Learning in Real Time: Theory and Empirical Evidence from the Term Structure of Survey Forecasts*

Andrew J. Patton
University of Oxford

Allan Timmermann
University of California San Diego

20 February 2008

Abstract

We develop an econometric framework for understanding how agents form expectations 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 and also accounts for agents' learning in real time about past, current and future values of economic variables. We use the model to develop insights into the properties of optimal forecasts across forecast horizons (the "term structure" of forecasts), and test its implications on a data set comprising survey forecasts of annual GDP growth and inflation with horizons ranging from 1 to 24 months. The model is found to closely match the term structure of forecast errors for consensus beliefs and is able to replicate the cross-sectional dispersion in forecasts of GDP growth but not for inflation - the latter appearing to be too high in the data at short horizons. Our analysis also suggests that agents systematically underestimated the persistent component of GDP growth but overestimated it for inflation during most of the 1990s.

Keywords: fixed-event forecasts, Kalman filtering, optimal updating, dispersion in beliefs.

J.E.L. Codes: E37, C53, C32.

*We thank Roy Batchelor, Steve Cecchetti, Jonas Dovern, Mike McCracken, Hashem Pesaran, Shaun Vahey, Michela Verardo, Mark Watson and seminar participants at the Board of Governors of the Federal Reserve, Cambridge, City University London, Duke, European Central Bank, London School of Economics, NBER Summer Institute, Oxford, Universite Libre Bruxelles (ECARES), Tilburg and Stanford (SITE workshop) for helpful comments and suggestions. Patton: Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, UK. Email: andrew.patton@economics.ox.ac.uk. Timmermann: Department of Economics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0553, USA. Email: atimmerm@ucsd.edu.

1 Introduction

How agents form expectations and learn about the state of the economy plays an important role in modern macroeconomic analysis. In recent remarks, Chairman of the Federal Reserve Ben Bernanke quotes academic work as showing that "the process of [the public's] learning can affect the dynamics and even the potential stability of the economy." He goes on to state that "Undoubtedly, the state of inflation expectations greatly influences actual inflation and thus the central bank's ability to achieve price stability." (Bernanke (2007)).

The rate at which uncertainty about macroeconomic variables is resolved through time is an important part of agents' learning process, due to the irreversibility and lags in many economic decisions ('time to build'; Kydland and Prescott (1982) and Dixit and Pindyck (1994)) and 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. Macroeconomic uncertainty has also been shown to be an important determinant of asset prices and volatility in financial markets.[1]

Agents' expectations and their 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". As a simple and powerful illustration, consider Figure 1, which shows 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).[2] 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

[1] Ederington and Lee (1996) and Andersen et al (2003) find that macroeconomic announcements have a significant effect on T-bill futures and exchange rates 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. Veronesi (1999) links stock prices to uncertainty about the state of the economy.

[2] The data is described in further detail in Section 3.1.

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%.

Theoretical models such as Lucas (1973) and Townsend (1983) suggest that heterogeneity in agents' beliefs affect the dynamics of the economy and uncertainty about the state of the economy is reflected not only in the size of the consensus forecast errors but also in the degree of cross-sectional dispersion in agents' beliefs. Indeed, authors such as Mankiw, Reis and Wolfers (2003, p.2) have recently suggested the possibility that ".. disagreement may be a key to macroeconomic dynamics." Cross-sectional dispersion in beliefs is thus an integral part of agents' learning process and Figure 1 shows that this measure saw similarly dramatic changes in 2001 and 2002: Prior to 9/11 the dispersion across forecasters was close to 0.7%. Dispersion then more than doubled to 1.6% in October through December of 2001, before falling back to its normal level once again.

The data in Figure 1 represents only one historical episode, but it illustrates several of the features of professional forecasters' real-time updating processes that survey data can shed light on: Rapid adjustment to news (reflected both in the consensus belief and in the cross-sectional dispersion) accompanied by decreasing forecast errors and declining cross-sectional dispersion as the forecast horizon is reduced.

Survey data such as that presented in Figure 1 have only been used to a very limited extent to shed light on agents' learning process.[3] To address this lacuna, our paper proposes a new approach for extracting information about how agents learn about the state of the economy and why they disagree. In particular, our paper develops and implements a framework for studying panels of forecasts containing numerous different forecast horizons ("large H"), and presents empirical results that shed new light on the speed at which the uncertainty surrounding realizations of macroeconomic variables is resolved. Our analysis exploits the rich information available by studying how forecasts of a variable measured at a low frequency (e.g., annual GDP growth) are updated at a higher frequency (monthly, in our case). The "large H" nature of our panel enables us to answer

[3] Exceptions include Davies and Lahiri (1995), Clements (1997), Swidler and Ketcher (1990) and Chernov and Müller (2007). The first three 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 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.

a number of interesting questions that are intractable with forecasts of just one or two different horizons.[4] We do so by modeling agents' learning problem and then matching the properties of forecasts and forecast dispersions across various horizons (the "term structures" of forecasts and dispersion) with the moments implied by our model. To our knowledge, such an approach has not been considered in the literature before.

Our analysis offers a variety of methodological contributions. From a modeling perspective, we develop a model that incorporates heterogeneity in agents' prior beliefs and information sets while accounting for measurement errors and the overlapping nature of the forecasts for various horizons. Allowing for the presence of a persistent latent component in the outcome, we employ a Kalman filter that characterizes how agents simultaneously backcast, nowcast and forecast past, current and future values of the variable of interest. This approach is made necessary by the partial overlap in forecasts recorded at different horizons. We develop a simulation-based method of moments (SMM) framework for estimating the parameters of our model in a way that accounts for agents' filtering problem. We view the shape of the mean squared forecast errors at different horizons as the objects to be fitted and use GMM or SMM estimation to account for the complex covariance patterns arising in forecasts recorded at different (overlapping) horizons. We test the validity of the model through its ability to fit the term structure of expectations under the assumption that agents use information efficiently.

Our analysis accomplishes several objectives. First, it allows us to characterize the degree of uncertainty surrounding key macroeconomic variables such as output growth and inflation−i.e. the relative importance of predictable and unpredictable components in these variables−as perceived by professional forecasters. We estimate the rate at which this uncertainty is resolved across forecast horizons and thus characterize how quickly forecasters learn. Second, we measure the size and degree of persistence in a persistent component in output growth and inflation as reflected in the joint distribution of the predicted and actual outcomes. Third, we compare the extracted estimates of agents' beliefs to observed outcomes in order to characterize the type of forecast errors agents made during the sample. Fourth, with some further structure and extensions, we identify the key sources of disagreement, i.e. heterogeneity in prior beliefs or in information sets, leading to cross-sectional dispersion in agents' predictions through the different effect that these have on dispersion

[4] Forecasts of a given variable from various forecast horizons are sometimes called "fixed event forecasts", see Nordhaus (1987) and Clements (1997) for example.

in beliefs at different horizons. Moreover, we propose a new way to model the effect of economic state variables on time-variation in the cross-sectional dispersion.

We find many interesting empirical results. During our sample, uncertainty about GDP growth at both long and short horizons was higher than uncertainty about inflation. Moreover, inflation was more persistent and less affected by measurement error and so was more predictable. Consistent with a simple model containing a persistent component in the predicted variable, uncertainty falls at a slower rate in the "next-year" forecasts (horizon$\geq$12 months) than in the "current-year" forecasts (horizon$<$12 months), following a concave pattern. Comparing the filtered estimates of the persistent component in output growth to the observed outcomes, our analysis reveals that the panel of forecasters were consistently surprised by the strong GDP growth during most of the 1990s. Conversely, the forecasters were consistently surprised by the low and declining inflation during the 1990s, with the estimated persistent component generally coming in above the realized values of inflation. In the final part of the sample the persistent component in inflation expectations is more in line with realized inflation. The estimates from our model, while closely matching the mean and autocorrelation in output growth and inflation are not able to match the very low volatility of these variables over our sample. These observations suggest that agents took some time before forecasters adjusted their views on both the level and volatility of long-run inflation, or did not believe that the lower volatility associated with the "Great Moderation" was permanent.

We also find that heterogeneity in agents' priors appear more important than heterogeneity in their information in explaining the cross-sectional dispersion in beliefs about GDP growth. Moreover, we find that GDP growth forecast dispersion has a strong and significant counter-cyclical component, whereas inflation forecast dispersion appears only weakly counter-cyclical. Finally, we find "excess dispersion" in inflation forecasts at short horizons: the disagreement between agents' predictions of inflation is high relative both to the prediction of our model, and relative to the objective degree of uncertainty about inflation at horizons of less than 6 months.

The plan of the paper is as follows. Section 2 presents our econometric framework for modelling how rational forecasters update their predictions as the forecast horizon shrinks, and presents an extension of the model to explain cross-sectional dispersion among forecasters, by allowing for heterogeneity in agents' information and their prior beliefs. Section 3 presents our empirical results on the consensus forecasts, using data from Consensus Economics over the period 1990-2004. Section 4 presents our empirical results for the model of cross-sectional dispersion, and Section 5 concludes. Technical derivations and details are provided in appendices.

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 makes use of the rich information available in high frequency revisions of forecasts of a lower frequency variable, e.g., monthly revisions of forecasts of annual GDP growth. 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 additional complications, and also means that the timing of the forecasts has to be carefully considered.

Our analysis assumes that agents choose their forecasts to minimize the mean squared forecast error, $e_{t,t-h} \equiv z_t - \hat{z}_{t,t-h}$, 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. Under this loss function, the optimal $h-$period forecast is simply the conditional expectation of z_t given information at time $t-h, \mathcal{F}_{t-h}$:[5]

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

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, but this need not

[5] 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 infrequently, see Carroll (2003).

imply that forecasters are getting better if, as is widely believed, the volatility of US output growth has also come down (Kim and Nelson (1999) and McConnell and Perez-Quiros (2000)).

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.

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 :

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

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:

$$\begin{aligned} y_t &= x_t + u_t \\ x_t &= \phi x_{t-1} + \varepsilon_t, \qquad -1 < \phi < 1 \\ u_t &\sim iid\ (0, \sigma_u^2), \\ \varepsilon_t &\sim iid\ (0, \sigma_\varepsilon^2) \\ E[u_t \varepsilon_s] &= 0\ \forall\ t, s. \end{aligned} \tag{3}$$

ϕ measures the persistence of x_t, while u_t and ε_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 $\mathcal{F}_t = \sigma\left(\left[x_{t-j}, y_{t-j}\right], j = 0, 1, 2, ...\right)$. Less parametric approaches are possible but are unlikely to be empirically successful given the relatively short time series typically available for survey data.

Our approach of using survey expectations to shed light on the process whereby agents learn and form expectations in real time can be compared with a more structural approach (e.g. Primiceri (2006)) which extracts expectations from a dynamic model of the economy. 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 (2006)), and the lack of information about which models individual agents actually use. While our analysis makes some simplifying assumptions, it does not otherwise require making strong assumptions about the structure of the economy.

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. Furthermore, the simplicity of the benchmark model allows a complete analytic 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 given the information set $\mathcal{F}_t = \sigma\left(\left[x_{t-j}, y_{t-j}\right], j = 0, 1, 2, ...\right)$. Then the optimal forecast of $z_t = \sum_{j=0}^{11} y_{t-j}$ given information at time $t-h$ is given by:*

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

The mean squared forecast error as a function of the forecast horizon is given by:

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

The proof of Proposition 1 is in Appendix A, and an example of this result is presented in Figure 2. 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 the second part 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 $h \to \infty$, 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 :

$$\lim_{h\to\infty} E\left[e^2_{t,t-h}\right] = 12\sigma^2_u + \frac{\sigma^2_\varepsilon}{(1-\phi)^2}\left(12 - 2\frac{\phi\left(1-\phi^{12}\right)}{1-\phi} + \frac{\phi^2\left(1-\phi^{24}\right) + \phi^2\left(1-\phi^{12}\right)^2}{1-\phi^2}\right). \qquad (4)$$

To illustrate Proposition 1, Figure 2 plots the root mean squared error (RMSE) for $h = 1, 2, ..., 24$ using parameters similar to those we obtain in the empirical analysis for U.S. GDP growth. Holding the unconditional variance of annual GDP growth, σ^2_z, and the ratio of the transitory component variance to the persistent component variance, σ^2_u/σ^2_x, fixed we show the impact of varying the persistence parameter, ϕ. 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 ϕ. Notice also how the change in RMSE gets smaller at the longest horizons, irrespective of the value of ϕ.

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 $h \to 0$. 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.

2.2 Introducing 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 in our empirical work in Sections 3 and 4 we use a state-space model and Kalman filtering. While this approach yields a realistic model for the signal extraction problem facing forecasters in practice, such an approach does not lend itself to easily interpretable formulas for the term structure of forecast errors. For this reason, and to gain intuition, we first consider a simplified model which captures the spirit of the measurement error problem. We assume here that the persistent component, x_t, is perfectly observable while the predicted variable, y_t, is observable only with noise:

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

We further assume here that the measurement error is mean zero and uncorrelated with all other innovations, i.e. for all t, s, $E[\eta_t] = E[\varepsilon_s \eta_t] = E[u_s \eta_t] = 0$, although this can be extended to account for other types of errors, cf. Aruoba (2007). This model nests the "no noise" model for $\sigma_\eta^2 = 0$.

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 $\tilde{\mathcal{F}}_t = \sigma([x_{t-j}, \tilde{y}_{t-j}], j = 0, 1, 2, ...)$.*

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

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

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

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

(3) The mean squared forecast error is given by

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

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 the optimal estimate of the true value of y_t combines information from these two sources according to their relative accuracy. The second part of the proposition reiterates that the measurement error does not affect $\hat{z}^*_{t,t-h}$ or $E\left[e^2_{t,t-h}\right]$ 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. If $\sigma_\eta^2 \to 0$ or $\sigma_u^2 \to 0$, then this term vanishes, as in either of those cases y_t is perfectly observable.

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 term 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. In the absence of measurement errors the MSE will converge to zero as $h \to 0$, whereas in the presence of measurement error the MSE will converge to some positive quantity. 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 (ϕ) and measurement error (σ_η^2) parameters are separately identified by jointly considering short and long ends of the term structure of MSE-values.

2.3 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. To be able to understand the cross-sectional dispersion in beliefs, we next extend our baseline model for consensus beliefs to incorporate heterogeneity. 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

$$d_{t,t-h}^2 \equiv \frac{1}{N} \sum_{i=1}^{N} \left(\hat{z}_{i,t,t-h} - \bar{z}_{t,t-h} \right)^2 \tag{6}$$

where $\bar{z}_{t,t-h} \equiv N^{-1} \sum_{i=1}^{N} \hat{z}_{i,t,t-h}$ 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 is the number of forecasters.

To capture heterogeneity in the forecasters' information, we assume that each forecaster observes a different signal of the current value of y_t, denoted $\tilde{y}_{i,t}$:

$$\begin{aligned} \tilde{y}_{i,t} &= y_t + \eta_t + \nu_{i,t} \\ \eta_t &\sim iid\left(0, \sigma_\eta^2\right) \ \forall \ t \\ \nu_{i,t} &\sim iid\left(0, \sigma_\nu^2\right) \ \forall \ t, i \\ E\left[\nu_{i,t}\eta_s\right] &= 0 \ \forall \ t, s, i. \end{aligned} \tag{7}$$

Individual forecasters' measurements of y_t are contaminated with a common source of noise, denoted η_t as in the model for consensus views, 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.[6] 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:

$$\tilde{y}_{t-1} = y_{t-1} + \eta_{t-1} \tag{8}$$

[6] 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 σ_ν close to zero.

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

$$\hat{z}^*_{i,t,t-h} \equiv E\left[z_t|\tilde{\mathcal{F}}_{i,t-h}\right], \qquad \tilde{\mathcal{F}}_{i,t-h} = \{\tilde{y}_{i,t-h-j}, \tilde{y}_{t-h-1-j}\}_{j=0}^{t-h}. \tag{9}$$

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 $h \to \infty$ 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 $h \to \infty$. 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 μ_i. We assume that forecaster i shrinks the optimal forecast based on his information set $\mathcal{F}_{i,t-h}$ towards his prior belief about the unconditional mean of z_t. The degree of shrinkage is governed by a parameter $\kappa^2 \geq 0$, with low values of κ^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 κ^2 implying a high weight on $\hat{z}^*_{i,t,t-h}$. As $\kappa^2 \to 0$ the forecaster places all weight on his prior beliefs and none on the data; as $\kappa^2 \to \infty$ the forecaster places no weight on his prior beliefs.

$$\begin{aligned}
\hat{z}_{i,t-h,t} &= \omega_h \mu_i + (1-\omega_h) E\left[z_t|\mathcal{F}_{i,t-h}\right], \qquad (10)\\
\omega_h &= \frac{E\left[e^2_{i,t,t-h}\right]}{\kappa^2 + E\left[e^2_{i,t,t-h}\right]}\\
e_{i,t,t-h} &\equiv z_t - E\left[z_t|\mathcal{F}_{i,t-h}\right].
\end{aligned}$$

Notice that we allow the weights placed on the prior and the optimal expectation $E\left[z_t|\tilde{\mathcal{F}}_{i,t-h}\right]$ to vary across the forecast horizons in a manner consistent with standard forecast combinations: as $\hat{z}^*_{i,t,t-h} \equiv E\left[z_t|\tilde{\mathcal{F}}_{i,t-h}\right]$ becomes more accurate (i.e., as $E\left[e^2_{i,t,t-h}\right]$ 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.[7] Furthermore, note that

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

For analytical tractability, and for better finite sample identification of κ^2, we impose that κ^2 is constant across all forecasters.[8]

The additional term μ_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 μ_i term would generally be time-varying, but we leave that possibility aside for now.

In Figure 4, we plot the theoretical term structures of dispersion coming from our empirical model, for various values of σ_μ, setting the other parameters to resemble those obtained for US GDP growth. This figure shows that for low values of σ_μ, the dispersion term structure is almost flat, while for larger values dispersion is quite high for long horizons and declines sharply as the forecast horizon shrinks towards zero.

3 Empirical Results: The Term Structure of Consensus Forecasts

This section presents empirical results for the consensus forecasts. After describing the data source, we 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 shed light on agents' beliefs about the persistent and transitory components of GDP growth and inflation over our sample period.

3.1 Data

Our data is taken from the Consensus Economics Inc. forecasts which comprise polls of more than 600 private sector forecasters and are widely considered by organizations such as the IMF

[7] 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.

[8] 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.

and the U.S. Federal Reserve. 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. Thus the H dimension of our data set is unusually large. However, the T dimension of our data set is rather short, with just 14 annual observations. We take the estimation problem here very seriously, but we have only limited data, and so our GMM and SMM standard errors and J-test statistics should be interpreted with this in mind. Additional details of the structure of the data are provided in Appendix B. Naturally our 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. To measure the realized value of the target variable (GDP growth or inflation), we use second release data.[9]

As a prelude to our analysis of the term structure of forecast errors, we initially undertook a range of statistical tests that check for biases and serial correlation in the forecast errors. We tested for biases in the forecasts, for individual forecast horizons and jointly across all horizons, by testing whether the forecast errors were mean zero and by estimating "Mincer-Zarnowitz" (1969) regressions. Few of these tests suggested lack of optimality for our forecasts, and hence we shall proceed to estimate the parameters of our model under the assumption that forecasters use information efficiently. Full details of these test results are available from the authors upon request.

3.2 Parameter Estimates and Tests

The simple benchmark model contains just three free parameters, namely the variance of the innovations in the temporary (σ_u^2) and persistent (σ_ε^2) components, and the persistence parameter, ϕ, for the predictable component. The expressions for the MSE as a function of h, stated in Proposition 1 for the benchmark model and in Section 3.3 and Appendix C for the model that employs a Kalman filter, enable us to use GMM to estimate the unknown parameters given a panel of forecast errors measured at various horizons. These parameters are not separately identifiable if forecasts for a single horizon are all that is available so access to multi-horizon forecasts is crucial

[9] Results are very similar when the first release is used instead as recommended by Corradi, Fernandez and Swanson (2007).

to our analysis. In contrast, since the variance of the h-period forecast error grows linearly in σ_u^2 while σ_ε^2 and ϕ 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:

$$\hat{\theta}_T \equiv \arg\min_{\theta} \; g_T(\theta)' W_T g_T(\theta) \tag{11}$$

$$g_T(\theta) \equiv \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} e_{t,t-1}^2 - MSE_1(\theta) \\ e_{t,t-2}^2 - MSE_2(\theta) \\ \vdots \\ e_{t,t-24}^2 - MSE_{24}(\theta) \end{bmatrix} \tag{12}$$

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

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 a 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.[10]

Figure 5 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 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

[10] We simulated 10,000 non-overlapping "years" of data from the model to compute the covariance matrix of the sample moments. Details are provided in the appendix.

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 ($h \geq 12$) than in current-year forecasts ($h < 12$).[11]

The fitted values from the models with and without measurement error, also shown in Figure 5, 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.

3.3 Allowing for 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 unchanged:

$$
\begin{aligned}
\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & \phi \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ \varepsilon_t \end{bmatrix} \qquad (13) \\
\begin{bmatrix} u_t \\ \varepsilon_t \end{bmatrix} &\sim iid \left(0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} \right),
\end{aligned}
$$

[11] Note that Figure 5 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 economy between formal announcement dates.

while the measurement equation becomes:

$$\begin{bmatrix} \tilde{y}_t \\ \tilde{x}_t \end{bmatrix} = \begin{bmatrix} y_t \\ x_t \end{bmatrix} + \begin{bmatrix} \eta_t \\ \psi_t \end{bmatrix}, \tag{14}$$

$$\begin{bmatrix} \eta_t \\ \psi_t \end{bmatrix} \sim iid \left(0, \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\psi^2 \end{bmatrix} \right).$$

This is a very simple state-space system.[12] Unfortunately, however, it does not yield a formula for the term structure of MSEs that is readily interpretable. The key difficulty that arises is best illustrated by considering "current-year" forecasts ($h \leq 12$). When producing a current-year forecast at time $t-h$, economic agents must use past and current information to backcast realizations dated at time $t-11, ..., t-h-1$; they must also produce a "nowcast" for the current month and, finally, must predict future realizations, $y_{t-h+1}, .., y_t$. 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 η_{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, ε_t, η_t and ψ_t, which are serially uncorrelated. We show how to accomplish this in Appendix C.

This extended model introduces two further parameters which reflect the magnitude of measurement errors (σ_η^2 and σ_ψ^2). To reduce the set of unknown parameters to a tractable number, we do two things. First, as a normalization we set $\sigma_\psi^2 \to \infty$, 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 σ_η^2 and σ_u^2 are well-identified in theory, in practice they are difficult to estimate separately. We therefore set σ_η to be proportional to $\sigma_u : \sigma_\eta = k \cdot \sigma_u$ and estimate 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) is generally robust for $1 \leq k \leq 4$. We set $k = 2$ in the estimation.

[12] Faust et al. (2005) find that revisions to U.S. GDP figures are essentially unpredictable, motivating the simple *iid* noise structure used above. 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.

Table 1 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.[13] 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.

Figure 5 shows that the specification with measurement errors 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.

3.4 Moments of Actual Data

The parameter estimates reported in Table 1 are chosen to match as closely as possible the observed term structure of mean-squared forecast errors. Naturally, these estimates also imply a set of fitted values and thus moments of the predicted variables. To see if these estimates match up with the actual outcomes, we simulated 50,000 months of data from our Kalman-filter based model and used these to compute the standard deviation and autocorrelation at lags of 1, 2 and 3 months. In a second step, we formed 90% confidence bands by breaking the simulated series into blocks of 14 years of data and computed 5% and 95% quantiles of the simulated distribution. The results show that the autocorrelation of the model-implied series match the observed outcomes, whose moments fall well within the simulated 90% confidence intervals.

In contrast, the standard deviation of output growth and inflation implied by our model estimates are not consistent with the data. In particular, the sample estimates from the real data fall well below the lower bound of the simulated confidence intervals. That is, forecasters were behaving as though they thought the standard deviation was higher than it really was. In a longer historical context, the standard deviation estimates for annual GDP growth and inflation during our sample of 1.6% and 0.7%, respectively, were very low and so it seems that agents did not adjust their views to this low level of volatility. This would be consistent with the difficulty in detecting structural breaks such as the "Great Moderation" in relatively low frequency data such as inflation

[13] The estimates in Table 1 are for the monthly frequency. The implied first-order autocorrelation coefficients for quarterly (annual) GDP growth and inflation are 0.80 and 0.88 (0.60 and 0.69) respectively.

and GDP growth.

3.5 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, but with the estimated parameter vector and the panel of forecasts we are able to infer the forecasters' estimated values of these variables. Without a model for the persistent and transitory components of these series it would be impossible to extract estimates of the forecasters' beliefs about the persistent components of GDP growth and inflation. 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 Appendix C.

In Figure 6 we present for each month in our sample the estimated persistent components of GDP growth and inflation, as implied by the observed consensus forecasts and the parameters of our model. 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 forecasters in our panel estimated the level of GDP growth in the early 1990s quite well, but were 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, the forecasters in our panel were 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 Empirical Results: Cross-Sectional Dispersion

This section studies disagreements among survey participants' forecasts of GDP growth and inflation. Rather than analyzing the forecasts reported by the individual survey participants, we study the cross-sectional dispersion. This measure matches our theoretical model in Section 2 and has the important advantage that it is not affected by incomplete data records due to the entry, exit and re-entry of individual forecasters.[14] Our analysis exploits the multi-horizon feature of the data: if cross-sectional dispersion was only available for a single horizon it would not be possible to infer the relative magnitude of priors versus information signals underlying the cross-sectional dispersion.

4.1 Estimation

Equations (7) to (10) in Section 2 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:

$$\begin{aligned} d_{t,t-h}^2 &= \delta_h^2 \cdot \lambda_{t,t-h} \\ E\left[\lambda_{t,t-h}\right] &= 1 \\ V\left[\lambda_{t,t-h}\right] &= \sigma_\lambda^2, \end{aligned} \tag{15}$$

where $d_{t,t-h}^2$ is the observed value of the cross-sectional dispersion. In particular, we assume that the residual, $\lambda_{t,t-h}$, 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, σ_μ^2, σ_ν^2 and κ^2, relate to the term structure of

[14] Entry and exit is a large problem for most survey data and makes it difficult to get long track records for individual forecasters.

forecast dispersions—i.e. how the dispersion changes as a function of the forecast horizon, h—while the fourth parameter, σ_λ^2, relates to the variance of forecast dispersions through time.

To estimate σ_λ^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 our model generates 72 moment conditions and contains 8 unknown parameters. Similar to 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 SMM (see Appendix C for details):

$$\hat{\theta}_T \equiv \arg\min_{\theta} \; g_T(\theta)' g_T(\theta), \tag{16}$$

$$g_T(\theta) \equiv \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} e_{t,t-1}^2 - MSE_1(\theta) \\ \vdots \\ e_{t,t-24}^2 - MSE_{24}(\theta) \\ d_{t,t-1}^2 - \delta_1^2(\theta) \\ \vdots \\ d_{t,t-24}^2 - \delta_{24}^2(\theta) \\ \left(d_{t,t-1}^2 - \delta_1^2(\theta)\right)^2 - \delta_1^4(\theta)\left(\exp\left(\sigma_\lambda^2\right) - 1\right) \\ \vdots \\ \left(d_{t,t-24}^2 - \delta_{24}^2(\theta)\right)^2 - \delta_{24}^4(\theta)\left(\exp\left(\sigma_\lambda^2\right) - 1\right) \end{bmatrix}. \tag{17}$$

Panel A of Table 2 reports parameter estimates for this model. Compared with Table 1, the estimates of the parameters defining the dynamics of the target variables are essentially unchanged. The estimates of κ and σ_μ suggest considerable heterogeneity across forecasters in our panel, whereas the estimates of σ_ν 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 σ_ν 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 ν_{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 the 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. Contrast 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, or indeed any model assuming a quadratic penalty for forecast errors and efficient use of information, and thus poses a puzzle.

4.2 Time-varying dispersion

There is a growing amount of theoretical and empirical work on the relationship between the uncertainty facing economic agents and the economic environment. Veldkamp (2006) and van Nieuwerburgh and Veldkamp (2006) propose endogenous information models where agents' participation in economic activity leads to more precise information about unobserved economic state variables such as (aggregate) technology shocks. In these models the number of signals observed by agents is proportional to the economy's activity level so more information is gathered in a good state of the economy than in a bad state. Recessions are therefore times of greater uncertainty which in turn means that dispersion among agents' forecasts can be expected to be wider during such periods. This is a natural consequence of our model to the extent that information signals contain a common component and so lead to more similar beliefs during periods where more information is available.

To address such issues, we consider one final 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 and 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, σ_ν^2, or through the variance of the prior beliefs about the long-run values of the series, σ_μ^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

$$\log \sigma_{\mu,t}^2 = \beta_0^\mu + \beta_1^\mu \log S_t, \tag{18}$$

where S_t is the default spread in month t. In this model, if $\beta_1^\mu > 0$ then 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 2. The fit of the models were not much changed by this extension. Interestingly, the results reveal a positive relationship between default spreads and σ_μ, 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^{th} 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^{th} 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^{th} or 95^{th} percentile value.

We conclude that 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 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 heterogeneity in forecasters' information signals and in their prior beliefs. Though highly parsimonious, our proposed models succeed in capturing the level, slope and curvature of the term structure of forecast errors, and shed light on the primary sources of the cross-sectional dispersion among forecasters.

Our finding of a persistent component in agents' forecast errors–indicated by the fact that our sample of survey forecasters underpredicted US output growth and overpredicted inflation for much of the 1990s–lends empirical credibility to models with gradual learning. This could be important since many macroeconomic series appear more persistent than standard dynamic stochastic general equilibrium models can match and, as shown by Milani (2006), allowing for agents' learning bring these models closer to matching the data. Slow learning is also consistent with the apparent failure of agents to realize how low volatility of both output growth and inflation had become after the "Great Moderation" whose effect was only established well into our historical sample.

Our empirical findings suggest that forecast dispersion is primarily driven by heterogeneity in beliefs about long-run values of GDP growth and inflation, as opposed to heterogeneity in information about the current state of the economy. Differences in beliefs about GDP growth appear to be strongly counter-cyclical (increasing during bad states of the world) whereas differences in agents' inflation forecasts are less state dependent. Moreover, while our model can match the dispersion observed among survey participants' forecasts of GDP growth, it fails to match the high dispersion in inflation forecasts observed at short horizons. Why professional forecasters' views of inflation at short horizons displays such "excess dispersion" is difficult to understand and poses a puzzle to any model based on agents' efficient use of information.

6 Appendix A: 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, $\mathcal{F}_{t-h} = \{x_{t-h}, y_{t-h}, x_{t-h-1}, y_{t-h-1}, ...\}$, requires accounting for the persistence in x. Note that

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

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

$$\begin{aligned}
z_t &= \sum_{j=0}^{11} x_{t-j} + \sum_{j=0}^{11} u_{t-j} \qquad (19) \\
&= \frac{\phi(1-\phi^{12})}{1-\phi} x_{t-12} + \frac{1}{1-\phi}\sum_{j=0}^{11}(1-\phi^{12-j})\varepsilon_{t-12+1+j} + \sum_{j=0}^{11} u_{t-j}.
\end{aligned}$$

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

$$\hat{z}^*_{t,t-h} \equiv E\left[z_t|\mathcal{F}_{t-h}\right] = \sum_{j=0}^{11} E\left[y_{t-j}|\mathcal{F}_{t-h}\right] = \sum_{j=0}^{11} E\left[x_{t-j}|\mathcal{F}_{t-h}\right] = \sum_{j=0}^{11} \phi^{h-j} x_{t-h},$$

$$\text{so} \quad \hat{z}^*_{t,t-h} = \frac{\phi^{h-11}\left(1-\phi^{12}\right)}{1-\phi} x_{t-h}\,, \quad \text{for} \quad h \geq 12.$$

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:

$$\hat{z}^*_{t,t-h} = \sum_{j=0}^{11} E\left[y_{t-j}|\mathcal{F}_{t-h}\right] = \sum_{j=h}^{11} y_{t-j} + \sum_{j=0}^{h-1} E\left[x_{t-j}|\mathcal{F}_{t-h}\right] = \sum_{j=h}^{11} y_{t-j} + \sum_{j=0}^{h-1} \phi^{h-j} x_{t-h},$$

$$\text{so} \quad \hat{z}^*_{t,t-h} = \sum_{j=h}^{11} y_{t-j} + \frac{\phi\left(1-\phi^{h}\right)}{1-\phi} x_{t-h}\,, \quad \text{for} \quad h < 12.$$

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 $h \geq 12$,

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

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

$$\begin{aligned} E\left[e^2_{t,t-h}\right] &= \sum_{j=0}^{11} E\left[u^2_{t-j}\right] + \sum_{j=0}^{11} \frac{\left(1-\phi^{j+1}\right)^2}{\left(1-\phi\right)^2} E\left[\varepsilon^2_{t-j}\right] + \sum_{j=12}^{h-1} \frac{\phi^{2j-22}\left(1-\phi^{12}\right)^2}{\left(1-\phi\right)^2} E\left[\varepsilon^2_{t-j}\right] \\ &= 12\sigma^2_u + \frac{\sigma^2_\varepsilon}{\left(1-\phi\right)^2}\sum_{j=0}^{11}\left(1-\phi^{j+1}\right)^2 + \frac{\left(1-\phi^{12}\right)^2}{\left(1-\phi\right)^2}\sigma^2_\varepsilon \sum_{j=12}^{h-1}\phi^{2j-22} \\ &= 12\sigma^2_u + \frac{\sigma^2_\varepsilon}{\left(1-\phi\right)^2}\left(12 - 2\frac{\phi\left(1-\phi^{12}\right)}{1-\phi} + \frac{\phi^2\left(1-\phi^{24}\right)}{\left(1-\phi^2\right)}\right) \\ &+ \frac{\phi^2\left(1-\phi^{12}\right)^2\left(1-\phi^{2h-24}\right)}{\left(1-\phi\right)^3\left(1+\phi\right)}\sigma^2_\varepsilon \end{aligned}$$

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

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

■

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*

$$
\begin{aligned}
E\left[y_{t-j-h}|\tilde{\mathcal{F}}_{t-h}\right] &= x_{t-j-h} + E\left[u_{t-j-h}|\tilde{\mathcal{F}}_{t-h}\right] \\
&= x_{t-j-h} + E\left[u_{t-j-h}|\tilde{y}_{t-j-h}, x_{t-j-h}\right] . \\
\text{Recall } \tilde{y}_{t-j-h} &= x_{t-j-h} + u_{t-j-h} + \eta_{t-j-h}. \\
\text{Define } \tilde{u}_{t-j-h} &\equiv \tilde{y}_{t-j-h} - x_{t-j-h} = u_{t-j-h} + \eta_{t-j-h}. \\
\text{Then } \; E\left[u_{t-j-h}|\tilde{y}_{t-j-h}, x_{t-j-h}\right] &= E\left[u_{t-j-h}|\tilde{u}_{t-j-h}\right].
\end{aligned}
$$

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 $h < 12$:

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

(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

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

$$\begin{aligned}
\text{Note that } y_{t-j} - E\left[y_{t-j}|\tilde{\mathcal{F}}_{t-h}\right] &= y_{t-j} - \left(\frac{\sigma_\eta^2}{\sigma_u^2+\sigma_\eta^2}\right) x_{t-j} - \left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_\eta^2}\right) \tilde{y}_{t-j,t-h} \\
&= x_{t-j} + u_{t-j} - \left(\frac{\sigma_\eta^2}{\sigma_u^2+\sigma_\eta^2}\right) x_{t-j} - \left(\frac{\sigma_u^2}{\sigma_u^2+\sigma_\eta^2}\right) \left(x_{t-j} + u_{t-j} + \eta_{t-j}\right) \\
&= \frac{\sigma_\eta^2}{\sigma_u^2+\sigma_\eta^2} u_{t-j} - \frac{\sigma_u^2}{\sigma_u^2+\sigma_\eta^2} \eta_{t-j}. \\
\text{Hence } V\left[y_{t-j} - E\left[y_{t-j}|\tilde{\mathcal{F}}_{t-h}\right]\right] &= \frac{\sigma_u^2\sigma_\eta^4}{\left(\sigma_u^2+\sigma_\eta^2\right)^2} + \frac{\sigma_u^4\sigma_\eta^2}{\left(\sigma_u^2+\sigma_\eta^2\right)^2} = \frac{\sigma_u^2\sigma_\eta^2}{\sigma_u^2+\sigma_\eta^2}.
\end{aligned}$$

As stated in the proposition, the variance of the extra terms in the MSE is then

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

■

7 Appendix B: The Structure of the Consensus Economics Target Variables

Our analysis takes the target variable, z_t, as the December-on-December change in the log-level of US real GDP or the Consumer Price Index, which can of course be written as the sum of the month-on-month changes in the log-levels of these series, denoted y_t, as in equation (2). In the Consensus Economics survey, however, the target variables are defined as

$$\begin{aligned} \text{GDP growth} \quad z_t^{GDP} &\equiv \frac{\bar{P}_t}{\bar{P}_{t-1}} - 1 \\ \text{Inflation} \quad z_t^{INF} &\equiv \frac{\overline{\overline{P}}_t}{\overline{\overline{P}}_{t-1}} - 1, \end{aligned}$$

where $\bar{P}_t \equiv \frac{1}{4}\sum_{j=0}^{3} P_{t-3j}^{GDP}$, $\overline{\overline{P}}_t \equiv \frac{1}{12}\sum_{j=0}^{11} P_{t-j}^{INF}$ and P_t^{GDP} and P^{INF} denote the level of real GDP and the Consumer Price Index respectively. For reasonable values of y_t, the variables z_t^{GDP} and z_t^{INF} can be shown to be accurately approximated as a linear combination of $(y_t, y_{t-1}, ..., y_{t-23})$:[15]

$$\begin{aligned} z_t(\mathbf{w}) &\equiv \sum_{j=0}^{23} w_j y_{t-j} \qquad (20) \\ w_j^{GDP} &= \begin{cases} \frac{1+\lfloor \frac{j}{3} \rfloor}{4}, & 0 \le j \le 11 \\ \frac{3-\lfloor \frac{j-12}{3} \rfloor}{4}, & 12 \le j \le 23 \end{cases}, \quad j = 0, 1, ..23 \\ w_j^{INF} &= \begin{cases} \frac{j+1}{12}, & 0 \le j \le 11 \\ \frac{23-j}{12}, & 12 \le j \le 23 \end{cases}, \quad j = 0, 1, .., 23 \\ w_j^{*} &= \begin{cases} 1, & 0 \le j \le 11 \\ 0, & 12 \le j \le 23 \end{cases}, \quad j = 0, 1, ..23 \end{aligned}$$

where $\lfloor a \rfloor$ rounds a down to the nearest integer and $\mathbf{w}$ is the vector of weights on the individual observations. The inflation weights are "tent-shaped" and thus similar to the Bartlett (1946) kernel, while the GDP weights have "flat segments" relative to the inflation weights, reflecting the averaging over quarterly observations rather than monthly observations.

Writing the Consensus Economics target variables in this way allows us to generalize the results given in Section 2, at the cost of simplicity and interpretability. For example, the key variables in

[15] Details are available from the authors on request.

Proposition 1 in this general case can be shown to be

$$\begin{aligned}
\hat{z}^*_{t,t-h}(\mathbf{w}) &= \sum_{j=h}^{23} w_j y_{t-j} + \sum_{j=0}^{h-1} w_j \phi^{h-j} x_{t-h}, \\
V\left[z_t(\mathbf{w})\right] &= \sigma_u^2 \mathbf{w}'\mathbf{w} + \sigma_\varepsilon^2 \left\{ \frac{\phi^2}{1-\phi^2} \left(\sum_{j=0}^{23} w_j \phi^{23-j} \right)^2 + \sum_{k=0}^{23} \left(\sum_{j=0}^{k} w_j \phi^{k-j} \right)^2 \right\}, \\
E\left[e^2_{t,t-h}(\mathbf{w})\right] &= \sigma_u^2 \sum_{j=0}^{h-1} w_j^2 + \sigma_\varepsilon^2 \sum_{k=0}^{h-1} \left(\sum_{j=0}^{k} w_j \phi^{k-j} \right)^2 .
\end{aligned}$$

The representation in equation (20) further enables us to relate our model, which has the simple weight vector denoted $\mathbf{w}^*$ above, to the data. Since both $\mathbf{w}^{GDP}$ and $\mathbf{w}^{INF}$ allocate weight to the previous year's monthly growth rates ($w_j > 0$ for some $j \geq 12$) a target variable defined with either of these weights will tend to be smoother and more autocorrelated than a target variable defined using $\mathbf{w}^*$, ceteris paribus. Thus if the Consensus Economics survey respondents are actually targeting $z_t\left(\mathbf{w}^{GDP}\right)$, for example, while our model assumes $z_t(\mathbf{w}^*)$ as the target, then our estimates of ϕ may be upward biased, since our model assumes that any predictability in z_t is the result of predictability in y_t, whereas in such a case some of the predictability would be purely due to the smoothed nature of the target variable.

As we show in the empirical section, the flexibility of our model, even with just a few parameters, means that we are able to fit the observed data even when fixing $\mathbf{w} = \mathbf{w}^*$. Making this assumption simplifies the presentation of our model and results, at the cost of some precision of the parameter estimates. The precision of these estimates is limited in the first instance due to the short time series of data we have (T is just 14), and so we choose to retain the simple, interpretable model and set $\mathbf{w} = \mathbf{w}^*$ in the paper.[16]

8 Appendix C: Kalman Filter Implementation

This appendix presents a detailed description of our implementation of the Kalman filter for our econometric model of the term structures of forecast errors and forecast dispersions.

[16] Moreover, it is not clear that the survey participants report a weighted forecast rather than the simple year-on-year forecast that we assume in our analysis. In private correspondence, Consensus Economics noted that "When we survey panelists for an annual forecast, i.e. real GDP growth in 2007, we are asking for the percentage change from 2006.", which leaves the precise definition of the target variable open to interpretation.

8.1 Details on the Kalman filter model for consensus RMSE

We first describe the model for the consensus forecasts, using notation similar to that 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 $\xi_t = [y_t, x_t]'$, $F = [\mathbf{0}, \phi\boldsymbol{\iota}]$, where $\mathbf{0}$ and $\boldsymbol{\iota}$ are 2×1 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

$$\tilde{y}_t = H'\xi_t + w_t. \tag{23}$$

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:

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

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

$$E\left[v_t w_s'\right] = 0\ \forall\ s, t. \tag{25}$$

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:

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

Following Hamilton (1994),

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

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:

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

Updating of the estimates is done via

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

while the multi-step prediction error uses:

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

The "current year" forecasts require "smoothed" estimates of current and past values, sometimes known as "nowcasts" and "backcasts". The smoothed estimates are obtained from:

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

where

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

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

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

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, ξ_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:

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

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

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

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

The first term arises from being unable to observe ξ_{t-h} : if ξ_{t-h} were observable without error then $P_0^*=0$ and this term would vanish. The second term is the combined impact of predicting ξ_{t-j}

over the measurement period ($j = t-11, t-10, ..., t$). This will be larger or smaller depending on how predictable the series is, which is determined completely by $\left(\sigma_u^2, \sigma_\varepsilon^2, \phi\right)$. The third term is the combined impact of having to predict the intermediate values of ξ_{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

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

while the general form for the backcast errors is:

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

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

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

Iterating backwards, we have

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

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

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

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:

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

Then the backcast errors are

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

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, $\left\{\left[v_{t-j}', w_{t-j}'\right]', j=0,1,...,10\right\}$. All of these elements are independent of each other and so the MSE for $h<12$ is directly obtained from the above expressions.

8.2 Details on the Kalman filter model for dispersion

The state equations for the individual forecaster are

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

while the measurement equations are

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

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 η_t, $E\left[\ddot{\eta}_t^2\right]=E\left[\eta_t^2\right]=\sigma_\eta^2$.

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

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

Our model for dispersion is based on δ_h^2. Unfortunately, a closed-form expression for δ_h^2 is not available and so we resort to simulations to evaluate δ_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, μ_i, are *iid* $N\left(0,\sigma_\mu^2\right)$ 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 (10). We then compute the cross-sectional variance of the individual forecasts to obtain $d^2_{t,t-h}$ and average these across time to obtain $\hat{\delta}^2_h$.

8.3 Details on the estimation of the models

To obtain P^*_1, P^*_0,...,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^{th} 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. We have also estimated the models presented in this paper using the full set of 24 moment conditions and the results were qualitatively similar.

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 ε_t as an *iid* $N\left(0, \sigma^2_\varepsilon\right)$ 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* $N\left(0, \sigma^2_u\right)$ time series of the same length as ε_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 8.1. 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 $t-24$. To match the sampling frequency of our data, we drop all but every twelfth time series observation from the simulated data, leaving us with $1,000$ non-overlapping years of data. Finally, for each observation we compute the 'term structure' of squared forecast errors, $\left(z_t - \hat{z}^*_{t,t-1}\right)^2, ..., \left(z_t - \hat{z}^*_{t,t-24}\right)^2$. We then compute the matrix of moment conditions, using the expression in equation (12), and from this we estimate the covariance matrix of the moments for the horizons used in estimation.[17]

A closed-form expression for δ_h^2 is not available and so we use simulations to obtain an estimate of it. For each evaluation of the objective function, we simulated 50 non-overlapping years of data for 30 forecasters to estimate δ_h^2.[18,19] The priors for each of the 30 forecasters, μ_i, were simulated as $iid\ N\left(0, \sigma_\mu^2\right)$. The only difference in the dispersion simulation is that we must simulate the residual term, $\lambda_{t,t-h}$. We multiply the estimated δ_h^2 series by $\lambda_{t,t-h}$, defined in equation (15), which is $iid \log N\left(-\frac{1}{2}\sigma_\lambda^2, \sigma_\lambda^2\right)$. 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 and third set of 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, μ_i^*, were simulated as $iid\ N(0,1)$, and then the actual "prior" for each time period, $\mu_{i,t}$, was set as $\mu_i^* \times \sigma_{\mu,t}$, where $\sigma_{\mu,t} = \exp\left\{\left(\beta_0^\mu + \beta_1^\mu \log S_t\right)/2\right\}$. 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 $\delta_h^2\left(\sigma_{\mu,t}\right)$, which is simply the sample mean of $d_{t,t-h}^2$ in the constant dispersion model, so that we can compute the dispersion residual. It was not computationally feasible to simulate $\delta_h^2\left(\sigma_{\mu,t}\right)$ 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 $\tilde{\delta}_h^2\left(\sigma_{\mu,t}\right)$. We checked the accuracy of this approximation for values in between these nodes and the errors were very small. We then use $\tilde{\delta}_h^2\left(\sigma_{\mu,t}\right)$, and the data, to compute the dispersion residuals and used these in the GMM estimation of the parameters of the model.

[17] 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.

[18] The actual number of forecasters in each survey exhibited some variation across t and h,with values between 22 and 32. In the simulations we set $N = 30$ for all t, h for simplicity.

[19] 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.

8.4 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

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

$$
\begin{aligned}
\text{Thus } F^{13} F^{(11)} \hat{\xi}_{1,1} &= \begin{bmatrix} 0 & \frac{\phi^{13}(1-\phi^{12})}{1-\phi} \\ 0 & \frac{\phi^{13}(1-\phi^{12})}{1-\phi} \end{bmatrix} \hat{\xi}_{1,1}. \\
\textit{Let } e_1' &\equiv [1, 0] \\
\hat{z}_{25,1} &= e_1' \hat{w}_{25,1} \\
&= e_1' F^{13} F^{(11)} \hat{\xi}_{1,1} \\
&= \frac{\phi^{13}\left(1-\phi^{12}\right)}{1-\phi} \hat{\xi}_{1,1}^{[2]} \\
&= \frac{\phi^{13}\left(1-\phi^{12}\right)}{1-\phi} E\left[x_1 | \tilde{\mathcal{F}}_1\right], \\
\text{so } E\left[x_1 | \tilde{\mathcal{F}}_1\right] &= \frac{1-\phi}{\phi^{13}\left(1-\phi^{12}\right)} \cdot \hat{z}_{25,1}.
\end{aligned}
$$

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:

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

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 ϕ; 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:

$$\begin{aligned}
\hat{w}_{25,14} &= \sum_{j=0}^{11} \hat{\xi}_{25-j,14} = \left(\sum_{j=0}^{11} F^{11-j}\right)\hat{\xi}_{14,14} \equiv F^{(11)}\hat{\xi}_{14,14}, \\
\text{and } \hat{z}_{25,14} &= e_1'\hat{w}_{25,14} = e_1'F^{(11)}\hat{\xi}_{14,14} \\
&= \begin{bmatrix} 1 & \frac{\phi\left(1-\phi^{11}\right)}{1-\phi} \end{bmatrix}\hat{\xi}_{14,14} \\
&= E\left[y_{14}|\tilde{\mathcal{F}}_{14}\right] + \frac{\phi\left(1-\phi^{11}\right)}{1-\phi}E\left[x_{14}|\tilde{\mathcal{F}}_{14}\right].
\end{aligned}$$

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)

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

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:

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

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 $\hat{\xi}^{[2]}_{15,15} \equiv E\left[x_{15}|\tilde{\mathcal{F}}_{14}\right]$. Thus there is only one unknown on the right-hand side above, namely $\hat{\xi}^{[1]}_{15,15} \equiv E\left[y_{15}|\tilde{\mathcal{F}}_{14}\right]$, 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:

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

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 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:

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

and the expressions derived above can be re-interpreted a expressions for $E\left[x_t - \mu|\tilde{\mathcal{F}}_t\right]$ and $E\left[y_t - \mu|\tilde{\mathcal{F}}_t\right]$. 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 $E\left[x_t|\tilde{\mathcal{F}}_t\right]$ and $E\left[y_t|\tilde{\mathcal{F}}_t\right]$ as before, and then add back the means to the estimates. This corresponds to estimating the parameter μ by GMM, using simply the sample mean of the z_t series.

References

[1] Andersen, Torben G., Tim Bollerslev, Francis X. Diebold and Clara Vega, 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange. American Economic Review 93, 38-62.

[2] Aruoba, Boragan, 2007, Data Revisions are not Well-Behaved. Forthcoming in Journal of Money, Credit and Banking.

[3] Bartlett, Maurice S., 1946, On the theoretical specification of sampling properties of autocorrelated time series, Journal of the Royal Statistical Society, Supplement 8, 27–41.

[4] Beber, Alessandro and Michael W. Brandt, 2006, Resolving Macroeconomic Uncertainty in Stock and Bond Markets. Mimeo, HEC Lausanne and Duke University.

[5] Bernanke, Ben S., 2007, Inflation Expectations and Inflation Forecasting. Remarks by Chairman Ben S. Bernanke, Federal Reserve Board.

[6] Carroll, Christopher, 2003, Macroeconomic Expectations of Household and Professional Forecasters. Quarterly Journal of Economics, 118(1), 269-298.

[7] Chernov, Mikhail and Philippe Müller, 2007, The Term Structure of Inflation Forecasts, work-in-progress, London Business School and Columbia Business School.

[8] Clements, Michael P., 1997, Evaluating the Rationality of Fixed-Event Forecasts. Journal of Forecasting 16, 225-239.

[9] Corradi, Valentina, Andres Fernandez and Norman R. Swanson, 2007, Information in the Revision Process of Real-time Data. Mimeo, Rutgers.

[10] Croushore, Dean, 2006, Forecasting with Real-Time Macroeconomic Data. Pages 961-982 in Graham Elliott, Clive Granger and Allan Timmermann (eds.) Handbook of Economic Forecasting, North Holland: Amsterdam.

[11] Croushore, Dean and Tom Stark, 2001, A Real-Time Data Set for Macroeconomists. Journal of Econometrics 105, 111-130.

[12] Davies, Anthony and Kajal Lahiri, 1995, A New Framework for Analyzing Survey Forecasts using Three-Dimensional Panel Data, Journal of Econometrics, 68, 205-227.

[13] Dixit, Avinash X. and Robert S. Pindyck, 1994, Investment under Uncertainty. Princeton University Press: Princeton N.J.

[14] Ederington, Louis H. and Jae Ha 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.

[15] Faust, Jon, John H. Rogers and Jonathan H. Wright, 2005, News and Noise in G-7 GDP Announcements, Journal of Money, Credit, and Banking, 37(3), 403-417.

[16] Giannone, Domenico, Lucrezia Reichlin and David Small, 2007, Nowcasting GDP: The Real Time Informational Content of Macroeconomic Data Releases, working paper, Université Libre de Bruxelles.

[17] Grier, Kevin and Mark J. Perry, 1998, On Inflation and Inflation Uncertainty in the G-7 Countries. Journal of International Money and Finance 17.

[18] Hamilton, James D., 1994, Time Series Analysis, Princeton University Press, Princeton, New Jersey.

[19] Hansen, Lars P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, Econometrica, 50, 1029-1054.

[20] Kim, Charles-Jim and Charles 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.

[21] Kydland, Finn E. and Edward C. Prescott, 1982, Time to Build and Aggregate Fluctuations. Econometrica 50, 1345-1370.

[22] Lahiri, Kajal and Xugang Sheng, 2006, Evolution of Forecast Disagreement and Uncertainty in a Heterogeneous Agents Model, working paper, SUNY-Albany.

[23] Lucas, Robert E., Jr. 1973, Some International Evidence on Output-Inflation Tradeoffs. American Economic Review, 63, 326-334.

[24] Mahadeva, Lavan and Alex 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] Mankiw, N. Gregory, Ricardo Reis and Justin Wolfers, 2003. Disagreement about inflation expectations. NBER Macroeconomics Annual.

[26] McConnell, Margaret M. and Gabriel Perez-Quiros, 2000, Output Fluctuations in the United States: What Has Changed Since the Early 1980s?, American Economic Review, 90, 1464-1476.

[27] Milani, Fabio, 2006, Expectations, Learning and Macroeconomic Persistence. Forthcoming in the Journal of Monetary Economics.

[28] Mincer, Jacob, and Victor Zarnowitz, 1969, The Evaluation of Economic Forecasts, in J. Mincer (ed.) *Economic Forecasts and Expectations*, National Bureau of Economic Research, New York.

[29] Nordhaus, William D., 1987, Forecasting Efficiency: Concepts and Applications. Review of Economics and Statistics 69, 667-674.

[30] Pesaran, M. Hashem and Martin Weale, 2006, Survey Expectations. Pages 715-776 in G. Elliott, C. Granger and A. Timmermann (eds.) Handbook of Economic Forecasting, North Holland: Amsterdam.

[31] Politis, Dimitris N., and Romano, Joseph P., 1994, The Stationary Bootstrap, Journal of the American Statistical Association, 89, 1303-1313.

[32] Primiceri, Giorgio E., 2006, Why Inflation Rose and Fell: Policy Makers' Beliefs of US Postwar Stabilization Policy. Forthcoming in Quarterly Journal of Economics.

[33] Ramey, Garey and Valerie Ramey, 1995, Cross-Country Evidence on the Links between Volatility and Growth. American Economic Review 85, 1138–51.

[34] Stock, James H. and Mark W. Watson, 2002, Macroeconomic Forecasting using Diffusion Indexes. Journal of Business and Economic Statistics 20, 147-162.

[35] Stock, James H. and Mark W. Watson, 2006, Forecasting with Many Predictors. Pages 515-554 in Graham Elliott, Clive W.J. Granger and Allan Timmermann (eds.), Handbook of Economic Forecasting, Elsevier.

[36] Swidler, Steve and David Ketcher, 1990, Economic Forecasts, Rationality, and the Processing of New Information over Time. Journal of Money, Credit, and Banking 22, 65-76.

[37] Tauchen, George, 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.

[38] Townsend, Robert M., 1983. Forecasting the forecasts of others. The Journal of Political Economy 91, 546-588.

[39] Van Nieuwerburgh, Stijn and Laura Veldkamp, 2006, Learning Asymmetries in Real Business Cycles. Journal of Monetary Economics 53, 753-772.

[40] Veldkamp, Laura, 2006, Slow Boom, Sudden Crash. Journal of Economic Theory 124(2), 230-257.

[41] Veronesi, Pietro, 1999, Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model. Review of Financial Studies 12(5), 975-1007.

Table 1: **GMM parameter estimates of the consensus forecast model**

	σ_u	σ_ε	ϕ	σ_η	J p-val
GDP growth	0.063	0.054	0.936	0.126	0.403
	(0.012)	(0.013)	(0.034)	(——)	
Inflation	0.000	0.023	0.953	0.000	0.935
	(——)	(0.007)	(0.047)	(——)	

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 variable. The parameter σ_η was fixed at $2\sigma_u$ and is reported here for reference only.

Table 2: SMM parameter estimates of the joint consensus forecast and dispersion model

Panel A: constant forecast dispersion

	σ_u	σ_ε	ϕ	σ_η	σ_ν	κ	σ_μ	J p-val
GDP growth	0.063	0.054	0.936	0.126	0.692	1.414	0.672	0.857
	(0.012)	(0.013)	(0.034)	(——)	(1.00)	(0.941)	(0.394)	
Inflation	0.000	0.023	0.953	0.000	0.045	0.493	0.509	0.000
	(——)	(0.007)	(0.046)	(——)	(0.168)	(0.167)	(0.127)	

Panel B: time-varying forecast dispersion

	σ_u	σ_ε	ϕ	σ_η	σ_ν	κ	β_0^μ	β_1^μ	J p-val
GDP growth	0.063	0.054	0.936	0.126	0.692	1.413	-0.560	3.075	0.906
	(0.012)	(0.013)	(0.034)	(——)	(0.869)	(0.738)	(1.10)	(1.832)	
Inflation	0.000	0.023	0.953	0.000	0.044	0.493	-1.318	0.178	0.000
	(——)	(0.007)	(0.046)	(——)	(0.151)	(0.156)	(0.546)	(2.371)	

Notes: This table reports SMM 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 variable. The parameter σ_η was fixed at $2\sigma_u$ and is reported here for reference only.

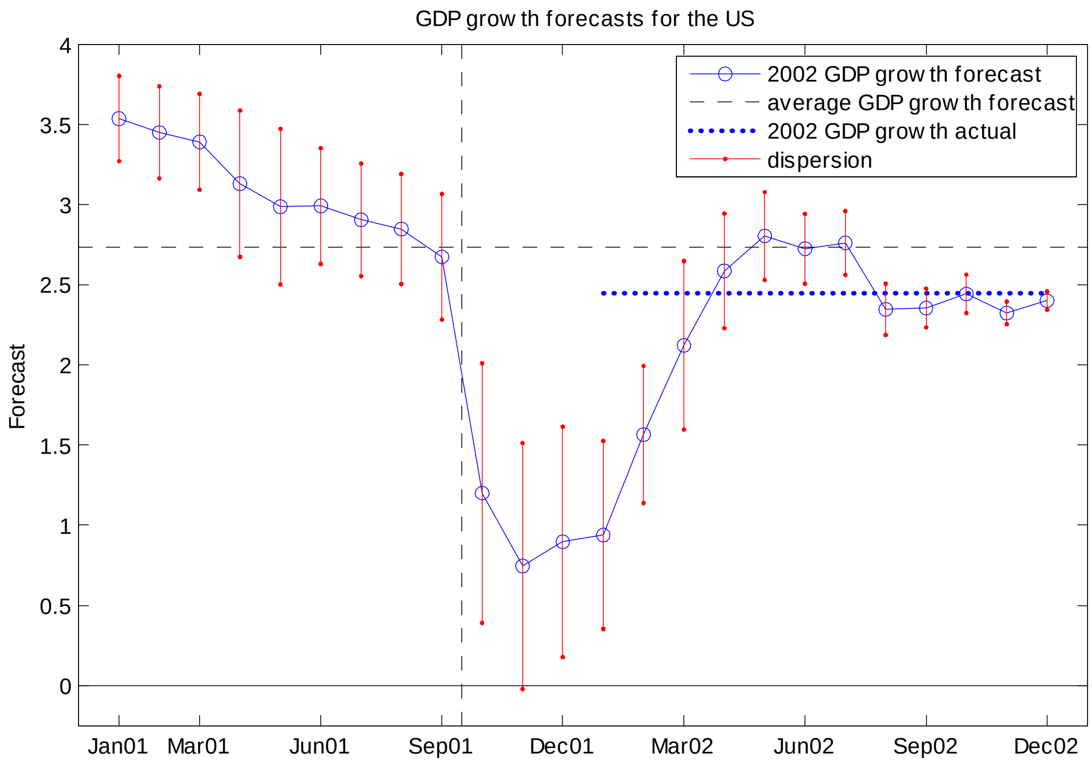

Figure 1: Evolution in consensus forecasts and forecast dispersions for US GDP growth in 2002, for horizons ranging from 24 months (January 2001) to 24 months (December 2002). The vertical lines plotted here are the consensus forecasts plus/minus the dispersion (measured in standard deviations).

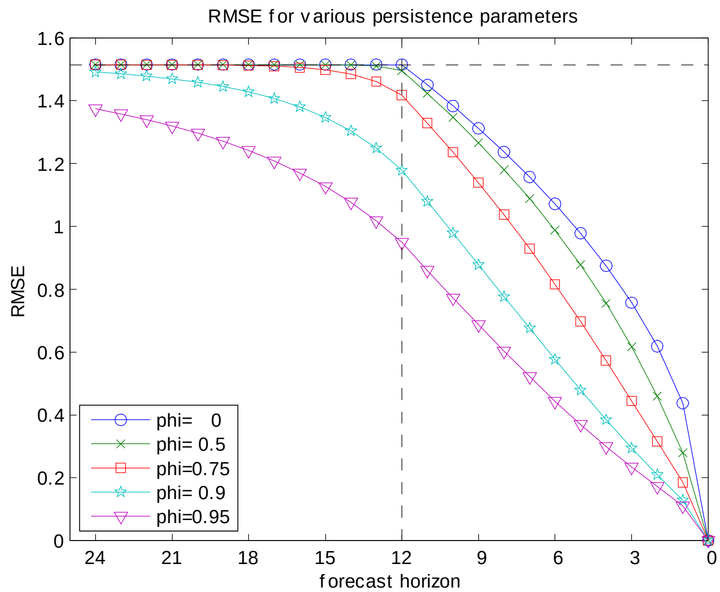

Figure 2: *Term structure of root-mean squared forecast errors for various degrees of persistence (ϕ) in the predictable component*

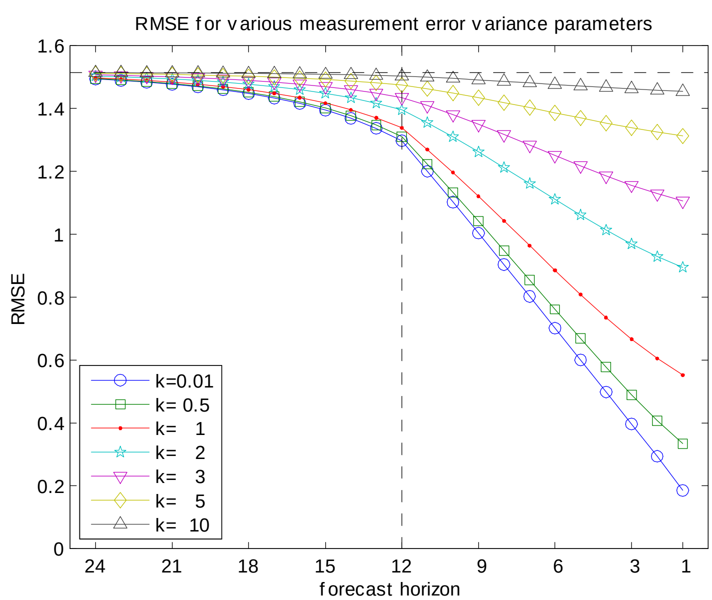

Figure 3: *Term structure of root-mean squared forecast errors for various degrees of measurement error in the predicted variable. In this example, the degree of measurement error is described as $\sigma_\eta = k\sigma_u$, where σ_u is the standard deviation of the unpredictable component of y_t.*

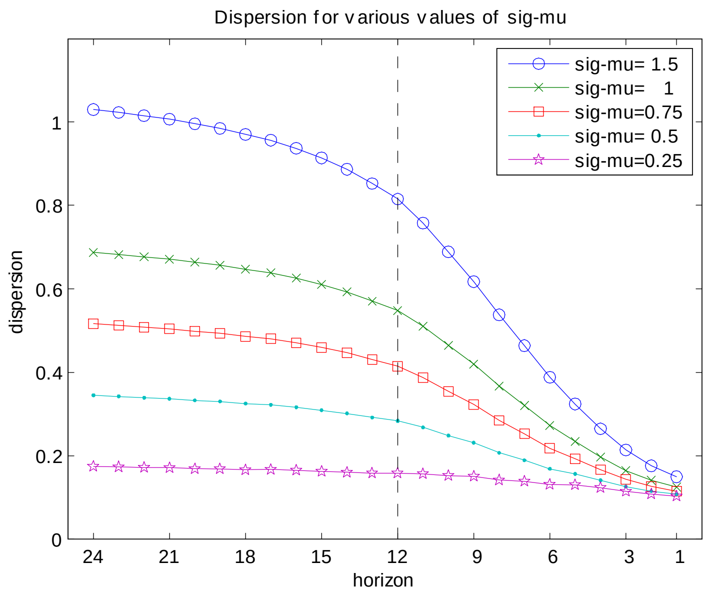

Figure 4: *Term structure of forecast dispersion for various levels of disagreement in beliefs about the long-run value of the target variable, measured by* σ_μ.

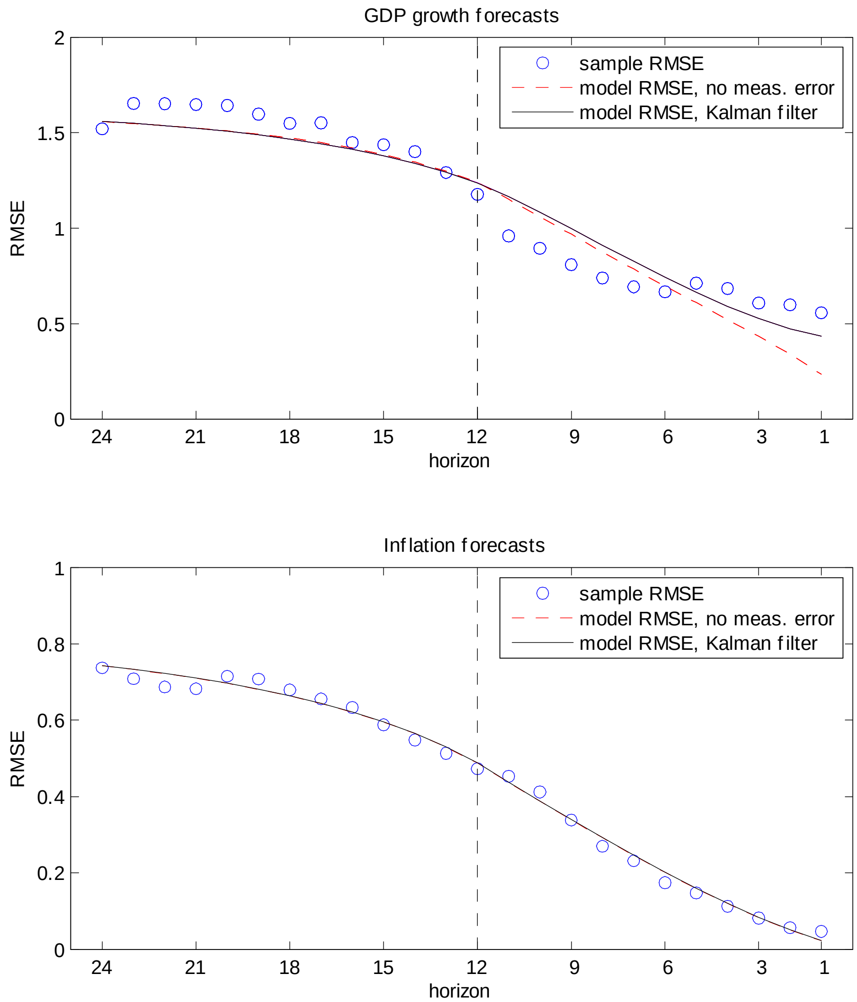

Figure 5: *Root mean squared forecast errors for GDP growth and Inflation in the U.S.*

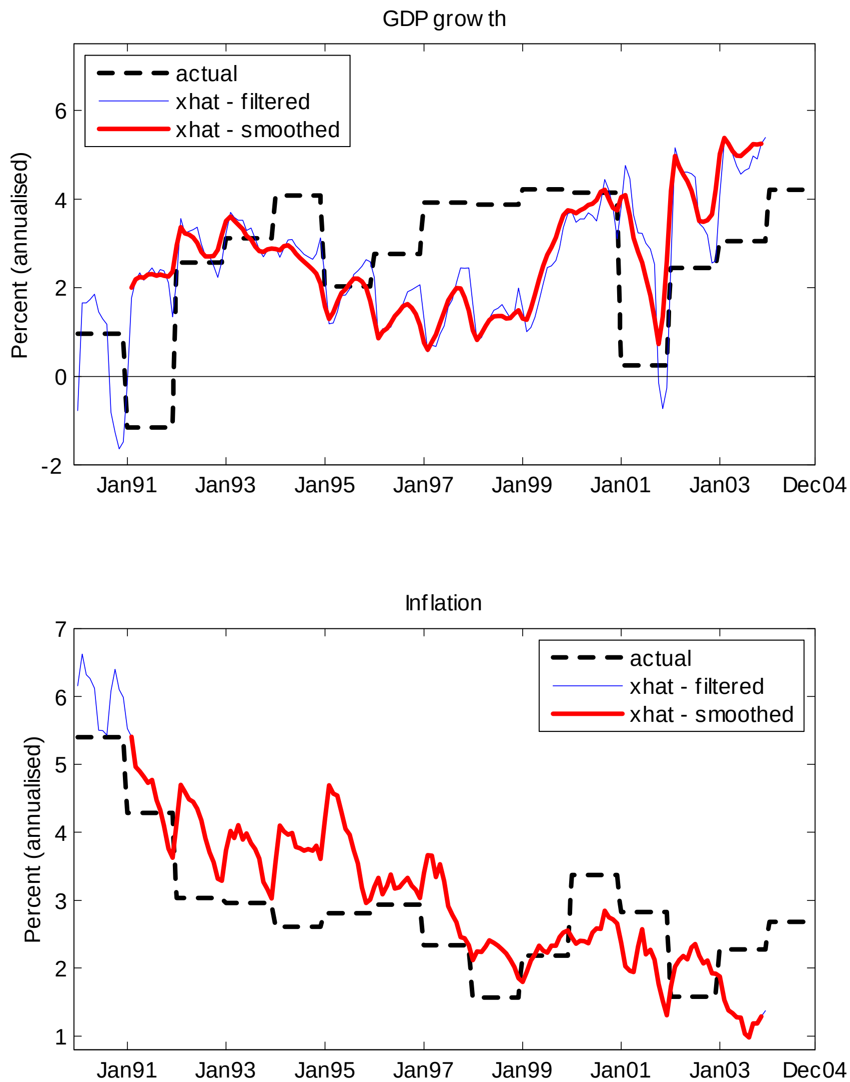

Figure 6: Estimates of the persistent component (xhat) of GDP growth and inflation for each month in the sample period, as implied by the observed forecasts and the estimated model for the term structure of forecast errors.

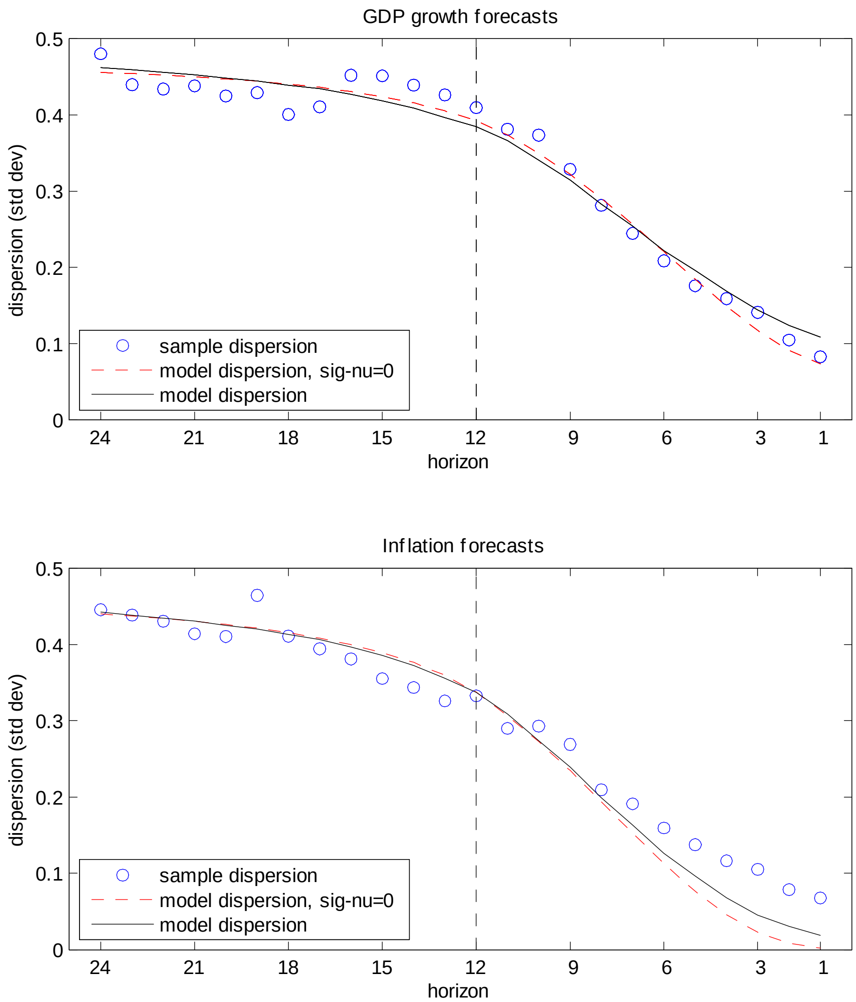

Figure 7: *Cross-sectional dispersion (standard deviation) of forecasts of GDP growth and Inflation in the U.S.*

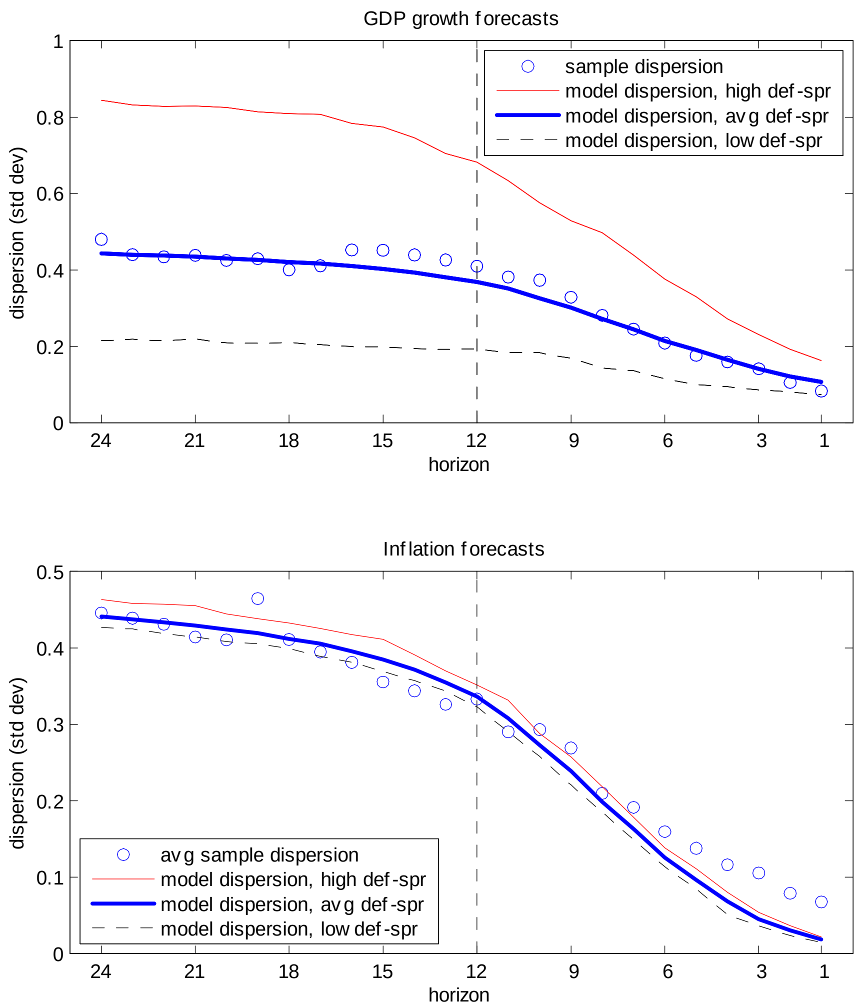

Figure 8: *Cross-sectional dispersion (standard deviation) of forecasts of GDP growth and Inflation in the U.S, when the default spread is equal to its sample average, its 95th percentile or its 5th percentile.*