smaller* than the RMSE of the consensus forecast. Examining this ratio as  $h$  goes from 24 down to 1 month we find that dispersion/RMSE ranges from 0.32 to 0.15 for GDP growth, while it ranges from 0.60 to 1.45 for inflation. So dispersion/RMSE decreases slightly from long to short horizons for GDP growth, whereas it *rises* substantially for inflation. This is difficult to explain within the confines of our model.



## 4.2 Time-varying dispersion

There is a growing amount of theoretical and empirical work on the relationship between uncertainty, somehow defined, and the economic environment. In this section we present an extension of our model to allow for time-varying dispersion of forecasts. Given our focus on GDP growth and inflation, a business cycle indicator is a natural variable to consider as a determinant of forecast dispersion. However, as our sample only runs from 1991 to 2004, such a variable would likely not exhibit sufficient variability. Instead we employ the default spread (the difference in average yields of corporate bonds rated by Moody's as BAA vs. AAA), which is known to be strongly counter-cyclical: the default spread increases during economic downturns. Over our sample period, for example, the default spread ranges from 55 basis points in September and November 1997 to 141 basis points in January 1991 and January 2002.

The most natural way to allow the default spread to influence dispersion in our model is through the variance of the individual signals received by the forecasters,  $\sigma_\nu^2$ , or through the variance of the prior beliefs about the long-run values of the series,  $\sigma_\mu^2$ . Given that the former variable explained very little of the (unconditional) dispersion term structure, we focus on the latter channel. We specify our model as

$$\log \sigma_{\mu,t}^2 = \beta_0^\mu + \beta_1^\mu \log S_t \quad (19)$$

where  $S_t$  is the default spread in month  $t$ . In this model, increases in the default spread coincide with increased differences in beliefs about the long-run value of the series, which in turn lead to an increase in the observed dispersion of forecasts.

Leaving the rest of the model unchanged, we estimated this extension and present the results in Panel B of Table 3. The fit of the models were not much changed by this extension. Interestingly, the results reveal a positive relationship between default spreads and  $\sigma_\mu$ , as evidenced by the signs of  $\hat{\beta}_1^\mu$  in Table 3. This parameter is not significantly different from zero for the inflation forecast model, but is significant at the 10% level for the GDP growth forecast model.

In Figure 8 we plot the estimated dispersions as a function of the level of default spreads. When the default spread is equal to its sample 95<sup>th</sup> percentile (131 basis points), GDP growth forecast dispersion is approximately double what it is when the default spread is equal to its sample average (83 basis points). Similarly, when the default spread is equal to its 5<sup>th</sup> percentile (58 basis points) GDP growth forecast dispersion is approximately one-half of the average figure. In contrast, the

dispersion of inflation forecasts is only weakly affected by the default spread, with changes of approximately no more than 10% when the default spread moves from its average value to its 5<sup>th</sup> or 95<sup>th</sup> percentile value. Thus GDP growth forecast dispersion has a strong and significant counter-cyclical component, whereas inflation forecast dispersion appears only weakly counter-cyclical.

## 5 Conclusion

This paper studied how macroeconomic uncertainty, measured through consensus forecast errors and the cross-sectional dispersion in forecasters' beliefs, is resolved over time. To this end we considered fixed event forecasts of macroeconomic variables which hold the event date constant, while reducing the length of the forecast horizon. We proposed a new model for the evolution in the consensus forecast which accounts for measurement errors and incorporates the forecasters' filtering problem, and developed a model for the cross-sectional dispersion among forecasters that accounts for differences in forecasters' information signals and differences among their prior beliefs. Though highly parsimonious, our simple models succeed in capturing the level, slope and curvature of the term structure of forecast errors, and shed some light on the primary sources of the cross-sectional dispersion among forecasters. Consistent with several previous studies, we find that measurement errors are an important source of forecast error for GDP growth forecasts, while they are essentially negligible for inflation forecasts. We also find significant evidence that forecast dispersion is primarily driven by differences in beliefs about long-run values of GDP growth and inflation, as opposed to differences in information about the current state of the economy.

## 6 Appendix: Proofs

**Proof of Proposition 1.** Since  $z_t = \sum_{j=0}^{11} y_{t-j}$  and  $y_t = x_t + u_t$ , where  $x_t$  is the persistent component, forecasting  $z_t$  given information  $h$  months prior to the end of the measurement period,  $\mathcal{F}_{t-h} = \{x_{t-h}, y_{t-h}, x_{t-h-1}, y_{t-h-1}, \dots\}$ , requires accounting for the persistence in  $x$ . Note that

$$\begin{aligned} x_{t-h+1} &= \phi x_{t-h} + \varepsilon_{t-h+1} \\ x_{t-h+2} &= \phi^2 x_{t-h} + \phi \varepsilon_{t-h+1} + \varepsilon_{t-h+2} \\ x_{t-h+3} &= \phi^3 x_{t-h} + \phi^2 \varepsilon_{t-h+1} + \phi \varepsilon_{t-h+2} + \varepsilon_{t-h+3} \\ &\vdots \\ x_t &= \phi^h x_{t-h} + \phi^{h-1} \varepsilon_{t-h+1} + \phi^{h-2} \varepsilon_{t-h+2} + \dots + \phi \varepsilon_{t-1} + \varepsilon_t \end{aligned}$$

Adding up these terms we find that, for  $h \geq 12$ ,

$$\begin{aligned} z_t &= \sum_{j=0}^{11} x_{t-j} + \sum_{j=0}^{11} u_{t-j} \\ &= \frac{\phi(1-\phi^{12})}{1-\phi} x_{t-12} + \frac{1}{1-\phi} \sum_{j=0}^{11} (1-\phi^{12-j}) \varepsilon_{t-12+1+j} + \sum_{j=0}^{11} u_{t-j}. \end{aligned} \tag{20}$$

Thus the optimal forecast for  $h \geq 12$  is

$$\begin{aligned} \hat{z}_{t,t-h}^* &\equiv E[z_t | \mathcal{F}_{t-h}] = \sum_{j=0}^{11} E[y_{t-j} | \mathcal{F}_{t-h}] = \sum_{j=0}^{11} E[x_{t-j} | \mathcal{F}_{t-h}] = \sum_{j=0}^{11} \phi^{h-j} x_{t-h}, \\ \text{so } \hat{z}_{t,t-h}^* &= \frac{\phi^{h-11} (1-\phi^{12})}{1-\phi} x_{t-h}, \text{ for } h \geq 12. \end{aligned}$$

For the current year forecasts ( $h < 12$ ) the optimal forecast of  $z_t$  makes use of those realizations of  $y$  that have already been observed. Thus the optimal forecast is:

$$\begin{aligned} \hat{z}_{t,t-h}^* &= \sum_{j=0}^{11} E[y_{t-j} | \mathcal{F}_{t-h}] = \sum_{j=h}^{11} y_{t-j} + \sum_{j=0}^{h-1} E[x_{t-j} | \mathcal{F}_{t-h}] = \sum_{j=h}^{11} y_{t-j} + \sum_{j=0}^{h-1} \phi^{h-j} x_{t-h}, \\ \text{so } \hat{z}_{t,t-h}^* &= \sum_{j=h}^{11} y_{t-j} + \frac{\phi(1-\phi^h)}{1-\phi} x_{t-h}, \text{ for } h < 12. \end{aligned}$$

Using these expressions for the optimal forecasts we can derive the forecast error,  $e_{t,t-h} \equiv z_t - \hat{z}_{t,t-h}^*$ , as a function of the forecast horizon. For  $h \geq 12$ ,

$$\begin{aligned} e_{t,t-h} &= \sum_{j=0}^{11} u_{t-j} + \sum_{j=0}^{11} x_{t-j} - \frac{\phi^{h-11} (1 - \phi^{12})}{1 - \phi} x_{t-h} \\ &= \sum_{j=0}^{11} u_{t-j} + \sum_{j=0}^{11} \frac{1 - \phi^{j+1}}{1 - \phi} \varepsilon_{t-j} + \sum_{j=12}^{h-1} \frac{\phi^{j-11} (1 - \phi^{12})}{1 - \phi} \varepsilon_{t-j} \end{aligned}$$

In computing the variance of  $e_{t,t-h}$  we exploit the fact that  $u$  and  $\varepsilon$  are *iid* through time and independent of each other at all lags. For  $h \geq 12$ ,

$$\begin{aligned} E [e_{t,t-h}^2] &= \sum_{j=0}^{11} E [u_{t-j}^2] + \sum_{j=0}^{11} \frac{(1 - \phi^{j+1})^2}{(1 - \phi)^2} E [\varepsilon_{t-j}^2] + \sum_{j=12}^{h-1} \frac{\phi^{2j-22} (1 - \phi^{12})^2}{(1 - \phi)^2} E [\varepsilon_{t-j}^2] \\ &= 12\sigma_u^2 + \frac{\sigma_\varepsilon^2}{(1 - \phi)^2} \sum_{j=0}^{11} (1 - \phi^{j+1})^2 + \frac{(1 - \phi^{12})^2}{(1 - \phi)^2} \sigma_\varepsilon^2 \sum_{j=12}^{h-1} \phi^{2j-22} \\ &= 12\sigma_u^2 + \frac{\sigma_\varepsilon^2}{(1 - \phi)^2} \left( 12 - 2 \frac{\phi (1 - \phi^{12})}{1 - \phi} + \frac{\phi^2 (1 - \phi^{24})}{(1 - \phi^2)} \right) \\ &\quad + \frac{\phi^2 (1 - \phi^{12})^2 (1 - \phi^{2h-24})}{(1 - \phi)^3 (1 + \phi)} \sigma_\varepsilon^2 \end{aligned}$$

as presented in the proposition. For  $h < 12$  we have:

$$\begin{aligned} e_{t,t-h} &= \sum_{j=0}^{11} y_{t-j} - \sum_{j=h}^{11} y_{t-j} - \frac{\phi (1 - \phi^h)}{1 - \phi} x_{t-h} \\ &= \sum_{j=0}^{h-1} u_{t-j} + \sum_{j=0}^{h-1} \frac{1 - \phi^{j+1}}{1 - \phi} \varepsilon_{t-j} \\ \text{so } E [e_{t,t-h}^2] &= \sum_{j=0}^{h-1} E [u_{t-j}^2] + \sum_{j=0}^{h-1} \frac{(1 - \phi^{j+1})^2}{(1 - \phi)^2} E [\varepsilon_{t-j}^2] \\ &= h\sigma_u^2 + \frac{\sigma_\varepsilon^2}{(1 - \phi)^2} \left( h - 2 \frac{\phi (1 - \phi^h)}{1 - \phi} + \frac{\phi^2 (1 - \phi^{2h})}{1 - \phi^2} \right). \end{aligned}$$

■

**Proof of Proposition 2.** (1) We first need to derive the optimal backcast of  $y_{t-j-h}$  given

$\tilde{\mathcal{F}}_{t-h}$ , for  $j \geq 0$ . Since  $u_{t-j-h}$  is *iid*

$$\begin{aligned} E \left[ y_{t-j-h} | \tilde{\mathcal{F}}_{t-h} \right] &= x_{t-j-h} + E \left[ u_{t-j-h} | \tilde{\mathcal{F}}_{t-h} \right] \\ &= x_{t-j-h} + E \left[ u_{t-j-h} | \tilde{y}_{t-j-h}, x_{t-j-h} \right]. \end{aligned}$$

$$\text{Recall } \tilde{y}_{t-j-h} = x_{t-j-h} + u_{t-j-h} + \eta_{t-j-h}.$$

$$\text{Define } \tilde{u}_{t-j-h} \equiv \tilde{y}_{t-j-h} - x_{t-j-h} = u_{t-j-h} + \eta_{t-j-h}.$$

$$\text{Then } E \left[ u_{t-j-h} | \tilde{y}_{t-j-h}, x_{t-j-h} \right] = E \left[ u_{t-j-h} | \tilde{u}_{t-j-h} \right].$$

Finally, since

$$\begin{aligned} \begin{bmatrix} u_{t-j-h} \\ \eta_{t-j-h} \end{bmatrix} &\sim N \left( 0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \right) \\ E \left[ u_{t-j-h} | \tilde{u}_{t-j-h} \right] &= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \tilde{u}_{t-j-h} \\ \text{so } E \left[ y_{t-j-h} | \tilde{\mathcal{F}}_{t-h} \right] &= x_{t-j-h} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \tilde{u}_{t-j-h} \\ &= \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} x_{t-j-h} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \tilde{y}_{t-j-h}, \end{aligned}$$

as claimed. If  $[u_t, \eta_t]'$  are not jointly normal, this forecast is interpretable as the optimal linear projection of  $y_{t-j-h}$  on elements of  $\tilde{\mathcal{F}}_{t-h}$ .

(2) We next use these results to find the optimal forecast of  $z_t$ . Note that the presence of measurement error in  $y_t$  does not affect optimal forecasts for  $h \geq 12$  because such predictions rely only on the measured values of  $x_t$ . Thus we only need to modify the forecasts for  $h < 12$ :

$$\begin{aligned} E \left[ z_t | \tilde{\mathcal{F}}_{t-h} \right] &= \sum_{j=0}^{11} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] \\ &= \sum_{j=0}^{h-1} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] + \sum_{j=h}^{11} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right], \quad \text{where} \\ \sum_{j=0}^{h-1} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] &= \sum_{j=0}^{h-1} E \left[ x_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] = \sum_{j=0}^{h-1} \phi^{h-j} x_{t-h} = \frac{\phi(1-\phi^h)}{1-\phi} x_{t-h} \\ \sum_{j=h}^{11} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] &= \sum_{j=h}^{11} \left( \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} x_{t-j} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \tilde{y}_{t-j} \right) \end{aligned}$$

(3) The presence of measurement error adds an extra term to the forecast errors for  $h < 12$ , corresponding to the difference between the true value of  $y$  and the forecaster's best estimate of  $y$ ,

$\sum_{j=h}^{11} y_{t-j} - \sum_{j=h}^{11} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right]$ . The forecast error is therefore

$$e_{t,t-h} = \sum_{j=0}^{h-1} u_{t-j} + \sum_{j=0}^{h-1} \frac{1 - \phi^{j+1}}{1 - \phi} \varepsilon_{t-j} + \sum_{j=h}^{11} \left( y_{t-j} - E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] \right)$$

$$\begin{aligned} \text{Note that } y_{t-j} - E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] &= y_{t-j} - \left( \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} \right) x_{t-j} - \left( \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \right) \tilde{y}_{t-j,t-h} \\ &= x_{t-j} + u_{t-j} - \left( \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} \right) x_{t-j} - \left( \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \right) (x_{t-j} + u_{t-j} + \eta_{t-j}) \\ &= \frac{\sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2} u_{t-j} - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\eta^2} \eta_{t-j}. \end{aligned}$$

$$\text{Hence } V \left[ y_{t-j} - E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] \right] = \frac{\sigma_u^2 \sigma_\eta^4}{(\sigma_u^2 + \sigma_\eta^2)^2} + \frac{\sigma_u^4 \sigma_\eta^2}{(\sigma_u^2 + \sigma_\eta^2)^2} = \frac{\sigma_u^2 \sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2}.$$

The variance of the extra terms in the MSE is then

$$V \left[ \sum_{j=h}^{11} \left( y_{t-j} - E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] \right) \right] = \frac{(12-h) \sigma_u^2 \sigma_\eta^2}{\sigma_u^2 + \sigma_\eta^2}$$

as stated in the proposition. ■

## 7 Technical appendix: Kalman filter implementation details

### 7.1 Details on the Kalman filter model for RMSE

We first describe the model for the consensus forecasts, using notation similar to those in Hamilton (1994). We assume that our forecaster knows the form and the parameters of the data generating process for  $z_t$  but does not observe this variable. Instead he only observes  $\tilde{y}_t$  which is a noisy estimate of  $y_t$ . We further assume that he uses the Kalman filter (KF) to optimally predict (forecast, “nowcast” and “backcast”) the values of  $y_t$  needed for the forecast of  $z_t$ .

The (scalar) variable of interest is

$$z_t \equiv \sum_{j=0}^{11} y_{t-j} \tag{21}$$

Letting  $\xi_t = [y_t, x_t]'$ ,  $F = [\mathbf{0}, \phi \iota]$ , where  $\mathbf{0}$  and  $\iota$  are  $2 \times 1$  vectors of zeros and ones, respectively, and  $v_t = [u_t + \varepsilon_t, \varepsilon_t]'$ , the state equation is

$$\xi_t = F \xi_{t-1} + v_t, \tag{22}$$

and the measurement equation is

$$\tilde{y}_t = H' \xi_t + w_t. \quad (23)$$

In our application the measurement variable is a scalar,  $\tilde{y}_t = y_t + \eta_t$ , but we will present our theoretical framework for the general case that  $\tilde{y}_t$  is a vector. The properties of the innovations to the state and measurement equations are:

$$\begin{aligned} v_t &\sim iid N(0, Q) \\ Q &= \begin{bmatrix} \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 \end{bmatrix} \\ w_t &\sim iid N(0, R), \end{aligned} \quad (24)$$

where in our application  $R = \sigma_\eta^2$ . Further, we assume

$$E[v_t w_s'] = 0 \quad \forall s, t \quad (25)$$

We also assume that the forecaster has been using the KF long enough that all updating matrices are at their steady-state values. This is done simply to remove any “start of sample” effects that may or may not be present in our actual data. But we still need to initialize the KF, which requires the following:

$$\begin{aligned} \tilde{\mathcal{F}}_t &= \sigma(\tilde{y}_t, \tilde{y}_{t-1}, \dots, \tilde{y}_1) \\ \hat{\xi}_{t|t-1} &\equiv E[\xi_t | \tilde{\mathcal{F}}_{t-1}] \equiv E_{t-1}[\xi_t] \\ \hat{y}_{t|t-1} &\equiv E[\tilde{y}_t | \tilde{\mathcal{F}}_{t-1}] \equiv E_{t-1}[\tilde{y}_t] = E_{t-1}[\xi_t]. \end{aligned}$$

Following Hamilton (1994),

$$\begin{aligned} E\left[\left(\xi_t - \hat{\xi}_{t|t-1}\right)\left(\tilde{y}_t - \hat{y}_{t|t-1}\right)'\right] &= E\left[\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)'\right] \quad \text{in our case} \\ &\equiv P_{t+1|t} = (F - K_t)P_{t|t-1}(F' - K_t') + K_t R K_t' + Q \\ &\rightarrow P_1^* \\ K_t &\equiv F P_{t|t-1} (P_{t|t-1} + R)^{-1} \rightarrow K^* \\ P_{t|t} &\equiv E\left[\left(\xi_t - \hat{\xi}_{t|t}\right)\left(\xi_t - \hat{\xi}_{t|t}\right)'\right] \\ &= P_{t|t-1} - P_{t|t-1} (P_{t|t-1} + R)^{-1} P_{t|t-1} \\ &\rightarrow P_1^* - P_1^* (P_1^* + R)^{-1} P_1^* \equiv P_0^* \neq P_1^* \end{aligned}$$

The convergence of  $P_{t|t-1}$ ,  $P_{t|t}$  and  $K_t$  to their steady-state values relies on  $|\phi| < 1$ , see Hamilton (1994), Proposition 13.1, and we impose this in the estimation.

To initialize these matrices we use:

$$\begin{aligned} P_{1|0} &\equiv E [(\xi_t - E[\xi_t]) (\xi_t - E[\xi_t])'] \\ &= \begin{bmatrix} \frac{\sigma_\varepsilon^2}{1-\phi^2} + \sigma_u^2 & \frac{\sigma_\varepsilon^2}{1-\phi^2} \\ \frac{\sigma_\varepsilon^2}{1-\phi^2} & \frac{\sigma_\varepsilon^2}{1-\phi^2} \end{bmatrix} \\ \text{and } \hat{\xi}_{1|0} &= E[\xi_t] = [0, 0]' \end{aligned}$$

Updating of the estimates is done via

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + P_{t|t-1} (P_{t|t-1} + R)^{-1} (\tilde{y}_t - \hat{\xi}_{t|t-1}),$$

while the multi-step prediction error uses:

$$\begin{aligned} \hat{\xi}_{t+s|t} &= F^s \hat{\xi}_{t|t} \\ P_{t+s|t} &\equiv E \left[ \left( \xi_{t+s} - \hat{\xi}_{t+s|t} \right) \left( \xi_{t+s} - \hat{\xi}_{t+s|t} \right)' \right] \\ &= F^s P_{t|t} (F')^s + F^{s-1} Q (F')^{s-1} + \dots + F^{s-2} Q (F')^{s-2} \dots \\ &\quad + F Q F' + Q \\ &= F^s P_{t|t} (F')^s + \sum_{j=0}^{s-1} F^j Q (F')^j \\ &\rightarrow P_s^*, \text{ for } s \geq 1. \end{aligned}$$

The “current year” forecasts of require “smoothed” estimates of current and past GDP growth and inflation, sometimes known as “nowcasts” and “backcasts”. The smoothed estimates are obtained from:

$$\begin{aligned} \hat{\xi}_{t|T} &= \hat{\xi}_{t|t} + J_t \left( \hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t} \right) \\ \text{where } J_t &\equiv P_{t|t} F' P_{t+1|t}^{-1} \rightarrow P_0^* F' (P_1^*)^{-1} \equiv J^* \\ P_{t|T} &= P_{t|t} + J_t (P_{t+1|T} - P_{t+1|t}) J_t' \\ &\rightarrow P_0^* + J^* (P_{t+1-T}^* - P_1^*) (J^*)' \\ &\equiv P_{t-T}^*, \text{ for } t \leq T \end{aligned}$$



where

$$\begin{aligned}
P_{T-1|T} &= P_{T-1|T-1} + J_{T-1} (P_{T|T} - P_{T|T-1}) J'_{T-1} \\
&\rightarrow P_0^* + J^* (P_0^* - P_1^*) (J^*)' \equiv P_{-1}^* \\
P_{T-2|T} &= P_{T-2|T-2} + J_{T-2} (P_{T-1|T} - P_{T-1|T-2}) J'_{T-2} \\
&\rightarrow P_0^* + J^* (P_{-1}^* - P_1^*) (J^*)' \equiv P_{-2}^* \\
\text{so } P_{T-3|T} &\rightarrow P_{-3}^* \\
&\equiv P_0^* + J^* (P_{-2}^* - P_1^*) (J^*)' \text{ etc.}
\end{aligned}$$

Using these results, forecasts of  $z_t$  can now be computed from

$$\hat{z}_{t,t-h} \equiv E \left[ z_t | \tilde{\mathcal{F}}_{t-h} \right] = \sum_{j=0}^{11} E \left[ y_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] \quad (26)$$

For horizons  $h \geq 12$  these predictions only involve forecasts (no “nowcasts” or “backcasts”) and these are relatively straight-forward to handle. To illustrate, consider the  $h = 12$  case in particular. We will look at forecasting the entire state vector,  $\xi_t$ , and then just focus on the (1, 1) element of the MSE matrix, which corresponds to the MSE of the prediction of  $z_t$ :

$$\begin{aligned}
E \left[ \xi_{t-j} | \tilde{\mathcal{F}}_{t-h} \right] &\equiv \hat{\xi}_{t-j|t-h} \\
\xi_{t-11} - \hat{\xi}_{t-11|t-12} &= F \left( \xi_{t-12} - \hat{\xi}_{t-12|t-12} \right) + v_{t-11} \\
\xi_{t-10} - \hat{\xi}_{t-10|t-12} &= F^2 \left( \xi_{t-12} - \hat{\xi}_{t-12|t-12} \right) + F v_{t-11} + v_{t-10} \\
&\dots \\
\xi_{t-1} - \hat{\xi}_{t-1|t-12} &= F^{11} \left( \xi_{t-12} - \hat{\xi}_{t-12|t-12} \right) + F^{10} v_{t-11} + F^9 v_{t-10} + \dots + F v_{t-2} + v_{t-1} \\
\xi_t - \hat{\xi}_{t|t-12} &= F^{12} \left( \xi_{t-12} - \hat{\xi}_{t-12|t-12} \right) + F^{11} v_{t-11} + F^{10} v_{t-10} + \dots \\
&\quad + F^2 v_{t-2} + F v_{t-1} + v_t \\
\text{so } \sum_{j=0}^{11} \left( \xi_{t-j} - \hat{\xi}_{t-j|t-12} \right) &= \left( \sum_{j=0}^{11} F^{j+1} \right) \left( \xi_{t-12} - \hat{\xi}_{t-12|t-12} \right) \\
&\quad + \left( \sum_{j=0}^{11} F^j \right) v_{t-11} + \left( \sum_{j=0}^{10} F^j \right) v_{t-10} + \left( \sum_{j=0}^9 F^j \right) v_{t-9} \dots \\
&\quad + \left( \sum_{j=0}^2 F^j \right) v_{t-2} + \left( \sum_{j=0}^1 F^j \right) v_{t-1} + v_t
\end{aligned}$$

Define  $F^{(k)} \equiv \sum_{j=0}^k F^j$ , then

$$\begin{aligned} \sum_{j=0}^{11} \left( \xi_{t-j} - \hat{\xi}_{t-j|t-12} \right) &= FF^{(11)} \left( \xi_{t-12} - \hat{\xi}_{t-12|t-12} \right) + \sum_{k=0}^{11} F^{(k)} v_{t-k}, \\ Cov \left[ \sum_{j=0}^{11} \left( \xi_{t-j} - \hat{\xi}_{t-j|t-12} \right) \right] &= FF^{(11)} P_0^* \left( FF^{(11)} \right)' + \sum_{k=0}^{11} F^{(k)} Q \left( F^{(k)} \right)'. \end{aligned}$$

Similarly, it can be shown that for  $h > 12$ , we have

$$\begin{aligned} Cov \left[ \sum_{j=0}^{11} \left( \xi_{t-j} - \hat{\xi}_{t-j|t-h} \right) \right] &= F^{h-11} F^{(11)} P_0^* \left( F^{h-11} F^{(11)} \right)' \\ &\quad + \sum_{k=0}^{11} F^{(k)} Q \left( F^{(k)} \right)' + \sum_{k=1}^{h-12} F^k F^{(11)} Q \left( F^k F^{(11)} \right)'. \quad (27) \end{aligned}$$

The first term arises from being unable to observe  $\xi_{t-h}$ : if  $\xi_{t-h}$  were observable without error then  $P_0^* = 0$  and this term would vanish. The second term is the combined impact of predicting  $\xi_{t-j}$  over the measurement period ( $j = t - 11, t - 10, \dots, t$ ). This will be larger or smaller depending on how predictable the series is, which is determined completely by  $(\sigma_u^2, \sigma_\varepsilon^2, \phi)$ . The third term is the combined impact of having to predict the intermediate values of  $\xi_{t-j}$ , between the current time,  $j = t - h$ , and the period just before start of the measurement period,  $j = t - 12$ . If  $h = 12$  then this term drops out, as there are no intermediate values to predict.

For horizons less than one year the prediction error will involve a combination of “backcast errors”, “nowcast error” and forecast errors. Each of these can be worked out in closed-form, though it is not trivial to combine each of these terms to form a final simple expression for any  $h < 12$ . We thus present the prediction error for general  $h < 12$  as a combination of backcast, nowcast and forecast errors, and then combine these terms in the estimation step:

$$z_t - \hat{z}_{t,t-h} = \sum_{j=h+1}^{11} \left( \xi_{t-j} - \xi_{t-j|t-h} \right) + \left( \xi_{t-h} - \xi_{t-h|t-h} \right) + \sum_{j=0}^{h-1} \left( \xi_{t-j} - \xi_{t-j|t-h} \right).$$

All prediction errors for  $h < 12$  can be expressed as a function of the nowcast error at time  $t - 11$  and the shocks to the system between time  $t - 10$  and time  $t$  inclusive. These shocks are the  $v_t$  and  $w_t$  terms which are *iid* and so we end up with an expression containing quantities that are uncorrelated, facilitating computation of the MSE.

Forecast errors take the general form

$$\xi_{t-j} - \xi_{t-j|t-h} = F^{h-j} \left( \xi_{t-h} - \xi_{t-h|t-h} \right) + \sum_{k=0}^{h-1} F^k v_{t-k}, \quad j \leq h, \quad (28)$$

while the general form for the backcast errors is:

$$\begin{aligned} \xi_{t-j} - \xi_{t-j|t-h} &= F^{h-j} \left( \xi_{t-j} - \xi_{t-j|t-j} \right) - \sum_{k=0}^{h-1} (J^*)^k \left( \hat{\xi}_{t-j+k|t-j+k} - \hat{\xi}_{t-j+k|t-j+k-1} \right), \quad j \geq h \\ &= F^{h-j} \left( \xi_{t-j} - \xi_{t-j|t-j} \right) + \sum_{k=0}^{h-1} (J^*)^k \left( \xi_{t-j+k} - \hat{\xi}_{t-j+k|t-j+k} \right) \\ &\quad - \sum_{k=0}^{h-1} (J^*)^k \left( \xi_{t-j+k} - \hat{\xi}_{t-j+k|t-j+k-1} \right). \end{aligned} \quad (29)$$

We also need a rule for expressing the current nowcast error as a function of previous nowcast errors plus the intervening shocks:

$$\begin{aligned} \xi_t - \hat{\xi}_{t|t} &= \left( I - P_1^* H (H' P_1^* H + R)^{-1} H' \right) F \left( \xi_{t-1} - \hat{\xi}_{t-1|t-1} \right) \\ &\quad + \left( I - P_1^* H (H' P_1^* H + R)^{-1} H' \right) v_t \\ &\quad - P_1^* H (H' P_1^* H + R)^{-1} w_t \\ &\equiv A \left( \xi_{t-1} - \hat{\xi}_{t-1|t-1} \right) + B v_t + C w_t. \end{aligned}$$

Iterating backwards, we have

$$\xi_t - \hat{\xi}_{t|t} = A^k \left( \xi_{t-k} - \hat{\xi}_{t-k|t-k} \right) + \sum_{j=0}^{k-1} A^j (B v_{t-j} + C w_{t-j}), \quad \text{for } k \geq 0 \quad (30)$$

The equivalent expression for a one-step forecast error is similarly obtained:

$$\xi_t - \hat{\xi}_{t|t-1} = F A^{k-1} \left( \xi_{t-k} - \hat{\xi}_{t-k|t-k} \right) + v_t + F \sum_{j=1}^{k-1} A^{j-1} (B v_{t-j} + C w_{t-j}), \quad \text{for } k \geq 1. \quad (31)$$

Finally, we express each prediction error as a function of the nowcast error at time  $t - 11$  and the intervening shocks. The forecast errors take the form:

$$\begin{aligned} \xi_{t-j} - \xi_{t-j|t-h} &= F^{h-j} \left( \xi_{t-h} - \xi_{t-h|t-h} \right) + \sum_{k=0}^{h-1} F^k v_{t-k}, \quad \text{for } j \leq h \\ &= \left( \sum_{k=j}^{h-1} F^{k-j} v_{t-k} \right) \\ &\quad + F^{h-j} \left\{ A^{11-j} \left( \xi_{t-11} - \xi_{t-11|t-11} \right) + \sum_{s=0}^{10-h} A^s (B v_{t-h-s} + C w_{t-h-s}) \right\}. \end{aligned}$$

Then the backcast errors:

$$\begin{aligned}
\xi_{t-j} - \xi_{t-j|t-h} &= F^{h-j} \left( \xi_{t-j} - \xi_{t-j|t-j} \right) - \sum_{k=0}^{h-1} (J^*)^k \left( \hat{\xi}_{t-j+k|t-j+k} - \hat{\xi}_{t-j+k|t-j+k-1} \right), \text{ for } j \geq h \\
&= A^{11-j} \left( \xi_{t-11} - \xi_{t-11|t-11} \right) \\
&\quad + \sum_{s=0}^{10-h} A^s (Bv_{t-j-s} + Cw_{t-j-s}) \\
&\quad - \sum_{i=1}^{j-h} (J^*)^i \left\{ v_{t-j+i} + FA^{10-j+i} \left( \xi_{t-11} - \xi_{t-11|t-11} \right) \right\} \\
&\quad - \sum_{i=1}^{j-h} (J^*)^i \left\{ F \sum_{s=1}^{10-j+i} A^{s-1} (Bv_{t-j+i-s} + Cw_{t-j+i-s}) \right\} \\
&\quad + \sum_{i=1}^{j-h} (J^*)^i \left\{ A^{11-j+i} \left( \xi_{t-11} - \xi_{t-11|t-11} \right) + \sum_{s=0}^{10-j+i} A^s (Bv_{t-j+i-s} + Cw_{t-j+i-s}) \right\}.
\end{aligned}$$

Each of these expressions is a function solely of the nowcast error at time  $t-11$ ,  $\left( \xi_{t-11} - \hat{\xi}_{t-11|t-11} \right)$ , and the shocks to the system between time  $t-10$  and  $t$ ,  $\left\{ \left[ v'_{t-j}, w'_{t-j} \right]', j = 0, 1, \dots, 10 \right\}$ . All of these elements are independent of each other and so the MSE for  $h < 12$  is directly obtained from the above expressions.

## 7.2 Details on the Kalman filter model for dispersion

The state equations for the individual forecaster are

$$\begin{bmatrix} y_t \\ x_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} 0 & \phi & 0 \\ 0 & \phi & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} u_t + \varepsilon_t \\ \varepsilon_t \\ 0 \end{bmatrix}, \quad (32)$$

while the measurement equations are

$$\begin{bmatrix} \tilde{y}_{it} \\ \tilde{y}_{it-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \ddot{\eta}_t + \nu_{it} \\ \eta_{t-1} \end{bmatrix}. \quad (33)$$

Ideally, we would have  $\ddot{\eta}_t = \eta_t$ , however that would lead to a violation that the innovation to the measurement equation is *iid* through time, since the first lag of the first element of the innovation vector would be correlated with the current value of the second element. To avoid this we specify  $\ddot{\eta}_t$  as an independent random variable, but with the same variance as  $\eta_t$ ,  $E[\ddot{\eta}_t^2] = E[\eta_t^2] = \sigma_\eta^2$ .

The MSE of the individual forecaster's prediction is obtained with just minor adjustments to the expressions presented above for the consensus forecast, and thus can be obtained in closed-form. The cross-sectional dispersion about the consensus forecast  $\bar{z}_{t,t-h} \equiv \frac{1}{N} \sum_{i=1}^N \hat{z}_{i,t,t-h}$  is computed as

$$d_{t,t-h}^2 \equiv \frac{1}{N} \sum_{i=1}^N (\hat{z}_{i,t,t-h} - \bar{z}_{t,t-h})^2.$$

$$\text{Let } \delta_h^2 \equiv \frac{1}{N} \sum_{i=1}^N E \left[ (\hat{z}_{i,t,t-h} - \bar{z}_{t,t-h})^2 \right]$$

Our model for dispersion is based on  $\delta_h^2$ . Unfortunately, a closed-form expression for  $\delta_h^2$  is not available and so we resort to simulations to evaluate  $\delta_h^2$ . We do this by simulating the state variables for  $T$  observations, and then generating a different  $\tilde{y}_{it}$  series for each of the  $N$  forecasters. We assume that the forecasters' priors,  $\mu_i$ , are *iid*  $N(0, \sigma_\mu^2)$  in the simulation. For each forecaster we obtain the optimal KF forecast and then combine this with the forecaster's prior to obtain his final forecast using equation (15). We then compute the cross-sectional variance of the individual forecasts to obtain  $d_{t,t-h}^2$  and average these across time to obtain  $\hat{\delta}_h^2$ .

### 7.3 Details of the estimation of the models

To obtain  $P_1^*$ ,  $P_0^*, \dots, P_{-11}^*$ ,  $K^*$  and  $J^*$  we simulate 100 non-overlapping years of data and update these matrices following Hamilton (1994). We use as estimates these matrices at the end of the 100<sup>th</sup> year.

We use only six forecast horizons ( $h = 1, 3, 6, 12, 18, 24$ ) in the estimation, rather than the full set of 24, in response to studies of the finite-sample properties of GMM estimates (Tauchen, 1986) which find that using many more moment conditions than required for identification leads to poor approximations from the asymptotic theory, particularly when the moments are highly correlated, as in our application.

To obtain the covariance matrix of the moments, used to compute standard errors and the test of over-identifying restrictions, we use the model-implied covariance matrix of the moments, based on the parameter estimate from the first-stage GMM parameter estimate. This matrix is not available in closed-form and so we simulate 1,000 non-overlapping years of data to estimate it. This is done as follows:

First, we simulate  $\varepsilon_t$  as an *iid*  $N(0, \sigma_\varepsilon^2)$  time series of length 12,024 periods. The  $x_t$  process

is then computed with  $x_0$  set equal to its unconditional mean. We simulate  $u_t$  as an *iid*  $N(0, \sigma_u^2)$  time series of the same length as  $\varepsilon_t$ .  $y_t$  is then computed as  $x_t + u_t$ , and  $z_t$  is the rolling 12-period sum of  $y_t$ , starting at  $t = 12$ . Next we compute the time series of forecasts: For each period, starting at  $t = 1$ , we compute the optimal forecasts of  $z_t$  for all horizons between one and 24 periods, using the formulas given in part (1) of Proposition 1 and in Section 7.1. We then drop the first 24 observations, so that every retained value of  $z_t$  has associated with it a full set of 24 forecasts, ranging from  $t - 1$  to  $t - 24$ . Next we create a  $12000 \times 25$  matrix of data, containing the simulated true value of  $z_t$  in the first column, and the 1- through 24-steps-ahead forecasts of that  $z_t$  in the remaining 24 columns. To match the sampling frequency of our data, we create a new matrix containing only every twelfth row of the previous matrix. This new matrix is then  $1000 \times 25$ . Finally, for each time period in the reduced matrix, we compute the ‘term structure’ of squared forecast errors,  $(z_t - \hat{z}_{t,t-1}^*)^2, \dots, (z_t - \hat{z}_{t,t-24}^*)^2$ . We then compute the matrix of moment conditions, using the expression in equation (8), and from this we estimate the covariance matrix of the moments for horizons  $h = 1, 3, 6, 12, 18, 24$ .<sup>18</sup>

Because a closed-form expression for  $\delta_h^2$  is not available, we use simulations to obtain an estimate of  $\delta_h^2$ . We simulated 50 non-overlapping years of data for 30 forecasters to estimate  $\delta_h^2$  for each value of the parameter vector.<sup>19</sup> The priors for each of the 30 forecasters,  $\mu_i$ , were simulated as *iid*  $N(0, \sigma_\mu^2)$ . The only difference in the dispersion simulation is that we must simulate the residual term,  $\lambda_{t,t-h}$ . We multiply the estimated  $\delta_h^2$  series by  $\lambda_{t,t-h}$ , defined in equation (16), which is *iid*  $\log N(-\frac{1}{2}\sigma_\lambda^2, \sigma_\lambda^2)$ . From this, we obtain ‘measured’ values of dispersion,  $d_{t,t-h} = \hat{\delta}_h^2 \cdot \lambda_{t,t-h}$ , and the squared dispersion residual,  $\lambda_{t,t-h}^2$ , which are used in the second six and final six moment conditions respectively. From these, combined with the MSEs, we compute the sample covariance matrix of the moments.

The model with time-varying dispersion was estimated in a similar way, with the following changes. We used the stationary bootstrap of Politis and Romano (1994), with average block length of 12 months, to “stretch” the default spread time series,  $S_t$ , to be 50 years in length for the simulation. The “standardized priors” for each of the 30 forecasters,  $\mu_i^*$ , were simulated as

---

<sup>18</sup>We examined the sensitivity of this estimate to changes in the size of the simulation and to re-simulating the model, and found that when 1000 non-overlapping years of data are used the changes in the estimated covariance matrix are negligible.

<sup>19</sup>Simulation variability for this choice of  $N$  and  $T$  was small, particularly relative to the values of the time-series variation in  $d_{t,t-h}^2$  that we observed in the data.

*iid*  $N(0, 1)$ , and then the actual “prior” for each time period,  $\mu_{i,t}$ , was set as  $\mu_i^* \times \sigma_{\mu,t}$ , where  $\sigma_{\mu,t} = \exp\{(\beta_0^\mu + \beta_1^\mu \log S_t)/2\}$ . Following this step the remainder of the simulation was the same as for the constant dispersion case above. In the estimation stage we need to compute the value of  $\delta_h^2(\sigma_{\mu,t})$ , which is simply the sample mean of  $d_{i,t-h}^2$  in the constant dispersion model, so that we can compute the dispersion residual. It was not computationally feasible to simulate  $\delta_h^2(\sigma_{\mu,t})$  for each unique value of  $\sigma_{\mu,t}$  in our sample, and so we estimated it for  $\sigma_{\mu,t}$  equal to its sample minimum, maximum and its  $[0.25, 0.5, 0.75]$  sample quantiles, and then used a cubic spline to interpolate this function, obtaining  $\tilde{\delta}_h^2(\sigma_{\mu,t})$ . We checked the accuracy of this approximation for values in between these nodes and the errors were very small. We then use  $\tilde{\delta}_h^2(\sigma_{\mu,t})$ , and the data, to compute the dispersion residuals and used these in the GMM estimation of the parameters of the model.

#### 7.4 Extracting estimates of the components of the target variable

For concreteness, we will consider estimates based on the first row in our panel, so the annual target variable is  $z_{25}$ , and the first forecast is  $\hat{z}_{25,1}$ . Let

$$\begin{aligned}
 w_{25} &\equiv \sum_{j=0}^{11} \xi_{25-j} = \begin{bmatrix} \sum_{j=0}^{11} y_{25-j} \\ \sum_{j=0}^{11} x_{25-j} \end{bmatrix} \\
 \text{then } \hat{w}_{25,1} &= \sum_{j=0}^{11} \hat{\xi}_{25-j,1} = \left( \sum_{j=0}^{11} F^{24-j} \right) \hat{\xi}_{1,1} \equiv F^{13} F^{(11)} \hat{\xi}_{1,1} \\
 \text{where } F^{(k)} &\equiv \sum_{j=0}^k F^j \\
 &= \begin{bmatrix} 1 & \frac{\phi(1-\phi^k)}{1-\phi} \\ 0 & \frac{1-\phi^{k+1}}{1-\phi} \end{bmatrix}, \text{ since } F = \begin{bmatrix} 0 & \phi \\ 0 & \phi \end{bmatrix} \\
 \text{and } F^k &= \begin{bmatrix} 0^k & \phi^k - 0^k \\ 0^{1+k} & \phi^k - 0^{1+k} \end{bmatrix} \\
 \text{So } F^j F^{(k)} &= \begin{bmatrix} 0^j & \frac{0^j \phi + \phi^j - \phi^{1+j+k} - 0^j}{1-\phi} \\ 0^{1+j} & \frac{0^{1+j} \phi + \phi^j - 0^{1+j} - \phi^{1+j+k}}{1-\phi} \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
\text{Thus } F^{13}F^{(11)}\hat{\xi}_{1,1} &= \begin{bmatrix} 0 & \frac{\phi^{13}(1-\phi^{12})}{1-\phi} \\ 0 & \frac{\phi^{13}(1-\phi^{12})}{1-\phi} \end{bmatrix} \hat{\xi}_{1,1} \\
\text{let } e'_1 &\equiv [1, 0] \\
\hat{z}_{25,1} &= e'_1 \hat{w}_{25,1} \\
&= e'_1 F^{13}F^{(11)}\hat{\xi}_{1,1} \\
&= \frac{\phi^{13}(1-\phi^{12})}{1-\phi} \hat{\xi}_{1,1}^{[2]} \\
&= \frac{\phi^{13}(1-\phi^{12})}{1-\phi} E[x_1 | \tilde{\mathcal{F}}_1] \\
\text{so } E[x_1 | \tilde{\mathcal{F}}_1] &= \frac{1-\phi}{\phi^{13}(1-\phi^{12})} \cdot \hat{z}_{25,1}
\end{aligned}$$

Since the 24-month forecast is proportional to the “nowcast” of the predictable component, with the proportionality constant being a simple function of the parameter of the data generating process (DGP), we can back out the forecaster’s “nowcast” of the predictable component from the forecast. This same steps hold for all “long horizon” forecasts:

$$\begin{aligned}
\hat{w}_{25,25-h} &= F^{h-11}F^{(11)}\hat{\xi}_{25-h,25-h}, \text{ for } h \geq 12 \\
F^{h-11}F^{(11)} &= \begin{bmatrix} 0 & \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi} \\ 0 & \frac{\phi^{h-11}(1-\phi^{12})}{1-\phi} \end{bmatrix}, \text{ for } h \geq 12 \\
\text{and } E[x_{25-h} | \tilde{\mathcal{F}}_{25-h}] &= \frac{1-\phi}{\phi^{h-11}(1-\phi^{12})} \cdot \hat{z}_{25,25-h}, \text{ for } h \geq 12
\end{aligned}$$

Thus using the long-horizon forecasts we can extract the filtered estimate of the predictable component of the target variable. This is, of course, available monthly, which is more frequently than data is available on GDP growth, although some inflation series are available monthly.

Interesting to note, the *only* parameter that affects our estimate of the predictable component is  $\phi$ ; the other parameters of the DGP and the parameters describing the measurement equation do not enter this expression.

To extract the filtered estimate of the unpredictable component we have to work with the short-horizon forecasts. These forecasts are a combination of pure forecasts, nowcasts and backcasts, and



are a bit trickier to handle. Consider the  $h = 11$  case:

$$\begin{aligned}\hat{w}_{25,14} &= \sum_{j=0}^{11} \hat{\xi}_{25-j,14} = \left( \sum_{j=0}^{11} F^{11-j} \right) \hat{\xi}_{14,14} \equiv F^{(11)} \hat{\xi}_{14,14} \\ \text{and } \hat{z}_{25,14} &= e'_1 \hat{w}_{25,14} = e'_1 F^{(11)} \hat{\xi}_{14,14} \\ &= \begin{bmatrix} 1 & \frac{\phi(1-\phi^{11})}{1-\phi} \end{bmatrix} \hat{\xi}_{14,14} \\ &= E \left[ y_{14} | \tilde{\mathcal{F}}_{14} \right] + \frac{\phi(1-\phi^{11})}{1-\phi} E \left[ x_{14} | \tilde{\mathcal{F}}_{14} \right]\end{aligned}$$

We have an estimate of the second term above from the long-horizon forecast of the following year's annual target variable (the  $h = 23$  forecast for the following year)

$$E \left[ x_{14} | \tilde{\mathcal{F}}_{14} \right] = \frac{1-\phi}{\phi^{12}(1-\phi^{12})} \cdot \hat{z}_{37,14}$$

and so we can combine this with the short-horizon forecast of this year's annual target variable to back out an estimate of the unpredictable component:

$$\begin{aligned}E \left[ y_{14} | \tilde{\mathcal{F}}_{14} \right] &= \hat{z}_{25,14} - \frac{\phi(1-\phi^{11})}{\phi^{12}(1-\phi^{12})} \cdot \hat{z}_{37,14} \\ \text{and thus } E \left[ u_{14} | \tilde{\mathcal{F}}_{14} \right] &= E \left[ y_{14} | \tilde{\mathcal{F}}_{14} \right] - E \left[ x_{14} | \tilde{\mathcal{F}}_{14} \right] \\ &= \hat{z}_{25,14} - \frac{\phi(1-\phi^{11})}{\phi^{12}(1-\phi^{12})} \cdot \hat{z}_{37,14} - \frac{1-\phi}{\phi^{12}(1-\phi^{12})} \cdot \hat{z}_{37,14} \\ &= \hat{z}_{25,14} - \frac{\phi(1-\phi^{11}) + 1-\phi}{\phi^{12}(1-\phi^{12})} \cdot \hat{z}_{37,14}\end{aligned}$$

Next consider the  $h = 10$  case:

$$\hat{w}_{25,15} = F^{(10)} \hat{\xi}_{15,15} + \hat{\xi}_{14,14} + J \left( \hat{\xi}_{15,15} - F \hat{\xi}_{14,14} \right)$$

From the  $h = 11$  data we have  $\hat{\xi}_{14,14}$ , and from the long horizon forecast of next year's variable we have  $\hat{\xi}_{15,15}^{[2]} \equiv E \left[ x_{15} | \tilde{\mathcal{F}}_{14} \right]$ . Thus there is only one unknown on the right-hand side above, namely  $\hat{\xi}_{15,15}^{[1]} \equiv E \left[ y_{15} | \tilde{\mathcal{F}}_{14} \right]$ , which we can obtain using  $\hat{z}_{25,15}$ . The estimates of  $E \left[ y_t | \tilde{\mathcal{F}}_t \right]$  obtained from the forecasts for  $h = 9$  down to  $h = 1$  can be obtained similarly, recursively using the estimates

from the longer horizons. The general expression for  $0 < h < 12$  is:

$$\begin{aligned}
\hat{z}_{25,25-h} &= \left( F^{(h)} + JJ^{(10-h)} \right) \hat{\xi}_{25-h,25-h} \\
&+ \left( I - JF + JJ^{(10-h)} (I - F) \right) \hat{\xi}_{24-h,24-h} \\
&+ \left( I - JF + JJ^{(9-h)} (I - F) \right) \hat{\xi}_{23-h,23-h} \\
&+ \dots \\
&+ (I - JF + J(I - F)) \hat{\xi}_{15,15} \\
&+ (I - JF) \hat{\xi}_{14,14} \\
&= \left( F^{(h)} + JJ^{(10-h)} \right) \hat{\xi}_{25-h,25-h} \\
&+ \left( \sum_{j=1}^{10-h} \left( I - JF + JJ^{(j-1)} (I - F) \right) \right) \hat{\xi}_{14+j,14+j} \\
&+ (I - JF) \hat{\xi}_{14,14}
\end{aligned}$$

By working from the longer horizons down to the shorter horizons, each of these expressions will have just a single unknown variable that can be obtained by solving the expression for that variable.

Our estimate of the unpredictable component relies on the matrix  $J = P_0 F' P_1$  in steady-state. The matrices  $P_0$  and  $P_1$  depend on the variances of the innovations to the state equation and on the parameters describing the measurement equation, and so our estimates of the unpredictable component rely on the full specification of the model, unlike our estimate of the predictable component which only depended upon  $\phi$ .

Note that we can also provide standard errors on our estimates of the predictable and unpredictable components of the target variable. The matrix  $P_0$  measures the MSE of the “nowcast” of the state equation:

$$\begin{aligned}
P_0 &\equiv E \left[ \left( \xi_t - \hat{\xi}_{t,t} \right) \left( \xi_t - \hat{\xi}_{t,t} \right)' \right] \\
\text{and so } V[x_t - \hat{x}_{t,t}] &= P_0^{[2,2]} \\
V[y_t - \hat{y}_{t,t}] &= P_0^{[1,1]} \\
\text{and } u_t &= y_t - x_t \\
\text{so } V[u_t - \hat{u}_{t,t}] &= P_0^{[1,1]} + P_0^{[2,2]} - 2P_0^{[1,2]}
\end{aligned}$$

It should be noted that our model for the MSE term structure assumed, without loss of generality in that application (as the inputs to the model were simply the forecast errors), that all variables

have zero mean. This of course is not true in reality, and does have implications for our estimates of  $x_t$  and  $u_t$ . If we modify our specification of the state equation to allow for a non-zero mean we obtain:

$$\xi_t = \begin{bmatrix} y_t - \mu \\ x_t - \mu \end{bmatrix} = \begin{bmatrix} 0 & \phi \\ 0 & \phi \end{bmatrix} \begin{bmatrix} y_{t-1} - \mu \\ x_{t-1} - \mu \end{bmatrix} + \begin{bmatrix} u_t + \varepsilon_t \\ \varepsilon_t \end{bmatrix}$$

Then the expressions derived above can be re-interpreted expressions for  $E[x_t - \mu | \tilde{\mathcal{F}}_t]$  and  $E[y_t - \mu | \tilde{\mathcal{F}}_t]$ .

The forecasts would become

$$w_{25} \equiv \sum_{j=0}^{11} \xi_{25-j} = \begin{bmatrix} \sum_{j=0}^{11} (y_{25-j} - \mu) \\ \sum_{j=0}^{11} (x_{25-j} - \mu) \end{bmatrix} = \begin{bmatrix} \sum_{j=0}^{11} y_{25-j} \\ \sum_{j=0}^{11} x_{25-j} \end{bmatrix} - 12\mu$$

Thus we can simply de-mean the forecasts (using, for example, one-twelfth the sample mean of the  $z_t$  series), compute  $E[x_t | \tilde{\mathcal{F}}_t]$  and  $E[y_t | \tilde{\mathcal{F}}_t]$  as before, and then add back the means to the estimates. This corresponds to estimating the parameter  $\mu$  by GMM, using simply the sample mean of the  $z_t$  series.

## References

- [1] Andersen, T.G., T. Bollerslev, F.X. Diebold and C. Vega, 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review* 93, 38-62.
- [2] Beber, A. and M.W. Brandt, 2006, Resolving Macroeconomic Uncertainty in Stock and Bond Markets. Mimeo, HEC Lausanne and Duke University.
- [3] Carroll, C., 2003, Macroeconomic Expectations of Household and Professional Forecasters. *Quarterly Journal of Economics*, 118(1), 269-298.
- [4] Chernov, M. and P. Müller, 2007, The Term Structure of Inflation Forecasts, work-in-progress, London Business School and Columbia Business School.
- [5] Clements, M.P., 1997, Evaluating the Rationality of Fixed-Event Forecasts. *Journal of Forecasting* 16, 225-239.
- [6] Croushore, D., 2006, Forecasting with Real-Time Macroeconomic Data. Pages 961-982 in G. Elliott, C. Granger and A. Timmermann (eds.) *Handbook of Economic Forecasting*, North Holland: Amsterdam.
- [7] Croushore, D. and T. Stark, 2001, A Real-Time Data Set for Macroeconomists. *Journal of Econometrics* 105, 111-130.

- [8] Davies, A. and K. Lahiri, 1995, A New Framework for Analyzing Survey Forecasts using Three-Dimensional Panel Data, *Journal of Econometrics*, 68, 205-227.
- [9] Dixit, A.X. and R.S. Pindyck, 1994, *Investment under Uncertainty*. Princeton University Press: Princeton N.J.
- [10] Ederington, L.H. and J.H. Lee, 1996, The creation and resolution of market uncertainty: The impact of information releases on implied volatility. *Journal of Financial and Quantitative Analysis* 31, 513-539.
- [11] Faust, J., J.H. Rogers and J.H. Wright, 2005, News and Noise in G-7 GDP Announcements, *Journal of Money, Credit, and Banking*, 37(3), 403-417.
- [12] Giannone, D., L. Reichlin and D. Small, 2007, Nowcasting GDP: The Real Time Informational Content of Macroeconomic Data Releases, working paper, Université Libre de Bruxelles.
- [13] Grier, K. and M.J. Perry, 1998, On Inflation and Inflation Uncertainty in the G-7 Countries. *Journal of International Money and Finance* 17.
- [14] Hamilton, J.D., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
- [15] Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, 50, 1029-1054.
- [16] Hansen, P.R., 2005, A Test for Superior Predictive Ability. *Journal of Business and Economic Statistics* 23, 365-380.
- [17] Isiklar, G., K. Lahiri, and P. Loungani, 2006, How Quickly Do Forecasters Incorporate News? Evidence from Cross-Country Surveys, *Journal of Applied Econometrics*, 21, 703-725.
- [18] Isiklar, G. and K. Lahiri, 2007, How Far Ahead Can We Forecast? Evidence from Cross-Country Surveys, *International Journal of Forecasting*, 23, 167-187.
- [19] Kim, C-J. and C.R. Nelson, 1999, Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle, *Review of Economics and Statistics*, 81, 608-616.
- [20] Kydland, F.E. and E.C. Prescott, 1982, Time to Build and Aggregate Fluctuations. *Econometrica* 50, 1345-1370.
- [21] Lahiri, K. and X. Sheng, 2006, Evolution of Forecast Disagreement and Uncertainty in a Heterogeneous Agents Model, working paper, SUNY-Albany.
- [22] Lamont, O.A., 2002, Macroeconomic Forecasts and Microeconomic Forecasters, *Journal of Economic Behaviour and Organization*, 48, 265-280.
- [23] Laster, D., P. Bennet and I.S. Geoum, 1999, Rational Bias in Macroeconomic Forecasts, *Quarterly Journal of Economics*, 114(1), 293-318
- [24] Mahadeva, L. and A. Muscatelli, 2005, National Accounts Revisions and Output Gap Estimates in a Model of Monetary Policy with Data Uncertainty, External MPC Unit Discussion Paper 14, Bank of England.

- [25] McConnell, M.M. and G. Perez-Quiros, 2000, Output Fluctuations in the United States: What Has Changed Since the Early 1980s?, *American Economic Review*, 90, 1464-1476.
- [26] Nordhaus, W.D., 1987, Forecasting Efficiency: Concepts and Applications. *Review of Economics and Statistics* 69, 667-674.
- [27] Ottaviani, M. and P.N. Sørensen, 2006, The Strategy of Professional Forecasting, *Journal of Financial Economics*, 81(2), 441-466.
- [28] Pesaran, M.H. and M. Weale, 2006, Survey Expectations. Pages 715-776 in G. Elliott, C. Granger and A. Timmermann (eds.) *Handbook of Economic Forecasting*, North Holland: Amsterdam.
- [29] Politis, D. N., and Romano, Joseph P., 1994, The Stationary Bootstrap, *Journal of the American Statistical Association*, 89, 1303-1313.
- [30] Ramey, G. and V. Ramey, 1995, Cross-Country Evidence on the Links between Volatility and Growth. *American Economic Review* 85, 1138-51.
- [31] Stock, J.H. and M.W. Watson, 2002, Macroeconomic Forecasting using Diffusion Indexes. *Journal of Business and Economic Statistics* 20, 147-162.
- [32] Stock, J.H. and M.W. Watson, 2005, Forecasting with Many Predictors, in Graham Elliott, Clive W.J. Granger and Allan Timmermann (eds.), *Handbook of Economic Forecasting*, Elsevier.
- [33] Swidler, S. and D. Ketcher, 1990, Economic Forecasts, Rationality, and the Processing of New Information over Time. *Journal of Money, Credit, and Banking* 22, 65-76.
- [34] Tauchen, G., 1986, Statistical Properties of Generalized Method of Moments Estimators of Structural Parameters obtained from Financial Market Data, *Journal of Business and Economic Statistics*, 4, 397-416.
- [35] Van Nieuwerburgh, S. and L. Veldkamp, 2006, Learning Asymmetries in Real Business Cycles. *Journal of Monetary Economics* 53, 753-772.
- [36] Veldkamp, L., 2006, Slow Boom, Sudden Crash. *Journal of Economic Theory* 124(2), 230-257.
- [37] Veronesi, P., 1999, Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model. *Review of Financial Studies* 12(5), 975-1007.
- [38] White, H., 2000, A Reality Check for Data Snooping. *Econometrica* 68, 1097-1126.
- [39] Zarnowitz, V. and L.A. Lambros, 1987, Consensus and Uncertainty in Economic Prediction. *Journal of Political Economy* 95, 591-621.

**Table 1: Testing rationality of consensus forecasts of US GDP growth and Inflation**

Horizon	Bias		MZ $p$ -values	
	GDP growth	Inflation	GDP growth	Inflation
1	0.11	0.01	0.99	1.00
2	0.07	0.02	0.99	1.00
3	0.03	0.04**	0.99	1.00
4	0.02	0.06**	0.99	1.00
5	0.01	0.08**	0.99	1.00
6	0.05	0.08*	0.99	1.00
7	0.07	0.11*	0.98	1.00
8	0.07	0.08	0.97	1.00
9	-0.06	0.01	0.98	0.99
10	-0.16	-0.03	0.98	0.97
11	-0.26	-0.05	0.97	0.96
12	-0.34	0.03	0.95	0.96
13	-0.33	0.11	0.95	0.96
14	-0.35	0.16	0.94	0.96
15	-0.29	0.22	0.95	0.97
16	-0.14	0.28	0.95	0.97
17	-0.10	0.26	0.99	0.98
18	-0.06	0.29	0.99	0.98
19	-0.07	0.33*	0.99	0.98
20	-0.07	0.36*	0.99	0.98
21	-0.08	0.33*	0.99	0.97
22	-0.07	0.32	0.99	0.97
23	-0.07	0.35*	0.99	0.97
24	-0.09	0.37*	0.98	0.98

Notes: \*\* and \* indicate that the bias is significant at the 1% and 5% levels respectively, based on bootstrap standard errors. The first two columns report the average bias in the forecast, for each variable and each horizon, which should be zero for a rational forecast. The final two columns give the bootstrap  $p$ -values from a joint test that  $\beta_0^h = 0 \cap \beta_1^h = 1$  in the Mincer-Zarnowitz regression of the realized value of the target variable on the forecast:  $y_t = \beta_0^h + \beta_1^h \hat{y}_{t,t-h} + \xi_{t,t-h}$ , for each horizon  $h$ .

**Table 2: GMM parameter estimates of the consensus forecast model**

	$\sigma_u$	$\sigma_\varepsilon$	$\phi$	$\sigma_\eta$	$J$ p-val
GDP growth	0.063 (0.012)	0.054 (0.013)	0.936 (0.034)	0.126 (--)	0.403
Inflation	0.000 (--)	0.023 (0.007)	0.953 (0.047)	0.000 (--)	0.935

Notes: The table reports the GMM estimates of the parameters of the Kalman filter model fitted to the consensus forecasts with standard errors in parentheses.  $p$ -values from the test of over-identifying restrictions are given in the row titled “ $J$  p-val”. The model is estimated using six moments from the MSE term structure for the consensus forecast for each country/variable. The parameter  $\sigma_\eta$  was fixed at  $2\sigma_u$  and is reported here for reference only.

**Table 3: GMM parameter estimates of the joint consensus forecast and dispersion model**

<i>Panel A: constant forecast dispersion</i>									
	$\sigma_u$	$\sigma_\varepsilon$	$\phi$	$\sigma_\eta$	$\sigma_\nu$	$\kappa$	$\sigma_\mu$	$J$ p-val	
GDP growth	0.063 (0.012)	0.054 (0.013)	0.936 (0.034)	0.126 (--)	0.692 (1.00)	1.414 (0.941)	0.672 (0.394)	0.857	
Inflation	0.000 (--)	0.023 (0.007)	0.953 (0.046)	0.000 (--)	0.045 (0.168)	0.493 (0.167)	0.509 (0.127)	0.000	
<i>Panel B: time-varying forecast dispersion</i>									
	$\sigma_u$	$\sigma_\varepsilon$	$\phi$	$\sigma_\eta$	$\sigma_\nu$	$\kappa$	$\beta_0^\mu$	$\beta_1^\mu$	$J$ p-val
GDP growth	0.063 (0.012)	0.054 (0.013)	0.936 (0.034)	0.126 (--)	0.692 (0.869)	1.413 (0.738)	-0.560 (1.10)	3.075 (1.832)	0.906
Inflation	0.000 (--)	0.023 (0.007)	0.953 (0.046)	0.000 (--)	0.044 (0.151)	0.493 (0.156)	-1.318 (0.546)	0.178 (2.371)	0.000

Notes to Table 3: This table reports GMM parameter estimates of the Kalman filter model of the consensus forecasts and forecast dispersions, with standard errors in parentheses.  $p$ -values from the test of over-identifying restrictions are given in the row titled “ $J$  p-val”. The model is estimated using six moments each from the MSE term structure for the consensus forecast and from the cross-sectional term structure of dispersion for each country/variable. The parameter  $\sigma_\eta$  was fixed at  $2\sigma_u$  and is reported here for reference only.

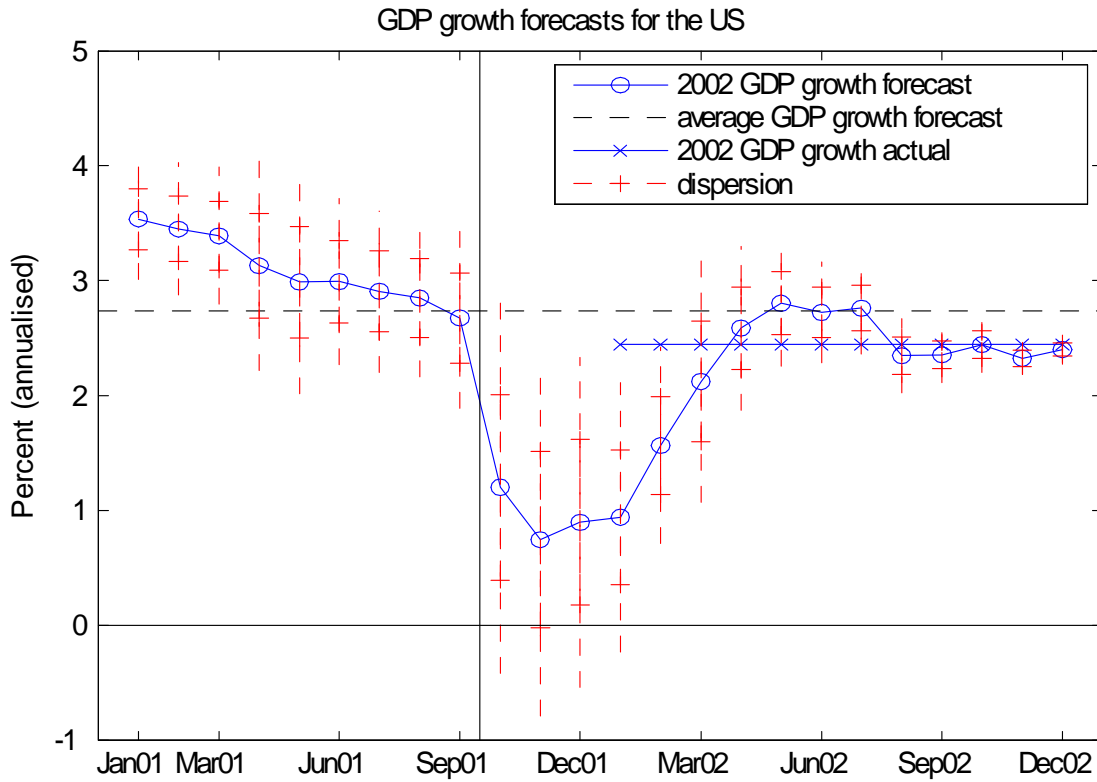


Figure 1: *Evolution in consensus forecasts and forecast dispersions for US GDP growth in 2002, for horizons ranging from 24 months (January 2001) to 24 months (December 2002).*



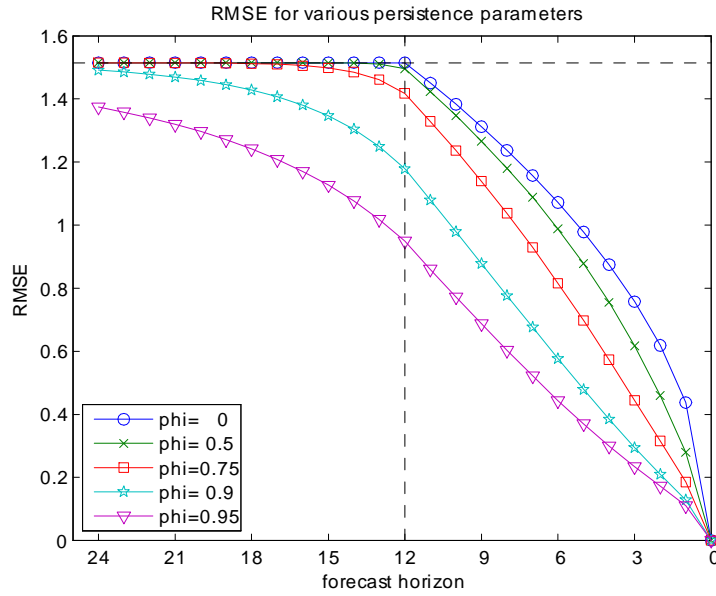


Figure 2: Term structure of root-mean squared forecast errors for various degrees of persistence ( $\phi$ ) in the predictable component

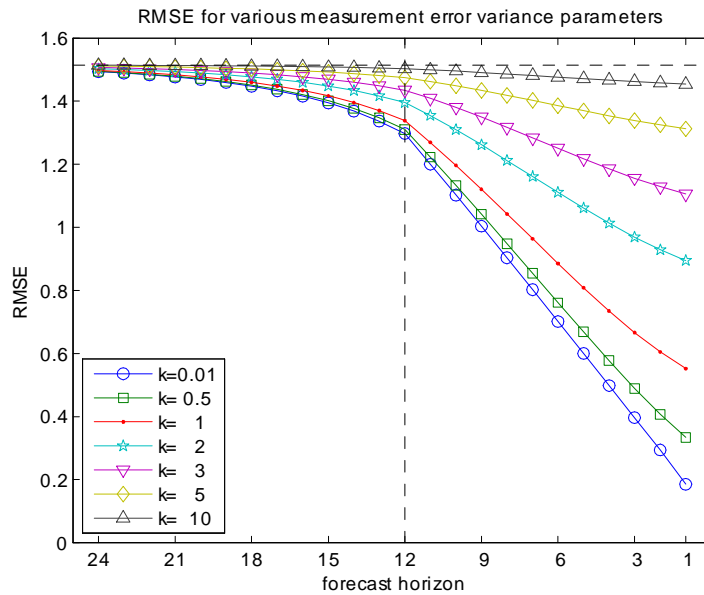


Figure 3: Term structure of root-mean squared forecast errors for various degrees of measurement error in the predicted variable. In this example, the degree of measurement error is described as  $\sigma_{\eta}^2 = k^2\sigma_u^2$ , where  $\sigma_u^2$  is the variance of the unpredictable component of  $y_t$ .

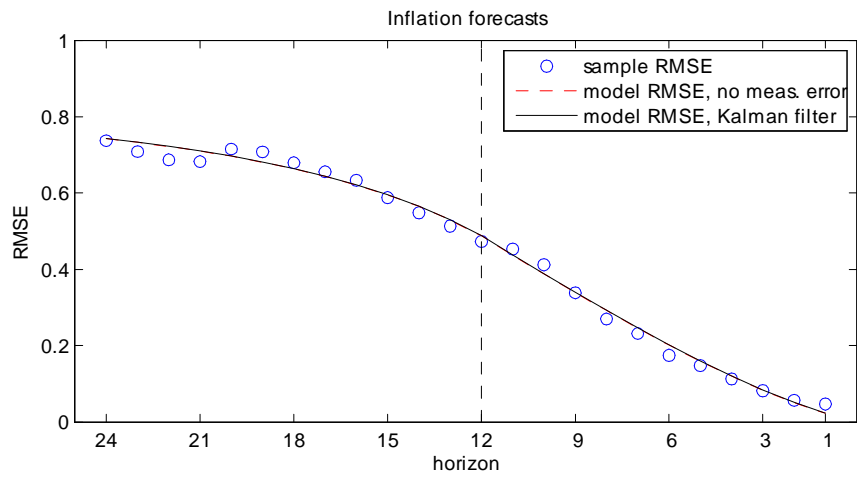
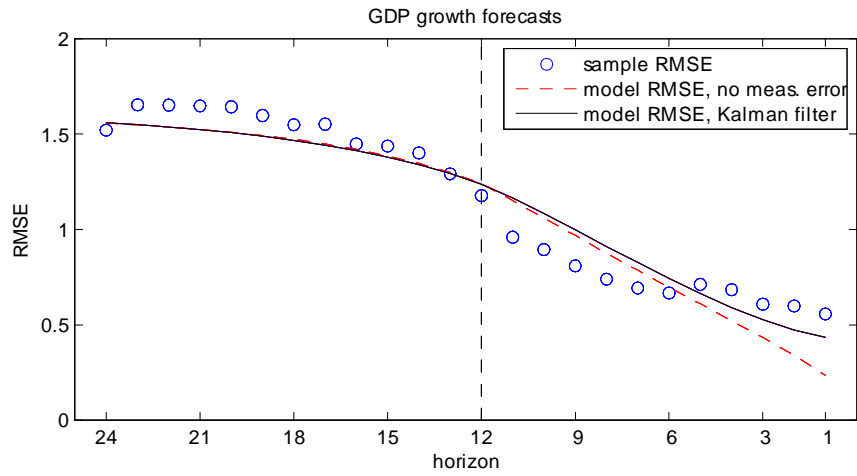


Figure 4: *Root mean squared forecast errors for GDP growth and Inflation in the U.S.*

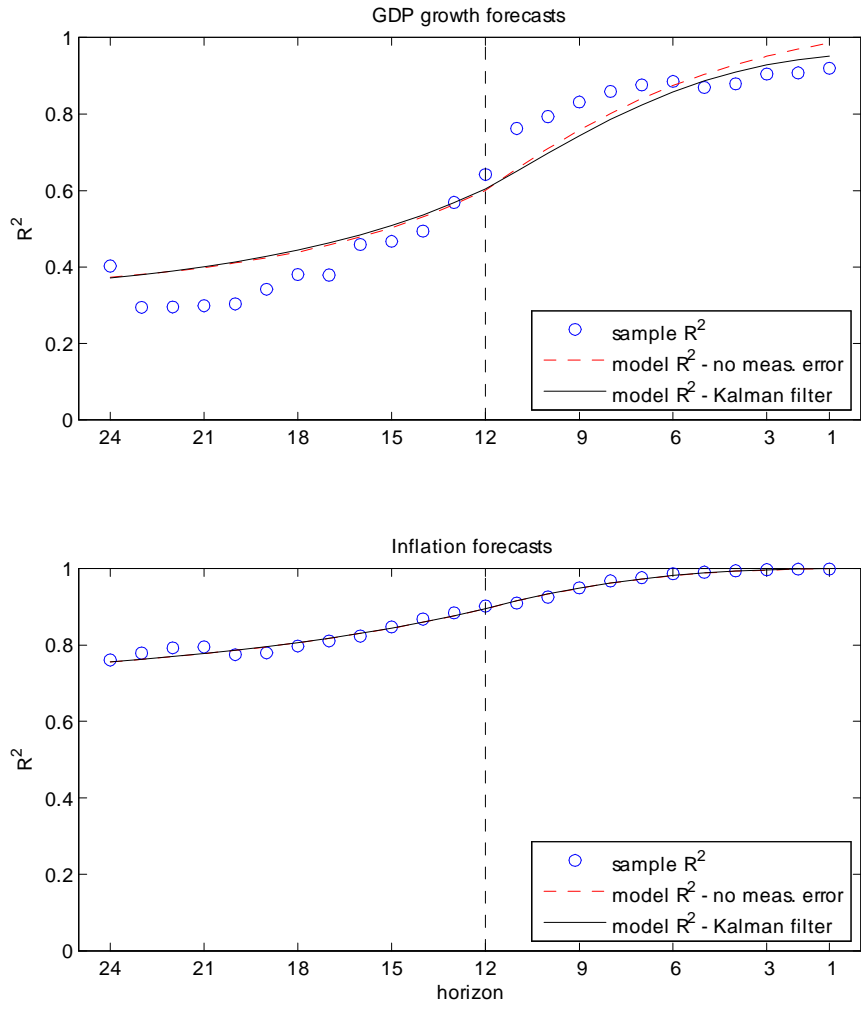


Figure 5: Empirical and model-implied  $R^2$  for forecasts of U.S. GDP growth and Inflation of various horizons.

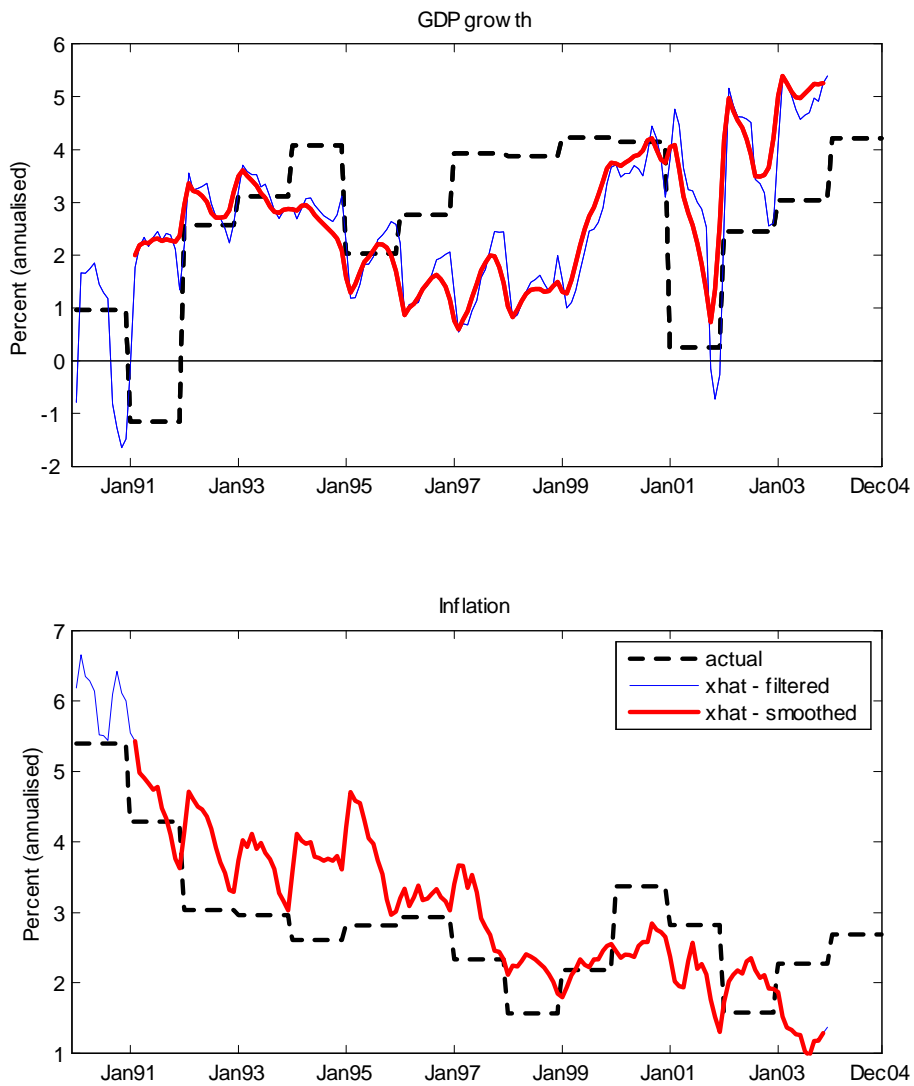


Figure 6: *Estimates of the persistent component ( $xhat$ ) of GDP growth and inflation for each month in the sample period, as implied by the observed forecasts and the estimated model for the term structure of forecast errors.*

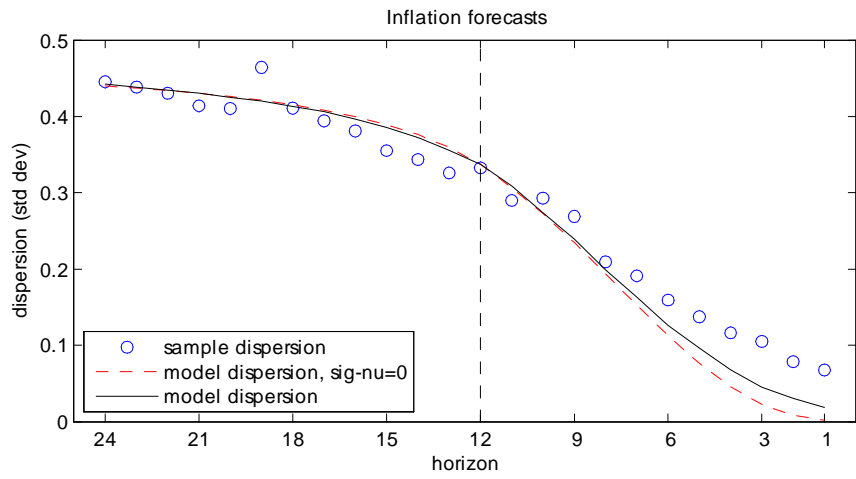
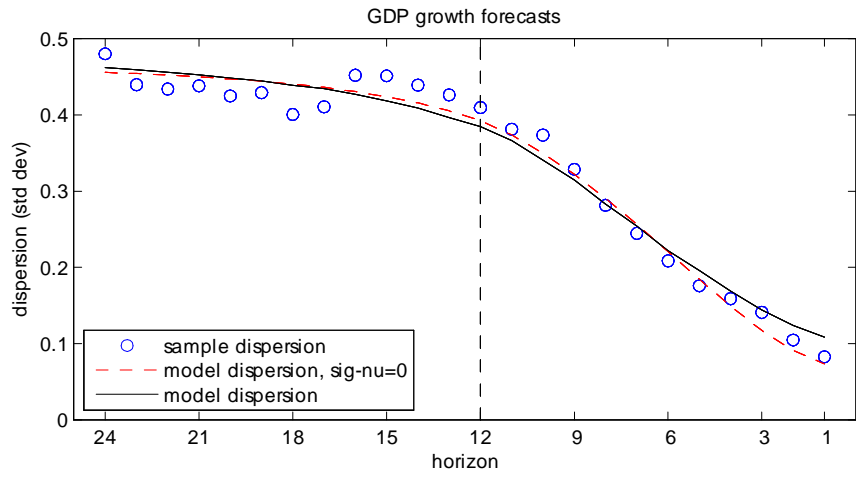


Figure 7: *Cross-sectional dispersion (standard deviation) of forecasts of GDP growth and Inflation in the U.S.*

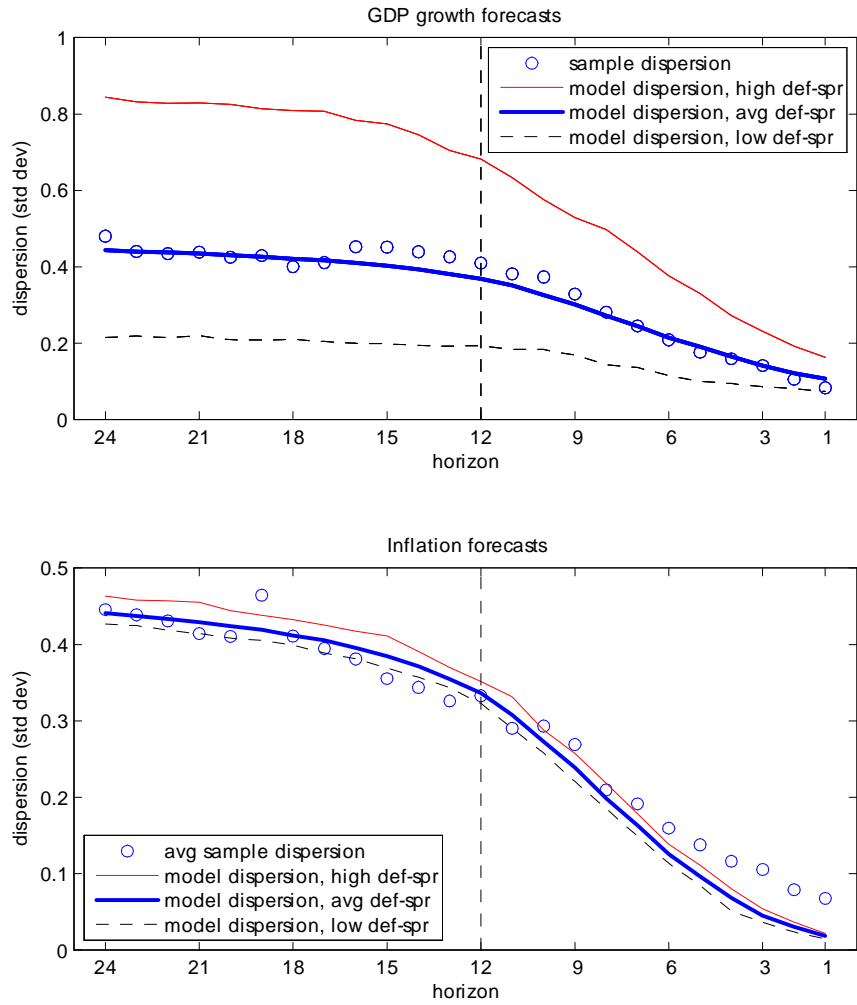


Figure 8: *Cross-sectional dispersion (standard deviation) of forecasts of GDP growth and Inflation in the U.S, when the default spread is equal to its sample average, its 95<sup>th</sup> percentile or its 5<sup>th</sup> percentile.*