University of Heidelberg

Department of Economics

Discussion Paper Series | No. 611

Forecast Performance, Disagreement, and Heterogeneous Signal-to-Noise Ratios

Jonas Dovern and Matthias Hartmann

March 2016

Forecast Performance, Disagreement, and Heterogeneous Signal-to-Noise Ratios

Jonas Dovern*[1,2] and Matthias Hartmann†[1,3]

[1]Alfred-Weber-Institute for Economics, Heidelberg University
[2]Department of Economics, RWTH Aachen University
[3]Department of Economics, University of Mannheim

February 20, 2016

Abstract

We propose an imperfect information model for the expectations of macroeconomic forecasters that explains differences in average disagreement levels across forecasters by means of cross sectional heterogeneity in the variance of private noise signals. We show that the forecaster-specific signal-to-noise ratios determine both the average individual disagreement level and an individuals' forecast performance: forecasters with very noisy signals deviate strongly from the average forecasts and report forecasts with low accuracy. We take the model to the data by empirically testing for this implied correlation. Evidence based on data from the *Surveys of Professional Forecasters* for the US and for the Euro Area supports the model for short- and medium-run forecasts but rejects it based on its implications for long-run forecasts.

JEL classification: E37, D80
Keywords: disagreement, expectations, imperfect information, signal-to-noise ratio.

1 Introduction

The dispersion of forecasts, of individual expectations, or of opinions in general has recently been in the focus of theoretical and empirical economics (Laster et al., 1999), finance (Harris and Raviv, 1993), and accounting (Barron et al., 1998). In particular, there is widespread evidence that macroeconomic forecasts differ widely across professional forecasters (see, e. g., Mankiw et al., 2003; Dovern, 2015). Analyzing disagreement in survey expectations may yield important insights about the expectation formation process. This, in turn, is important because expectations are key for understanding macroeconomic dynamics.

*Corresponding author: Bergheimer Str. 58, 69115 Heidelberg, Germany. Phone: +49-6221-54-2958. E-mail: jonas.dovern@awi.uni-heidelberg.de
†Financial support by the Fritz-Thyssen-Foundation is gratefully acknowledged.

Different mechanisms have been proposed in the literature to explain disagreement across forecasters. These include the use of different forecasting models (Branch, 2004), the existence of long-lasting beliefs due to historical experiences (Malmendier and Nagel, 2011), "sticky" information structures where agents do update their information sets only infrequently (Mankiw and Reis, 2002), or "imperfect" information structures where agents either receive (idiosyncratic) noisy signals about the state of the world or face limited cognitive capacities such that they are not able to process the full information available to them (Sims, 2003).

Kajal Lahiri has greatly contributed to this literature by suggesting models that incorporate several distinct explanations for disagreement across forecasters.[1] In Lahiri and Sheng (2008) and Lahiri and Sheng (2010a), disagreement across forecasters arises due to differences in individual forecasters' prior beliefs, differences in processing new information, and differences in the relative importance that forecasters attach to their priors and the new information. This framework is especially useful to describe situations in which a sequence of forecasts (with shrinking forecast horizon) is made for one particular random variable, such as an annual growth rate. Such forecasts are commonly referred to as "fixed event" forecasts.

For forecasts of the "fixed horizon" type which we analyze below, models with imperfect information structure have emerged as the most promising approach for modelling disagreement across forecasters (Coibion and Gorodnichenko, 2012; Andrade and Le Bihan, 2013).[2] Andrade et al. (2014) develop a non-stationary model with imperfect information that is able to account for the fact that forecasters disagree about the distant future. One attractive feature of this type of approach is that it can straightforwardly be introduced into structural dynamic stochastic general equilibrium (DSGE) models (see, e. g., Lorenzoni, 2009; Melosi, 2014; Nimark, 2014).

Existing models with imperfect information assume that agents are symmetric, i. e., all agents face the same information structure—in particular, they face homogeneous signal-to-noise ratios. Disagreement is generated only by the idiosyncratic flow of information. This implies that, on average, every agent deviates from the average forecast just about as far as any other agent. However, using a large data set with professional forecasts for different macroeconomic variables, Dovern (2015) finds that there is substantial persistence in the degree to which individual forecasters deviate from the average forecast (consensus) and also differences in the average level of disagreement across forecasters. To account for this fact, he suggests to generalize imperfect information models to allow for heterogeneous signal-to-noise ratios. The central hypothesis is that forecasters who receive very noisy signals about the state of the world form, on average, forecasts which deviate more strongly from the cross-sectional average forecast than predictions by individuals who face a high signal-to-noise ratio.

In this paper, we formally introduce such a model, describe its properties, and empirically test whether the idea of heterogeneous signal-to-noise rates finds support in the data. To this

[1] In general, Kajal Lahiri has paid much attention to the heterogeneity of forecasters in many of his papers (Davies and Lahiri, 1995; Lahiri and Liu, 2006; Lahiri and Sheng, 2010b; Lahiri et al., 2015).

[2] The alternative "sticky" information model of Mankiw and Reis (2002) has been empirically rejected as an appropriate model to describe the behavior of professional forecasters mainly due to the fact that the observed frequency of forecast updates is much higher than implied by this model (Dovern, 2013; Andrade and Le Bihan, 2013; Dovern et al., 2015).

end, we empirically test a direct implication of this type of model. This implication is the following: those forecasters with very noisy signals not only exhibit high levels of disagreement; the model also implies that they produce larger forecast errors than forecasters with high signal-to-noise ratios. Thus, the imperfect information model with heterogeneous signal-to-noise ratios implies that the cross-sectional correlation between forecasters' performance and their average level of disagreement is positive. This is the hypothesis that we test below.

Our main findings are as follows. First, we show that an imperfect information model with heterogeneous signal-to-noise ratios can generate substantial differences in individual average levels of disagreement. Second, we show that such model implies a strong positive correlation between these levels of disagreement and the individual forecast performance. Finally, we show empirically that this correlation is also observed in the data for short- and medium-run forecasts while the two measures corresponding to long-run forecasts are not significantly correlated.

The remainder of this paper is structured as follows. In Section 2, we describe theoretical models with imperfect information structure that allows us to analyze disagreement and forecast performance of individual forecasters under homogeneous as well as heterogeneous signal-to-noise ratios. In Section 3, we briefly describe the data that we use for the empirical analysis. In Section 4, we present empirical evidence based on the Surveys of Professional Forecasters from the US and the Euro Area. Section 5 concludes.

2 A Stylized Model with Imperfect Information

To describe the predictions of models with imperfect information, we use a simplified version of the one described in Coibion and Gorodnichenko (2012). We focus on a parsimonious data generating process (DGP) which can easily be extended to a model with richer time series dynamics. At first, we consider a model with homogeneous signal-to-noise ratios. In the next step, we allow for different signal-to-noise ratios across forecasters.

2.1 Homogeneous Signal-to-Noise Ratios

The fundamental (or true) DGP for the scalar random variable y_t is represented by the autoregressive process

$$y_t = \alpha y_{t-1} + \varepsilon_t, \tag{2.1}$$

where $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Due to information rigidities ("imperfect information"), each forecaster $i = 1,\dots, N$ observes a noisy signal about the state of the world which is given by

$$y_{i,t} = y_t + \eta_{i,t}, \tag{2.2}$$

where $\eta_{i,t} \sim \mathcal{N}(0, \sigma_\eta^2)$ denotes the idiosyncratic noise shocks with $\mathbb{E}[\eta_{it}\eta_{js}] = 0$ for $i \neq j$ or $s \neq t$. Furthermore, we assume that the noise shocks are independent of the fundamental shock ε_t. One can interpret the shock $\eta_{i,t}$ as representing differences in the information sets of forecasters. An alternative interpretation is that this term represents differences in forecasters' capabilities to filter/interpret publicly available information (Lahiri and Sheng, 2008). Note that in the case of homogeneous signal-to-noise ratios σ_η^2 is equal for all forecasters.

We denote by $y_{i,t|s}$ agent i's best estimate of y_t in the MSE-sense, conditional on information available in period s. Given the model defined by (2.1) and (2.2), agents optimally employ the Kalman filter (Kalman, 1960) to recursively update their estimates about y_t. Their predictions of the state variable y_t, conditional on information in $t-1$ are denoted $y_{i,t|t-1}$ and are recursively defined by

$$y_{i,t|t-1} = \alpha y_{i,t-1|t-1} \tag{2.3}$$

$$y_{i,t|t} = y_{i,t|t-1} + \left[P_{t|t-1}/\left(P_{t|t-1} + \sigma_\eta^2\right)\right]\left(y_{i,t} - y_{i,t|t-1}\right) \tag{2.4}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}^2/\left(P_{t|t-1} + \sigma_\eta^2\right) \tag{2.5}$$

$$P_{t+1|t} = \alpha^2 P_{t|t} + \sigma_\varepsilon^2. \tag{2.6}$$

Since agents are symmetric, the state variances do not differ across forecasters once the effect of initial values is washed out. The steady-state value of $P_{t+1|t} = P_{t|t-1} = \bar{P}$ is implicitly defined by the Riccati equation

$$P_{t+1|t} = \alpha^2\left(P_{t|t-1} - P_{t|t-1}^2/\left(P_{t|t-1} + \sigma_\eta^2\right)\right) + \sigma_\varepsilon^2, \tag{2.7}$$

which we obtain by using (2.5) and (2.6). In general, this forecast error variance is decreasing in the signal-to-noise ratio $(\sigma_\varepsilon^2/\sigma_\eta^2)$ and increasing in (the absolute value of) α. In what follows, we confine the index i to state forecasts $y_{i,t|t}$ which differ even in the steady state due to the idiosyncratic shocks $\eta_{i,t}$.

To analyze the relation between a forecaster's performance and her average level of disagreement, we compare the mean squared forecast error (MSFE) to the average squared deviation of an individual forecast from the average forecast. Defining the one-step-ahead forecast error as

$$e_{i,t+1|t} = y_{t+1} - y_{i,t+1|t} = y_{t+1} - \alpha y_{i,t|t}, \tag{2.8}$$

the expected MSFE in the model is given by the forecast error variance $\mathbb{E}[e_{i,t+1|t}^2] = \bar{P}$ for all i. Thus, obviously, the model predicts that forecasters have bad performance when the signal-to-noise ratio is low and when they have to forecast highly persistent processes.

To derive the model-implied unconditional disagreement, we start by formally defining individual disagreement in a particular period as

$$d_{i,t+1|t} = \left(y_{i,t+1|t} - (1/N)\sum_{j=1}^{N} y_{j,t+1|t}\right)^2. \tag{2.9}$$

Next, we show how the unconditional expectation of disagreement is related to the forecast performance. Consider (2.3) and abbreviate $\bar{\Phi} = \bar{P}/\left(\bar{P} + \sigma_\eta^2\right)$, such that

$$y_{i,t+1|t} = \alpha y_{i,t|t}$$

$$= \alpha \left(y_{i,t|t-1} + \bar{\Phi}(y_{i,t} - y_{i,t|t-1}) \right).$$

Rearranging and using (2.2) results in

$$y_{i,t+1|t} = \alpha \left((1-\bar{\Phi}) y_{i,t|t-1} + \bar{\Phi}(y_t + \eta_{i,t}) \right). \tag{2.10}$$

Employing the lag operator L that is defined such that $L^s y_{i,t+1|t} = y_{i,t+1-s|t-s}$, solve (2.10) and write

$$y_{i,t+1|t} = \frac{\alpha \bar{\Phi}}{1 - \alpha(1-\bar{\Phi})L}(y_t + \eta_{i,t}). \tag{2.11}$$

Combining (2.9) and (2.11) yields

$$\mathbb{E}[d_{i,t+1|t}] = \mathbb{E}\left[\frac{\alpha^2 \bar{\Phi}^2}{[1-\alpha(1-\bar{\Phi})L]^2} \left(y_t + \eta_{i,t} - (1/N) \sum_{j=1}^{N} (y_t + \eta_{j,t}) \right)^2 \right] \tag{2.12}$$

$$= \frac{\alpha^2 \bar{\Phi}^2}{[1-\alpha(1-\bar{\Phi})L]^2} \mathbb{E}\left[\frac{N^2 \eta_{i,t}^2 - 2N\eta_{i,t} \sum_{j=1}^{N} \eta_{j,t} + \left(\sum_{j=1}^{N} \eta_{j,t} \right)^2}{N^2} \right] \tag{2.13}$$

$$= \frac{\alpha^2 \bar{\Phi}^2}{[1-\alpha(1-\bar{\Phi})]^2} \frac{N-1}{N} \sigma_\eta^2 \tag{2.14}$$

where the last step follows since $\mathbb{E}[\eta_{it}\eta_{js}] = 0$ for $i \neq j$ or $s \neq t$. Replacing $\bar{\Phi}$ and assuming $(N-1)/N \approx 1$, we can write the relation between $\mathbb{E}[d_{i,t+1|t}]$ and $\mathbb{E}[e_{i,t+1|t}^2] = \bar{P}$ as

$$\mathbb{E}[d_{i,t+1|t}] = \frac{\alpha^2 \bar{P}/(\bar{P}+\sigma_\eta^2)\sigma_\eta^2}{\left[1 - \alpha\left(1 - \bar{P}/(\bar{P}+\sigma_\eta^2)\right)\right]^2}, \tag{2.15}$$

which is the same for all individuals in the homogeneous case. It can be shown that this expected level of disagreement is decreasing in the signal-to-noise ratio and increasing in α. Figure 1 visualizes this dependence.[3] Both effects are intuitive: if there is few information in the signals that forecasters receive, forecasts are far dispersed around the average forecast. Likewise, if forecasters have to deal with persistent processes, idiosyncratic noise has long-lasting effects on the forecasts and lead to a more dispersed forecast distribution. Note that in the extreme of a static DGP ($\alpha = 0$), disagreement disappears.

Taken together, the model with homogeneous signal-to-noise ratios predicts a positive relation between performance and disagreement across forecasting environments which differ in terms of the fundamental DGP and the signal-to-noise ratio. For each set of parameters, however, it predicts that every forecaster has the same expected performance and the same expected level of disagreement. The latter prediction is at odds with the empirical evidence provided in Dovern (2015) and we now turn to a version of the model with heterogeneous signal-to-noise ratios.

[3] In the simulation, α takes values from 0 to 0.95 and $\sigma_\varepsilon^2/\sigma_\eta^2$ ranges from 0.1 to 2.

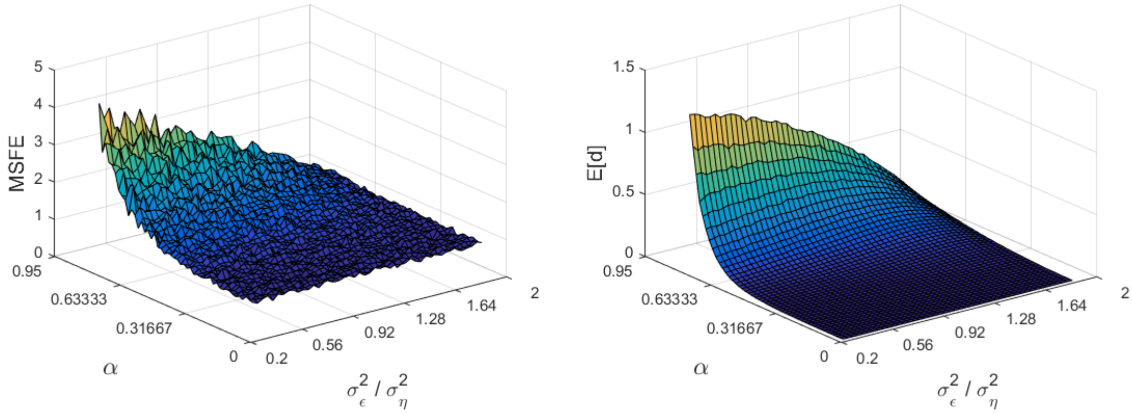

Figure 1: Effect of signal-to-noise ratio and persistence on MSFE and disagreement

2.2 Heterogeneous Signal-to-Noise Ratios

We make only one change to the model given by (2.1) and (2.2): the variance of the noise shocks are forecaster-specific in the alternative model, and we denote them by $\sigma^2_{i,\eta}$. Thus, there might be forecasters that, on average, receive small noise shocks (relative to the variance of the fundamental shocks) and those that receive large noise shocks. Since there is no closed-form solution for the relation between $\sigma^2_\varepsilon/\sigma^2_{i,\eta}$ and $\mathbb{E}[d_{i,t+1|t}]$, not the least because the expected degree of disagreement of each forecaster now depends on the signal-to-noise ratios of all forecasters, we use a simulation study to highlight the implications of the model. We are particularly interested in the correlation between $\mathbb{E}[e^2_{i,t+1|t}]$ and $\mathbb{E}[d_{i,t+1|t}]$.

To obtain a continuum of forecasters with varying signal-to-noise ratios, we assume that each forecaster draws an individual noise variance factor, δ_i, from a generalized beta distribution with mean 1 and variance σ^2_β, such that $\sigma^2_{i,\eta} = \delta_i \sigma^2_\eta$.[4] In practice, we use $\sigma_\beta = 1/3$ which is obtained by using a beta(2,6) distribution scaled by the factor of $(2+6)/2$.[5]

For $\alpha = 0.9$, $N = 40$, $\sigma^2_\varepsilon = 1$, and $\sigma^2_\eta = 1$, Figure 2 shows how different individual signal-to-noise ratios lead to different individual MSFEs and different individual levels of average disagreement.[6] The lower plot displays the relation between the individual signal-to-noise ratio, $\sigma^2_{i,\eta}$, and the average level of disagreement of each forecaster. The upper right plot displays the relation between $\sigma^2_{i,\eta}$ and the forecasters' MSFEs. Finally, the upper left plot displays the implied relationship between disagreement and forecast performance. It is evident that the model predicts a linear relation between those two variables.

So far, we have concentrated on one-step-ahead forecasts. But, of course, in reality we also observe forecasts with larger forecast horizons. So, which predictions does our model make for these? The first thing to note is that the optimal h-step-ahead forecast for each forecaster is

[4] By ensuring that the factors are equal to 1, on average, we can simulate cases which are comparable to the homogeneous case in the sense that the average signal-to-noise ratio is equal to that in the model with symmetric forecasters.

[5] Note that the scaling ensures that the mean is equal to 1. Results are the same qualitatively for other parameterizations of the beta distribution or other distributional assumptions.

[6] The simulation is based on $T = 5,000$ to ensure a good approximation of the expected moments.

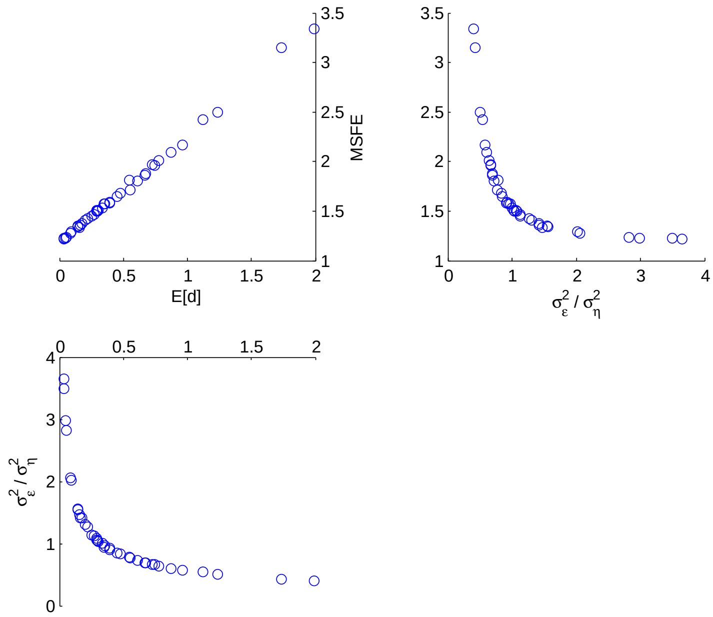

Figure 2: MSFEs and average disagreement levels under heterogeneous signal-to-noise ratios

given by $y_{t+h|it} = \alpha^h y_{t|it}$. If the DGP is stationary ($|\alpha| < 1$), the MSFEs increase in h up to a limiting value. This value is dervied by noting that for any h, the general form of (2.7) with heterogeneous signal-to-noise ratios is given by

$$P_{i,t+h|t} = \alpha^{2h} \left[P_{i,t|t-1} - P^2_{i,t|t-1} / \left(P^2_{i,t|t-1} + \sigma^2_{i,\eta} \right) \right] + \sigma^2_\varepsilon \sum_{j=0}^{h-1} \alpha^{2j}. \tag{2.16}$$

For long-run forecasts where $h \to \infty$,

$$P_{i,t+h|t} \to P_\infty = \frac{\sigma^2_\varepsilon}{1 - \alpha^2}, \tag{2.17}$$

the unconditional variance of the DGP. Note that this limit does not depend on $\sigma^2_{i,\eta}$. Thus, the model predicts that disagreement disappears for increasing forecast horizons. This can also be seen by noting that the long-run forecasts converge to the unconditional mean of the true DGP which, in our example, is given by 0. Figure 3 shows the relation between forecast performance and average individual disagreement at different forecast horizons for two values of α. We observe three things. First, the positive relationship between the two is confirmed by our simulations. Second, disagreement (and obviously MSFEs) is higher for persistent fundamental DGPs. Third, for increasing h all forecasters do indeed converge to the same performance level and disagreement converges to zero. Abstracting from the fact that all forecasters submit the same predictions if $h \to \infty$, the slope of the "lines" in the performance-disagreement plane does not depend on the horizon, i. e., the correlation between individual performance and disagreement is the same for all horizons.

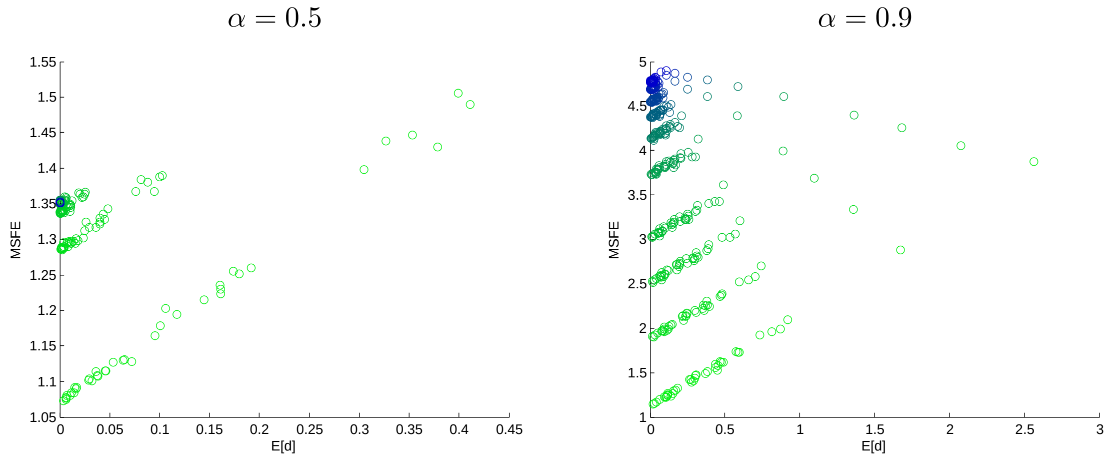

Figure 3: Distribution of MSFEs and levels of $\mathbb{E}[d_{i,h}]$ under heterogeneous signal-to-noise ratios (stationary DGP)

Empirical results of (Andrade et al., 2014; Dovern, 2015), however, suggest that long-run disagreement can be substantial and may even exceed short-run disagreement. Andrade et al. (2014) suggest to introduce a second informational rigidity into the model by assuming that there are two types of fundamental shocks—transitory shocks and permanent shocks (to the unconditional mean of y_t). To demonstrate what such an extension implies for the relation between individual disagreement and forecast performance, we replace (2.1) by

$$y_t = (1-\alpha)\mu_t + \alpha y_{t-1} + \varepsilon_t, \quad \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2) \tag{2.18}$$

$$\mu_t = \mu_{t-1} + e_t^\mu \quad e_t^\mu \sim \mathcal{N}(0, \sigma_\mu^2), \tag{2.19}$$

where ε and e_t^μ are independent of each other and we set $\mu_0 = 0$ as a starting value without loss of generality. In this model, forecasters have to infer whether changes in $y_{i,t}$ are due to noise shocks, temporary fundamental shocks, or permanent fundamental shocks. Since they come up with different estimates for μ_t and since their long-run forecasts will converge to these estimates, there is long-run disagreement in this model. How does this modification affect the cross-sectional correlation between individual disagreement and forecast performance? Figure 4 indicates how the latter changes relative to the stationary case.[7] Comparing Figure 4 to 3 shows that the strong positive relationship is preserved. Not surprisingly, forecast errors are higher, on average in the nonstationary case. Again, the slope of the horizon-specific “lines” in the performance-disagreement space does not differ across horizons.

3 Data

We primarily rely on data from the Survey of Professional Forecasters (SPF) for the US. In addition, we use long-run forecasts from the SPF for the Euro area (EA) to analyze the relation between individual disagreement and forecast performance for very large forecast horizons. Both

[7] We set $\sigma_\mu^2 = 0.5$.

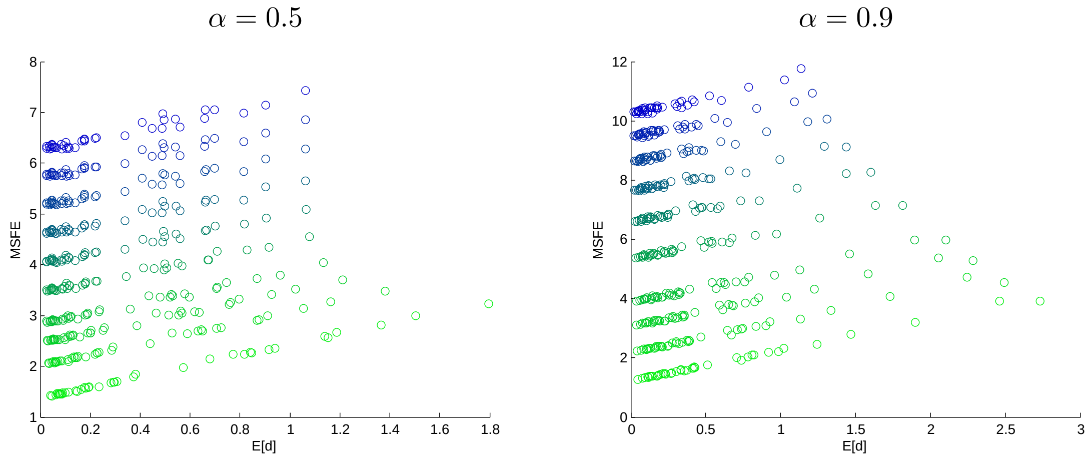

Figure 4: Distribution of MSFEs and levels of $\mathbb{E}[d_{i,h}]$ under heterogeneous signal-to-noise ratios (non-stationary DGP)

surveys are conducted at a quarterly frequency. Survey participant—mostly from research institutes and financial institutions—are questioned about their forecasts for the future development of a range of macroeconomic variables.

The US SPF had formerly been conducted by the American Statistical Association and the National Bureau of Economic Research and was taken over by the Federal Reserve Bank of Philadelphia in 1990. It covers a wide range of variables, but we concentrate on forecasts about the (annualized) growth rate of real gross domestic product (GDP)[8], the inflation rate, and the three-month treasury bill rate. In this paper, we focus on fixed-horizon forecasts for one to five quarters ahead. The sample periods range from 1968q4 to 2015q3 for GDP and the unemployment rate and from 1981q3 to 2015q3 for inflation and the interest rate.

The European SPF is conducted by the European Central Bank. It covers forecasts for real GDP growth, the inflation rate, and the unemployment rate. We focus on the those forecasts with a horizon of five years.[9] The sample covers the forecast periods from 1999q1 to 2015q3.

Table 1 displays some descriptive statistics for the different surveys, different variables, and different forecast horizons. The data from the US SPF offer a longer sample and a larger total number of forecasters in the cross-section compared to the EA SPF. The average number of forecasts in each survey (which determines how well disagreement measures can be estimated) is considerably larger in the EA SPF. Note that the pronounced differences in average disagreement and the average squared forecast error are due to the different nature of the forecasts: the data used from the US SPF refer to (annualized) quarter-on-quarter changes or quarterly averages, respectively, while the data from the EA SPF refer to annual averages.

[8]Survey waves before 1992q1 refer to gross national product (GNP) rather than GDP.

[9]Strictly speaking, the five-years-ahead forecasts are fixed-event forecasts made for the annual average of the forecast variables in a particular target year. This target year is changing in such a way that the forecast horizon varies between 21 and 18 quarters. Given the very long forecast horizon, it is unlikely that forecasts are affected by changes of the target year or small variations of the forecast horizon.

Table 1: Descriptive statistics for SPF data

	US					EA
GDP	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=20$
$\bar{y}_h$	2.4	2.7	3.0	3.2	3.2	2.1
$\bar{d}_h$	3.1	3.3	3.5	3.9	4.1	0.1
$\bar{e}_h$	11.6	12.6	13.9	15.7	14.7	6.1
T	188	188	188	188	183	63
$\emptyset N$	17.5	17.5	17.5	17.4	16.4	43.8
N	258	258	258	258	255	89
Inflation	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=20$
$\bar{y}_h$	2.8	2.8	2.9	3.0	3.0	1.9
$\bar{d}_h$	1.1	0.8	0.7	0.7	0.7	0.05
$\bar{e}_h$	9.5	7.9	7.5	7.7	8.3	1.2
T	137	137	137	137	137	63
$\emptyset N$	11.4	11.4	11.4	11.3	11.1	45.0
N	181	181	181	181	178	90
Unempl.	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=20$
$\bar{y}_h$	6.3	6.3	6.2	6.2	6.2	8.0
$\bar{d}_h$	0.02	0.05	0.1	0.1	0.2	0.6
$\bar{e}_h$	0.2	0.4	0.7	1.1	1.4	8.1
T	189	189	189	189	184	63
$\emptyset N$	17.8	17.8	17.8	17.7	16.7	39.8
N	260	260	260	260	258	81
Tbill	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=20$
$\bar{y}_h$	4.0	4.0	4.1	4.2	4.3	–
$\bar{d}_h$	0.06	0.15	0.2	0.3	0.4	–
$\bar{e}_h$	.5	.9	1.7	2.6	3.6	–
T	137	137	137	137	137	–
$\emptyset N$	11.2	11.2	11.2	11.2	10.8	–
N	182	182	182	182	180	–

Notes: The time-average of the consensus forecast ($\bar{y}_h$) is given by the *grand mean* across N and T, i. e., $\bar{y}_h = (1/NT)\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t+h|t}$.

4 Empirical Results

To see whether the positive correlation between individual disagreement and forecast performance, as implied by the imperfect information model with heterogeneous signal-to-noise ratios, is empirically supported or rejected, we start by looking at short- to medium-run forecasts from the US SPF. We only consider forecasters that submitted at least 4 forecasts.[10] For each of these forecasters (and separately for each of the variables and forecast horizons considered), we calculate the MSFE and the average level of disagreement, which we denote by $\widehat{\mathbb{E}[d_{i,h}]}$. We then run regressions of the form[11]

$$\widehat{MSFE}_{i,h} = \beta_0 + \beta_1 \widehat{\mathbb{E}[d_{i,h}]} + u_{i,h} \tag{4.1}$$

The results are given in Table 2. The results Most importantly, all estimated slope coefficients (β_1) are significantly different from 0 and often close to unity. This finding is robust across h and for all three variables and suggests that $\widehat{MSFE}_{i,h}$ and $\widehat{\mathbb{E}[d_{i,h}]}$ adjust proportionately—similar to what we find in the stylized theoretical model above. We also find a considerable degree of explanatory content of the disagreement statistics for the MSFE of all variables' forecasts. Furthermore, the R^2 statistics show that for inflation, unemployment and the tbill rate, the relation between $\widehat{MSFE}_{i,h}$ and $\widehat{\mathbb{E}[d_{i,h}]}$ becomes weaker as the forecast horizon increases. The pattern is less clear, however, in case of the R^2 for the regression based on GDP data. To see whether this observed trend causes disagreement and forecast performance based on very long-run forecasts to be unrelated to each other, we now turn to the EA SPF that provides information about five-year ahead forecasts.

Table 3 shows results based on the long-run forecasts from the EA SPF. Clearly, the correlation between $\widehat{MSFE}_{i,h}$ and $\widehat{\mathbb{E}[d_{i,h}]}$ is not significantly different from 0 here. This suggests that features other than informational imperfections as modelled above might be contributing to disagreement at very large forecast horizons.

5 Conclusion

In this paper, we show that a model with imperfect information structure and heterogeneous signal-to-noise ratios has a directly testable implication regarding the correlation between the individual average level of disagreement and the individual forecast performance. We confirm empirically and by means of simulation that forecasters with low signal-to-noise ratios deviate a lot from the consensus forecast and, at the same time, produce relative large forecast errors.

We conclude that imperfect information models with heterogeneous signal-to-noise ratios as proposed in Dovern (2015) are not rejected by the data based on the correlation between forecast performance and disagreement for short- to medium-run forecast. For very long-run forecasts,

[10] Results are robust against selecting a higher required number of observations.

[11] To limit the influence of outlier observations, we use the square root of $\widehat{MSFE}_{i,h}$ as the dependent variable in (4.1). The results based directly on $\widehat{MSFE}_{i,h}$ are qualitatively equivalent to those reported in the paper and are available from the authors upon request.

Table 2: Correlation between Performance and Disagreement (US SPF)

GDP	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
β_0	1.762***	1.779***	2.193***	1.860***	2.067***
	(14.76)	(13.60)	(16.53)	(14.85)	(14.10)
β_1	0.972***	0.955***	0.838***	1.168***	0.997***
	(14.82)	(13.74)	(12.43)	(18.50)	(13.95)
N	256	254	249	247	241
R^2	0.46	0.43	0.38	0.58	0.45
Inflation	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
β_0	1.481***	1.651***	1.811***	1.948***	2.007***
	(7.67)	(10.45)	(11.80)	(12.65)	(9.78)
β_1	1.309***	1.067***	0.879***	0.788***	0.921***
	(7.54)	(6.38)	(5.23)	(4.89)	(4.23)
N	181	179	178	177	173
R^2	0.24	0.19	0.13	0.12	0.09
Unemployment	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
β_0	0.287***	0.433***	0.503***	0.665***	0.719***
	(12.16)	(11.52)	(11.37)	(11.66)	(11.44)
β_1	1.227***	0.852***	1.059***	0.917***	0.982***
	(7.92)	(5.30)	(7.28)	(5.82)	(6.57)
N	258	256	251	249	244
R^2	0.20	0.10	0.18	0.12	0.15
Tbill	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
β_0	0.369***	0.453***	0.665***	0.627***	0.878***
	(7.82)	(8.70)	(9.44)	(6.08)	(6.64)
β_1	1.487***	1.528***	1.367***	1.928***	1.641***
	(10.71)	(13.66)	(11.08)	(11.55)	(8.84)
N	182	179	179	175	171
R^2	0.39	0.51	0.41	0.44	0.32

Notes: Dependent variables are the RMSFEs of the forecasters. The constant is denoted β_0, whereas β_1 refers to the correlation between the RMSFE and the disagreement of the forecasters.

Table 3: Correlation between Performance and Disagreement (EA SPF)

	Inflation	GDP	Unempl.
β_0	1.136***	2.515***	2.663***
	(17.37)	(14.04)	(11.09)
β_1	−0.409	−0.797	0.001
	(−1.41)	(−1.33)	(0.00)
N	85	83	78
R^2	0.02	0.02	0.00

Notes: Dependent variables are the RMSFEs of the forecasters. The constant is denoted β_0, whereas β_1 refers to the correlation between the RMSFE and the disagreement of the forecasters.

in contrast, the model's predictions are rejected by the data. Overall, this type of model remains promising for describing the expectation formation process in macroeconomic models.

We can think about a number of extensions of our work. First, one might use the information that the US SPF provides on the occupation of forecasters. In Lahiri and Sheng (2008), the authors briefly comment on how this sort of information may be employed to examine questions regarding the presence of strategic bias as described, e. g., by Laster et al. (1999). Second, the model-implied relation between the persistence of the fundamental DGP on the one hand and the level of disagreement on the other hand could be tested empirically in a cross-country setting. This might yield further insights about the way expectations are formed and about how disagreement is related to fundamental characteristics of macroeconomic dynamics. Finally, adapting the model to generate zero correlation between average individual disagreement and individual forecast performance for long-run forecasts remains a challenging task.

References

Andrade, P., Crump, R. K., Eusepi, S., and Moench, E. (2014). Fundamental disagreement. Working Papers 524, Banque de France.

Andrade, P. and Le Bihan, H. (2013). Inattentive professional forecasters. *Journal of Monetary Economics*, 60(8):967–982.

Barron, O. E., Kim, O., Lim, S., and Stevens, D. E. (1998). Using analysts' forecasts to measure properties of analysts' information environment. *Accounting Review*, pages 421–433.

Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal*, 114(497):592–621.

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116 – 159.

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

Dovern, J. (2013). When are GDP forecasts updated? Evidence from a large international panel. *Economics Letters*, 120(3):521–524.

Dovern, J. (2015). A multivariate analysis of forecast disagreement: Confronting models of disagreement with survey data. *European Economic Review*, 80(C):16–35.

Dovern, J., Fritsche, U., Loungani, P., and Tamirisa, N. T. (2015). Information rigidities: Comparing average and individual forecasts for a large international panel. *International Journal of Forecasting*, 31(1):144–154.

Harris, M. and Raviv, A. (1993). Differences of opinion make a horse race. *Review of Financial Studies*, 6(3):473–506.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45.

Lahiri, K. and Liu, F. (2006). Modelling multi-period inflation uncertainty using a panel of density forecasts. *Journal of Applied Econometrics*, 21(8):1199–1219.

Lahiri, K., Peng, H., and Sheng, X. (2015). Measuring uncertainty of a combined forecast and some tests for forecaster heterogeneity. CESifo Working Paper Series 5468, CESifo Group Munich.

Lahiri, K. and Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, 144(2):325–340.

Lahiri, K. and Sheng, X. (2010a). Learning and heterogeneity in GDP and inflation forecasts. *International Journal of Forecasting*, 26(2):265–292.

Lahiri, K. and Sheng, X. (2010b). Measuring forecast uncertainty by disagreement: the missing link. *Journal of Applied Econometrics*, 25(4):514–538.

Laster, D., Bennett, P., and Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics*, 114(1):293–318.

Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review*, 99(5):2050–2084.

Malmendier, U. and Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, 126(1):373–416.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics*, 117:1295–1328.

Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement on inflation expectations. *NBER Macroeconomics Annual*, pages 209–248.

Melosi, L. (2014). Estimating models with dispersed information. *American Economic Journal: Macroeconomics*, 6(1):1–31.

Nimark, K. P. (2014). Man-bites-dog business cycles. *American Economic Review*, 104(8):2320–2367.

Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.