

Measuring uncertainty and assessing its predictive power in the euro area

Pilar Poncela^{1,2} · Eva Senra³

Received: 12 February 2016 / Accepted: 8 September 2016 / Published online: 28 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Expectations and uncertainty play a key role in economic behavior. This paper deals with both, expectations and uncertainty derived from the European Central Bank Survey of Professional Forecasters. Given the strong turbulences that the euro area macroeconomic indicators observe since 2007, the aim of the paper is to check whether there is any room for improvement of the consensus forecast accuracy for GDP growth and inflation when accounting for uncertainty. We propose a new measure of uncertainty, alternative to the ad hoc equal weights commonly used, based on principal components. We test the role of uncertainty in forecasting macroeconomic performance in the euro area between 2005 and 2015. We also check the role of surprises in the considered forecasting sample.

The contents of this publication do not necessarily reflect the position or opinion of the European Commission. The work was initiated while the first author was still at Universidad Autónoma de Madrid and while the second author was visiting the Department of Economics at SUNY University at Albany. The authors gratefully acknowledge the comments received from Kajal Lahiri and an anonymous referee; however, any remaining errors are our own. Financial support from the Spanish Ministry of Economy and Competitiveness, project numbers ECO2015-70331-C2-1-R, ECO2015-66593-P and ECO2014-56676C2-2-P and Universidad de Alcalá is acknowledged.

✉ Pilar Poncela
pilar.poncela@jrc.ec.europa.eu

Eva Senra
eva.senra@uah.es

- ¹ European Commission, Joint Research Centre (JRC), Via Enrico Fermi 2749, 21027 Ispra, VA, Italy
- ² Universidad Autónoma de Madrid, Madrid, Spain
- ³ Facultad de CC Económicas, Empresariales y Turismo, Universidad de Alcalá, Plaza de la Victoria s/n, 28802 Alcalá de Henares, Spain

Keywords Survey of Professional Forecasters · Combination of forecasts · Uncertainty · Surprises · Principal components

1 Introduction

Expectations and uncertainty are two crucial issues in economic theory. Under general equilibrium theory, the rationality assumption makes households and firms to optimize their decisions according to expectations. Additionally, high uncertainty environments may delay households' and firms' consumption and investment and, for inflation, [Friedman \(1977\)](#) started an ongoing debate on the positive link between inflation and inflation uncertainty. The European Central Bank Survey of Professional Forecasters (ECB-SPF) was started in 1999 as a complement to the Eurosystem staff macroeconomic projections to assess risks to price stability. The ECB-SPF provides point and density forecasts for GDP, inflation and the unemployment rate for medium- and long-term horizons provided by a panel of more than one hundred forecasters. The ECB-SPF results a valuable tool to assess both, expectations and uncertainty.

This paper focuses on forecasting GDP growth and inflation in the euro area for the period 2005–2015 by means of the ECB-SPF. There is a large body of the literature reporting success in forecasting with combination of forecasts methods since the seminal work of [Bates and Granger \(1969\)](#). Later surveys on combination of forecasts can be found in [Menezes et al. \(2000\)](#), [Newbold and Harvey \(2002\)](#) and [Timmermann \(2006\)](#) or, more recently, [Wallis \(2011\)](#) and [Aiolfi et al. \(2011\)](#), among others. It is also known that the simple average of point forecasts, dubbed as *consensus*, is usually a benchmark difficult to beat ([Stock and Watson 2004](#), named this fact as the “*forecast combination puzzle*”). Regarding the ECB-SPF, two recent papers, [Genre et al. \(2013\)](#) and [Conflitti et al. \(2015\)](#), have stated the relative good performance of the consensus forecasts against more sophisticated techniques.

GDP growth and inflation in the euro area have shown high variability in the sample that we analyze, and high macroeconomic uncertainty is considered to be one of the main factors contributing to the protracted weakness of the euro area activity in recent years (see [ECB 2013](#), for instance). Therefore, though uncertainty is always a relevant concept, it has regained interest since the 2007–2009 recession.

Uncertainty is an unobserved variable, and there is not an unique way to measure it. [Bloom \(2009\)](#) uses as a proxy for uncertainty stock market realized and implied volatilities; [Bachman et al. \(2013\)](#) use a measure of cross-sectional dispersion of survey-based forecasts; [Baker et al. \(2013\)](#) refer to the frequency of newspaper references to economic policy uncertainty; [Jurado et al. \(2015\)](#) emphasize the link between uncertainty and predictability, and [Scotti \(2013\)](#) measures the size of the surprises when economic data are released. For the ECB-SPF, [Bowles et al. \(2007,2010\)](#), [Conflitti \(2012\)](#) and [Abel et al. \(2016\)](#) have evaluated uncertainty measures estimated from density and point forecasts and from forecast errors.

There are three approaches to uncertainty based on surveys of forecasts. First, *disagreement* from point forecasts is commonly used as a proxy to uncertainty. Second, [Zarnowitz and Lambros \(1987\)](#) define *uncertainty* as the equal weight average of individual perception of uncertainty. Third, the aggregate uncertainty coming from

the equal weight aggregation of the individual density forecasts. In this paper, we use these three measures of uncertainty, but we also propose new measures alternative to the ad hoc equal weights commonly used. Our weights are based on principal components over the original individual uncertainties, as this approach should help in better identifying the common uncertainty underlying to all the individuals. We use these weights to aggregate individual uncertainties, to aggregate density forecasts and to obtain also a new measure of aggregate uncertainty and disagreement.

Our main thesis is that as GDP growth and inflation in the euro area since 2008 are greatly affected by the business cycle, uncertainty should be a key element that could help improving one year ahead forecast accuracy. Both measures of uncertainty, traditional and principal component weights, show that uncertainty in the euro area increased significantly since year 2008 and that it still remains at high levels. However, including them in the forecasting exercise does not improve significantly the forecast accuracy. This also contrasts with the latest recovery of the euro area GDP growth or the low inflation, indicating that the improvement comes from external factors, like the oil price market shocks, instead than from a restoration of the confidence.

Finally, we look at the predictive power of ex post measures of uncertainty and surprises on the time period for the two variables analyzed. Surprises, defined as 1 year ahead consensus forecast errors, are not a measure of uncertainty, but they also indicate the degree of knowledge about the future outcome. It is well known in the literature that economic agents follow closely macrosurprises and that they affect the evolution of asset prices or may modify short-run expectations (see [Scotti \(2013\)](#)). We document that for the time period 2005–2015 the inclusion of surprises improves 1 year ahead forecast accuracy by 40 % in the out of sample exercise for GDP growth.

The structure of this paper is as follows: Sect. 2 introduces the ECB-SPF and analyzes its forecasting performance for 1 year ahead forecasts for GDP growth and inflation and compares consensus against some other common alternatives that worked well in other surveys and in the ECB-SPF. Section 3 defines uncertainty measures, introduces new alternative ones and estimates their evolution on the euro area GDP growth and inflation. Section 4 tests empirically the predictive power of uncertainty between 2005 and 2015. Section 5 introduces ex post uncertainty measures and surprises as relevant regressors and checks their forecasting ability, and finally, Sect. 6 concludes.

2 Consensus forecast behavior and alternative dimension reduction models performance

The ECB Survey of Professional Forecasters (ECB-SPF) is a quarterly panel of forecasts, starting in 1999, surveying on real GDP growth rate, HICP inflation and the unemployment rate expectations for the euro area as a whole. The panel covers institutions that are required to possess macroeconomic expertise relating to the euro area and not just to their own economy. Also, they required several years' experience in forecasting and publishing forecasts. Respondents provide point forecasts for rolling horizons (1 and 2 years ahead), fixed calendar year horizons (current year, next year and year after next) and longer-term expectations (5 years ahead). They also provide

the probability distributions that correspond to their point forecasts at all horizons. A total of 115 institutions have participated since the beginning. [García \(2003\)](#), [Bowles et al. \(2007,2010\)](#) and [ECB \(2014\)](#) summarize the main characteristics of the ECB-SPF participants since its inception.

We focus on the behavior of rolling 1 year ahead point forecasts for GDP growth and inflation. GDP growth is a quarterly indicator measured as the annual change in the level of GDP in quarter t compared with its value in quarter $t - 4$, and inflation is a monthly indicator (observed quarterly for the survey) defined as the annual rate of the Harmonized Consumer Price Index (HICP) in month t compared with its value in month $t - 12$. The time span we consider goes from the second quarter of 1999 (1999Q2) to the second quarter of 2015 (2015Q2) for GDP growth and from September 1999 (1999Q3) to September 2015 (2015Q3) for inflation, as that is the period that we can compare with the forecasts given by the ECB-SPF. There are two additional issues regarding the data: (1) The euro area for the time period analyzed has expanded; while in the year 2000 it was formed by 11 countries, in 2015 it includes 19 countries, and (2) GDP growth data are subject to important revisions and final data may not correspond with the figures that forecasters first observed in real time (see, for instance, [Lahiri and Sheng \(2010a\)](#)). The problem of revisions is smaller when looking at inflation. To account for this problems we use as observed data the second release of the indicators, as they are the first complete estimation. The data are published in the real-time database for the euro area developed by [Giannone et al. \(2012\)](#) that is updated in the ECB Web page.

Figure 1 shows GDP growth and inflation. GDP growth for the period considered has been on average 1.03 %, but it showed a very different behavior pre- and post-2008. Previous to year 2008, the average GDP growth was 1.88 %, and afterward, it has been 0.07 %. But this lower average GDP growth is also accompanied by higher fluctuations after 2008. While pre-2008 the standard deviation of GDP growth was 1.1 % points (p.p. henceforth), after 2008 it increased to 1.9 p.p. (1.7 times the previous value). Looking at inflation, it has been 1.88 % on average but also with different mean and standard deviation pre- and post-2008. Previous to 2008, inflation was 2.19 % slightly exceeding the ECB medium-term inflation objective while afterward it reduced to 1.56 % with negative values observed at several moments for the first time in the euro area inflation. The inflation standard deviation was multiplied by more than four, increasing from 0.3 p.p. pre-2008 to 1.2 p.p. afterward. These statistics highlight the difficulties faced by the forecasters in this time span.

Let y_{it} be the four period ahead forecast from individual i , $i = 1, \dots, N_t$, $t = 1, \dots, T$, with information up to $t - 4$ for the target variable y_t (GDP growth or inflation). We first define *Consensus* as the average point forecast at every moment in t , it is,

$$Consensus_t = \frac{1}{N_t} \sum_{i=1}^{N_t} y_{it} \quad (1)$$

where N_t is the number of forecasts available at moment t . This average is also the value that the ECB-SPF provides as consensus in its quarterly report. Figure 1 shows the consensus for GDP growth rate and inflation. We assess the consensus forecast

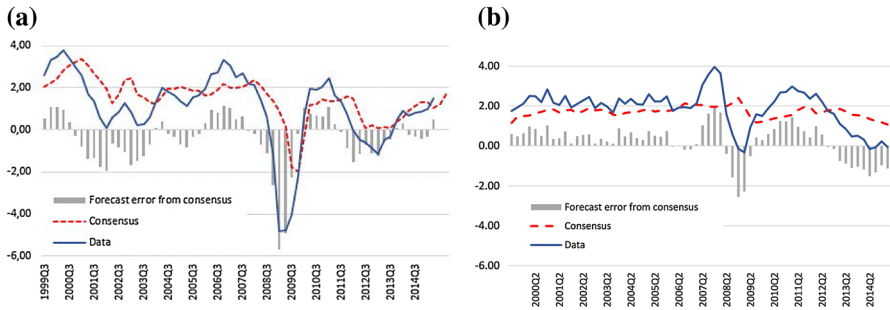


Fig. 1 GDP growth and inflation: observed data, consensus forecast and forecast errors. **a** GDP, **b** inflation

error ($e_t = y_t - Consensus_t$) by its mean forecast error ($MFE = \frac{1}{T} \sum_{t=1}^T e_t$) and its root mean square forecast error ($RMSFE = \sqrt{\frac{1}{T} \sum_{t=1}^T e_t^2}$). The MFE for GDP growth has been -0.46 p.p. for the whole sample, and it has remained negative both, pre- and post-2008 (-0.23 p.p. and -0.71 p.p., respectively). The MFE for inflation has been 0.19 for the whole sample. Nevertheless, it was positive (0.48 p.p.) for the sample pre-2008 and negative (-0.13 p.p.) from 2008 until the end of the sample.. To gain some insights into possible bias, we run Mincer–Zarnowitz regressions of the form

$$y_t = \alpha_0 + \alpha_1 Consensus_t + a_t, t = 1, \dots, T \tag{2}$$

and check the null hypothesis $H_0 : \alpha_0 = 0, \alpha_1 = 1$. Notice that the target variable (real GDP growth or inflation) is annual growth over four quarters, and forecasts are made every quarter for the annual growth on an rolling basis. This implies that Eq. (2) will have a $MA(h - 1)$ error structure, where h is the forecast horizon, and therefore, we have applied a robust Wald test using the Newey and West variance–covariance estimator. We find that we cannot reject unbiasedness for the whole sample for GDP growth (p -value for the null hypothesis is 0.16). For inflation, consensus forecast unbiasedness cannot be rejected for the whole sample (p -value equal to 0.60), but the hypothesis is hard to maintain pre-2008 given that the p -value is 0.00 , while post-2008 the possible bias seems to disappear as the p -value is 0.98 . This rises some bias concerns, but big forecast errors due to the crisis have an important effect on the estimation of both the parameters and their standard deviations, and special shocks like oil and food prices are determining the inflation pre-crisis. For inflation, Bowles et al. (2007) point out that a large part of this systematic error can be explained by the sequence of asymmetric and largely unpredictable shocks that hit the euro area over the period and that adjusting for these facts, there is far less evidence of a systematic underprediction or bias in the inflation expectations.

The RMSFE for the whole sample is 1.46 p.p. for GDP growth and 1.05 p.p. for inflation with a different behavior pre- and post-2008 where the accuracy of both indicators worsened by 1.7 and 1.9 times their pre-crisis levels, respectively.

2.1 Forecasting performance of alternative models

We are going to set consensus as our forecast target and we shall complement it with uncertainty and surprise measures in the following sections. However, this section checks the validity of consensus as benchmark despite some issues as the selection of regular respondents, bias concerns or the good performance of dimension reduction models found with a shorter sample by [Genre et al. \(2013\)](#).

The first concern is that we have an unbalanced panel of forecasts along time. Consensus is calculated with the available point forecasts at every moment in time. However, there is a number of individuals that participate only occasionally and it is common in the literature to consider only “regular” or “frequent” forecasters (see, for instance, [Capistrán and Timmermann 2009](#); [Lahiri et al. 2016](#) or more recently with ECB-SPF data, [Genre et al. 2013](#) and [Conflitti et al. 2015](#) among others). In our case, we are going to consider those individuals that have a participation rate $\geq 60\%$. This reduces the number of participants to 38 for GDP growth and to 33 for inflation point forecasts. We name *Consensus60* to the consensus forecast obtained with those frequent forecasters. For the remaining missing data, there are several ways to impute them, but to make fair comparisons among the models we do have to treat the missing observations uniformly across different combining procedures, as [Lahiri et al. \(2016\)](#) point out. In this regards, along the alternative models we do not impute the missing data explicitly and consider sample statistics with the available data at each moment (as it is done with the original consensus).

The second concern is about the apparent bias we find with consensus forecasts. We run an ordinary least squares (OLS¹) regression between the observed values and a constant and use the estimated parameters to build our forecast (*Rconsensus*)

$$Rconsensus_t = \hat{\alpha}_0 + \hat{\alpha}_1 Consensus60_t \quad (3)$$

where $\hat{\alpha}_0$ and $\hat{\alpha}_1$ are estimated with information only up to $t - 1$ from the equation:

$$y_\tau = \alpha_0 + \alpha_1 Consensus60_\tau + a_\tau, \tau = 1, \dots, t - 1 \quad (4)$$

and y_τ are the observed values for GDP growth or inflation, as it corresponds. Therefore, $Rconsensus_t$ is a valid out of sample forecast.

In two previous papers for the US Survey of Professional Forecasters, [Poncela and Senra \(2006\)](#) and [Poncela et al. \(2011\)](#) find that factor models estimated by principal components were a good alternative to improve consensus. [Genre et al. \(2013\)](#) also find that for inflation, factor models performed well against consensus with the ECB-SPF. As [Lahiri et al. \(2016\)](#) point out forecast combinations would work when there is heterogeneity among forecasters. Precisely, the principal components pick up this heterogeneity since they are the linear combinations of the forecasters, orthogonal among each other, in the direction of maximum variability. So the first r principal

¹ As in Eqs. (2), (4) the error term will show a MA structure. As we only need point estimators to make the forecasts, OLS is a consistent estimator and the point estimates coincide with GMM. This also applies to all the regressions that correspond to Table 1 and columns (3) to (9) in Table 2.

Table 1 Comparison forecast of alternative models

GDP growth	RMSFE consensus	Ratio to RMSFE consensus				
		Consensus60	Rconsensus	Common factors		
				F_1	F_1, F_2	F_1, F_2, F_3
	1.46	0.99	1.02	1.02	0.98	0.93
Inflation	1.05	1.00	1.08	1.09	1.07	1.07

Table 2 Forecasting results

GDP growth	Obs.	Consensus		Ratio to RMSFE consensus					
		MFE	RMSFE	Disagreement		Uncertainty		Ag. uncert.	
				D	D *	U	U*	U_{agg}	U_{agg}^*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
2005–2015	2.03	-0.43	1.46	1.20	1.10	0.96	0.96	1.10	0.94
2005	1.42	-0.52	0.58	0.87	0.44	0.50	0.49	0.61	0.48
2006	2.66	0.81	0.86	0.97	1.49	1.44	1.47	1.30	1.48
2007	2.61	0.54	0.66	0.71	1.32	1.31	1.32	1.16	1.33
2008	0.73	-1.16	1.47	0.83	0.84	0.92	0.66	0.85	0.78
2009	-3.94	-3.25	3.92	1.24	1.05	0.89	0.89	1.09	0.85
2010	1.63	0.78	0.79	2.28	2.26	1.25	1.74	2.10	1.58
2011	1.53	0.09	0.73	1.38	1.37	1.16	1.30	1.20	1.31
2012	-0.50	-1.12	1.16	0.50	0.48	1.19	0.98	0.62	1.00
2013	-0.37	-0.54	0.72	0.79	1.02	0.61	0.70	0.95	1.01
2014	0.81	-0.17	0.34	1.59	2.14	1.65	1.34	1.58	1.50
2015	1.24	0.07	0.40	1.76	1.99	1.05	1.02	1.65	1.10
Inflation									
2005–2015	1.74	0.03	1.05	1.03	1.16	1.00	1.04	0.99	1.06
2005	2.24	0.48	0.51	0.32	0.44	0.31	0.42	0.32	0.36
2006	2.09	0.29	0.43	0.86	0.79	0.89	0.78	0.90	0.77
2007	2.26	0.18	0.53	1.00	0.96	0.98	0.96	0.99	0.97
2008	3.19	1.22	1.54	0.79	0.85	0.74	0.85	0.74	0.85
2009	0.25	-1.73	1.90	0.99	1.13	0.70	1.10	0.63	1.07
2010	1.79	0.53	0.57	1.58	1.28	2.13	1.75	1.61	1.71
2011	2.78	1.21	1.22	0.93	0.91	0.91	0.98	0.38	0.35
2012	2.46	0.67	0.70	0.95	0.81	1.55	1.17	1.42	1.46
2013	1.32	-0.46	0.58	1.12	1.23	0.87	0.99	1.42	1.77
2014	0.28	-1.20	1.21	1.18	1.22	1.20	1.19	1.60	1.61
2015	0.03	-1.14	1.15	1.19	1.13	1.35	1.18	1.51	1.38

components are the best (in the sense that capture the maximum possible variability) representation of the data in a r -dimensional subspace. In a sense, for a problem with N forecasters, if one picks up $r < N$ principal components, you assume that the remaining $N - r$ principal components are associated with noise variability and, therefore, it is not worth to estimate them. The implications of estimating more principal components is that you consider that there is information useful for forecasting in the successive principal components and are not just noise.

To compute the forecasts by principal components, we proceed in two steps. First we compute the variance-covariance matrix \sum_y with elements σ_{yij} calculated with all the common observations between the point forecasts for individual i and j , and compute the r first principal components associated with the three eigenvectors with highest eigenvalues. Second, we run three different regressions like (4) with the first (F_1), the first and the second (F_1 and F_2) and the three first principal component common factors (F_1 , F_2 and F_3) as regressors, respectively. Finally, we use the estimated parameters with information up to time $t - 1$ to forecast as in Eq. (3).

For the out of sample forecasting exercise, we first consider an initial sample from the third quarter of 1999 until the fourth quarter of 2003. Second, with the estimated parameters and the values of the consensus and the principal component common factors obtained with the SPF individual 1 year ahead forecasts in the first quarter of 2004, we give alternative forecasts for the first quarter of 2005. Then, we use an expanding window by adding one more observation to the estimation sample and make a new forecast. The forecast sample goes from the first quarter in 2005 until the second quarter of 2015 for GDP growth and until September 2015 for inflation.

Table 1 shows the original consensus RMSFE for the whole sample for GDP growth and inflation, as well as the ratio between the alternative models RMSFE and the consensus RMSFE. A value of the ratio equal to one indicates the same forecasting performance, < 1 an improvement over the consensus benchmark and over 1 a worse performance. None of the alternatives significantly improves the consensus accuracy: Consensus performance does not change if we consider either all the respondents or only frequent respondents; the bias correction made by regressing on the observed values and a constant does not improve the forecast accuracy and it performs 2 and 8 % worse than just the original consensus for GDP growth and inflation, respectively; factor models slightly improve forecast accuracy for GDP growth but not for inflation. This contradicts the results by Genre et al. (2013), as the good performance of factor models for inflation deteriorates greatly after 2013.

3 Measuring uncertainty in the euro area: a new proposal

3.1 Measuring uncertainty

We focus on uncertainty measures based on Surveys of Forecasts and define uncertainty as the variance of the future outcome of the target indicator conditional to the available information as in Zarnowitz and Lambros (1987), Lahiri et al. (1988), Giordani and Söderlind (2003), Wallis (2005), Lahiri and Sheng (2010b), Rich and Tracy (2010), among others, and Conflitti (2012) and Abel et al. (2016) specifically for the ECB-SPF.

The ECB-SPF provides both, point and density forecasts. Let's denote by $pdf(y_t|i, t - 4)$, $i = 1, \dots, N$ the individual density forecast given by forecaster i in moment $t - 4$ and assume they have common expected mean μ_t (consensus) and variance U_t (uncertainty). Let μ_{it} and U_{it} be the estimated mean and variance from the individual forecasts. Then, taking $v_i = 1/N$, we can compute μ_t and U_t by

$$\mu_t = \sum_{i=1}^N v_i \mu_{it} \tag{5}$$

$$U_t = \sum_{i=1}^N v_i U_{it}. \tag{6}$$

Equation (5) is equivalent to the measure of consensus from point forecasts in Eqs. (1) and (6) is the measure of individual uncertainty defined by Zarnowitz and Lambros (1987). Alternatively, we can derive a measure of aggregate uncertainty (U_{aggt}) from the resulting distribution after pooling individual density forecasts

$$pdf(y_t|t - 4)_t = \sum_{i=1}^N v_i pdf(y_t|i, t - 4). \tag{7}$$

Denoting by μ_{aggt} and U_{aggt} the mean and variance of this new distribution, we can easily find that there is no difference between μ_{aggt} and μ_t and that the following relationship holds for the variance:

$$U_{aggt} = \sum_{i=1}^N v_i U_{it} + \sum_{i=1}^N v_i (\mu_{i,t} - \mu_{aggt})^2. \tag{8}$$

Equation (8) is common in the literature (see Lahiri et al. 1988; Giordani and Söderlind 2003; Wallis 2005; Conflitti 2012 for equal weights and Hall and Mitchell 2007 for general v_i). It shows that the global uncertainty U_{aggt} is the sum of uncertainty and the average of the squared differences between individual means and the aggregated mean, that is, disagreement

$$D_t = \sum_{i=1}^N v_i (\mu_{i,t} - \mu_{aggt})^2. \tag{9}$$

The most common practice is to use Eqs. (5)–(9) with equal weights $v_i = 1/N$, to average first- and second-order moments as well as the individual density forecasts.

As we focus on uncertainty, we propose alternative weights to better summarize the common information in individual uncertainties. As with point forecasts in the previous section and Genre et al. (2013), Poncela and Senra (2006) and Poncela et al. (2011) for point forecasts, we apply principal components to find the linear combination

of the individual uncertainties that best represents the information contained in the uncertainties in a subspace of dimension 1.

Therefore, we compute the variance–covariance matrix Σ_U of the individual variances, and consider the first eigenvector associated to the maximum eigenvalue, as the weights to compute U_t in Eq. (6). We apply the same weights to aggregate μ_t in Eq. (5), and all the terms in Eqs. (7)–(9).

3.2 Uncertainty in the euro area

To estimate the alternative measures of uncertainty, we need to estimate the first- and second-order moments of the individual distributions. In the ECB-SPF, forecasters assign probabilities to a range of intervals at each moment in time. The number of intervals has changed along time, and currently, they range from $<-6\%$ to $>5\%$ for GDP growth and <-2 to $>4\%$ for inflation with 0.5 p.p. interval width, except for the first and the last intervals that are open. We have considered these intervals for the whole sample and assigned 0 probability when they were not available. In most of the cases forecasters place their probabilities to just a few intervals and the normality assumption has been tested and rejected by [Conflitti \(2012\)](#). Therefore, we do not assume any specific distribution for the probabilistic beliefs and compute means and variances by locating the probabilities to the midpoints of the intervals ([D'Amico and Orphanides 2008](#) find that there are only small differences with midpoint, uniform or normal approaches). The midpoint approach is an approximation that also depends on the width of the intervals, and it produces bias upwards estimates of the individual variances. To compensate for this, we apply the Sheppard correction to variance estimates (see [Kendall and Stuart 1977](#)). Naming x_k to the midpoint of each interval, we close the open intervals from the extremes by assuming they have the same length than the closed intervals.² Denoting by p_{tki} , $k = 1, \dots, K$, with $K = 24$ for GDP growth and $K = 15$ for inflation,³ to the probability assigned by individual i at time t to interval k , we can calculate the first and second moments of the distribution at each period t , $t = 1, \dots, T$ and for each individual i , $i = 1, \dots, N$.

$$\begin{aligned}\mu_{i,t} &= \sum_{j=1}^K p_{tji} x_j \\ U_{i,t} &= \sum_{j=1}^K p_{tji} (x_j - \mu_{i,t})^2\end{aligned}\quad (10)$$

² The choice of the end points, though arbitrary, is not restrictive as individuals usually assign negligible probability to outer class intervals. There is only one exception in the first quarter of 2009 for GDP growth when, due to the sudden downwards revision of the forecasts as a consequence of the crisis, the individuals concentrated 64 % of the probability in the first interval. In that occasion we considered as the midpoint of the lower interval $x_1 = -2.2$ to match the level of uncertainty observed in the following quarter, when intervals were extended until $<-6\%$. However, main results remain unaltered.

³ To facilitate the notation we do not distinguish between GDP growth of inflation, but probabilities are different for each variable.

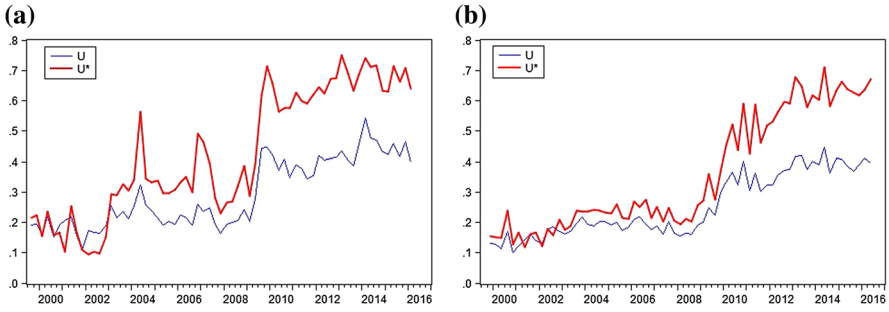


Fig. 2 Uncertainty (U) and weighted uncertainty (U^*): (a) GDP, (b) inflation

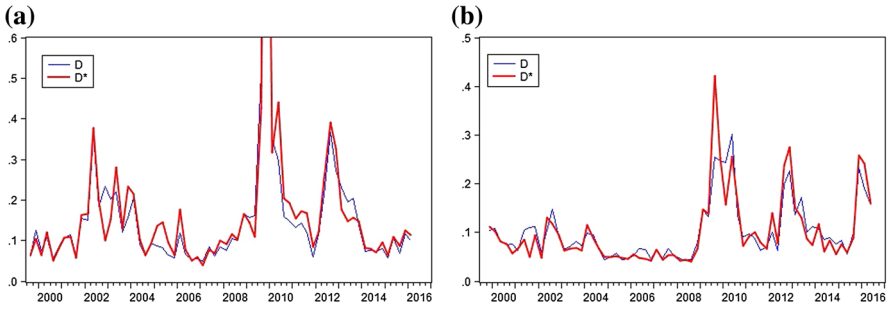


Fig. 3 Disagreement (D) and weighted disagreement (D^*): (a) GDP, (b) inflation

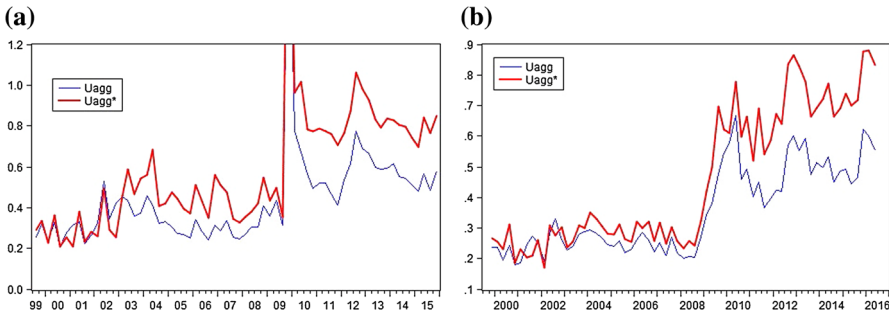


Fig. 4 Aggregate uncertainty (U_{agg}) and weighted aggregated uncertainty (U_{agg}^*): (a) GDP, (b) inflation

We follow the same strategy as before and keep only frequent forecasters that have participated at least 60 % of the time. This criterion leaves us with 31 individuals for GDP growth and 29 for inflation. Despite this restriction, the number of participants is still not constant along time and both mean and variance from the density histograms are unbalanced panels of data. For the remaining missing data, we define N_t as the number of respondents in moment t and substitute N by N_t in Eqs. (5)–(9). Regarding the variance–covariance matrices needed to estimate the weights by principal components, we build the elements σ_{uij} of the variance–covariance matrix \sum_U with all the common observations between the density variances for individuals i and j .

Figures 2, 3 and 4 show the alternative uncertainty measures estimated for the euro area by both, equal and principal component weights, from density forecasts, disagreement and forecast errors. All the measures of uncertainty rise during the crisis.

Figure 2a plots the uncertainty in GDP growth estimated by averaging the individual uncertainties with equal weights (U) and our alternative principal components estimation (U^*). After the financial crisis average uncertainty (U) of the individual density forecasts increased by 1.7 times its pre-crisis level, and it has remained high since then. The weighted average uncertainty (U^*) estimated by principal components explains 48.5 % of the total set of individual uncertainties and shows a high correlation (0.96) with the equal weights U . However, U^* depicts a different picture, as it shows that U has underestimated the level of uncertainty previous to the crisis and also that the permanent increase in the level of uncertainty after the financial crisis has been multiplied by two, bigger also than that estimated by U . There are no signs of receding uncertainty in any of the two measures. Disagreement measures in GDP growth 1 year ahead forecasts are shown in Fig. 3a. The highest heterogeneity among the forecasters happened in the 1 year ahead forecasts for 2009Q4, when both measures D and D^* showed a sudden increase in more than 12 times their past values. Despite that moment, disagreement D and D^* show high correlation (0.90) and both measures seem to have returned to much lower pre-crisis values. Aggregated uncertainty measures (U_{agg} and U_{agg}^*) are derived from the aggregated density forecast and are the sum of common uncertainty and disagreement. On average, uncertainty (U and U^*) represents around 70 % of total aggregate uncertainty, being the share around 0.5 p.p. greater in the principal components weight case. Therefore, aggregate uncertainty, as it is shown in Fig. 4a, is mainly dominated by uncertainty but it also shows the high peaks of disagreement. The correlation between U_{agg} and U_{agg}^* is 0.92, and as a consequence of the financial crisis, both measures increased permanently. Aggregate uncertainty estimated by principal components weights are superior to equal weights along the whole sample and also the magnitude of the increase is bigger with U_{agg}^* (2.1 times vs 1.8 with U_{agg}).

Figure 2b plots the uncertainty in inflation estimated by averaging the individual uncertainties with equal weights (U) and our alternative estimation (U^*). After the financial crisis average uncertainty (U) of the individual density forecasts increased by 1.8 times its pre-crisis level, and it has remained high since then. The weighted average uncertainty (U^*) estimated by principal components explains 66.2 % of the total set of individual uncertainties and shows a high correlation (0.99) with the equal weights U . However, U^* depicts a different picture as it shows that U has underestimated the levels of uncertainty previous to the crisis and that this difference has increased in time. In total, the permanent increase in the level of uncertainty after the financial crisis has been multiplied by 2 with U^* . There are no signs of receding uncertainty in any of the two measures, and they even seem to follow an upward trend. Disagreement, D and D^* , measures in 1 year ahead forecasts for inflation are shown in Fig. 3b. There is high heterogeneity among the forecasters in 1 year ahead forecasts for 2009Q3, 2010Q2, 2012Q4 and 2015Q4. D and D^* show high correlation (0.88), and both measures seem not to return to pre-crisis values. Aggregate uncertainty (U_{agg} and U_{agg}^*) are again mainly dominated by uncertainty (U and U^*) that represent around 71 and 76 % of total aggregate uncertainty, respectively. Therefore, aggregate uncertainty, as it is

shown in Fig. 4b, is mainly dominated by uncertainty but also shows the high peaks in disagreement. The correlation between U_{agg} and U_{agg}^* is 0.96, and since the financial crisis they show an upward trend.

4 Forecasting performance of uncertainty measures

In this section, we check empirically whether the uncertainty estimated by density forecasts helps in forecasting GDP growth and inflation. The inclusion of uncertainty or second-order moments in the mean equation can only be justified if the data are non-normal. Lahiri and Teigland (1987) and the references therein discuss whether the shape of the distribution has an impact on forecast precision and find that, for the ASA-NBER quarterly surveys, distributions vary significantly over time and the assumption of normality is rejected quite often. Paloviita and Viren (2014) also deal with the relationship between first and second moments of inflation and output growth forecasts from the ECB-SPF, finding that in the recent crisis individual forecasters have reacted to increasing uncertainty by adopting a completely different distribution.

We run a pseudo-out of sample forecasting exercise for the period 2005Q1–2015Q2 for GDP growth and 2005Q1–2015Q3 for inflation. The forecasting procedure is two steps and replicates Eqs. (3) and (4) to the current setup with the square root of the uncertainty measures analyzed in the previous sample used as alternative regressors. In the first step, we estimate the parameters of an OLS (see footnote 1) relationship between the target variables, the constant, consensus and each of the uncertainty measures with all the information available in moment $t - 1$. In the second step the estimated parameters are used as weights to give a forecast for the target variable given the values of the uncertainty measures.

Table 2 summarizes the results. Column (1) shows the average observed data for the whole forecasting sample and the average of the observed data in every natural year for GDP growth and for inflation. Columns (2) and (3) summarize the forecasting performance of consensus for the whole sample and for each calendar year by means of the MFE and the RMSFE (for every calendar year MFE and RMSFE are calculated only with the forecast errors available for that year). Finally, columns (4) to (9) show the ratio between the RMSFE obtained with the inclusion of the proposed uncertainty measure and the consensus RMSFE presented in column (3). Therefore, a value smaller than 1 indicates better behavior of the alternative model against consensus and greater than 1 otherwise. We have also calculated this ratio yearly, considering only the forecast errors in every calendar year.⁴

At first sight we can see that none of the uncertainty measures improves consensus forecast accuracy for the whole forecasting sample (rows corresponding to 2005–2015) for GDP growth or inflation. We can also see that there are only marginal differences between equal weights or principal component weights (compare D and D^* in columns (4) and (5), U and U^* in columns (6) and (7) and so on).

⁴ We acknowledge that this yearly magnitudes are not statistically significant, but we include them for better understanding the role of uncertainty.

Given the business cycle complexity of the time span used for the forecasting exercise, it is interesting to look at the yearly performance. Looking at GDP growth we can see that uncertainty measures improve the forecasting accuracy in years 2005, 2008 and 2009 and 2012 and 2013, when GDP growth is decelerating or negative. On the contrary, uncertainty measures seem highly misleading in forecasting the years or accelerating GDP growth. Regarding inflation, we can observe some similar but dual behavior to GDP growth. Uncertainty seems to help forecasting inflation in years 2005 to 2008 and 2011 and 2012 when inflation was over the ECB inflation target, but again uncertainty seems to mislead forecasts when inflation is low. This behavior of uncertainty is therefore in line with expected according to economic theory.

5 A look at ex post measures of uncertainty and surprises

All the measures of uncertainty used in the previous sections indicate the level of ex ante uncertainty shared by respondents. In this section, we are going to look at the predictive accuracy as a proxy for ex post uncertainty. The underlying hypothesis assumes that episodes associated with low ex post predictive forecast errors are indicative of a low level of uncertainty and the opposite when the forecast errors are high. Despite uncertainty, surprises also play a key role in the process of expectation formation. They are defined as the unforecastable part of daily news, and they impact on high frequency asset prices, exchange rates, interest rates or government bond yields [see, for instance, Andersen et al. (2003, 2007), Faust et al. (2007) or Altavilla et al. (2014)]. Jurado et al. (2015) highlight the role of the unpredicted variations and Scotti (2013) builds both, a surprise and an uncertainty index based on the new information contained in the data. The importance of surprises in traditional short-run time series forecasting is also well known, where the use of the forecast errors allow for high adaptability to the most recent characteristics. However, it is not so frequent to use surprises for forecasting macroeconomic indicators at medium-term horizon, specifically in our case, for 1 year ahead forecasts.

The reasons behind the use of surprises in forecasting GDP growth and inflation in the euro area for the period 2005–2015 are twofold: On one hand, there are some concerns on bias and the inclusion of a constant term in the regressions does not seem to help; on the other hand, uncertainty measures (that are always positive) seem to be adding predictive content only in certain specific years, looking as if the sign of the surprise matters.

We define *Surprise* as the forecast error made in moment t by the consensus of the forecasters 1 year in advance:

$$Surprise_t = y_t - Consensus_t \quad (11)$$

Therefore, the absolute value of the surprise at time t (U_{Sur_t}) is an ex post measure of uncertainty:

$$U_{Sur_t} = |Surprise_t|. \quad (12)$$

Figure 1 (bar lines) plots the surprises realized for GDP growth and inflation. We assess the predictive content of ex post uncertainty and surprises for the period 2005–

2015 in the euro area by running two pseudo-real-time forecasting exercises. As in the previous sections, we proceed in two steps and adapt Eqs. (3) and (4) adding the new regressors. However, there are two differences: First, we have to lag the new regressors ($U_{Sur_{t-1}}$ or $Surprise_{t-1}$, respectively) to build valid out of sample forecasts; second, the new regressors will be correlated with the error term in Eq. (4), given that both are 1 year ahead forecast errors on a quarterly basis and will have a MA structure, and we have to estimate using an instrumental variable estimator (as in [Kinal and Lahiri 1988](#); [Rich et al. 1992](#)).

For GDP growth we have used as instrument for the lagged surprises the consensus lagged one period, that is the consensus forecasts made for period $t - 1$ with information available for the forecasters up to $t - 5$. Notice that the first difference of the consensus gives an idea of both, how the forecasters perceive the target variable is changing and what have they learned from the previous forecasts. As the consensus made for time t is already an exogenous regressor, we just use its lagged value as an instrument in the case of GDP growth rate. When we checked for the strength of the instrument, we found that the R^2 of the auxiliary regression of the endogenous regressor (surprises at $t - 1$) over the consensus at t and the consensus at $t - 1$ is 0.56 and the t-statistic for the coefficient associated with the instrument is -4.85 when we use robust standard errors derived from the Newey and West variance-covariance estimator. For the lagged uncertainty we have used as instrument the absolute value of the change in consensus in moment t with respect to $t - 1$ and the statistics are similar to those shown for surprises.

The forecasting results show that the RMSFE ratio with respect to the consensus benchmark are 1.07 and 0.59 with the use of the ex post measure of uncertainty and the surprises, respectively.

Finally, we could not use this approach for inflation. The R^2 of the auxiliary regression of the endogenous regressor (surprises at $t - 1$) over the consensus at t and the consensus at $t - 1$ is only 0.02 and the t-statistic for the coefficient associated with the instrument is around 0.49, meaning that the proposed instrument does not have any explicative power over the endogenous regressor (lagged 1 period surprises). We have tried other possible instruments like lagged (4 periods) values of the observed data or the surprises, but they were not appropriate. A closer look at inflation data (see [Fig. 1](#)) might give an insight of what is going on. Forecasters are under-predicting inflation for the first part of the sample. So the information conveyed in surprises is that they should increase their forecasts of inflation. Then you get to the maximum and the surprise tells you that you should overcorrect more and in the next periods inflation suffers a drop of four points. This leads to the absence of correlation between the forecast errors (surprises) and consensus. The information contained in the sign of the surprises is totally misleading.

6 Concluding remarks

Forecasting 1 year ahead GDP growth and inflation in the euro area for the period 2005–2015 is a challenging task, as the macroeconomic performance has been strongly affected by business cycle fluctuations and exogenous shocks coming mainly from the price of commodities, specially oil.

We have examined expectation and uncertainty measures derived from the ECB-SPF. We have analyzed disagreement and traditional measures of uncertainty based on equal weights and proposed alternative ones based on principal component weights extracted from individual uncertainties. We find that while disagreement has diminished since the crisis, uncertainty measures based on subjective individual uncertainties have remained at high levels.

Regarding the forecast accuracy, consensus from ECB-SPF is hard to beat by more complicated alternatives and we have found that uncertainty does not help to improve forecast accuracy in the period 2005–2015 for either GDP growth or inflation. We have also noted that the bad/good relative performance of more sophisticated models than consensus or including uncertainty can be classified according to the yearly behavior of the indicator. Alternative models seem to improve consensus performance during the years of the crisis while inflation do the opposite.

Finally, we have checked whether an ex post measure of uncertainty based on forecast error and surprises in this specific time span played any role in improving the forecast performance 1 year ahead. The inclusion of surprises improves by 41 % the forecasting performance of consensus in GDP growth. Though acknowledging the specific characteristics of the time span considered for the forecast exercise 2005–2015 in the euro area, the role of surprises in 1 year ahead forecasts could be a matter of further research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abel J, Rich R, Song J, Tracy J (2016) The measurement and behavior of uncertainty: evidence from the ECB Survey of Professional Forecasters. *J Appl Econom* 31(3):533–550
- Aiolfi M, Capistran C, Timmermann A (2011) Forecast combinations. In: Clements M, Hendry D (eds) *The Oxford handbook of economic forecasting*. Oxford University Press, Oxford, pp 355–388
- Altavilla C, Giannone D, Modugno M. (2014) The low frequency effects of macroeconomic News on government bond yields. Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 2014-52
- Andersen TG, Bollerslev T, Diebold FX, Vega C (2003) Micro effects of macro announcements: real-time price discovery in foreign exchange. *Am Econ Rev* 93:38–62
- Andersen TG, Bollerslev T, Diebold FX, Vega C (2007) Real-time price discovery in stock, bond and foreign exchange markets. *J Int Econ* 73:251–277
- Bates JM, Granger CW (1969) The combination of forecasts. *Op Res Q OR* 20(4):451–468
- Bachman R, Estner S, Sims E (2013) Uncertainty and economic activity: evidence from business survey data. *AEJ Macroecon* 5:217–249
- Baker SR, Bloom N, Davis SJ (2013) Measuring economic policy uncertainty. *Chic Booth Res Pap* 13–02
- Bloom N (2009) The impact of uncertainty shocks. *Econometrica* 77(3):623–685
- Bowles C, Fritz R, Genre V, Kenny G, Meyler A, Rautanen T (2007) The ECB Survey of Professional Forecasters (SPF). A review after eight's years experience. *Occasional Paper Series* 59 April 2007
- Bowles C, Fritz R, Genre V, Kenny G, Mayler A, Rautanen T (2010) An evaluation of the growth and unemployment forecasts in the ECB Survey of Professional Forecasters. *OECD J J Bus Cycle Meas Anal* 2:63–90

- Capistrán C, Timmermann A (2009) Forecast combination with entry and exit of experts. *J Bus Econ Stat* 27:428–440
- Clemen RT (1989) Combining forecasts: a review and annotated bibliography. *Int J Forecast* 5(4):559–583
- Conflitti C (2012) Measuring uncertainty and disagreement in the European survey of professional forecasters. *OECD J J Bus Cycle Meas Anal* 2011(2):69–103
- Conflitti C, De Mol C, Giannone D (2015) Optimal combination of survey forecasts. *Int J Forecast* 31(4):1096–1103
- D'Amico S, Orphanides A (2008) Uncertainty and disagreement in economic forecasting. Finance and Economics Discussion Series. Board of Governors of the Federal Reserve System 2008-56
- De Menezes LM, Bunn DW, Taylor JW (2000) Review of guidelines for the use of combined forecasts. *Eur J Oper Res* 120(1):190–204
- Diebold F, Mariano RS (1995) Comparing predictive accuracy. *J Bus Econ Stat* 13(3):253–263
- ECB (2013) How has macroeconomic uncertainty in the euro area evolved recently?. *ECB Mon Bull Oct* :44–48
- ECB (2014) Fifteen years of the ECB survey of professional forecasters. *ECB Mon Bull January* 2014:55–67
- Faust J, Rogers JH, Wang SYB, Wright JH (2007) The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *J Monet Econ* 54(4):1051–1068
- Friedman M (1977) Nobel lecture: inflation and unemployment. *J Polit Econ* 85(3):451–472
- García JA (2003) An Introduction to the ECB's Survey of Professional Forecasters", *ECB Occasional Paper Series* 8 September 2003
- Genre V, Kenny G, Meyler A, Timmermann A (2013) Combining expert forecasts: can anything beat the simple average? *Int J Forecast* 29(1):108–121
- Giannone D, Henry J, Lalik M, Modugno M (2012) An area-wide real-time database for the euro area. *Rev Econ Stat* 94(4):1000–1013
- Giordani P, Söderlind P (2003) Inflation forecast uncertainty. *Eur Econ Rev* 47:1037–1059
- Hall SG, Mitchell J (2007) Combining density forecasts. *Int J Forecast* 23:1–13
- Jurado K, Ludvigson SC, Ng S (2015) Measuring uncertainty. *Am Econ Rev* 105(3):1177–1216
- Kendall M, Stuart A (1977) *The advanced theory of statistics, vol 1, 4th edn.* Macmillan, New York
- Kinal T, Lahiri K (1988) A model for ex ante real interest rates and derived inflation forecasts. *J Am Stat Assoc* 83(403):665–673
- Lahiri K, Sheng X (2010a) Learning and heterogeneity in GDP and inflation forecasts. *Int J Forecast* 26:265–292
- Lahiri K, Sheng X (2010b) Measuring forecast uncertainty by disagreement: the missing link. *J Appl Econom* 25:514–538
- Lahiri K, Peng H, Zhao Y (2016) Online learning and forecast combination in unbalanced panels. *Econom Rev Forthcom.* doi:[10.1080/07474938.2015.1114550](https://doi.org/10.1080/07474938.2015.1114550)
- Lahiri K, Teigland C (1987) On the normality of probability distributions of inflation and GNP forecasts. *Int J Forecast* 3:269–279
- Lahiri K, Teigland C, Zaporowski M (1988) Interest rates and the subjective probability distribution of inflation forecasts. *J Money Credit Bank* 20(2):233–248
- Newbold P, Harvey DI (2002) Forecast combination and encompassing. In: Clements MP, Hendry DF (eds) *A companion to economic forecasting.* Wiley, Hoboken, pp 268–283
- Paloviita M, Viren M (2014) Inflation and output growth uncertainty in individual survey expectations. *Empirica* 41:69–81
- Poncela P, Senra E (2006) A two factor model to combine US inflation forecasts. *Appl Econ* 38(18):2191–2197
- Poncela P, Rodríguez J, Sánchez-Mangas R, Senra E (2011) Forecast combination through dimension reduction techniques. *Int J Forecast* 27(2):224–237
- Rich RW, Raymond JE, Butler JS (1992) The relationship between forecast dispersion and forecast uncertainty: evidence from a survey data–arch model. *J Appl Econom* 7(2):131–148
- Rich R, Tracy J (2010) The relationships between expected inflation, disagreement and uncertainty: evidence from matched point and density forecasts. *Rev Econ Stat* 92(1):200–207
- Scotti C (2013) Surprise and uncertainty indexes: real-time aggregation of real-activity macro surprises. *FRB Int Fin Discuss Pap* 1093
- Stock JH, Watson MW (2004) Combination forecasts of output growth in a seven-country data set. *J Forecast* 23(6):405–430

- Timmermann A (2006) Forecast combinations. In: Elliott G, Granger CWJ, Timmermann A (eds) Handbook of economic forecasting. Elsevier, Amsterdam, pp 135–196
- Wallis KF (2005) Combining density and interval forecasts: a modest proposal. *Oxf B Econ Stat* 67:983–994
- Wallis KF (2011) Combining forecasts-forty years later. *Appl Fin Econ* 21(1–2):33–41
- Zarnowitz V, Lambros LA (1987) Consensus and uncertainty in economic prediction. *J Polit Econ* 95(3):591–621