

Conrad, Christian

Conference Paper

When does information on forecast variance improve the performance of a combined forecast?

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Time Series, No. D17-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Conrad, Christian (2017) : When does information on forecast variance improve the performance of a combined forecast?, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Time Series, No. D17-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/168200>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

When does information on forecast variance improve the performance of a combined forecast?

Christian Conrad* and Matthias Hartmann†

February 28, 2017

Abstract

In this paper, we show that the consensus forecast can be biased if some forecasts minimize an asymmetric loss function and the DGP features conditional heteroscedasticity. This result still holds if cross sectional heterogeneity in the loss function is allowed for, including the case where a share of the cross section of forecasts is produced under a symmetric (squared) loss objective. In this setting, the time-varying bias depends on the variance of the process. As a consequence, the information from the ex-ante variation of forecasts can be used to improve the predictive accuracy of the combined forecast. We consider two widely employed measures for the ex-ante forecast variance, namely the average over the variances of individual cross sectional units on the one hand and the cross sectional dispersion of point forecasts (“disagreement”) on the other hand. Both statistics are shown to be informative. The average individual variance provides the largest predictive content. Forecast survey data from the Euro area and the U.S. confirm the implications of the theoretical model.

JEL classification: C51, C53.

Keywords: Forecast uncertainty, Disagreement, Forecast combination, Linex loss.

*Department of Economics, Heidelberg University, Bergheimer Straße 58, 69115 Heidelberg, Germany, Email: christian.conrad@awi.uni-heidelberg.de; Phone: +49/6221/54/3173.

†Department of Economics, Heidelberg University, Bergheimer Straße 58, 69115 Heidelberg, Germany, Email: matthias.hartmann@awi.uni-heidelberg.de; Phone: +49/6221/54/2908. Financial assistance from the Fritz Thyssen Foundation is gratefully acknowledged.

1 Introduction

Macroeconomic forecasts often violate the assumption of unbiasedness in the sense of Mincer and Zarnowitz (1969). The presence of a bias term is often interpreted as a deviation from rationality (Conlisk, 1996). However, an alternative explanation that does not require the specification of a particular form of non-rationality is that forecasters attempt to minimize an asymmetric loss function. It has been shown by Granger (1969) how biased forecasts can arise as a result of asymmetric loss. Moreover, Christoffersen and Diebold (1997) show that in the presence of conditional heteroscedasticity, a bias that is generated by asymmetric loss can be time-varying due to the dependence of the forecasts on the conditional variance.

In this paper, we study if alternative measures of forecast variance are useful to improve the predictive accuracy of a combined forecast. To this end, we outline a model framework where a principal elicits the forecasts of a number of survey participants. These agents predict the realizations of a dynamic data generating process (DGP) that features conditionally heteroscedastic disturbances. Moreover, each individual agent might or might not be characterized by asymmetric loss. This is expressed by means of a relatively general loss function that is called the “linex-” criterion. This function nests the symmetric mean squared error function (Varian, 1974; Zellner, 1986). Consequently, agents’ forecasts contain a time-varying bias that depends on the disturbances’ conditional variance. Extending results of Granger (1969) Christoffersen and Diebold (1997), Capistrán and Timmermann (2009) and Issler and Lima (2009), we note that such a bias is also present when an equally weighted combination (“consensus forecast”) of the individual forecasts is computed.

In general, it is unclear how the unobservable ex-ante uncertainty of a forecast should be measured. By discussing distinct assumptions regarding the conditional variance process, we are able to compare the scope of alternative measures of forecast variance to improve the predictive accuracy of the biased consensus forecast. In particular, we explore various means to introduce heterogeneity into the conditional mean and the con-

ditional variance of forecasters' models and derive corresponding representations for two commonly employed ex-ante measures of forecast uncertainty. The first metric of forecast uncertainty is the average over individual agents' forecast variance. Zarnowitz and Lampros (1987) and Lahiri and Sheng (2010) discuss under which circumstances this statistic can be interpreted as the ex-ante uncertainty of the consensus forecast. The average individual forecast variance is typically regarded as the theoretically most appealing measure of uncertainty, however, it is only available if forecasters report density- or variance forecasts. This is the case, e.g., in the Survey of Professional Forecasters (SPF) that is conducted by the US-FED or the ECB. Another widely used uncertainty measure is the cross-sectional variance of agents' point predictions. Though the relation of this statistic to the average individual uncertainty is not always clearly stated, disagreement is widely employed if forecast survey data are examined that only elicit point forecasts.

This study contributes to the literature in the following ways. First, we extend the framework of related studies such as Capistrán and Timmermann (2009) or Issler and Lima (2009) to examine which measure of forecast variance should be preferred as a means to reduce the time-varying bias of the consensus forecast. In particular, we discuss under which conditions disagreement can serve as a substitute for the average individual forecast variance in this context. Second, we extend the results derived in Christoffersen and Diebold (1997), Batchelor and Peel (1998) and Capistrán and Timmermann (2009) by explicitly considering the consensus forecast in addition to analyzing individual forecasts. From the viewpoint of rationality testing, this is important since many related studies do not employ individual data but aggregate statistics such as the consensus forecast to test for the biasedness of forecasts by means of model-free procedures like the Mincer-Zarnowitz test. Apart from these considerations, we also contribute to a recent discussion which addresses the information content of density forecasts. Evaluating measures of uncertainty, Clements (2010) documents that the informative content of density forecasts for the purpose of deriving ex-ante uncertainty statistics that align with realized squared forecast errors is rather marginal. In this study, we evaluate the predictive content of

higher moments of density forecasts such as average individual uncertainty from the perspective of a distinct model framework. Moreover, similar as in Issler and Lima (2009), we propose the bias-adjusted consensus forecast as an improved forecast combination that is feasible in “large- N ” situations where the number of forecasts that are combined can exceed the time-series dimension. In this respect, we contribute to the rapidly evolving literature on combining forecasts in “big-data” environments. The augmented consensus is an alternative to computationally intensive approaches from the emerging machine learning literature that have been recently proposed for such problems.

Our empirical strategy is based on an examination of both individual and consensus forecasts. Thereby, we provide a detailed account of the heterogeneity at the individual level, the degree to which biases of single forecasts carry over to the consensus forecast and if such a bias can be reduced by commonly employed measures of uncertainty. By using the information from matched point- and density forecasts, we estimate the time-varying bias in terms of the deviation of the midpoint of the empirical distribution and the point forecast. This metric is not based on realized values and might therefore be less affected by distortions that arise merely as a result of a mismatch between the choice of the data vintage that is employed to compute realized forecast errors and the one that is actually targeted by individual forecasters or other statistical artifacts.

We find that both average individual uncertainty and disagreement are useful in that they help to improve the accuracy of the consensus forecasts. This is first derived theoretically. Then, we document by means of empirical data from the SPF of the U.S.-FED and the ECB that this claim holds for both inflation and GDP growth forecasts. Moreover, the predictive content of the uncertainty measures is documented both from an in-sample and an out-of-sample point of view. Furthermore, we highlight the practical relevance of the considered theoretical setting by examining several of its testable implications. This includes the biasedness of forecasts that is found in most cases for the consensus statistics and also for a considerable share of the individual forecasts for inflation and GDP growth. Another implication of the theoretical setting that is confirmed by the data is

that the mean squared error of the consensus forecast does not necessarily increase with the forecast horizon. This is a result of the combination of asymmetric loss and conditional heteroscedasticity, a finding that has been first derived by Patton and Timmerman (2007). However, to our knowledge, we are the first to document this behavior empirically.

The remainder of this paper is structured as follows. In Section 2, the theoretical model is introduced. Here, we distinguish between individual forecasters (agents) and the principal who uses the information they provide, which we also call a “researcher” in the following. Furthermore, we state assumptions regarding their respective information sets and forecast loss functions. Next, we compare the MSE of the combined forecasts to the MSE of the combination that is augmented by measures of forecast variance. To test the model’s implications, we employ data from the SPF that is introduced in Section 3. In Section 4, we report and discuss the results from the empirical investigation. Section 5 summarizes and concludes.

2 Forecasting environment

This section introduces the framework in which consensus forecasts are obtained. We consider a situation where a user of a forecast survey such as, e.g., the SPF observes a sequence of individual forecasts for the target variable y_t that is predicted with a forecast horizon of $\ell = 1$. In the following, we first describe the assumptions regarding the DGP, the model of survey participants and the loss functions of the survey user on the one hand and survey participants on the other hand. The distinction between the producers and the users of forecasts becomes particularly important if it is possible that these groups are characterized by distinct loss functions (Elliott et al., 2005). Second, alternative measures of ex-ante forecast variance are introduced. Third, we summarize and discuss the implications of this model framework. Towards the end of this section, we discuss some alternative specifications of the model framework.

2.1 Assumptions and notation

Assumption 1 (DGP) *Let y_t be generated by the process*

$$y_t = \mathbf{E}[y_t | \mathcal{F}_{t-1}] + \eta_t, \quad (1)$$

$$\eta_t = \sqrt{h_t \tau} Z_t, \quad (2)$$

where $Z_t \stackrel{iid}{\sim} \mathcal{N}(0, 1)$ and \mathcal{F}_{t-1} denotes the information set at $t-1$. Moreover, η_t represents a conditionally heteroscedastic process where h_t is measurable with respect to \mathcal{F}_{t-1} . We further assume that $\mathbf{E}[h_t] = 1$, such that $\mathbf{E}[\eta_t^2] = \tau$ with τ denoting a constant term.

The variance process in (2) could be described e.g. by means of a GARCH specification, e.g., $h_t = (1 - \alpha - \beta) + \alpha \frac{\eta_{t-1}^2}{\tau} + \beta h_{t-1}$, whereas the constant τ can be thought of as a scaling factor for the conditional variance of y_t .

Assumption 2 *The survey panel consists of $i = 1, \dots, N$ forecasters who believe in their own “model” of the DGP. The model of individual i reads as*

$$y_{i,t} = \mathbf{E}[y_t | \mathcal{F}_{t-1}] + \eta_{i,t}, \quad (3)$$

$$\eta_{i,t} = \sqrt{h_t \tau_i} Z_t, \quad (4)$$

where h_t and τ_i are measurable with respect to the information set \mathcal{F}_{t-1} .

Assumption 3 *The components h_t , τ_i and Z_t are independent of each other for all t and i . Moreover, $\tau_i \stackrel{iid}{\sim} \mathcal{D}(\tau, \sigma_\tau^2)$ where “ \mathcal{D} ” stands for some generic distribution.*

Assumptions 1 and 2 imply that, conditional on information available to forecaster i in period $t-1$, $\mathbf{E}[\eta_{i,t}^2 | \mathcal{F}_{t-1}] = h_t \tau_i$. This means that a forecaster is over(under-) confident relative to τ if $\tau_i < \tau$ ($\tau_i > \tau$). Next, we characterize the prediction problem of individual forecasters by a flexible, asymmetric loss function that is a generalization of (14):

Assumption 4 (Linex loss) *The “linex” function (Varian, 1974; Zellner, 1986) is given*

by

$$L_{i,t}^{linex}(e_{i,t|t-1}, \phi_i) = \frac{1}{\phi_i^2} [\exp(\phi_i e_{i,t|t-1}) - \phi_i e_{i,t|t-1} - 1], \quad (5)$$

where $e_{i,t|t-1} = y_t - \hat{y}_{i,t|t-1}$ is individual i 's one-step ahead forecast error and the parameter ϕ_i determines the degree of asymmetry.

In (5), positive (negative) values of $e_{i,t|t-1}$ lead to larger increases in the individual loss $L_{i,t}^{linex}(\cdot)$ if $\phi_i > 0$ ($\phi_i < 0$). Moreover, $L_{i,t}^{linex}(\cdot)$ converges to the squared loss $(e_{i,t|t-1})^2$ as $\phi_i \rightarrow 0$. The shape of $L_{i,t}^{linex}(e_{i,t|t-1}, \phi_i)$ for some representative values of ϕ_i is shown in Figure 1.

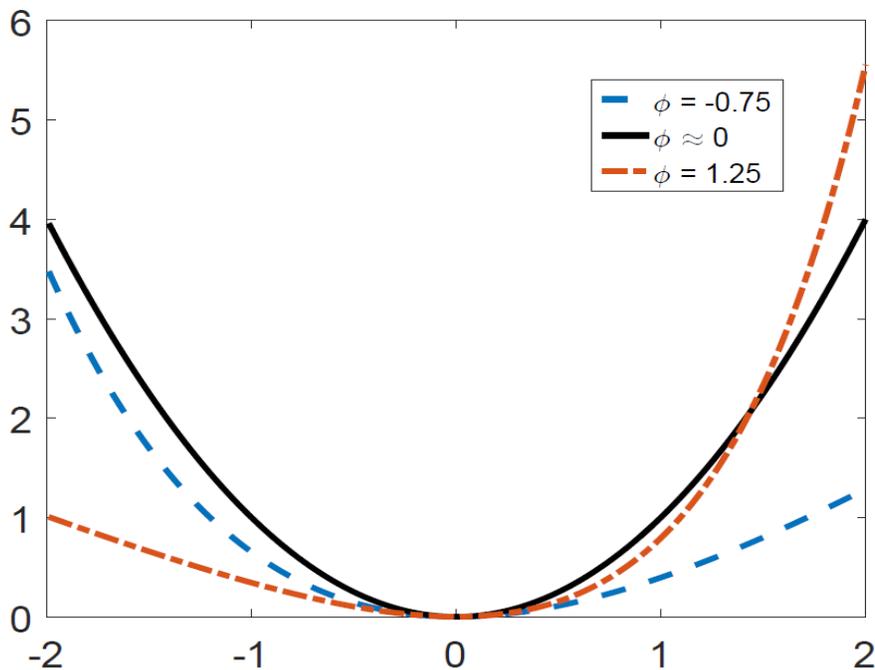


Figure 1: Linex loss function. If $\phi > 0$ ($\phi < 0$), under-predictions (over-predictions) are more costly. The solid line depicts the limiting case of symmetric (squared) loss, where $\phi \rightarrow 0$.

Individual survey forecasts $\hat{y}_{i,t|t-1}$ are often combined to obtain a single aggregated forecast. Such aggregates are often found to be more accurate than single forecasts from the underlying cross section (Timmermann, 2006).

Definition 1 *The equally weighted average of individual forecasts given by*

$$\hat{y}_{t|t-1}^{(C)} = \frac{1}{N} \sum_{i=1}^N \hat{y}_{i,t|t-1} \quad (6)$$

is referred to as the consensus forecast.

2.2 Individual forecasts and the consensus

If forecasters minimize (5), the optimal forecast depends on both the individual's conditional expectation and individually perceived conditional variance, i.e.

$$\hat{y}_{i,t|t-1} = \mathbf{E}[y_t | \mathcal{F}_{t-1}] + \frac{\phi_i}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{t-1}], \quad (7)$$

where the last term on the right hand side enters (7) only if $\phi_i \neq 0$, i.e. if the loss function is asymmetric. The forecast error associated with (7) is given by

$$e_{i,t|t-1} = y_t - \mathbf{E}[y_{i,t} | \mathcal{F}_{t-1}] - \frac{\phi_i}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{t-1}]. \quad (8)$$

which shows that the linex-optimal forecast contains a time-varying individual-specific bias term. According to Definition 1, this implies for the consensus that

$$\hat{y}_{t|t-1}^{(C)} = \frac{1}{N} \sum_{i=1}^N \mathbf{E}[y_t | \mathcal{F}_{t-1}] + \frac{1}{N} \sum_{i=1}^N \frac{\phi_i}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{t-1}]. \quad (9)$$

Hence, we have the following result.

Proposition 1 *Under the Assumptions 1, 2 and 4, the error of the consensus forecast is given by*

$$\begin{aligned} e_{t|t-1}^{(C)} &= y_t - \frac{1}{N} \sum_{i=1}^N \mathbf{E}[y_t | \mathcal{F}_{t-1}] - \frac{1}{N} \sum_{i=1}^N \frac{\phi_i}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{t-1}], \\ &= \eta_t - \frac{1}{2} h_t \overline{\phi \tau}, \end{aligned} \quad (10)$$

i.e., $e_{t|t-1}^{(C)}$ contains a time-varying bias term.

Apparently, this bias term depends on the average ex-ante uncertainty (i.e. forecast variance) of the individual forecasts $\hat{y}_{i,t|t-1}$. In the next step, we state two distinct ways how uncertainty can be measured. To facilitate the discussion in terms of notation, we consider the following special case.

Assumption 5 (*Homogeneity in the loss function*)

Let $\phi_i = \phi$ for all i in (5), i.e. all forecasters share the same form of asymmetric loss.

2.3 Ex-ante uncertainty and disagreement

Definition 2 *The average over individual forecaster's ex-ante forecast variance,*

$\mathbf{Var}[y_{i,t}|\mathcal{F}_{t-1}] = h_t\tau_i$, denoted as

$$\mathcal{U}_{t|t-1} = \frac{1}{N} \sum_{i=1}^N h_t\tau_i = h_t\bar{\tau} \quad (11)$$

is referred to as the average individual uncertainty.

For large cross sections, $\bar{\tau} \xrightarrow{p} \tau$, so that $\text{plim } \mathcal{U}_{t|t-1} = \mathbf{Var}[y_t|\mathcal{F}_{t-1}] = h_t\tau$. The metric $\mathcal{U}_{t|t-1}$ has been suggested as a quantification of the uncertainty of a consensus forecast by Zarnowitz and Lambros (1987) or Lahiri and Sheng (2010). However, it is only available as an empirical measure of forecast uncertainty if, say, a survey data set provides density- or variance forecasts. To obtain a proxy for $\mathcal{U}_{t|t-1}$ in the absence of such information, the cross-sectional variance of individual point forecasts is often employed.

Definition 3 *The cross-sectional variance of individual point forecasts is called disagreement. This measure is given by*

$$D_{t|t-1} = \frac{1}{N-1} \sum_{i=1}^N (\hat{y}_{i,t|t-1} - \hat{y}_{t|t-1}^{(C)})^2. \quad (12)$$

By plugging (7) and (9) into (12) and under Assumptions 1 to 5, we obtain

$$D_{t|t-1} = \frac{1}{N-1} \sum_{i=1}^N (\hat{y}_{i,t|t-1} - \hat{y}_{t|t-1}^{(C)})^2 = \frac{N}{N-1} \frac{\phi^2 h_t^2}{4} \left(\overline{\tau_t^2} - (\overline{\tau})^2 \right). \quad (13)$$

For sufficiently large cross sections, $\frac{N}{N-1} \approx 1$, hence we neglect this factor in (13) in the following. Note that $\mathcal{U}_{t|t-1}$ does not depend on the asymmetry parameter ϕ , but $D_{t|t-1}$ does.

2.4 The accuracy of the consensus forecast

Owing to the well documented success of this simple forecast combination approach in many empirical applications, users might be interested in the performance of the consensus forecast.¹

Assumption 6 *The researcher does not observe (3) and (4) but only the point forecasts $\hat{y}_{i,t|t-1}$ and density forecasts, $f_{i,t|t-1}$ that are elicited from the individual survey participants. Moreover, the user has a squared error loss function and minimizes the expected loss.*

$$\begin{aligned} L_t^{sq}(e_{t|t-1}^{(C)}) &= (y_t - \hat{y}_{t|t-1}^{(C)})^2 \\ &= \left(e_{t|t-1}^{(C)} \right)^2, \end{aligned} \quad (14)$$

where $e_{t|t-1}^{(C)} = y_t - \hat{y}_{t|t-1}^{(C)}$ denotes the consensus forecast error.

The symmetric criterion in (14) is widely employed in the forecasting literature. Although the consensus forecast may promise to deliver relatively accurate predictions of y_t , the researcher might figure that further improvements of L_t^{sq} are feasible, for example if information regarding forecast uncertainty is employed. As outlined above, such information is provided by the survey participants, for example, through the density forecasts $f_{i,t|t-1}$. As we will show in the following, the accuracy of $\hat{y}_{t|t-1}^{(C)}$ can be improved if not all fore-

¹Examples include combined forecasts for macroeconomic variables such as GDP growth and inflation (Stock and Watson, 2004), FX rate returns (Wright, 2008) or stock market volatility (Becker and Clements, 2008).

casting agents are confronted with the same symmetric loss function as the researcher. Based on the previously described setting, we can state the MSE of the consensus forecast as a next step. This serves as a reference point to which the augmented consensus forecasts that make use of uncertainty statistics are compared in subsequent sections. Given Assumption 5, the MSE of the consensus forecast can be written as

$$\begin{aligned}
\mathbf{E} \left[\left(e_{t|t-1}^{(C)} \right)^2 \right] &= \mathbf{E} \left[\left(\eta_t - \frac{\phi}{2} h_t \bar{\tau} \right)^2 \right] \\
&= \mathbf{E}[\eta_t^2] - 2 \mathbf{E}[\eta_t \frac{\phi}{2} h_t \bar{\tau}] + \mathbf{E} \left[\frac{\phi^2}{4} h_t^2 (\bar{\tau})^2 \right] \\
&= \mathbf{E}[\eta_t^2] + \frac{\phi^2}{4} \mathbf{E}[h_t^2] \mathbf{E}[(\bar{\tau})^2], \tag{15}
\end{aligned}$$

where the last equality follows from Assumptions 1 and 3. Due to the dependence of (10) on the average uncertainty $\frac{1}{N} \sum_{i=1}^N \mathbf{Var}[y_{i,t} | \mathcal{F}_{t-1}]$, this term is incorporated in the MSE of the consensus forecast.

2.5 Augmenting the consensus forecast

Notwithstanding the high predictive accuracy of the consensus forecast, improvements in forecast performance are still feasible under certain assumptions. In the following, we show how such accuracy gains can be achieved in the forecasting framework that is outlined above. Moreover, we show under which circumstances the uncertainty metric $\mathcal{U}_{t|t-1}$ provides at least as much information as $D_{t|t-1}$. The consensus forecast that is augmented by $\mathcal{U}_{t|t-1}$ is given by

$$\hat{y}_{t|t-1}^{(U)} = \hat{y}_{t|t-1}^{(C)} + \gamma^{(U)} \mathcal{U}_{t|t-1}.$$

Consequently, the MSE of the error from the combined prediction augmented by $\mathcal{U}_{t|t-1}$ can be written as

$$\begin{aligned}
\text{MSE}^{(\mathcal{U})} &= \mathbf{E} \left[\left(e_{t|t-1}^{(\mathcal{U})} \right)^2 \right] \\
&= \mathbf{E} \left[\left(\eta_t - \frac{\phi}{2} h_t \bar{\tau} - \gamma^{(\mathcal{U})} \mathcal{U}_{t|t-1} \right)^2 \right] \\
&= \mathbf{E}[\eta_t^2] + \frac{\phi^2}{4} \mathbf{E}[h_t^2(\bar{\tau})^2] + \gamma^{(\mathcal{U})} \phi \mathbf{E}[h_t^2(\bar{\tau})^2] + (\gamma^{(\mathcal{U})})^2 \mathbf{E}[h_t^2(\bar{\tau})^2]. \quad (16)
\end{aligned}$$

Next, consider $D_{t|t-1}$ as a predictor variable. In analogy to the case above, this yields

$$\begin{aligned}
\text{MSE}^{(D)} &= \mathbf{E} \left[\left(e_{t|t-1}^{(D)} \right)^2 \right] \\
&= \mathbf{E} \left[\left(\eta_t - \frac{\phi}{2} h_t \bar{\tau}_t - \gamma^{(D)} D_{t|t-1} \right)^2 \right] \\
&= \mathbf{E}[\eta_t^2] + \frac{\phi^2}{4} \mathbf{E}[h_t^2(\bar{\tau})^2] + \gamma^{(D)} \mathbf{E}[\phi h_t \bar{\tau}_t D_{t|t-1}] + (\gamma^{(D)})^2 \mathbf{E}[(D_{t|t-1})^2]. \quad (17)
\end{aligned}$$

Proposition 2 *The MSE-optimal choice of $\gamma^{(D)}$ in (17) results in*

$$\begin{aligned}
\text{MSE}_*^{(\mathcal{U})} &= \mathbf{E}[\eta_t^2] \quad \text{and} \\
\text{MSE}_*^{(D)} &= \mathbf{E}[\eta_t^2] + c, \quad c \geq 0. \quad (18)
\end{aligned}$$

With the result from (15), this implies that $\text{MSE}^{(C)} \geq \text{MSE}_^{(D)} \geq \text{MSE}_*^{(\mathcal{U})} = \text{MSE}^{(\text{optimal})}$.*

The proof of Proposition 2 is in the Appendix.

2.6 Generalized forecasting framework

In the following, we open up the model framework from the previous section and explore the implications of certain generalizations. First, we examine the MSE of the consensus at anticipation horizons larger than unity. The second extension we investigate are more general forms of loss functions than the linex criterion. Third, we consider the case where

the homogeneity assumption $\phi_i = \phi$ for all i is replaced by $\phi_i \neq \phi_j$ for $i \neq j$, while $\tau_i = \tau$ for all i .

2.6.1 Multi-horizon forecasts

In related studies that characterize professional forecasters, the comparison of outcomes across distinct anticipation horizons is a routine step. Since data sets such as the SPF provide forecasts not only for the forthcoming quarter but also several steps ahead, it is tempting to examine the explanatory content of models such as the one described in Assumption 1 for predictions at distinct horizons. To illustrate our arguments, we consider an example for the conditionally heteroscedastic process in (1). Thereby, we can derive a statement about the behavior of the consensus forecast at different horizons that is empirically evaluated in the subsequent section.

Assumption 7 (*Unit variance GARCH*) *The conditional variance h_t is specified by a GARCH(1,1) process, i.e.*

$$h_t = (1 - \alpha - \beta) + \alpha \frac{\eta_{t-1}^2}{\tau} + \beta h_{t-1} \quad (19)$$

under the restrictions $\alpha > 0$, $\beta \geq 0$ and $\alpha + \beta < 1$. Moreover, let $h_{t|t-\ell}$ denote the ℓ -step ahead forecast of the conditional variance based on (19), where $\ell \geq 1$.

Assumption 8 *The process y_t is weakly stationary and has the Wold representation*

$$y_t = \sum_{j=0}^{\infty} \psi_j \eta_{t-j}, \text{ with } \sum_{j=0}^{\infty} \psi_j^2 < \infty \text{ and } \psi_0 = 1.$$

Under assumptions 1, 4 and 5, the individual ℓ -step ahead forecast is given by

$$\hat{y}_{i,t|t-\ell} = \mathbf{E}[y_t | \mathcal{F}_{t-\ell}] + \frac{\phi}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{i,t-\ell}], \quad (20)$$

cf. Christoffersen and Diebold (1997). As for $\ell = 1$, the consensus forecast $\hat{y}_{t|t-\ell}^{(C)}$ takes the same form as (20).

Proposition 3 *The MSE of $\hat{y}_{t|t-\ell}^{(C)}$ can be either increasing or decreasing in ℓ .*

Proof. The MSE of $\hat{y}_{t|t-\ell}^{(C)}$ can be written as

$$\begin{aligned} \text{MSE}^{(C)} &= \mathbf{E} \left[\left(y_t - \mathbf{E}[y_t | \mathcal{F}_{t-\ell}] + \mathbf{E}[y_t | \mathcal{F}_{t-\ell}] - \hat{y}_{t|t-\ell}^{(C)} \right)^2 \right] \\ &= \mathbf{E} \left[(y_t - \mathbf{E}[y_t | \mathcal{F}_{t-\ell}])^2 \right] + \mathbf{E} \left[\left(\mathbf{E}[y_t | \mathcal{F}_{t-\ell}] - \hat{y}_{t|t-\ell}^{(C)} \right)^2 \right]. \end{aligned} \quad (21)$$

From (9) and given Assumption 5, it can be seen that the second term in (21) equals $-\frac{1}{N} \sum_{i=1}^N \frac{\phi}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{i,t-\ell}]$. Using Assumptions 2 and 8, we have

$$\text{MSE}^{(C)} = \sum_{j=0}^{\ell-1} \psi_j^2 \mathbf{E}[\eta_{t-j}^2] + \mathbf{E} \left[\left(-\frac{\phi}{2} h_{t|t-\ell} \bar{\tau} \right)^2 \right]. \quad (22)$$

Given Assumption 3, the second term in (22) can be written as

$$\mathbf{E} \left[\left(-\frac{\phi}{2} \mathbf{Var}[y_t | \mathcal{F}_{t-\ell}] \right)^2 \right] = \frac{\phi^2}{4} \mathbf{E}[h_{t|t-\ell}^2] \mathbf{E}[(\bar{\tau})^2].$$

Next, using Assumption 7 we obtain an expression for the ℓ -step ahead forecast of the conditional variance, i.e.

$$h_{t|t-\ell} = 1 - (\alpha + \beta)^\ell + \alpha(\alpha + \beta)^{\ell-1} \frac{\eta_{t-\ell}^2}{\tau} + \beta(\alpha + \beta)^{\ell-1} h_{t|t-1}.$$

Abbreviating $\alpha_\ell = \alpha(\alpha + \beta)^{\ell-1}$ and $\beta_\ell = \beta(\alpha + \beta)^{\ell-1}$, this leads to

$$\begin{aligned} \mathbf{E}[h_{t|t-\ell}^2] &= [1 - (\alpha + \beta)^\ell] + 2 [1 - (\alpha + \beta)^\ell] (\alpha_\ell + \beta_\ell) \\ &\quad + \frac{(1 - \alpha - \beta)(1 + \alpha + \beta)(3\alpha_\ell^2 + \beta_\ell^2 + 2\alpha_\ell\beta_\ell)}{1 - 3\alpha^2 - \beta^2 - 2\alpha\beta} \\ &\rightarrow 1 \text{ if } \ell \rightarrow \infty. \end{aligned}$$

With the parameter restrictions on α and β from Assumption 7, this shows that the second term in (22) is decreasing in ℓ . It can also be seen from (22) that the first term, in contrast, is increasing in ℓ . Hence the influence of ℓ on the MSE of $\hat{y}_{t|t-\ell}^{(C)}$ can point in either direction, as it is stated in Proposition 3. ■

2.6.2 Skewness in the loss function

In this section, we generalize Assumption 4 such that higher moments of y_t may affect the loss of individual forecasters. Thereby, it is possible to model the dependence of forecasts on time-varying higher moments such as, for example, conditional skewness. Time-variation in the third moment can be relevant because, for example, the probability of negative inflation shocks might be higher during recession periods. Similar as Christofferson and Diebold (1997), we consider an optimal forecast of general form. This prediction may depend on higher moments up to order three, say. Denoting the first three time-conditional moments of the DGP by $\mu_{t+1|t}$, $\sigma_{t+1|t}$ and $\mathcal{S}_{t+1|t}$, respectively, this forecast reads as

$$\hat{y}_{t+1|t} = G(\mu_{t+1|t}, \sigma_{t+1|t}, \mathcal{S}_{t+1|t}), \quad (23)$$

i.e. the optimal forecast depends on the first three conditional moments and where the function G is continuously differentiable. Note that, in contrast to (7), the relation in (23) will in general not be linear. A first-order Taylor series expansion of $\hat{y}_{t+1|t}$ around the unconditional moments yields

$$\hat{y}_{t+1|t} \approx G(\mu, \sigma, \mathcal{S}) + G'(\cdot) \begin{pmatrix} \mu_{t+1|t} - \mu \\ \sigma_{t+1|t} - \sigma \\ \mathcal{S}_{t+1|t} - \mathcal{S} \end{pmatrix} = \tilde{y}_{t+1|t}. \quad (24)$$

Defining $\nu = G(\mu, \sigma, \mathcal{S}) - G'(\mu, \sigma, \mathcal{S})(\mu + \sigma + \mathcal{S})$, $\beta_1 = G'(\mu, \sigma, \mathcal{S})$, $\beta_2 = G'(\mu, \sigma, \mathcal{S})$ and $\beta_3 = G'(\mu, \sigma, \mathcal{S})$, the approximation in (24) can be written as

$$\tilde{y}_{t+1|t} = \nu + \beta_1 \mu_{t+1|t} + \beta_2 \sigma_{t+1|t} + \beta_3 \mathcal{S}_{t+1|t}. \quad (25)$$

2.6.3 An alternative way to introduce heterogeneity

To consider a complementary case with respect to the one in Assumption 5, we alter this assumption in the following way.

Assumption 9 (*Homogeneity in the conditional variance, heterogeneity in ϕ_i*)

Let $\tau_i = \tau$ for all i . In contrast to Assumption 5, let $\phi_i \neq \phi_j$ for all $i \neq j$.

Under Assumption 9, uncertainty and disagreement are given by

$$\mathcal{U}_{t|t-1} = h_t \tau \tag{26}$$

and

$$D_{t|t-1} = \frac{N}{N-1} \frac{h_t^2 \tau^2}{4} \left(\overline{\phi_i^2} - (\bar{\phi})^2 \right), \tag{27}$$

respectively. Comparing (27) to (12) shows that under Assumption 9, $D_{t|t-1}$ is based on the cross sectional variation in the asymmetry parameter ϕ_i .

Proposition 4 *Given Assumption 9, we obtain*

$$\begin{aligned} MSE_*^{(U)} &= \mathbf{E}[\eta_t^2] \quad \text{and} \\ MSE_*^{(D)} &= \mathbf{E}[\eta_t^2] + c, \quad c \geq 0. \end{aligned} \tag{28}$$

With the result from (15), this implies that $MSE^{(C)} \geq MSE_*^{(D)} \geq MSE_*^{(U)} = MSE^{(optimal)}$.

This means that replacing assumption 5 by assumption 9 does not essentially affect the conclusions regarding the MSE of the augmented forecasts. The proof is found in the appendix.

3 Data

In the following empirical analysis, two data sets with survey forecasts are employed. The SPF is conducted both for the U.S. by the Federal Reserve Bank of Philadelphia and for the Euro area by the ECB. We refer to these data sets as U.S.-SPF and ECB-SPF. The

quarterly survey forecast data covers the period between 1992Q1 and 2013Q4 in case of the U.S.-SPF and the period between 1999Q1 and 2016Q2 in case of the ECB-SPF.

Both the U.S.-FED and the ECB provide both point- and density forecasts. The survey participants provide density forecasts for a range of potential outcomes regarding the growth rate of the GDP deflator. To this end, they assign probability values to a set of histogram bins which cover a range that is prespecified by the survey questionnaire.

The manual of the U.S.-SPF states that the reported data “gives the mean responses for the probabilities that the annual-average over annual-average percent change in the level of the GDP price index falls into a number of alternative ranges.” Here, the annual-average level is the average of the quarterly levels over the four quarters of a calendar year. In the case of the ECB-SPF, inflation and GDP growth point- and density forecasts refer to the year-on-year growth rates of the “harmonized index of consumer prices” (HICP) as the inflation metric and real GDP for the Euro area. A particularity of the density forecasts in the U.S.-SPF is that forecasts that are reported during successive quarters of a certain year relate to the growth rate of prices or GDP of that or the next entire year. This means that the forecast horizon declines with each quarter until a new target year starts. This is referred to as a “fixed-event” structure in the related literature. For this reason, we restrict the empirical analysis of the U.S.-SPF to forecasts from the 1st quarter of each year. Thereby, we obtain forecasts with a horizon of approximately one year. In contrast, the ECB-SPF provides one-year ahead forecasts at each quarter. After excluding forecasters who report less than 6 predictions, the total number of individual forecasts is 46 (48) for inflation (GDP growth) forecasts from U.S.-SPF, whereas the ECB-SPF comprises 65 (67) individual forecasts for inflation and GDP growth forecasts, respectively. Forecast errors are based on “real-time” data, i.e. we employ the earliest data vintage that is available. For the U.S., the data source is the real-time database of the Federal Reserve Bank of Philadelphia. For the Eurozone, data are drawn from the ECB real-time data set.

4 Empirical results

Since the SPF predictions are among the few data sources that provide density forecasts for macroeconomic variables, tests for the dependency of the consensus forecast on forecast uncertainty can be readily implemented. Thus, employing the SPF data provides an ideal empirical framework for testing the hypotheses that are implied by the theoretical model.

In this section, we collect evidence for biases in SPF forecasts. We then document to which extent such biases might be explained by the theoretical claims regarding asymmetric loss in the previous sections. We first examine the evidence for bias in individual and aggregate forecasts by means of conventional diagnostics that are based on forecast errors. As a second step, an introduction of the empirical measure by means of which we evaluate the relation between the consensus forecast and forecast variation follows. Here we compare the point forecasts to the ones that are derived as the midpoint of density forecasts. Third, we test if the theoretical findings regarding the predictive content of the distinct metrics of ex-ante variance are confirmed empirically by comparing the relative accuracy of the resulting augmented forecast combinations. Fourth, the examination of the MSE of the consensus forecast across anticipation horizons provides an assessment of the relevance of the model framework with asymmetric loss and conditional heteroscedasticity. Finally, we test if extensions of the baseline model are justified in terms of their explanatory content for the SPF data.

4.1 The evidence on bias in SPF forecasts

In this section, we examine two of the central claims of the theoretical model, namely that individual point forecasts and the consensus forecast are biased. We first consider individual data and then test for biasedness of the consensus forecast. The presence of a bias term in the individual and aggregate predictions is an implication of asymmetry in the loss function. This can be seen from equation (8). The presence of forecast bias is routinely tested in the framework of Mincer and Zarnowitz (1969), where the realization of the target variable is regressed on the forecast and a constant term, i.e. $y_t = \delta_{i0} + \delta_{i1}\hat{y}_{i,t|t-1} + \epsilon_{it}$.

The null hypothesis of unbiasedness is then specified as $H_0 : (\delta_{i0}, \delta_{i1})' = (0, 1)'$. Figure

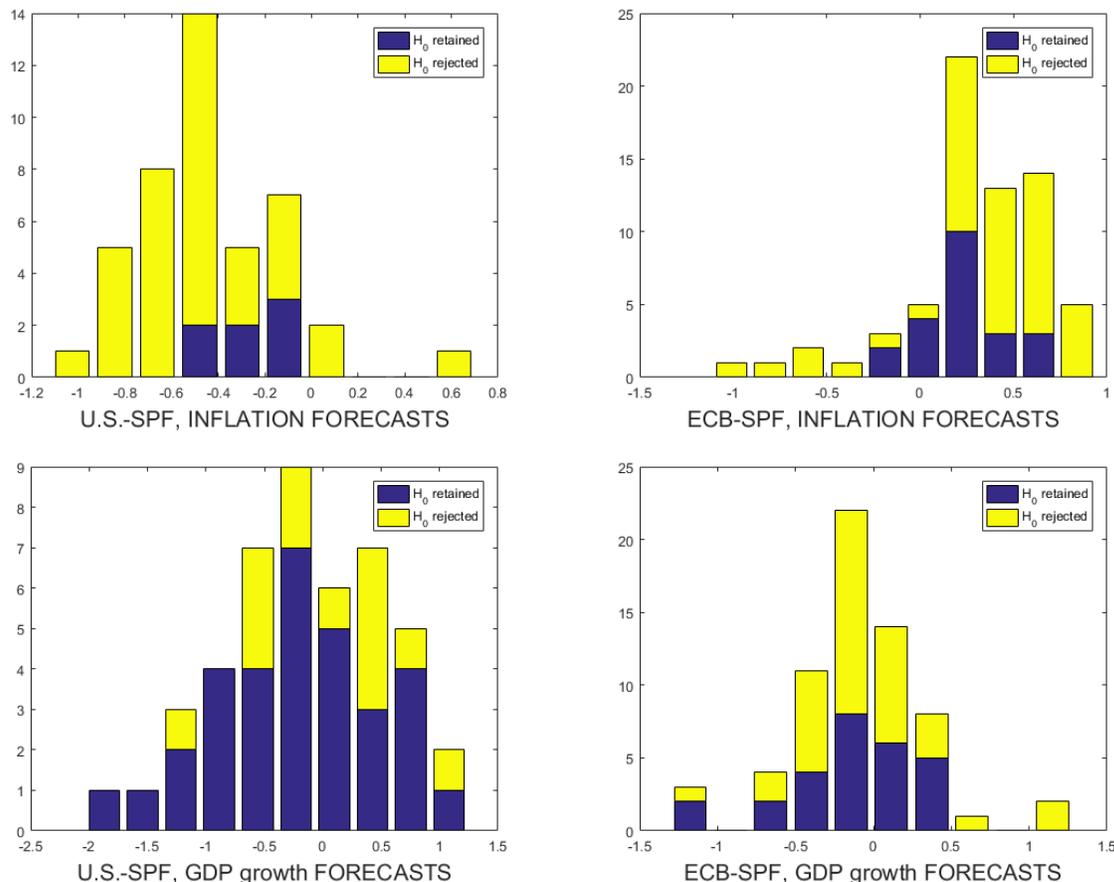


Figure 2: Histogram of individual average one-year-ahead forecast errors. Forecasts which are diagnosed as biased at the 10%-level by Mincer-Zarnowitz tests are marked in yellow.

2 contains histograms for the individual mean forecast errors. Similar as in Capistrán and Timmermann (2009), we distinguish between those average forecast errors where the Mincer-Zarnowitz test indicates the presence of a bias term and those where the respective H_0 cannot be rejected. The former cases are shown as yellow in the graphs. Comparing the findings for the U.S. to the ones for the Euro area, we find that the number of negative bias terms is larger except in the case of the inflation forecasts in the ECB-SPF. This means that in the majority of cases, inflation and GDP growth are more often over- than underpredicted. For example, a relatively large number of individual forecasts for inflation predictions in the U.S. are diagnosed to be, on average, significantly higher than the realized values. In contrast, the histograms for the GDP growth forecasts appear

slightly more symmetric, though the number of negative bias terms is also larger than the number of positive ones in this case. On average across forecasters, the share of biased forecasts seems to be larger in the U.S. sample than in the Euro area sample. Both for the inflation- and the GDP growth forecasts, the number of large (negative) bias terms is higher for the participants of the U.S.-SPF. Thus, there is relatively strong evidence for biasedness, both in inflation- and GDP growth forecasts. This is the case for both the U.S. and the Euro area sample and is in line with the theoretical result in equation (8).

Testing for biasedness in the consensus forecast proceeds in the same way as for the individual data. The results are summarized in Table 1.

Table 1: Tests of unbiasedness of consensus forecasts

Target variable	Data set	Mean error	δ_0	δ_1	F -statistic [p -value]
Inflation	U.S.-SPF	-0.44	1.21 (0.46)	0.25 (0.20)	17.94 [0.00]
	ECB-SPF	0.08	0.10 (1.28)	0.98 (0.77)	0.20 [0.81]
GDP growth	U.S. SPF	-0.35	3.96 (1.39)	-0.49 (0.48)	7.61 [0.00]
	ECB-SPF	-0.33	0.72 (0.49)	0.19 (0.28)	5.15 [0.01]

NOTE: Cell entries show 1) Mean errors and 2) coefficient estimates and test results from Mincer-Zarnowitz regressions for one-year-ahead forecasts. Newey-West HAC standard errors of coefficient estimates are reported in parentheses.

The coefficient estimates for δ_0 and δ_1 in the table show that according to the F -statistic for $H_0 : (\delta_0, \delta_1)' = (0, 1)'$, all forecasts except the inflation predictions for the Euro area are diagnosed as biased. This result is in line with the graphical display of the inflation forecasts in the Euro area in Figure 3. As the graph on the upper right hand side of the figure shows, a large number of forecasters underpredicted inflation until the years of the financial crisis (i.e. most forecast errors are positive). In contrast, overpredictions occurred more frequently afterwards. This may be the reason why no significant bias can be detected for inflation forecasts from the Euro area if coefficient estimates are computed based on information from both periods. This can also be seen from the average error statistic in the third column of Table 1, which is relatively close to zero. Moreover, the

estimates of the asymmetry parameter that are shown in Figure 5 also shows that the distribution of individual ϕ_i coefficients is most symmetric in the case of the Euro area inflation data. These estimates will be described in more detail below. A more clear-cut tendency to overpredict is found for the remaining forecasts, irrespective of the target variable or the data set that is employed. In these cases, we find more overpredictions and the average size of underpredictions is larger than the size of overpredictions.

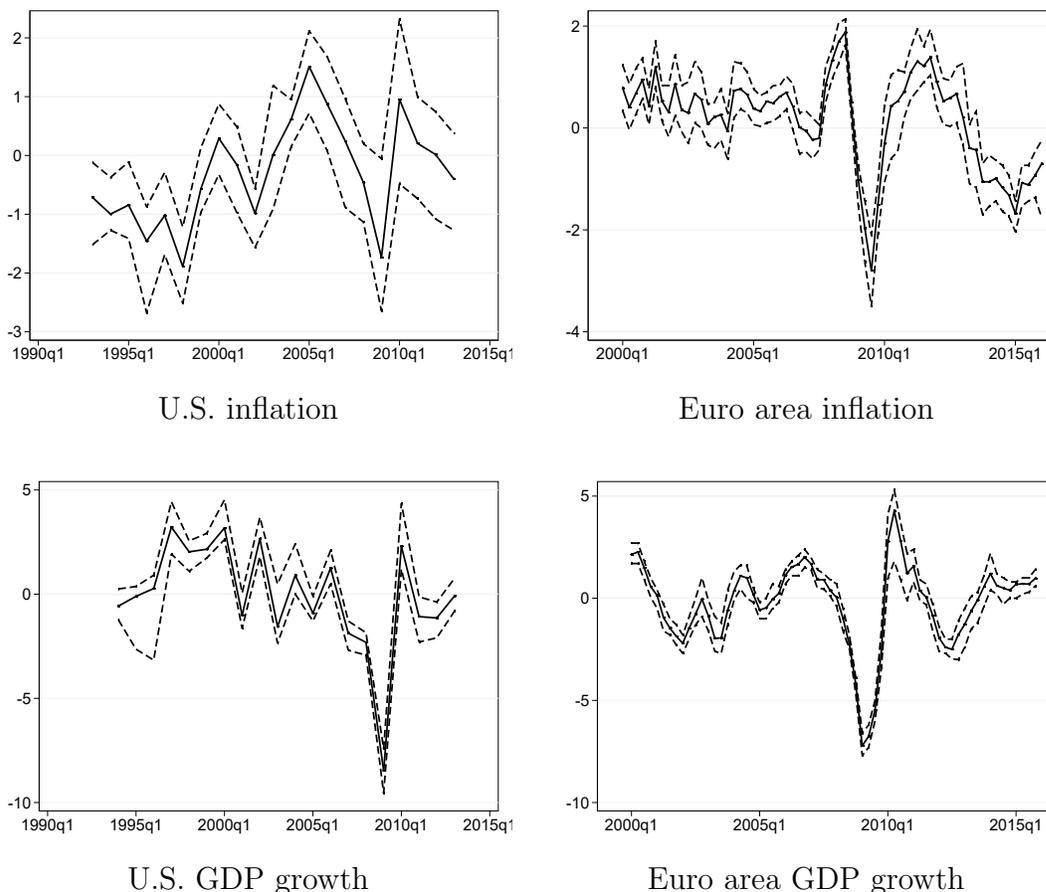


Figure 3: The graphs show one-year-ahead forecast errors of the consensus (solid line) and the 5%- and 95% quantiles of the cross section of individual forecasts (dashed lines).

Interestingly, the slope coefficients are rather small and in all cases insignificant. However, the inflation forecasts for the Eurozone are distinct also in this case. For these forecasts, the estimate of δ_1 is not significantly different from unity. The deviations from unity of the remaining coefficient estimates for δ_1 are significant at the 5% level in all three cases. In contrast, the assumption that any one of these slope coefficient estimates equals zero

cannot be rejected. This might be understood by recalling that the point forecasts include both $\mathbf{E}[y_t|\mathcal{F}_{t-1}]$ and $\mathbf{Var}[y_{i,t}|\mathcal{F}_{t-1}]$. Hence, in a Mincer-Zarnowitz regression, the point forecast may be regarded as an imprecisely measured (and biased) metric of $\mathbf{E}[y_t|\mathcal{F}_{t-1}]$. Pagan (1984) has shown that the coefficient estimates of such mismeasured variables will be biased towards zero.

4.2 Construction of empirical measures

In the following empirical analysis, we make use of the matched point- and density forecasts from the SPF data sets provided by the U.S. FED and ECB, which contain point forecasts $\hat{y}_{i,t|t-1}$ besides the density forecasts that are denoted as $f_{i,t|t-1}$. From individual density forecasts, the expected value as denoted $\mathbf{E}[\widehat{y_{i,t}}|\mathcal{F}_{t-1}]$ is obtained as an alternative to $\hat{y}_{i,t|t-1}$. It has been documented e.g. by Engelberg et al. (2009) that many forecasters in the SPF report $\hat{y}_{i,t|t-1}$ and $\mathbf{E}[\widehat{y_{i,t}}|\mathcal{F}_{t-1}]$ that do not coincide. Engelberg et al. (2009) conjecture that this can be due to asymmetric preferences which underly the point forecasts. If such a bias is present only in $\hat{y}_{i,t|t-1}$, the difference between the two forecasts should reveal this term. The histogram-based expected value $\mathbf{E}[\widehat{y_{i,t}}|\mathcal{F}_{t-1}]$ is computed as $\mathbf{E}[\widehat{y_{i,t}}|\mathcal{F}_{t-1}] = \sum_{k=1}^K P_{i,t|t-1}(k)\mu_{i,t|t-1}(k)$, where K denotes the number of histogram bins, $P_{i,t|t-1}(k)$ stands for the probability weight forecaster i attaches to bin k and $\mu_{i,t|t-1}(k)$ represents the midpoint of that bin.² Following Capistrán and Timmermann (2009) *inter alia*, we proceed by treating the leftmost and rightmost bin of the histogram as closed intervals with the same width as the intermediate intervals. Similar to Clements (2014), we define the difference between the mean of the density forecast and the point prediction as

$$\xi_{i,t|t-1} = \mathbf{E}[\widehat{y_{i,t}}|\mathcal{F}_{t-1}] - \hat{y}_{i,t|t-1}, \quad i = 1, \dots, N.$$

²Other methods that have been applied to derive density forecasts from the SPF questionnaires include the fitting of a normal- or a beta distribution (cf. Engelberg et al, 2009). The results that are documented in the following are qualitatively robust with respect to the consideration of these alternative methods.

If the term $\frac{\phi_i}{2} \mathbf{Var}[y_{i,t} | \mathcal{F}_{t-1}]$ is, as in the discussion of the Mincer-Zarnowitz results in section 4.1 interpreted as a measurement error within $\hat{y}_{i,t|t-1}$, then, according to Clements (2014), a forecast error $e_{i,t|t-1}$ can be regarded as a noisy proxy variable for $\xi_{i,t|t-1}$. Moreover, working with $\xi_{i,t|t-1}$ does not require to specify which data vintage of the target variable y_t is forecasted by the survey participants. A further advantage of $\xi_{i,t|t-1}$ in comparison to the forecast error is that it is unaffected by persistent shocks in y_t that generate biased forecast errors which might be mistakenly interpreted as the result of asymmetric loss (Clements, 2014). An alternative method to estimate asymmetry parameters that is based on forecast errors and the assumption that the conditional variance of the DGP and of the individual models can both be consistently estimated is proposed by Franses et al. (2016). In the empirical analysis of Franses et al. (2016) that is carried out for sales data, standardized forecast errors are employed to estimate the parameters of certain asymmetric loss functions, including linex, by means of OLS. In this case, the standardization might mitigate the aforementioned issue of persistence in the disturbance process. By employing model-based estimates of the conditional variance, the analysis of Franses et al. (2016) offers a view that is complementary to our empirical approach that uses survey data to compute the conditional variance.

To examine the scope of information from the histogram forecasts to improve the consensus, we extract the variances of individual density forecasts as a measure of ex-ante uncertainty. From the histogram forecasts, we obtain $\hat{u}_{i,t|t-1} = \mathbf{Var}[\widehat{y_{i,t} | \mathcal{F}_{t-1}}] = \sum_{k=1}^K P_{i,t|t-1}(k) \left(\mu_{i,t|t-1}(k) - \mathbf{E}[\widehat{y_{i,t} | \mathcal{F}_{t-1}}] \right)^2$. The second metric that is related to forecast variation is the disagreement statistic $D_{t|t-1}$. As in the case of uncertainty, we distinguish between the theoretical and the empirical measure and denote the empirical disagreement measure as $\hat{D}_{t|t-1}$ in the following.

4.3 Testing for a time-varying bias in SPF forecasts

The following section consists of two parts. After first examining the evidence for a time-varying bias in individual forecasts, we turn to aggregate statistics in the second step. If

Table 2: Empirical measures of forecasts and forecast variance

Statistic	Symbol	Source	Formula
<i>Consensus forecast</i>	$\hat{y}_{t t-1}^{(C)}$	$\hat{y}_{i,t t-1}$	$\frac{1}{N} \sum_{i=1}^N \hat{y}_{i,t t-1}$
<i>Avg. cond. expectation</i>	$\mathbf{E}[\widehat{y_t \mathcal{F}_{t-1}}]$	$f_{i,t t-1}$	$\frac{1}{N} \sum_{i=1}^N \mathbf{E}[\widehat{y_{i,t} \mathcal{F}_{t-1}}]$
<i>Histogram-point deviation</i>	$\xi_{t t-1}$	$\hat{y}_{i,t t-1}, f_{i,t t-1}$	$\mathbf{E}[\widehat{y_t \mathcal{F}_{t-1}}] - \hat{y}_{t t-1}^{(C)}$
<i>Avg. indiv. uncertainty</i>	$\hat{U}_{t t-1}$	$f_{i,t t-1}$	$\frac{1}{N} \sum_{i=1}^N \mathbf{Var}[\widehat{y_{i,t} \mathcal{F}_{t-1}}]$
<i>Disagreement</i>	$\hat{D}_{t t-1}$	$\hat{y}_{i,t t-1}$	$\frac{1}{N-1} \sum_{i=1}^N (\hat{y}_{i,t t-1} - \hat{y}_{t t-1}^{(C)})^2$
<i>Avg. indiv. skewness</i>	$\hat{S}_{t t-1}$	$f_{i,t t-1}$	$\frac{1}{N-1} \sum_{i=1}^N \mathbf{Skew}[\widehat{y_{i,t} \mathcal{F}_{t-1}}]$

(7) holds and $\hat{y}_{i,t|t-1}$ entails a bias due to asymmetric loss and conditional heteroscedasticity, we have that

$$\xi_{i,t|t-1} = -\frac{\phi_i}{2} \mathbf{Var}[\widehat{y_{i,t}|\mathcal{F}_{t-1}}].$$

To obtain a graphical impression of $\xi_{i,t|t-1}$, consider Figure 4 which depicts the average deviation, $\xi_{t|t-1}$ for the distinct data sets. It can be seen that the deviation between conditional expectation and the point forecasts is negative most of the time. This holds, with some exceptions, for both the inflation and the GDP growth forecasts. In case of the inflation forecasts in the Euro area, $\xi_{i,t|t-1}$ reverts its sign around the year 2009, shortly after the outbreak of the financial crisis. While $\mathbf{E}[\widehat{y_t|\mathcal{F}_{t-1}}]$ exceeds $\hat{y}_{t|t-1}^{(C)}$ until this date, the reverse holds for the rest of the sample. In the following analysis, we examine potential reasons for this behavior that is not found in this pronounced form for the other data sets. In the remaining cases, positive values of $\xi_{i,t|t-1}$ are limited to shorter periods, either at the start of the sample period as in the case of U.S. GDP growth or for single time instances around the most turbulent period after the Lehman bankruptcy incident in case of U.S. inflation or GDP growth in the Eurozone. The box plots for individual deviations, $\xi_{i,t|t-1}$, that are found on the right-hand side next to these graphs show that these findings are in most cases not generated by outliers, with the exception of GDP growth in the

Euro area, where the large spike in $\xi_{i,t|t-1}$ in 2008 suggests that for some individuals, the conditional expectation has been considerably larger than the point forecast. These forecasters issued considerably more accurate point forecasts during the most turbulent period. This is evidently not reflected in their histogram forecasts since $\xi_{i,t|t-1}$ takes high values in this particular time period. To examine the adjustment behavior of individual forecasters in more detail, we employ the following regression model:

$$\xi_{i,t|t-1} = \gamma_i \widehat{\mathbf{Var}[y_{i,t}|\mathcal{F}_{t-1}]} + \varepsilon_{i,t}. \quad (29)$$

From the estimates $\hat{\gamma}_i$ in (29), we obtain the asymmetry parameter as $\hat{\phi}_i = -2\hat{\gamma}_i$. Figure 5 shows a histogram of these coefficient estimates. Those estimates that are significant at least at the 10% level are shown in yellow. From the figure it can be seen that the asymmetry parameter is positive in the majority of cases. This finding also holds if only the significant estimates are considered. Given the definition of the linex loss function, this implies that most forecasters consider underpredictions as more costly. For both the inflation and GDP growth forecasts, coefficient estimates vary between -2 in the lowest cases up to a value of about 5. Computing cross sectional averages of the asymmetry parameter, we obtain the following numbers:

	U.S. Inflation	Euro area inflation	U.S. GDP growth	Euro area GDP growth
$\bar{\phi}$	0.73	0.04	1.12	0.32

This shows that the asymmetry parameter estimates, on average, suggest that underpredictions are considered as more costly. A value close to zero is obtained only for the inflation forecasts in the Eurozone. This also underlines the findings regarding the unbiasedness of these forecasts in contrast to the remaining data sets that are reported in Table 1 for the Mincer-Zarnowitz regressions. To consider an example for the loss that is implied by such number for the asymmetry parameter, a forecaster with $\phi = 1$ that produces a positive forecast error (i.e. an underprediction) equal to one percentage point in inflation experiences a loss of $L_{i,t}^{linex} \approx 0.72$, whereas $\phi = 5$ results in $L_{i,t}^{linex} \approx 5.69$.

4.4 The explanatory content of $\widehat{\mathcal{U}}_{t|t-1}$, $\widehat{\mathcal{D}}_{t|t-1}$ and $\widehat{\mathcal{S}}_{t|t-1}$

In the next step, we analyze the scope of density forecasts to improve the consensus forecast $\hat{y}_{t|t-1}^{(C)}$. Thus, we focus on aggregate statistics in the following. This has the advantage that individual forecasters who deliver very few observations can not exert an undue influence on the empirical findings. As noted by Elliott et al. (2008), findings from the U.S. SPF obtained at the level of individual forecasters have to be interpreted with care because for several forecasters, only a relatively small number of observations are available. The aggregate statistics that are derived from individual point and density forecasts are listed in Table 2.

Given the evidence of biasedness in individual forecasts, we now investigate the theoretical hypothesis that aggregate measures of forecast variance can explain the wedge between $\mathbf{E}[\widehat{y}_t | \mathcal{F}_{t-1}]$ and $\hat{y}_{t|t-1}^{(C)}$. For this purpose, we employ a time series regression model to reveal the relative explanatory content of the ex-ante measures of forecast variance $\widehat{\mathcal{U}}_{t|t-1}$ and $\widehat{\mathcal{D}}_{t|t-1}$. Moreover, to allow for the presence of higher moments in a (general) loss function, we examine if the average conditional skewness can also help to explain variation in $\xi_{t|t-1}$. The regression equation reads as

$$\xi_{t|t-1} = \gamma \widehat{\mathcal{U}}_{t|t-1} + \lambda \widehat{\mathcal{S}}_{t|t-1} + \varepsilon_t. \quad (30)$$

An analogously specified model is employed to examine the predictive content of $\widehat{\mathcal{D}}_{t|t-1}$ instead of $\widehat{\mathcal{U}}_{t|t-1}$. The results from (30) for the four distinct data sets are summarized in the tables 3 to 6. The coefficient estimates $\hat{\phi} = -2\hat{\gamma}_i$ for $\widehat{\mathcal{U}}_{t|t-1}$ are positive in all cases. This is in line with the findings for individual forecasts and the general level of $\xi_{t|t-1}$ in the majority of cases as it is depicted in Figure 4. The same holds in general also for the estimates obtained if $\widehat{\mathcal{D}}_{t|t-1}$ appears on the right hand side of (30), however, estimates are insignificant in two cases. We find that the estimates of $\widehat{\mathcal{U}}_{t|t-1}$ are throughout significant at the 1% level, irrespective of which data set is considered. Thus, $\widehat{\mathcal{U}}_{t|t-1}$ is related to the bias in the consensus forecast and should, consequently, increase the predictive accuracy

of $\hat{y}_{t|t-1}^{(C)}$. In contrast, $\hat{D}_{t|t-1}$ is insignificant if inflation in the U.S. or GDP growth in the Eurozone is considered. This is in line with the theoretical finding that $\hat{U}_{t|t-1}$ is more directly related to the time-varying bias term within $\hat{y}_{t|t-1}^{(C)}$ than $\hat{D}_{t|t-1}$.

As a next step, we turn to extensions of the baseline model. As outlined in section 2.6.2, the consideration of more general loss functions than (5) may lead to cases where the conditional skewness influences the point forecasts. For this reason, the empirical model in (30) allows for the influence of $\hat{\mathcal{S}}_{t|t-1}$.

The corresponding results are reported in column III of Tables 3 to 6. We find that the conditional skewness, $\hat{\mathcal{S}}_{t|t-1}$, has a significant explanatory content for $\xi_{t|t-1}$ for both inflation and GDP growth, yet only for the Euro area data set if $\hat{\mathcal{S}}_{t|t-1}$ is included as the only regressor. This specification should, however, be affected by an omitted variable bias if the conditional variance component is an important determinant of $\xi_{t|t-1}$, as it has been documented in columns I and II. Turning to the outcomes reported in column IV, however, shows that the joint consideration of $\hat{U}_{t|t-1}$ and $\hat{\mathcal{S}}_{t|t-1}$ leads to significant results for both variables. Interestingly, in all tables, the size of the coefficient of $\hat{U}_{t|t-1}$ is larger in column IV than in column II, where $\hat{U}_{t|t-1}$ is the only explanatory variable. From this we conclude that the consideration of a broader class of loss functions makes sense from an empirical point of view, yet this generalization strengthens the initial claim regarding the explanatory content of $\hat{U}_{t|t-1}$. The positive sign of most of the coefficient estimates for $\hat{\mathcal{S}}_{t|t-1}$ can be explained by noting that forecasters should be aware that the conditional skewness and the conditional expectation are positively related. For example, a positive skewness coefficient typically results in an expected value that exceeds the median. Given this effect, forecasters might tend to report point predictions that are closer to the conditional median than to the conditional mean, i.e. we expect a negative sign for the coefficient β_3 in (25). As a consequence, the coefficient of $\hat{\mathcal{S}}_{t|t-1}$ in a regression with $\xi_{t|t-1}$ on the left hand side should be positive as long as $\hat{\mathcal{S}}_{t|t-1}$ takes positive values. In this sense, the negative value of λ for the case of U.S. GDP growth makes sense since, in contrast to the other data sets, the time-average of $\hat{\mathcal{S}}_{t|t-1}$ in this data set is negative

and equals a value of -0.21.

Table 3: U.S. inflation

Col. no.	I	II	III	IV
Sample period: 1993Q1 to 2013Q1				
$\widehat{D}_{t t-1}$	-0.35** (0.15)			
$\widehat{U}_{t t-1}$		-0.28** (0.13)		-1.06*** (0.36)
$\widehat{S}_{t t-1}$			-0.26 (0.19)	1.69*** (0.52)
No. obs.	21	21	21	21

Note: Newey-West HAC standard errors in parentheses. Asterisks [***, **, *] denote significance at the 1%, 5% and 10% level, respectively.

Table 4: Euro area inflation

Col. no.	I	II	III	IV
Sample period: 1999Q1 to 2016Q1				
$\widehat{D}_{t t-1}$	-0.09* (0.05)			
$\widehat{U}_{t t-1}$		-0.04** (0.02)		-0.05*** (0.02)
$\widehat{S}_{t t-1}$			0.23*** (0.04)	0.25*** (0.04)
No. obs.	69	69	69	69

Note: Newey-West HAC standard errors in parentheses. Asterisks [***, **, *] denote significance at the 1%, 5% and 10% level, respectively.

4.4.1 An out-of-sample forecast competition

A further test of the theory outlined in this study is to examine out-of-sample forecasts of inflation and GDP growth. Candidate predictions that are compared are the average conditional mean, the consensus forecast and distinct extensions of the consensus forecast. As it is described in Proposition 2, we expect the lowest overall MSE for $\mathbf{E}[\widehat{y}_t | \widehat{\mathcal{F}}_{t-1}]$, and,

Table 5: U.S. GDP growth

Col. no.	I	II	III	IV
Sample period: 1993Q1 to 2013Q1				
$\widehat{D}_{t t-1}$	-1.09*			
	(0.56)			
$\widehat{U}_{t t-1}$		-0.38*		-0.92***
		(0.19)		(0.28)
$\widehat{S}_{t t-1}$			-0.10	-1.28***
			(0.31)	(0.49)
No. obs.	21	21	21	21

Note: Newey-West HAC standard errors in parentheses. Asterisks [***, **, *] denote significance at the 1%, 5% and 10% level, respectively.

Table 6: Euro area GDP growth

Col. no.	I	II	III	IV
Sample period: 1999Q1 to 2016Q1				
$\widehat{D}_{t t-1}$	0.05			
	(0.13)			
$\widehat{U}_{t t-1}$		-0.11***		-0.11***
		(0.03)		(0.03)
$\widehat{S}_{t t-1}$			1.18**	1.16**
			(0.58)	(0.51)
No. obs.	69	69	69	69

Note: Newey-West HAC standard errors in parentheses. Asterisks [***, **, *] denote significance at the 1%, 5% and 10% level, respectively.

moreover, that the consideration of $\widehat{D}_{t|t-1}$ and $\widehat{U}_{t|t-1}$ helps to increase the predictive accuracy of $\hat{y}_{t|t-1}^{(C)}$. In the following, we discuss the outcomes of a horse-race between these alternative forecasts. For this purpose, we compare the MSE of $\mathbf{E}[\widehat{y}_t | \mathcal{F}_{t-1}]$ and $\hat{y}_{t|t-1}^{(C)}$ to the MSE of augmented consensus forecasts that are computed, for example as

$$\hat{y}_{t|t-1} = \hat{y}_{t|t-1}^{(C)} + \hat{\gamma}\widehat{U}_{t|t-1} + \hat{\lambda}\widehat{S}_{t|t-1}, \quad (31)$$

where the estimates of γ and λ are obtained from regressions such as the one in (30). The findings of the forecast competition are found in Table 7. As expected, the average conditional mean is the most successful forecast. The finding of a high predictive accuracy of $\mathbf{E}[\widehat{y_t|\mathcal{F}_{t-1}}]$ is in line with Proposition 1 since this forecast does not include the bias term $\frac{\phi}{2}\widehat{\mathcal{U}}_{t|t-1}$. Consequently, the consensus that includes this term is throughout among the forecasts with the lowest forecast accuracy. Exceptions from this pattern are the forecasts for inflation in the Eurozone. A reason for this might be that individual asymmetry parameters are located on both sides of the cross sectional average $\bar{\phi}$, such that the bias terms cancel each other. From Figure 5 and the table for $\bar{\phi}$ on page 24, it can be seen that the distribution of individual estimates ϕ_i is rather symmetric in this case, whereas it is more tilted to one side for U.S. inflation and GDP growth in the Euro area. Here, $\mathbf{E}[\widehat{y_t|\mathcal{F}_{t-1}}]$ delivers more accurate forecasts than $\hat{y}_{t|t-1}^{(C)}$. Moreover, we find that the forecast $\hat{y}_{t|t-1}^{(C)}$ can be improved by either $\widehat{\mathcal{U}}_{t|t-1}$ or $\widehat{\mathcal{D}}_{t|t-1}$, though the former measure provides a higher information content. Again, the most pronounced results obtain for predictions of U.S. inflation. However, the forecast variance statistics $\widehat{\mathcal{U}}_{t|t-1}$ or $\widehat{\mathcal{D}}_{t|t-1}$ typically improve predictive accuracy, whereas the conditional skewness delivers mixed results that improve forecasts to a considerable extent only for U.S. inflation when it is considered in combination with $\widehat{\mathcal{U}}_{t|t-1}$. Overall, we conclude that the information from density forecasts can be employed fruitfully to improve the accuracy of the consensus. This is possible in particular in those cases where a considerable part of the cross section can be characterized by a similar asymmetry in the forecast loss function. These results clearly support the implications of the theoretical model that is described in this work.

4.5 Multi-step forecasts

A further testable implication of the considered model is that the MSE of the consensus forecast need not be an increasing function of the forecast horizon. Patton and Timmermann (2007) were among the first to show that such a behavior can be the result of asymmetric loss and a DGP with conditionally heteroscedastic disturbances. However,

Table 7: Forecast Evaluation.

$\bullet =$	$\widehat{\text{MSE}}^{(\bullet)}$	$\widehat{\text{MSE}}^{(E)} - \widehat{\text{MSE}}^{(\bullet)}$	$\widehat{\text{MSE}}^{(C)} - \widehat{\text{MSE}}^{(\bullet)}$	$\widehat{\text{MSE}}^{(U)} - \widehat{\text{MSE}}^{(\bullet)}$	$\widehat{\text{MSE}}^{(D)} - \widehat{\text{MSE}}^{(\bullet)}$	$\widehat{\text{MSE}}^{(S)} - \widehat{\text{MSE}}^{(\bullet)}$
U.S. sample: 1993Q1 to 2013Q1. Inflation						
$\mathbf{E} [y_{i,t} \widehat{\mathcal{F}}_{t-1}]$	0.62	0
$\hat{y}_{t t-1}^{(C)}$	0.87	-0.2438	0	.	.	.
\widehat{U}	0.82	-0.1955	0.0483	0	.	.
\widehat{D}	0.87	-0.2470	-0.0032	-0.0515	0	.
\widehat{S}	0.87	-0.2533	-0.0095	-0.0578	-0.0063	0
$\widehat{U} + \widehat{S}$	0.62	-0.0048	0.2390	0.1907	0.2422	0.2485
Euro area sample: 1999Q1 to 2016Q1. Inflation						
$\mathbf{E} [y_{i,t} \widehat{\mathcal{F}}_{t-1}]$	0.85	0
$\hat{y}_{t t-1}^{(C)}$	0.86	-0.0069	0	.	.	.
\widehat{U}	0.86	-0.0078	-0.0009	0	.	.
\widehat{D}	0.86	-0.0088	-0.0019	-0.0010	0	.
\widehat{S}	0.85	-0.0012	0.0057	0.0066	0.0077	0
$\widehat{U} + \widehat{S}$	0.86	-0.0018	0.0051	0.0060	0.0070	-0.0006
U.S. sample: 1993Q1 to 2013Q1. GDP growth						
$\mathbf{E} [y_{i,t} \widehat{\mathcal{F}}_{t-1}]$	6.76	0
$\hat{y}_{t t-1}^{(C)}$	6.31	0.4456	0	.	.	.
\widehat{U}	6.28	0.4802	0.0347	0	.	.
\widehat{D}	6.20	0.5616	0.1160	0.0814	0	.
\widehat{S}	6.36	0.4035	-0.0421	-0.0768	-0.1581	0
$\widehat{U} + \widehat{S}$	6.76	0.0031	-0.4425	-0.4771	-0.5585	-0.4004
Euro area sample: 1999Q1 to 2016Q1. GDP growth						
$\mathbf{E} [y_{i,t} \widehat{\mathcal{F}}_{t-1}]$	4.27	0
$\hat{y}_{t t-1}^{(C)}$	4.31	-0.0389	0	.	.	.
\widehat{U}	4.30	-0.0294	0.0095	0	.	.
\widehat{D}	4.30	-0.0292	0.0097	0.0002	0	.
\widehat{S}	4.32	-0.0446	-0.0058	-0.0152	-0.0154	0
$\widehat{U} + \widehat{S}$	4.31	-0.0348	0.0041	-0.0054	-0.0056	0.0099

little is known about the empirical relevance of this argument. To the best of the author's knowledge, the current study is the first which documents that a non-increasing behavior of the MSE of a consensus forecast can be found for SPF data. Figure 6 depicts the MSE of the consensus forecast over the forecast horizon. The data sets represented by the

graphs are forecasts for GDP growth in the U.S. and inflation and GDP growth for the Euro area. For the U.S.-SPF, long-term inflation forecasts with a horizon of more than 2 years are only available for the most recent years which precludes to employ this data for such a purpose. We find a similar pattern for the MSE of all the consensus forecasts. The MSE of all forecasts initially increases with the forecast horizon. For horizons between six and less than twelve quarters, the increase is less pronounced. This might reflect the sparsity of relevant information at these horizons. At a horizon of 12 quarters, all forecasts reach a maximum value for the MSE. Beyond this horizon, the employed data sets do not provide forecasts except for the very long term. Comparing the MSE of these predictions with a horizon of approximately five years to the remaining ones, it turns out that the long-term MSE is only slightly larger than the value obtained for two-year ahead forecasts and is considerably smaller than the one for three-year ahead predictions. Such a reduction of the MSE cannot be explained by typical dynamic processes but requires an argument such as the one provided by the decline in the conditional variance, as it is outlined in Proposition 3. An initial increase in the MSE seems to be largely driven by the accumulation of squared shocks that are contained in the Wold representation of the process. Beyond horizons at which the net increase of the MSE due to such terms is negligible, the decline in the MSE becomes visible due to the more pronounced convergence of the conditional variance towards the unconditional variance.

5 Summary and conclusions

By means of a model which assumes a generalization of the mean squared error criterion as the relevant loss for individual forecasters, we show that the performance a combined forecast can be improved by augmenting a consensus forecast by measures of ex-ante forecast variance. An empirical illustration, which is based on density forecasts from the SPF confirms the implications of the theoretical model. We find that while point forecasts are biased and therefore less accurate than the conditional mean predictions, the former can be improved by employing the information from the SPF to compute the average

individual uncertainty and the variation across point predictions. As suggested by our model, the former leads to higher improvements in the augmented forecast. A number of plausibility checks based on testable implications of the theoretical model underline the empirical relevance of the considered framework.

References

- Batchelor, R., Peel, D.A., 1998. Rationality testing under asymmetric loss. *Economics Letters* 61, 49-54.
- Becker, R., Clements, A.E., 2008. Are combination forecasts of S&P 500 volatility statistically superior? *International Journal of Forecasting* 122-133.
- Capistrán, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking* 41, 365-395.
- Christoffersen, P.F., Diebold, F.X., 1997. Optimal prediction under asymmetric loss. *Econometric Theory* 13, 808-817.
- Clements, M.P., 2010. Explanations of the inconsistencies in survey respondents' forecasts. *European Economic Review* 54, 536-549.
- Clements, M.P., 2014. US inflation expectations and heterogeneous loss functions, 1968-2010. *Journal of Forecasting* 33, 1-14.
- Conlisk, J., 1996. Why bounded rationality? *Journal of Economic Literature* 34, 669-700.
- Elliott, G., Komunjer, I., Timmermann, A., 2005. Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies* 72, 1107-1125.
- Elliott, G., Komunjer, I., Timmermann, A., 2008. Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association* 6, 122-157.

- Engelberg, J., Manski, C.F., Williams, J., 2009. Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics* 27, 30-41.
- Franses, P.H., Legerstee, R., Paap, R., 2016. Estimating loss functions of experts. *Applied Economics*, forthcoming.
- Granger, C.W.J., 1969. Prediction with a generalized cost of error function. *Operations Research Quarterly* 20, 199-207.
- Issler, J.V., Lima, L.R., 2009. A panel data approach to economic forecasting: The bias-corrected average forecast. *Journal of Econometrics* 152, 153-164.
- Lahiri, K., Sheng, X., 2010. Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics* 25, 514-538.
- Mincer, J.A., Zarnowitz, V. 1969. The evaluation of economic forecasts. In: J. Mincer (ed.) *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*. NBER, New York.
- Pagan, A., 1984. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review* 25, 221-247.
- Patton, A.J, Timmermann, A., 2007. Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics* 140, 884-918.
- Stock, J.H., Watson, M.W., 2004. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting* 23, 405-430.
- Timmermann, A., 2006. Forecast combinations. In: G. Elliott, C.W.J. Granger, A. Timmermann (eds.) *Handbook of Economic Forecasting*. Elsevier, Amsterdam.
- Varian, H.R., 1974. A Bayesian approach to real estate assessment. In: S.E. Fienberg and A. Zellner (eds.) *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*. North-Holland, Amsterdam.

Wright, J.H. 2008. Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics* 146, 329-341.

Zarnowitz, V., Lambros, L., 1987. Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95, 591-621.

Zellner, A., 1986. Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association* 81, 446-451.

A Proof of Proposition 2

Let $\gamma_*^{(\mathcal{U})}$ denote the solution of

$$\begin{aligned} \frac{\partial \text{MSE}^{(\mathcal{U})}}{\partial \gamma^{(\mathcal{U})}} &\stackrel{!}{=} 0. \\ \Rightarrow \gamma_*^{(\mathcal{U})} &= -\frac{\phi}{2} \end{aligned} \quad (32)$$

$$\Rightarrow \text{MSE}_*^{(\mathcal{U})} = \mathbf{E}[\eta_t^2] = \text{MSE}^{(optimal)}. \quad (33)$$

Next, the solution for the case when the consensus is augmented by disagreement is given by setting

$$\begin{aligned} \frac{\partial \text{MSE}^{(D)}}{\partial \gamma^{(D)}} &\stackrel{!}{=} 0. \\ \Rightarrow \gamma_*^{(D)} &= -\frac{\mathbf{E}[\phi h_t \bar{\tau} D_{t|t-1}]}{2 \mathbf{E}[(D_{t|t-1})^2]} \\ \Rightarrow \text{MSE}_*^{(D)} &= \mathbf{E}[\eta_t^2] + \mathbf{E}\left[\frac{\phi^2}{4} h_t^2 (\bar{\tau})^2\right] - \frac{(\mathbf{E}[\phi h_t \bar{\tau} D_{t|t-1}])^2}{4 \mathbf{E}[(D_{t|t-1})^2]} \\ &= \mathbf{E}[\eta_t^2] + \frac{\phi^2}{4} \mathbf{E}[h_t^2] \mathbf{E}[(\bar{\tau})^2] - \frac{\phi^2 \left(\mathbf{E}\left[h_t^3 \bar{\tau} \left(\bar{\tau}^2 - (\bar{\tau})^2\right)\right]\right)^2}{4 \mathbf{E}\left[h_t^4 \left(\bar{\tau}^2 - (\bar{\tau})^2\right)^2\right]}. \end{aligned} \quad (34)$$

Comparing (34) to (15) shows that $\text{MSE}_*^{(D)}$ is smaller than the MSE of the combined forecast without disagreement, i.e. $\text{MSE}^{(C)} \geq \text{MSE}_*^{(D)}$. Further rearrangement of (34) leads to

$$\begin{aligned} \text{MSE}_*^{(D)} &= \mathbf{E}[\eta_t^2] + \frac{\phi^2 \mathbf{E}[h_t^2 (\bar{\tau})^2] \mathbf{E}\left[h_t^4 \left(\bar{\tau}^2 - (\bar{\tau})^2\right)^2\right] - \left(\mathbf{E}\left[h_t^3 \bar{\tau} \left(\bar{\tau}^2 - (\bar{\tau})^2\right)\right]\right)^2}{4 \mathbf{E}\left[h_t^4 \left(\bar{\tau}^2 - (\bar{\tau})^2\right)^2\right]} \\ &= \mathbf{E}[\eta_t^2] + \\ &\quad \frac{\phi^2 \mathbf{E}[h_t^2] \mathbf{E}[h_t^4] \mathbf{E}[(\bar{\tau})^2] \mathbf{E}\left[\left(\bar{\tau}^2 - (\bar{\tau})^2\right)^2\right] - (\mathbf{E}[h_t^3])^2 \left(\mathbf{E}\left[\bar{\tau} \left(\bar{\tau}^2 - (\bar{\tau})^2\right)\right]\right)^2}{4 \mathbf{E}[h_t^4] \mathbf{E}\left[\left(\bar{\tau}^2 - (\bar{\tau})^2\right)^2\right]} \end{aligned} \quad (35)$$

From the Cauchy-Schwarz-inequality, it follows that 1.) $\mathbf{E}[h_t^2] \mathbf{E}[h_t^4] \geq (\mathbf{E}[h_t^3])^2$ and 2.) $\mathbf{E}[(\bar{\tau})^2] \mathbf{E}\left[\left(\bar{\tau}^2 - (\bar{\tau})^2\right)^2\right] \geq \left(\mathbf{E}\left[\bar{\tau}\left(\bar{\tau}^2 - (\bar{\tau})^2\right)\right]\right)^2$, thus the second term in the last line of (35) is positive. Hence, we can compare (33) to (35), which shows that $\text{MSE}_*^{(\mathcal{U})} = \text{MSE}^{(optimal)} \leq \text{MSE}_*^{(D)}$. Together with the result from (34), Proposition 2 follows.

B Proof of Proposition 4: $\phi_i \neq \phi_j$ for $i \neq j$

First, consider $\mathcal{U}_{t|t-1}$ as a predictor variable. This yields

$$\begin{aligned}
\text{MSE}^{(\mathcal{U})} &= \mathbf{E}\left[\left(e_{t|t-1}^{(\mathcal{U})}\right)^2\right] \\
&= \mathbf{E}\left[\left(\eta_t - \frac{\bar{\phi}}{2}h_t\tau - \gamma^{(\mathcal{U})}\mathcal{U}_{t|t-1}\right)^2\right] \\
&= \mathbf{E}\left[\left(\eta_t - \frac{\bar{\phi}}{2}h_t\tau - \gamma^{(\mathcal{U})}\tau\right)^2\right] \\
&= \mathbf{E}[\eta_t^2] + \frac{(\bar{\phi})^2}{4}\mathbf{E}[h_t^2]\mathbf{E}[\tau^2] + \gamma^{(\mathcal{U})}\bar{\phi}\mathbf{E}[h_t^2]\mathbf{E}[\tau^2] + (\gamma^{(\mathcal{U})})^2\mathbf{E}[h_t^2]\mathbf{E}[\tau^2] \quad (36)
\end{aligned}$$

Let $\gamma_*^{(\mathcal{U})}$ denote the solution of

$$\begin{aligned}
\frac{\partial \text{MSE}^{(\mathcal{U})}}{\partial \gamma^{(\mathcal{U})}} &\stackrel{!}{=} 0. \\
\Rightarrow \gamma_*^{(\mathcal{U})} &= -\frac{\bar{\phi}}{2} \quad (37)
\end{aligned}$$

$$\Rightarrow \text{MSE}_*^{(\mathcal{U})} = \mathbf{E}[\eta_t^2]. \quad (38)$$

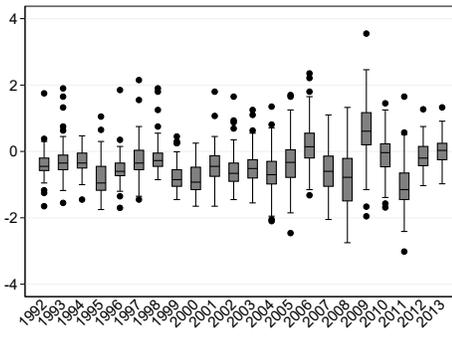
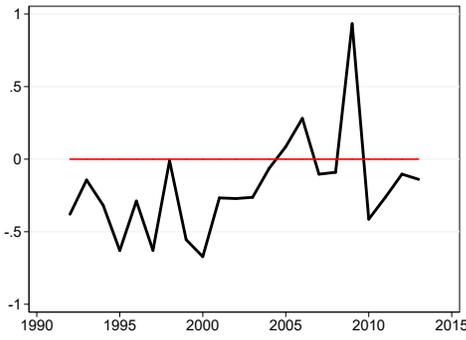
Next, a solution for $\text{MSE}^{(D)}$ is found by setting

$$\begin{aligned}
\frac{\partial \text{MSE}^{(D)}}{\partial \gamma^{(D)}} &\stackrel{!}{=} 0. \\
\Rightarrow \gamma_*^{(D)} &= -\frac{\bar{\phi}\mathbf{E}[h_t\tau D_{t|t-1}]}{2\mathbf{E}[(D_{t|t-1})^2]}.
\end{aligned}$$

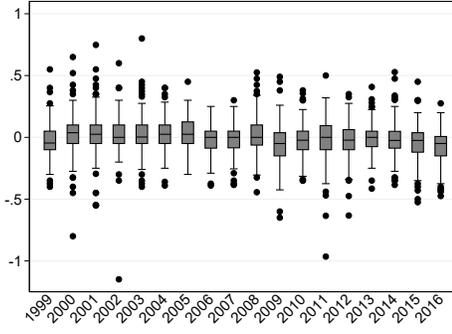
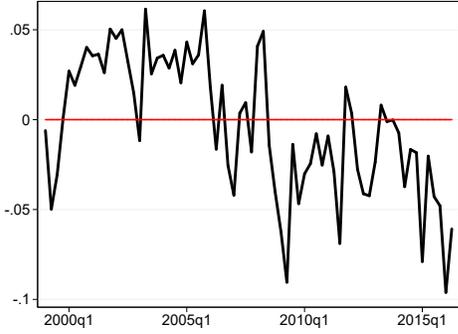
Then,

$$\begin{aligned}
\text{MSE}_*^{(D)} &= \mathbf{E}[\eta_t^2] + \frac{(\bar{\phi})^2}{4} \mathbf{E}[h_t^2] \mathbf{E}[\tau^2] - \frac{(\bar{\phi})^2 (\mathbf{E}[h_t \tau D_{t|t-1}])^2}{4 \mathbf{E}[(D_{t|t-1})^2]} \\
&= \mathbf{E}[\eta_t^2] + \frac{(\bar{\phi})^2}{4} \mathbf{E}[h_t^2] \mathbf{E}[\tau^2] - \frac{(\bar{\phi})^2 (\mathbf{E}[h_t^3])^2 (\mathbf{E}[\tau^3])^2}{4 \mathbf{E}[h_t^4] \mathbf{E}[\tau^4]} \\
&= \mathbf{E}[\eta_t^2] + \frac{(\bar{\phi})^2 (\mathbf{E}[h_t^2] \mathbf{E}[h_t^4] \mathbf{E}[\tau^2] \mathbf{E}[\tau^4] - (\mathbf{E}[h_t^3])^2 (\mathbf{E}[\tau^3])^2)}{4 \mathbf{E}[h_t^4] \mathbf{E}[\tau^4]} \\
&\geq \mathbf{E} [(y_t - \mathbf{E}[y_{i,t} | \mathcal{F}_{t-1}])^2] = \mathbf{E}[\eta_t^2]. \tag{39}
\end{aligned}$$

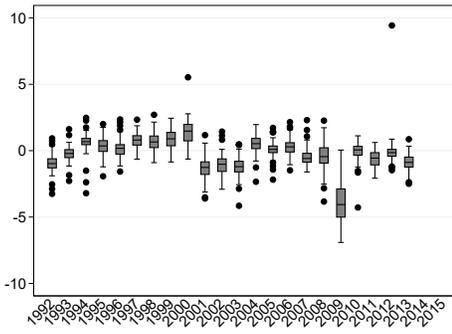
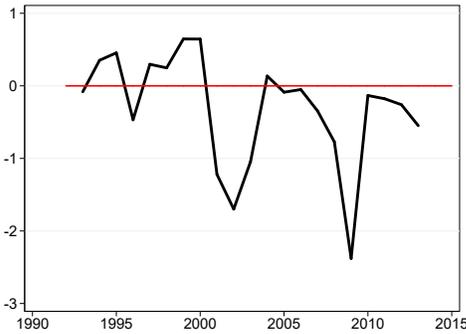
The inequality $\text{MSE}_*^{(D)} \leq \text{MSE}^{(C)}$ follows directly from the second line of (39). Moreover, as for (34), the positivity of the second term on the right-hand side of (39) follows from the Cauchy-Schwarz-inequality. This shows that $\text{MSE}_*^{(U)} \leq \text{MSE}_*^{(D)}$ and the result follows.



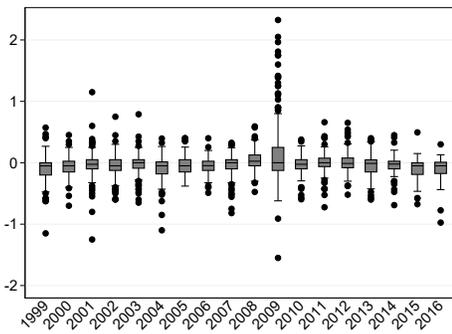
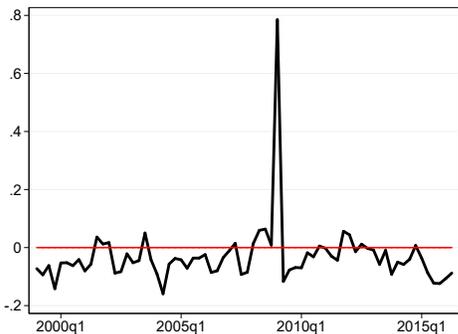
U.S. inflation



Euro area inflation



U.S. GDP growth



Euro area GDP growth

Figure 4: The left panel depicts the average deviation of the histogram mean from the point forecast consensus, i.e. $\xi_{t|t-1} = \mathbf{E}[\widehat{y_t} | \mathcal{F}_{t-1}] - \hat{y}_{t|t-1}^{(C)}$. The right panel depicts box plots of individual deviations, $\xi_{i,t|t-1}$.

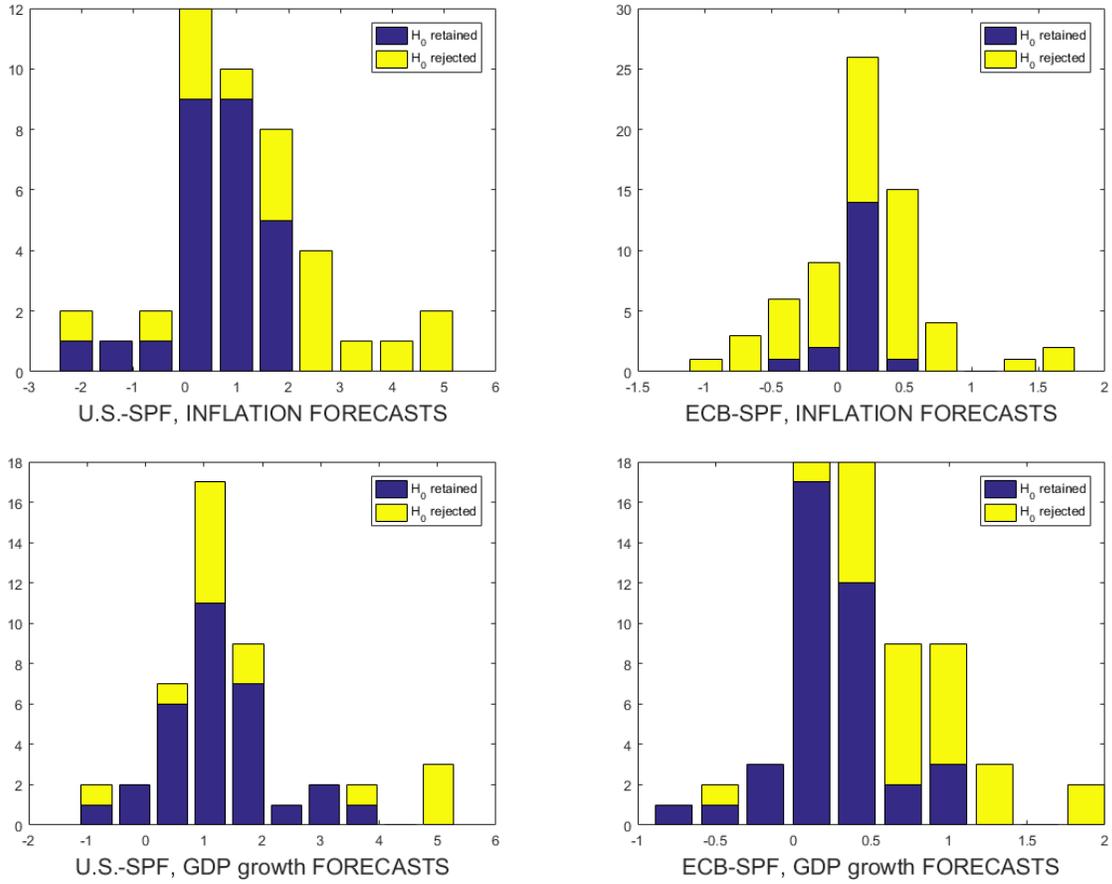
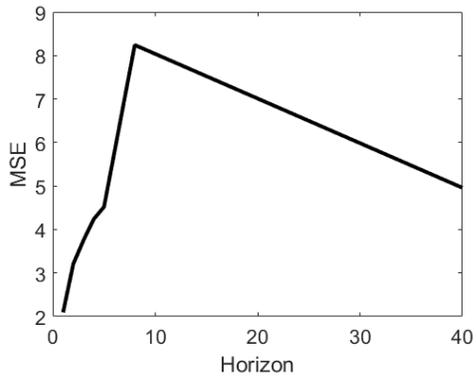
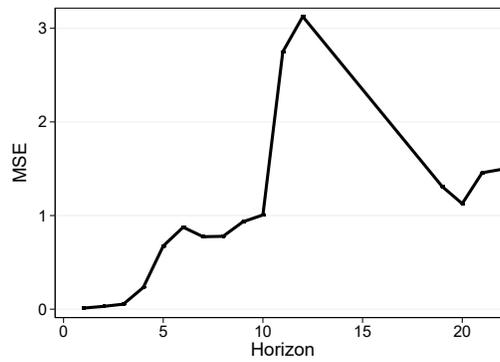


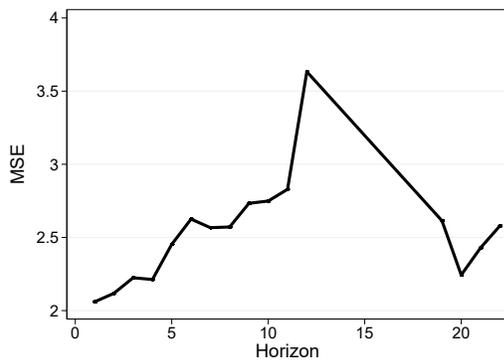
Figure 5: Histogram of individual asymmetry parameter estimates $\hat{\phi}_i$. Estimates that are significantly different from zero at the 10%-level are shown in yellow.



U.S. GDP growth



Euro area inflation



Euro area GDP growth

Figure 6: MSE of the consensus across the forecast horizon, in quarters.