Weighted Averages of Individual Measures of Uncertainty for the European Central Bank's Survey of Professional Forecasters

Víctor López-Pérez

First draft: 11 March 2014
This draft: 10 November 2015

Abstract

This paper estimates aggregate measures of uncertainty from individual density forecasts by professional forecasters. It improves on the existing literature along two dimensions. First, it controls for changes to the composition of the panel of respondents to the survey. And second, it assigns more weight to the information submitted by forecasters with better forecasting performance. Using data from the European Central Bank's Survey of Professional Forecasters from 1999 Q1 to 2014 Q3, the paper finds that the effects of changes in the composition of the panel on aggregate uncertainty can be large in a statistical and economic sense. It also finds that the estimates of aggregate uncertainty that use performance-based weights differ significantly from the simple averages used in the literature and their dynamics are more consistent with the dynamics displayed by uncertainty indicators computed from financial markets.

Author's affiliation: Departamento de Economía, Facultad de Ciencias de la Empresa, Universidad Politécnica de Cartagena, c/ Real 3, Cartagena, Spain. e-mail: victor.lopez@upct.es. Telephone number: (+34) 868071259. Fax number: (+34) 968325781.

1.   Introduction

Macroeconomic outcomes are the result of millions of decisions taken by economic agents worldwide, and the economic literature tries hard to understand the determinants of these decisions. One of these determinants is the degree of uncertainty in the economy, which is especially important for savings and investment decisions. If uncertainty is large, consumers are expected to save more for precautionary reasons (Caballero 1990) and risk-averse investors may delay irreversible investment plans (Leahy and Whited 1996).

The end of the Great-Moderation era has put the term uncertainty back to the front pages of newspapers and academic articles (Baker, Bloom and Davis 2013). Uncertainty played a major role in the freeze of credit markets worldwide after the fall of Lehman Brothers in 2008. It was also a key factor in the propagation or contagion of the sovereign debt crisis from Greece to other European countries in 2011.

The evolution of uncertainty over time is thereby of interest for academics and researchers, who need estimates of uncertainty to investigate the links between uncertainty and economic outcomes. It is also of interest for policy-makers, who need to closely monitor the available estimates of uncertainty, anticipate the effects on the economy of changes in uncertainty and take the appropriate policy actions to achieve their policy objectives (Bloom 2009).

The first place to look at for measures of risk and proxies for uncertainty is probably the financial market. The Chicago Board Options Exchange Volatility Index (VIX) is a well-known example (Basu and Bundick 2012). However, the non-conventional policies by many governments and central banks in developed economies during the recent past may have involuntary contributed to distort the signals extracted from financial markets in general and proxies for uncertainty in particular (Bekaert, Hoerova and Lo Duca 2013).

In this context, data from surveys may provide a more accurate picture of macroeconomic uncertainty than financial indicators. The density forecasts of macroeconomic variables by professional forecasters are particularly valuable for the estimation of the degree of macroeconomic uncertainty perceived by survey participants. They combine the expertise of highly-skilled professionals with the heterogeneity of views that naturally comes from survey methods. Consequently, measures of uncertainty from surveys like the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters, the European Central Bank's Survey of Professional Forecasters (ECB's SPF) and the Bank of England's Survey of External Forecasters have gained prominence in economic and policy discussions.

The existing literature has explored different measures of uncertainty constructed from survey data. Available measures of uncertainty are the standard deviation of point forecasts, typically known as "disagreement" (Neamtiu, Shroff, White and Williams 2014), the variance of the average density forecast (ECB 2014), the average standard deviation of the individual density forecasts (Giordani and Soederlind 2003), the root mean subjective variance or RMSV (Batchelor and Dua 1996), the implied RMSV (Boero, Smith and Wallis 2008), the average or median inter-quartile range of the individual density forecasts (Engelberg, Manski and Williams 2011, Abel, Rich, Song

and Tracy 2015), and the average entropy of the individual density forecasts (Rich and Tracy 2010, Wallis 2006).

All these measures exhibit two problems. The first is the *panel-composition problem*: the existing measures of aggregate uncertainty do not take into account that the panel of professional forecasters changes from one survey round to the next. Therefore, the evolution of these measures of uncertainty over time may be meaningless because it compounds true changes in uncertainty with artificial changes due to the variations in the panel of respondents to the survey (Engelberg, Manski and Williams 2011).

Figure 1 shows the number of professional forecasters that submitted density forecasts of GDP growth one and two years ahead in each ECB's SPF round. There are large variations in the pool of respondents from one survey round to the next. For instance, the ECB received 49 density forecasts of GDP growth two years ahead in 2005 Q2. In the following round, 2005 Q3, it received only 35. Let's assume that there were no true changes in uncertainty between 2005 Q2 and 2005 Q3. Let's also assume that the forecasters that replied in 2005 Q2 but not in 2005 Q3 perceived less (more) uncertainty on average than the forecasters that participated in 2005 Q2 and 2005 Q3. Then, the estimates of aggregate uncertainty in 2005 Q3 would increase (decrease) not because uncertainty has truly changed but because the ECB lost almost one third of its panel in 2005 Q3.

Abel, Rich, Song and Tracy (2015) partially addressed the panel-composition issue by dropping from their sample the panellists with less than a minimum number of responses. They considered five different scenarios, with the minimum number of responses by panellist set at 5, 10, 15, 20 or 25. While this procedure may reduce the impact of panel changes, their measures of uncertainty still mix data from panellists with different participation patterns and, as a result, are still contaminated by variations in the panel of respondents to the survey.

The second problem the existing measures of aggregate uncertainty used in the SPF literature suffer from is the *same-weight problem*: they all assign the same weight to the uncertainty perceived by each survey participant, without taking forecasting performance into consideration. Put differently, the existing estimates of aggregate uncertainty do not give more weight to the information submitted by a forecaster who always performs better than the average. Similarly, they do not give less weight to the responses by a forecaster that always underperforms. As a result, the unweighted estimates of aggregate uncertainty used in the literature may be biased because they may implicitly give a weight larger than optimal to underperforming forecasters.

There are reasons to believe that the simple average of the individual measures of uncertainty may be different than a weighted average based on forecasting performance. Kenny, Kostka and Masera (2014) found a positive relationship at the individual level between forecasting performance and the variance of the density forecasts in the ECB's SPF: the best-performing respondents submitted density forecasts with more probability in the tails (i.e. they perceived more uncertainty). If the highest weights are given to the best-performing forecasters when aggregating their individual estimates of uncertainty, the resulting estimates of aggregate uncertainty may differ from the unweighted estimates used in the SPF literature.

I am not aware of any study that used performance-based weights to construct estimates of aggregate uncertainty from the degree of uncertainty perceived at the individual level. The literature on forecast combination has explored the use of performance-based weights to obtain a weighted average of individual point forecasts (Stock and Watson 2004, Capistrán and Timmermann 2009, Smith and Wallis 2009, Genre, Kenny, Meyler and Timmermann 2013). Others have used weights to obtain an aggregate density forecast as a combination of individual density forecasts (Hall and Mitchell 2005 and 2007, Geweke and Amisano 2010, Jore, Mitchell and Vahey 2010, Kascha and Ravazzolo 2010, Billio, Casarin, Ravazzolo and van Dijk 2013, Conflitti, De Mol and Giannone 2015). But performance-based weights have never been used for the aggregation of the estimates of uncertainty perceived by each survey participant into an estimate of aggregate uncertainty.

This paper solves these two problems. It solves the *panel-composition problem* by estimating an aggregate measure of uncertainty from the data submitted by forecasters that replied to two consecutive survey rounds. Using data from the ECB's SPF from 1999 Q1 to 2014 Q3, the paper finds that the effects of changes to the composition of the panel on aggregate uncertainty can be large in a statistic and economic sense. Moreover, changes to the composition of the panel may alter the direction of change of aggregate uncertainty measures. In this regard, the current standard of aggregating the results from all the participants in each survey round, independently of their participation in previous rounds, may produce very misleading results.

The paper then solves the *same-weight problem* by using performance-based weights to obtain weighted averages of individual measures of uncertainty. Forecasting performance is assessed with a strictly proper scoring rule: the logarithmic score of the density forecasts (Gneiting and Raftery 2007). The forecasters with better scores are assigned higher weights. The paper finds that, in the ECB's SPF sample, the weighted estimates of aggregate uncertainty differ significantly from the simple averages used in the literature. The differences are statistically significant and economically relevant. In particular, weighted estimates indicate a much larger increase in uncertainty than the simple averages since the start of the financial crisis. Moreover, while the unweighted estimates of aggregate uncertainty have stayed rather flat in the euro area since 2010, the weighted estimates display significant variation. The latter are much more consistent with the shocks that have hit the euro area in the last five years, like the sovereign debt crisis and the recession in 2012. They are also much more in line with the volatility displayed by the uncertainty indicators from financial markets, like the VSTOXX index.

The remainder of the paper is organised as follows. Section 2 briefly presents the ECB's SPF and describes the data used in the analysis. Section 3 discusses how SPF-based measures of uncertainty that do not control for changes to the panel of respondents are affected by its unbalanced nature and proposes a practical solution to this problem. Section 4 presents the estimates of aggregate uncertainty when the individual measures of uncertainty are weighted according to each participant's forecasting performance. Section 5 concludes, discussing the implications of the findings and proposing directions for future research.

2.   The data

2.1   The ECB's SPF

The ECB's SPF surveys expectations by professional forecasters located in the European Union each quarter since 1999 Q1. Most forecasters work for financial institutions while others belong to universities, government agencies, labour unions and business organisations. 103 forecasters have participated at least once in the survey, although average participation is around 60 forecasters per round. The panel is unbalanced: many forecasters do not reply sometimes while others have left the panel and have been replaced with new panellists.

The ECB's SPF surveys density forecasts of the year-on-year inflation rate, the year-on-year GDP growth rate, and the unemployment rate, all for the euro area. The forecasts used in this paper are fixed-horizon expectations one and two years ahead. Fixed-event forecasts (e.g. expectations for the "current calendar year" or the "next calendar year") are not used because uncertainty is expected to shrink mechanically if the forecast horizon shortens from one survey round to the next.[1]

2.2   Treatment of open-ended bins

To produce a density forecast, forecasters are asked to distribute probabilities among a set of predefined bins for each variable.[2] The lowest and the highest bins are open ended (e.g. "less than 0%" or "more than 3.9%"). For the measurement of uncertainty, an assumption needs to be made on the probability placed in open-ended bins. These bins are much less informative than the rest because of their infinite width. Previous studies have typically assumed that open-ended bins have the same or double width than closed bins (e.g. Batchelor and Dua 1996). This assumption may lead to underestimate uncertainty, especially if open-ended bins contain relatively large probabilities. For instance, forecaster 52 in round 2009 Q1 assigned all the probability to GDP growth one year ahead to the "less than -1.0%" bin. Given that the width of the closed bins in the ECB's SPF is 0.5%, it seems hard to believe that he/she assigned all the probability either to the [-1.5%,-1.1%] bin or to the [-2.0%,-1.1%] bin when his/her point forecast was -2.9%.

To avoid drawing wrong inference from uninformative data, any density with at least one open-ended bin that cannot *reasonably* have the same width than the closed bins is removed from the sample. Specifically, any density with 50% more probability in an open-ended bin than in any other bin with non-zero probability is excluded. This resulted in the exclusion of 224 inflation densities, 492 GDP-growth densities and 165 unemployment densities for the sample used in this paper: 1999 Q1 – 2014 Q3.[3] An exception is made with the densities with less than 1% probability in open-ended bins.

---

[1] See http://www.ecb.europa.eu/stats/prices/indic/forecast/html/index.en.html for a full description of the survey.

[2] Details on the bins available to the SPF forecasters and on the forecast horizons surveyed in each SPF round can be obtained from the document "ECB Survey of Professional Forecasters (SPF): description of SPF dataset", available here:
http://www.ecb.europa.eu/stats/prices/indic/forecast/shared/files/dataset_documentation_csv.pdf??8b0b9b a730b2241d43fec92dacd2944d.

[3] For completeness, densities with 100% probability in open-ended bins are also excluded.

Otherwise, all computer-generated densities with support from -∞ to ∞ (e.g. a normal density function) would be excluded.

For the remaining density forecasts, 5265 for inflation, 4861 for GDP growth and 4847 for unemployment, open-ended bins are treated as having the same width than closed bins. Thus, the next sections do not make any distinction between closed and open-ended bins: all bins are assumed to be closed and assumed to have the same width.

2.3   The choice of the measure of uncertainty

The measure used to estimate the degree of uncertainty perceived by each forecaster is new: *the Gini index of the individual density forecast*. Borrowed from the literature on income and wealth inequality, the Gini index (Gini 1955) is based on the Lorenz curve (Lorenz 1905). This curve is typically used to represent the percentage of total wealth in the hands of the poorest *x*% of the population. The Lorenz curve may also be applied to the analysis of uncertainty by representing the cumulative probability allocated to the *x* least likely bins of a density forecast.

If a forecaster faced no uncertainty, her density forecast would have probability 1 in one bin and 0 in the rest. In this case, the Lorenz curve would be 0 from the first bin to the one before the last and then it would jump to 1 in the last interval. If a forecaster faced maximum uncertainty, her density forecast would have the same probability allocated to every bin. Then, the Lorenz curve would increase uniformly from the first bin to the last.

Figure 2 exemplifies the use of Lorenz curves with the ECB's SPF data. The density forecast of GDP growth two years ahead submitted by forecaster 103 in 2015 Q1 has six bins with positive probabilities, with 0.35 and 0.30 probability in the [1.5, 1.9%] and [1.0, 1.4%] bins respectively. This relatively low level of uncertainty pushes the Lorenz curve of this density forecast towards the bottom-right corner on the chart (blue line). The density forecast by forecaster 92 for the same variable and forecast horizon has positive probabilities in all the bins provided by the ECB. The highest probability placed in a single interval by this forecaster was 0.13. This relatively higher level of uncertainty yields a Lorenz curve, the purple line, much closer to the maximum-uncertainty line (the black line).

The individual Gini index of uncertainty is defined as the distance between the 45-degree line and the Lorenz curve of the individual density forecast divided by the area below the 45-degree line:

$$G = -\frac{\sum_{i=1}^{n}(x_i - lc_i)}{\sum_{i=1}^{n}x_i}$$ [1]

where *n* is the number of intervals, *x* is the *n*x1 vector of ordinates from the 45-degree line, $(1/n, 2/n,…, 1)'$, and *lc* is the *n*x1 vector of ordinates from the Lorenz curve of the

individual density forecast. As the original Gini index declines with uncertainty, the sign was changed to turn it into an index that increases with uncertainty.[4]

The Gini index has two advantages over the most frequently-used measure of uncertainty from density forecasts, the standard deviation of the individual density forecast (Batchelor and Dua 1996, Giordani and Soederlind 2003, Boero, Smith and Wallis 2008). First, the Gini index takes its maximum value when the density forecast is uniform, i.e. when the forecaster faces maximum uncertainty and all the bins are perceived as equally likely. The standard deviation of a density forecast reaches its maximum when the forecaster puts 0.5 probability in the lowest bin and the other 0.5 in the highest. Obviously, the formulation of the latter density forecast requires a lot of information, e.g. that the probability of an outcome located in the intermediate bins is zero. This amount of information is completely at odds with the notion of maximum uncertainty.

The second advantage of the Gini index over the standard deviation of a density forecast is that the computation of the standard deviation requires an assumption on how the probability is distributed within each bin.[5] Some studies assume that all the probability allocated to a bin is in the middle point of the bin (Rich and Tracy 2010) while others fit a continuous distribution to the histogram. The normal distribution (Boero, Smith and Wallis 2015) and the generalised beta distribution (Engelberg, Manski and Williams 2011, Abel, Rich, Song and Tracy 2015) are the most commonly used while the piecewise linear distribution (Conflitti 2011) and the skewnormal distribution (García and Manzanares 2007) have also been proposed.

How robust is the standard deviation to these assumptions? I computed the average standard deviation of the individual density forecasts of unemployment two years ahead under two different assumptions. Under the first assumption, all the probabilities are located in the middle point of each bin. Under the second assumption, a normal density is fitted to the each individual density forecast. The correlation coefficient between the quarter-on-quarter percentage changes from both series is 0.86 with a standard error of 0.03. This correlation is quite high but statistically different from 1. Kendall's rank correlation coefficient between the percentage changes from the two series is 0.66, which is high but far away from 1. I would use the standard deviation if all the alternatives needed such assumptions, but there are alternative measures that do not need them.

The Gini index of the density forecast does not require any assumption about the distribution of the probability inside each bin, making it more attractive than the standard deviation of the density forecast as a measure of individual uncertainty. An alternative measure that does not require this assumption either is the entropy of the individual density forecast (Rich and Tracy 2010, Wallis 2006). Why not using the entropy of a density forecast to measure uncertainty? As the Gini index, the entropy takes its maximum value when the density forecast is uniform and it does not need any assumption regarding the distribution of the probabilities within each bin. However, the

---

[4] This formula is the discrete approximation to the area between the 45-degree line and the Lorenz curve. If n were infinity, G would be bounded beween -1 and 0. As n in the ECB's SPF is large but not infinity, the Gini indices of uncertainty computed in this paper are bounded between $(1-n)/n$ and 0.

[5] Another popular measure of uncertainty in the SPF literature, the inter-quartile range, suffers from the same problem.

Gini index has an advantage over the entropy: the non-linear nature of the entropy implies that the effect on the entropy from a certain change in the probabilities of a density forecast depends on the initial values of these probabilities. In the context of a simple example with only two bins, Figure 3 shows that moving 0.1 probability from one bin to the other leads to larger absolute changes in the entropy when the initial probabilities allocated to the two bins are very different, i.e. when the level of entropy is smaller. The Gini index does not suffer from this drawback.[6]

This property of the entropy is relevant in the context of the ECB's SPF because some forecasters assign zero probability to too many intervals. This behaviour has been labelled "overconfidence" and worsens forecasting performance (Kenny, Kostka and Masera 2014).[7] The entropy of the density forecasts submitted by overconfident forecasters is smaller than the entropy of the density forecasts submitted by more "prudent", better-performing forecasters. As changes in the entropy are larger when the initial level of entropy is smaller, changes in the average entropy would be relatively more affected by changes in the behaviour of the overconfident forecasters. Thus, the entropy implicitly puts more weight on the worst forecasters.

The Gini index is immune to this problem as well. For these reasons, I choose the Gini index of the individual density forecast as the estimate of the level of uncertainty perceived by each SPF forecaster.

2.4   Realisation of the forecasted variables

To weight the individual perceptions of uncertainty by each forecaster based on the forecaster's performance, I need to compute scores by forecaster. To this end, I compare the density forecasts with the realisations of the forecasted variables. These realisations are retrieved from the ECB's Euro Area Real-Time Database.[8] This database collects vintages of many macroeconomic variables as they appeared in each issue of the ECB Monthly Bulletin (Giannone, Henry, Lalik and Modugno 2010).

Real-time data is used in this paper because ECB's SPF participants tried to forecast inflation, GDP growth and unemployment as defined by the statistical methodology existing at the time of the production of the forecast. If the latest vintage of data were used instead, differences between forecasts and realisations would not only arise because of forecast errors. They may also be caused by subsequent methodological changes to the calculation of the forecasted variable that led to backward revisions in the historical time series of the variable.

At the time of retrieving the data from the Real-Time Database (July 2015), inflation data was available until September 2014, real GDP growth data until 2014 Q3, and unemployment data until August 2014. Due to the length of the forecast horizons in the ECB's SPF, scores may be computed up to 2013 Q4 and 2012 Q4 for density forecasts of inflation one and two years ahead respectively, 2014 Q1 and 2013 Q1 for density

---

[6] As shown on Figure 3, this is always true with two bins. With three or more bins, the Gini index retains this property as long as the change in the probabilities does not alter the ordering of the bins in the Lorenz curve.

[7] Overconfidence is defined here as "excessive precision in one's beliefs" (Moore and Healy 2008).

[8] http://sdw.ecb.europa.eu/browse.do?node=4843526

forecasts of GDP growth one and two years ahead, and 2013 Q4 and 2012 Q4 for density forecasts of unemployment one and two years ahead.

3.   Measures of uncertainty for the ECB's SPF: solving the *panel-composition* problem.

The SPF literature combines the measures of individual uncertainty obtained from the individual density forecasts into a measure of aggregate uncertainty. This aggregation is done by taking the mean or the median of the measures of individual uncertainty for all the forecasters that participated in a given survey round. By proceeding in this way, the existing estimates of aggregate uncertainty do not take into account that the SPF panel of professional forecasters changes from one survey round to the next. Engelberg, Manski and Williams (2011) argued the evolution of these measures of uncertainty over time is mixing true changes in uncertainty with artificial changes due to the variations in the panel of respondents to the survey. In their analysis of the US SPF, they recommended to go beyond aggregate figures and to examine the changes to the individual responses of the survey. This is the avenue this section takes, trying to ascertain the effects on the aggregate measures of uncertainty caused by the entry and exit of forecasters from the ECB's SPF panel.

Figure 4 shows the average Gini indices of the density forecasts submitted by all the respondents to the ECB's SPF (black lines). Each index has been normalised to 100 in a base quarter.[9] These series are very similar to the ones used in the SPF literature and do not control for changes in the composition of the panel of respondents to the survey.

To control for entry to and exit from the pool of respondents, I select for each survey round the subset of respondents that replied to that survey *and* to the previous survey. Then, I compute the average Gini indices of the density forecasts submitted by this subset of forecasters in both survey rounds. The percentage change between these two indices is an estimate of how uncertainty has changed between the two survey rounds. This estimate is not affected by changes in participation by construction.

These percentage changes in uncertainty from one round to the next make it possible to construct an index of aggregate uncertainty that is not affected by changes to the panel of respondents to the survey. These indices are the green lines in Figure 4. A comparison between the green and black lines clearly shows that controlling for changes in the panel of respondents to the ECB's SPF is important for the estimation of macroeconomic uncertainty in the euro area. Look for example at panels a, b, e or f, where density forecasts of inflation or unemployment are used to estimate uncertainty. Aggregate uncertainty clearly falls from 2002 to 2007 when I control for entry and exit while uncertainty does not decline by much or at all when I use the data from all the respondents to the survey. The period from 2002 to 2007 was one of economic bonanza, characterised by declining uncertainty and volatility: Figure 5 shows the standardised 12-month and 24-month VSTOXX indices of financial market volatility. These indices are the euro-area versions of the VIX index, which has been used to proxy for macroeconomic uncertainty.[10] The VSTOXX indices were much lower in 2007 than in

---

[9] The base quarter is not the same for all panels in Figure 4. The reason will become clear in Section 4.
[10] The VSTOXX indices are based on Eurostoxx 50 real-time options prices and are designed to reflect the market expectations of short-term and long-term volatility by measuring the square root of the implied

2002, more in line with the estimates of macroeconomic uncertainty from the ECB's SPF that control for entry and exit.

Table 1 displays the correlation coefficient between changes in the aggregate Gini indices of uncertainty using the complete pool of respondents and those using the subset of respondents that submitted density forecasts during two consecutive rounds. The correlations are not close to 1, ranging from 0.50 to 0.71. Therefore, the use of a measure of uncertainty that aggregates the results from all the respondents may lead to misleading results because some of its variation is caused not by true changes in uncertainty but by changes in the composition of the panel of respondents.

4. Measures of uncertainty for the ECB's SPF: solving the *same-weight* problem.

The measures of aggregate uncertainty used in the SPF literature are unweighted averages that assign the same weight to the uncertainty perceived by each survey participant, without taking forecasting performance into consideration. They do not give more weight to the information submitted by a forecaster who always performs better than the average, or less weight to the responses by a forecaster that always underperforms. As a result, the unweighted estimates of aggregate uncertainty used in the literature may be biased because they may implicitly give a weight larger than optimal to underperforming forecasters.

4.1 Weighted averages of individual estimates of uncertainty

This section presents estimates of aggregate uncertainty computed with data from the ECB's SPF. These estimates are weighted averages of individual measures of uncertainty, instead of the simple unweighted averages used in the literature. The weight assigned to each forecaster is based on his/her forecasting performance, which is assessed by the logarithmic scoring rule. This rule is one of the four strictly proper scoring methods described by Gneiting and Raftery (2007).

The logarithmic score is one if not the most popular scoring rule (see Hall and Mitchell 2005 and Kascha and Ravazzolo 2010, for examples of uses of the logarithmic score with survey data). The logarithmic score by forecaster *i* in period *t* is:

$$S_{it} = \log p_{rit} \tag{2}$$

where p_{rit} is the probability assigned by the forecaster to the bin that includes the realisation of the forecasted variable. The logarithmic score takes a value of zero if the forecaster assigned all the probability to the bin where the realisation fell and takes a value of minus infinity if the forecaster placed zero probability in that bin. Each forecaster has a different score for each forecasted variable and forecast horizon, i.e. the score for inflation forecasts one year ahead by forecaster 1 is likely to differ from his/her score for inflation forecasts two years ahead.

---

variance across all options of a given time to expiration. The data is obtained from http://www.stoxx.com/download/historical_values/h_vstoxx.txt. The quarterly data shown on Figure 5 are average daily data over the quarter. Data for 2015 Q3 ends on 2 September 2015.

Individual weights will be assigned on the basis of each forecaster's average performance over a period of time. Average performance by forecaster i at time t is computed as the simple average of his/her individual scores from period t-W to period t.

$$\overline{S}_{it} = \frac{1}{1+W} \sum_{w=0}^{W} S_{it-w} \qquad \text{with } W \geq 0 \qquad [3]$$

For instance, if W=1, the individual average score is computed over the current and the previous survey rounds. Obviously, the individual average score cannot be computed for the first W survey rounds. If a forecaster has not participated in all the 1+W survey rounds, the average is computed over the rounds when he/she participated. The weight assigned to forecaster i in survey round t is assumed to be: [11]

$$W_{it} = \frac{e^{\overline{S}_{it-h}}}{\sum_{j=1}^{J} e^{\overline{S}_{jt-h}}} \qquad [4]$$

where J is the number of forecasters that participated in survey round t. This guarantees that the sum of the weights of the participating forecasters equals one. h is the length of the forecast horizon (in quarters) of the variable of interest: h=4 for one-year-ahead forecasts and h=8 for two-years-ahead forecasts. This guarantees that the weights can be computed in real time because the scores are available after four and eight quarters for one- and two-years ahead forecasts respectively.

Intuitively, the best forecasters, i.e. those with the highest average logarithmic scores, will receive the highest weights. These weights may then be used to obtain an estimate of aggregate uncertainty equal to the weighted average of the individual Gini indices of the density forecasts. This is the point at which this paper departs from the existing SPF literature, which uses *simple* averages. Figure 6 shows the resulting estimates of aggregate uncertainty for the expected inflation rate, the expected rate of growth of real GDP and the expected unemployment rate surveyed in the ECB's SPF.

Each chart shows a solid black line, which is the time series of the simple average of the individual Gini indices of uncertainty. The two dashed lines are the bounds of the 95% confidence interval around the simple average. The red line is the weighted average of the individual Gini indices of uncertainty. Note that I use here the data submitted by all forecasters and thereby I am not controlling for changes in the composition of the panel of participants. I will control for entry and exit in sub-section 4.3.

For the estimates of uncertainty based on density forecasts of inflation, the weighted and unweighted estimates are similar until the start of the financial crisis. Then, the weighted estimates suggest a much larger increase in uncertainty than the unweighted estimates. After the initial increase, the unweighted estimates stay rather flat, while the weighted estimates indicate a sharp decrease in uncertainty, followed by another, more moderate increase around 2011-2012 when fears of a potential break-up of the euro area mounted. After the president of the ECB announced in the summer of 2012 that the

---

[11] For the estimation of aggregate uncertainty under optimal weights, see sub-section 4.2.

institution would do "whatever it takes" to preserve the integrity of the euro area, the weighted estimates of uncertainty declined significantly. None of these movements can be observed in the unweighted estimates.

The estimates of uncertainty based on density forecasts of unemployment show similar dynamics. The unweighted estimates stay rather flat after the initial increase in 2008-2009. The weighted estimates show more interesting dynamics, with uncertainty declining sharply in 2010 when GDP in the euro area started to grow around 2% a year. Uncertainty increased again in 2012 when the euro area economy went into a recession. Finally, the weighted estimates of uncertainty fell again in 2014 when GDP growth climbed above zero again. Weighted estimates of uncertainty based on density forecasts of GDP growth show a similar picture, although those based on one year ahead forecasts are more volatile.

Interestingly, the weighted estimates of uncertainty track much better than the unweighted ones the dynamics of uncertainty extracted from financial market indicators. Figure 5 showed the standardised 12 and 24-month VSTOXX indices of stock market volatility in the euro area. The M-shaped dynamics of these indices between 2008 and 2013 are remarkably similar to those of the weighted estimates of uncertainty described above. The unweighted estimates are not able to replicate these fluctuations.

The main results described above are robust to the value of *W*. I have tried with values ranging from zero to eleven quarters. Very high values of *W* increase the chances of giving zero weight to most forecasters, while low values of *W* induce more volatility in the individual weights over time. The values used in Figure 6 try to balance these two effects: *W* is set equal to 1 for inflation forecasts one year ahead and unemployment forecasts two years ahead, 4 for unemployment forecasts two years ahead, 5 for inflation and GDP-growth forecasts two years ahead, and 6 for GDP-growth forecasts one year ahead. Results for different values of *W* are available from the author upon request.

Even with these relatively low values of *W*, there are a few occasions when the average score is minus infinity for all the participants in a survey round. These episodes are concentrated around the start of the financial crisis, which most participants did not foresee. The weighted estimates of uncertainty based on the logarithmic score cannot be computed in such occasions and are replaced by linear interpolations between the previous and the next weighted estimates available. Weighted estimates based on different proper scoring rules, like the Brier score, the rank-probability score or the spherical score do not suffer from this drawback and the main results of the paper are robust to the scoring rule used. Results based on these other scoring rules are available from the author upon request.

An interesting feature of the weighted estimates of aggregate uncertainty is that their volatility is higher than the volatility of the unweighted estimates. This is because of two reasons. First, relative performance is changing over time, especially during the most turbulent periods. It is during these periods when the best forecasters outperform the rest. Consequently, the weights assigned to the best forecasters tend to increase in more turbulent times. In these periods, the weighted estimate of aggregate uncertainty

deviates more from the simple average and moves closer to the individual measures of uncertainty by the best forecasters, which tend to be higher.

This effect can be seen in Figure 7, which shows the weights assigned to the best forecaster, who may change over time. The survey rounds when a weight is not shown are the rounds when the individual average logarithmic scores h quarters before were minus infinity for all the participants. In these cases, the denominator in equation [4] is zero.

The charts show that the weight assigned to the best forecaster increases in more turbulent periods, for example at the start of the financial crisis. These are the periods when the weighted estimate of aggregate uncertainty tends to deviate the most from the simple average (e.g. in 2009).

The second reason for the higher volatility of the weighted estimates compared to the unweighted estimates is that changes in participation from one round to the next cause changes in the weights. This is because the denominator in equation [4] varies, even if relative performances did not change much.

For instance, Figure 7 shows that the volatility of the weights is the highest for GDP growth forecasts (panels c and d in Figure 7), which translates into more volatile weighted estimates of uncertainty (panels c and d in Figure 6). On the contrary, the volatility of the weights is the lowest for inflation forecasts (panels a and b in Figure 7), which leads to more stable weighted estimates of uncertainty (panels a and b in Figure 6). To eliminate the mechanical effects from changes in participation on the weights and then on estimates of uncertainty, I will replicate the analysis using data from subsets of forecasters with the same participation patterns. As mentioned above, this exercise will be done in sub-section 4.3.

Another interesting result of the comparison between weighted and unweighted estimates of aggregate uncertainty is that, when the former deviates from the latter, it is mostly to indicate higher uncertainty. Figure 6 shows how frequently the weighted estimate of aggregate uncertainty crosses the upper bound of the 95% confidence interval around the simple average. However, the weighted estimate very rarely crosses the lower bound of the confidence interval. In other words, simple averages of individual measures of uncertainty computed with data from the ECB's SPF may produce estimates of aggregate uncertainty that are frequently too low. The reason is that the same-weight rule assigns too much weight to "overconfident" forecasters. These forecasters' predictions are typically worse than the average forecast, reducing the signal-to-noise ratio of the estimations of aggregate uncertainty. When the weight assigned to these forecasters is lowered according to their forecasting performance, the estimates of aggregate uncertainty are frequently significantly higher, both statistically and economically.

4.2   Estimates of aggregate uncertainty with optimal weights

The previous sub-section showed estimates of aggregate uncertainty changed with larger weights being given to the best forecasters. However, the functional form used for the calculation of the weights (equation [4]) was *ad-hoc*. This section checks if the main results presented above are robust to the use of optimal weights.

To compute the optimal weights, I follow Conflitti, De Mol and Giannone (CDG 2015) and minimise the Kullback-Leibler Information Criterion (KLIC) between the true density function of the forecasted variable and the combined density forecast submitted by SPF forecasters:

$$KLIC = E\left[\ln q_v - \ln p_v^{spf}(W_1, W_2, ..., W_I)\right] \tag{5}$$

where q_v is the true density function of variable v, and p_v^{spf} is the combined SPF density forecast with time-invariant individual weights W_1, W_2, ..., W_I. CDG assumed that the optimal weights are constant over time, which means that the optimal weights cannot be computed in real time, an important difference with respect to the previous sub-section. The true density function is unobservable but is independent of the weights. Therefore, the weights that minimise the KLIC are the weights that maximise the average logarithmic score of the combined density forecast.

The computation of the optimal weights requires a balanced panel. CDG replaced missing individual density forecasts with uniform densities. Unfortunately, this assumption makes non-respondents to perform better than many respondents, which results in disproportionately high weights for non-respondents. Instead, I removed from the sample the forecasters who have participated less than ten times since 1999 Q1. For the remaining forecasters, I assumed that the logarithmic score of the missing individual density forecasts in a given survey round is equal to the average logarithmic score of the density forecasts submitted by the remaining forecasters that participated in that survey round. This procedure is done separately for each forecasted variable (inflation, GDP growth and unemployment) and forecast horizon (one and two years ahead) because the optimal weights will be different for different variables and horizons.

Once the panel is balanced, I removed from the panel those forecasters whose logarithmic score is minus infinity in a period when at least another forecaster's score was better. This reduces the cross-sectional dimension of the optimisation problem, and the maximisation of the average logarithmic score would have assigned a weight equal to zero to these forecasters anyway. Interestingly, there are a few survey rounds when the logarithmic score is minus infinity for all the participating forecasters (e.g. in 2009 Q1 for GDP growth forecasts one year ahead). This happens for instance when the realised value of the forecasted variable does not belong to any of the available closed bins. To avoid penalising the forecasters who participated in a survey round with this technical deficiency of the questionnaire, I excluded these periods from the maximisation problem.

Two maximisation routines are conducted. The first is a grid search, where the average logarithmic score of the combined density forecast is computed for every possible combination of non-negative weights that sums one. The weights in the grid change in 0.01 steps. The second routine is the algorithm proposed by CDG. Both procedures yield the same results, and the number of non-zero elements in the vector of optimal weights turns out to be very small: it ranges from one, for density forecasts of inflation two years ahead, to three, for density forecasts of inflation one year ahead and unemployment two years ahead.

13

I use now the optimal weights to obtain the weighted averages of the individual estimates of uncertainty. Figure 8 shows the comparison between these weighted estimates and the simple unweighted estimates used in the literature. One thing to note is that there are some gaps in the weighted estimates. This happens when none of the participants with a positive optimal weight participated in a survey round (the panel was balanced for the optimisation but is back to its unbalanced form for the estimation of uncertainty).

Apart from this, the results obtained in the previous sub-section are robust to the use of optimal weights:

- The weighted estimates of uncertainty increased by much more than the unweighted estimates at the start of the financial crisis. With the exception of panel *b*, the estimated increase in uncertainty is several times larger using the weighted estimates. This effect is especially clear in panel *d* (estimates obtained from forecasts of GDP growth two years ahead): the unweighted estimate increased from around -0.85 to -0.80, while the weighted estimate surged from around -0.80 to above -0.60.

- The unweighted estimates stayed relatively flat after their increase at the beginning of the financial crisis, while the weighted estimates of uncertainty display significant variation. For instance, panel c (estimates obtained from forecasts of GDP growth one year ahead) shows that the weighted estimate of uncertainty declined significantly during 2009 and 2010 and then increased again in 2011. Again, the unweighted estimates of uncertainty fail to produce such swings.

- The weighted estimates of uncertainty tend to suggest higher levels of uncertainty than the unweighted estimates. All the panels in Figure 8 show that deviations of the weighted average from the simple average are mostly to the upside. Importantly, these deviations are statistically significant and economically important. As described above, this result suggests that the best forecasters submit forecasts with more probability in the tails.

All in all, the results with the optimal weights go hand in hand with the main results presented in sub-section 4.1. But are these results still valid when the different participation patterns by SPF panellists are taken into account? The next sub-section investigates this issue.

4.3.   Estimates of aggregate uncertainty using subsets of forecasters with the same participation pattern

The analysis presented in sub-sections 4.1 and 4.2 has ignored that the ECB's SPF panel dataset is unbalanced. In Section 3, I showed that the estimated changes in aggregate uncertainty from one round to the next may be significantly contaminated by changes in participation. To control for entry and exit, I proposed to estimate changes in aggregate uncertainty between two rounds by focusing on the replies by the forecasters that have participated in both rounds. This is the approach followed in this sub-section.

Figure 9 shows the resulting estimates of aggregate uncertainty when the same methodology described in Section 3 is applied here. Each chart shows three lines. The

black line is the unweighted average of the individual measures of uncertainty from all the respondents. It is the same index depicted in black in Figure 3. The green and the red lines are two indices of aggregate uncertainty obtained from the replies by forecasters who participated in two consecutive rounds. The green index assigns the same weight to each of these forecasters and is equal to the green line shown on Figure 3. The red index gives them different weights according to their forecasting performance.

The weights used for the calculation of the weighted estimates of aggregate uncertainty (the red lines in Figure 9) are the real-time weights computed using the methodolofgy described in sub-section 4.1. For the subset of forecasters who participated in rounds t-1 and t, the period-t weights are used to compute the weighted averages of their individual measures of uncertainty in periods t-1 and t. As stated above, the period-t weights are computed using the last $1+W$ quarters of available data. W varies across forecasted variables and forecast horizons. Therefore, the weighted estimates of aggregate uncertainty start at different dates in each chart because the survey data before 1999 Q1 is not available. The start date for the weighted estimates is defined as the base quarter for all the indices shown in each chart.[12]

The weights are re-scaled up to sum one, if necessary, because some forecasters who received a positive weight in round t may have not participated in round t-1. The gaps in the red lines in Figure 9 appear when zero weights are assigned to all the forecasters who participated in both rounds. In these cases, for the first value of the index after the gap, I used the data from the subset of forecasters who participated in the last round before the gap and in the first round after the gap. In one case, however, zero weights were still assigned to all the forecasters who participated in these rounds (panel b, round 2009 Q1) and the change in the index from its previous value was assumed to be equal to the change in the weighted estimate of aggregate uncertainty for all the respondents.

Four results are obtained from this analysis. First, the increase in aggregate uncertainty since the start of the financial crisis is estimated to be larger when higher weights are given to the best forecasters. For instance, the weighted estimate in panel *a* jumps by more than 14% between 2008 and 2009, while the unweighted indices did so by around 5% only. This finding confirms the results obtained in the previous sub-sections.

Second, the unweighted estimates tend to be rather flat after 2009, something not very realistic as explained above. The weighted estimates show significant variation after 2009. Again, this result confirms the findings presented in sub-sections 4.1 and 4.2. For instance, the decline in the weighted estimates of aggregate uncertainty in the second half of 2009 and 2010 (see panels *a*, *b* and *c*) and their increase from the second half of 2011 (see panels *b*, *c*, *d*, *e* and *f*) match very well with the movements in the VSTOXX indices shown on Figure 5.

Third, the weighted estimates of aggregate uncertainty computed from the forecasts of nominal variables (inflation) declined after 2009 and stabilised around relatively high levels. But the weighted estimates computed from the forecasts of real variables (GDP growth and unemployment) showed little signs of stabilisation and continued to increase until 2014. Higher and higher uncertainty may help explain why the ECB took a series

---

[12] This is the reason for the different base quarters across panels in Figure 4.

of non-conventional monetary-policy measures since 2010, which culminated in the announcement of the Public Sector Purchase Programme (or "Quantitative Easing") in 2015. While these policy measures have calmed financial markets, as suggested by the sharp decline in the VSTOXX indices since 2012, they do not seem to have had much success to reduce aggregate macroeconomic uncertainty, at least as perceived by professional forecasters. The developments in the unweighted estimates of aggregate uncertainty do not give any reason to take progressively more aggressive policy measures, as they remain stable or decline slightly since 2009.

Fourth, the weighted estimates of aggregate uncertainty obtained in this section are much less volatile than the weighted estimates presented in sub-sections 4.1 and 4.2. As expected, much of the volatility in the weighted estimates shown on Figures 6 and 9 was caused by changes to the panel of survey participants. When the forecasts by panellists with the same participation patterns are used, the noise generated by entry and exit of participants is eliminated from the estimates of aggregate uncertainty.

Overall, the results presented in this section confirm that unweighted estimates of aggregate uncertainty may mislead researchers and policy-makers because they implicitly assign higher-than-optimal weights to forecasters with very poor track records. The alternative consists on using estimates of aggregate uncertainty that give more weight to the best forecasters. This section has shown that these weighted estimates give rise to more reasonable and interesting dynamics of aggregate uncertainty than the unweighted estimates.

5.   **Conclusion**

The SPF literature has developed many different measures of aggregate uncertainty. I advocate that all of them suffer from two problems. The first problem is what I called the *panel-composition problem*: the existing measures of aggregate uncertainty do not take into account that the panel of professional forecasters changes from one survey round to the next. Therefore, the evolution of these measures of uncertainty over time is meaningless because it compounds true changes in uncertainty with artificial changes due to the variations in the panel of respondents to the survey.

The second problem is what I called the *same-weight problem*: the existing measures of aggregate uncertainty assign the same weight to the uncertainty perceived by each survey participant, without taking forecasting performance into consideration. They do not give more weight to the information submitted by a forecaster who always performs better than the average. As a result, the unweighted estimates of aggregate uncertainty used in the literature may be biased because they may implicitly give a weight larger than optimal to underperforming forecasters.

This paper deviates from the existing literature in two dimensions. It solves the *panel-composition problem* by estimating an aggregate measure of uncertainty from the data submitted by forecasters that replied to two consecutive survey rounds. Using ECB's SPF data from 1999 Q1 to 2014 Q3, the paper finds that the effects of changes in the composition of the panel on aggregate uncertainty can be large in a statistical and economic sense. In this regard, the current standard of aggregating the results from all

the participants in each survey round, independently of their participation in previous rounds, may produce very misleading results.

The paper then solves the *same-weight problem* by using performance-based weights to obtain weighted averages of individual measures of uncertainty. Forecasting performance is assessed with the logarithmic score of the density forecasts. The forecasters with better scores are assigned higher weights. The paper finds that the weighted estimates of aggregate uncertainty differ significantly from the simple averages used in the SPF literature. The differences are statistically significant and economically relevant. In particular, weighted estimates indicate a much larger increase in uncertainty than the simple averages since the start of the financial crisis. Moreover, while the unweighted estimates of aggregate uncertainty have stayed rather flat in the euro area since 2010, the weighted estimates display significant variation. The latter are much more consistent with the shocks that have hit the euro area in the last five years, like the sovereign debt crisis and the recession in 2012. They are also much more in line with the volatility displayed by the uncertainty indicators from financial markets, like the VSTOXX index.

Future research will revisit the link between uncertainty and macroeconomic outcomes using weighted estimates of aggregate uncertainty and controlling for changes in the panel of respondents to the survey. In particular, the robustness of the negative relationship between uncertainty and expected GDP growth may be explored when weighted estimates of aggregate uncertainty are used instead of simple averages.

References

Abel J, Rich R, Song J, Tracy J (2015) The measurement and behavior of uncertainty: evidence from the ECB Survey of Professional Forecasters. Forthcoming in the Journal of Applied Econometrics.

Baker SR, Bloom N, Davis SJ (2013) Measuring economic policy uncertainty. Chicago Booth Research Paper 13-02.

Basu S, Bundick B (2012) Uncertainty shocks in a model of effective demand. Boston College Working Paper 774.

Batchelor RA, Dua P (1996) Empirical measures of inflation uncertainty: a cautionary note. Applied Economics 28: 333-341.

Bekaert G, Hoerova M, Lo Duca M (2013) Risk, uncertainty and monetary policy. Journal of Monetary Economics 60: 771-788.

Billio M, Casarin R, Ravazzolo F, van Dijk HK (2013) Time-varying combinations of predictive densities using nonlinear filtering. Journal of Econometrics 177: 213-232.

Bloom N (2009) The impact of uncertainty shocks. Econometrica **77**: 623-685.

Boero G, Smith J, Wallis KF (2008) Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters. The Economic Journal 118: 1107-1127.

Boero G, Smith J, Wallis KF (2015) The measurement and characteristics of professional forecasters' uncertainty. Forthcoming in the Journal of Applied Econometrics.

Caballero RJ (1990) Consumption puzzles and precautionary savings. Journal of Monetary Economics 25: 113-136.

Capistrán C, Timmermann A (2009) Forecast combination with entry and exit of experts. Journal of Business and Economic Statistics 27: 428-440.

Conflitti C (2011) Measuring uncertainty and disagreement in the European Survey of Professional Forecasters. Journal of Business Cycle Measurement and Analysis 7: 69-103.

Conflitti C, De Mol C, Giannone D (2015) Optimal combination of survey forecasts. International Journal of Forecasting 31: 1096-1103.

ECB (2014) Fifteen years of the ECB Survey of Professional Forecasters. ECB Monthly Bulletin, January: 55-67.

Engelberg J, Manski CF, Williams J (2011) Assessing the temporal variation of macroeconomic forecasts by a panel of changing composition. Journal of Applied Econometrics 26: 1059-1078.

Garcia JA, Manzanares A (2007) What can probability forecasts tell us about inflation risks? ECB Working Paper 825.

Geweke J, Amisano G (2010) Optimal prediction pools. Journal of Econometrics 164: 130–141.

Genre V, Kenny G, Meyler A, Timmermann A (2013) Combining expert forecasts: can anything beat the simple average? International Journal of Forecasting 29: 108-121.

Giannone D, Henry J, Lalik M, Modugno M (2010) An area-wide real-time database for the euro area. ECB Working Paper 1145.

Gini C (1955) Memorie di metodologica statistica. EV Veschi, Rome.

Giordani P, Soederlind P (2003) Inflation forecast uncertainty. European Economic Review 47: 1037–59.

Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 477: 359-378.

Hall SG, Mitchell J (2005) Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR 'fan' charts of inflation. Oxford Bulletin of Economics and Statistics 67: 995-1033.

Hall SG, Mitchell J (2007) Combining density forecasts. International Journal of Forecasting 23: 1–13.

Jore AS, Mitchell J, Vahey SP (2010) Combining forecast densities from VARs with uncertain instabilities. Journal of Applied Econometrics 25: 621–634.

Kascha C, Ravazzolo F (2010) Combining inflation density forecasts. Journal of Forecasting 29: 231-250.

Kenny G, Kostka T, Masera F (2014) Density characteristics and density forecast performance: a panel analysis. Empirical Economics. doi: 10.1007/s00181-014-0815-9.

Leahy JV, Whited TM (1996) The effect of uncertainty on investment: some stylized facts. Journal of Money, Credit and Banking 28: 64-83.

Lorenz MO (1905) Methods of measuring the concentration of wealth. Publications of the American Statistical Association 9: 209–219.

Moore DA, Healy PJ (2008) The trouble with overconfidence. Psychological Review 115: 502–517.

Neamtiu M, Shroff N, White HD, Williams CD (2014) The impact of ambiguity on managerial investment and cash holdings. Journal of Business Finance & Accounting 41: 1071-1099.

Rich R, Tracy J (2010) The relationships among expected inflation, disagreement, and uncertainty: evidence from matched point and density forecasts. Review of Economics and Statistics 92: 200-207.

Smith J, Wallis KF (2009) A simple explanation of the forecast combination puzzle. Oxford Bulletin of Economics and Statistics 71: 331-355.

Stock JH, Watson MW (2004) Combination forecasts of output growth in a seven-country data set. Journal of Forecasting 26: 405-430.

Wallis, KF (2006) A note on the calculation of entropy from histograms. MPRA Paper 52856.

Charts and tables

Table 1 Correlation between the changes in the aggregate Gini index of uncertainty using the complete pool of respondents and using the subset of respondents that submitted their density forecasts during two consecutive rounds.

Variable	Forecast horizon	Correlation coefficient (1999:2-2014:3)
Inflation	1 year ahead	0.67
	2 years ahead	0.65
GDP growth	1 year ahead	0.71
	2 years ahead	0.59
Unemployment	1 year ahead	0.64
	2 years ahead	0.50

Fig. 1 Number of ECB's SPF participants that submitted density forecasts of GDP growth one and two years ahead in each survey round.
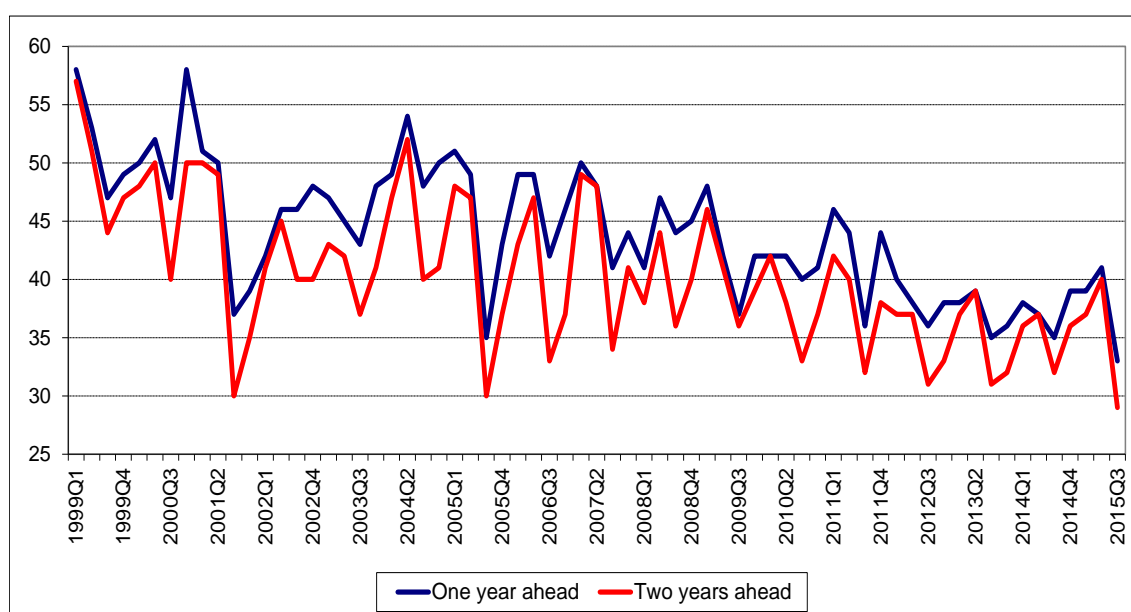
Fig. 2 Lorenz curves from density forecasts of GDP growth two years ahead from the 2015 Q1 SPF round.
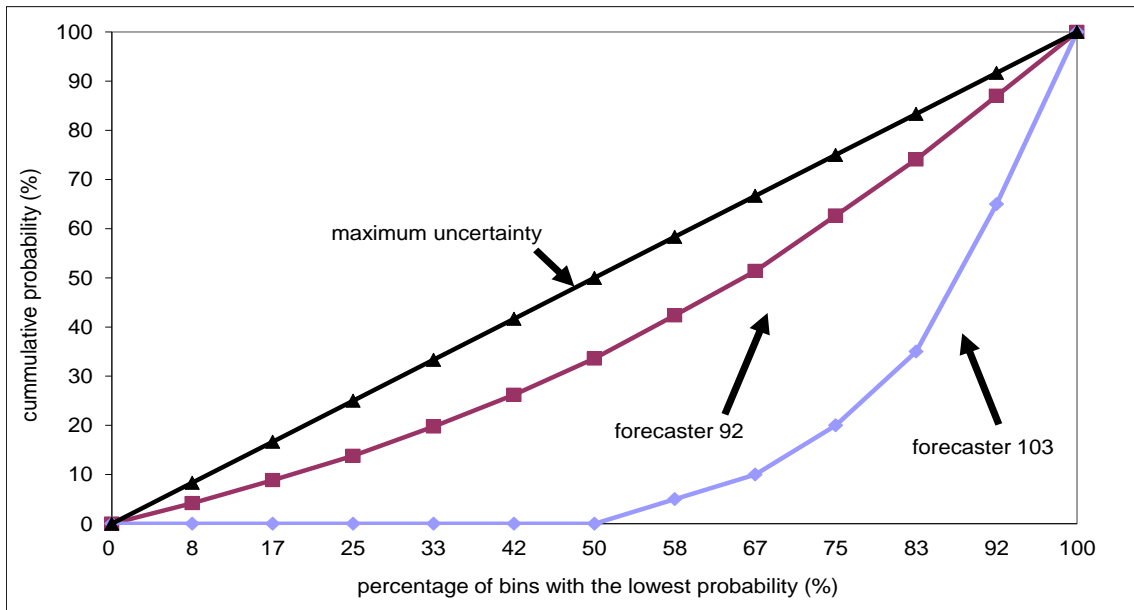
Fig. 3 Illustration of the absolute changes in the entropy and the Gini index when 0.1 probability is moved across bins (example with two bins).
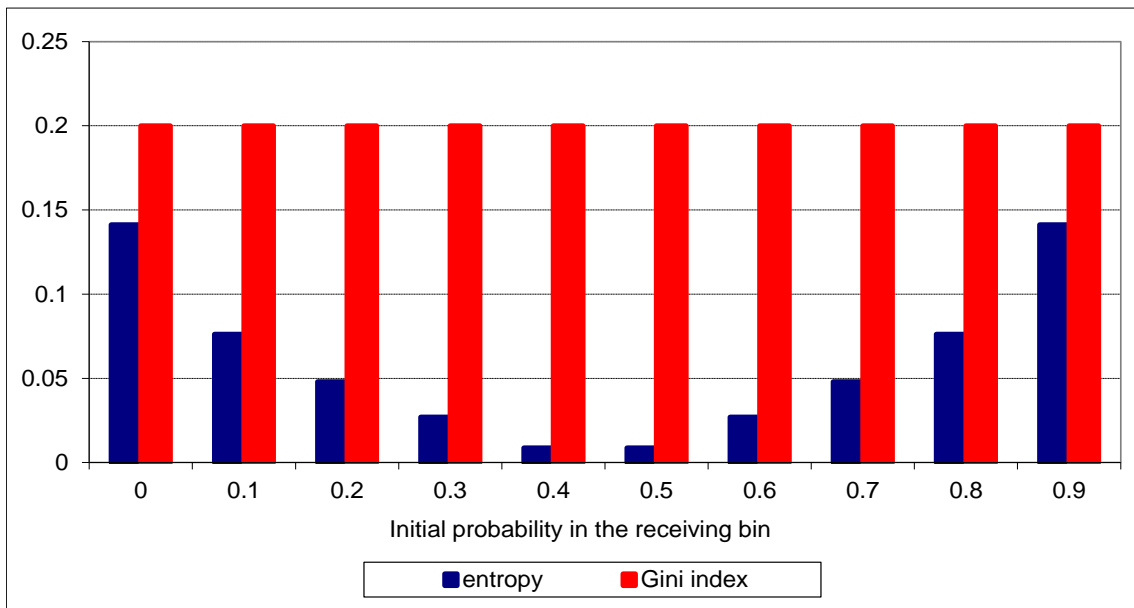
Fig. 4 Estimates of aggregate uncertainty for the ECB's SPF.

a) From density forecasts of inflation one year ahead (2000 Q1 = 100).
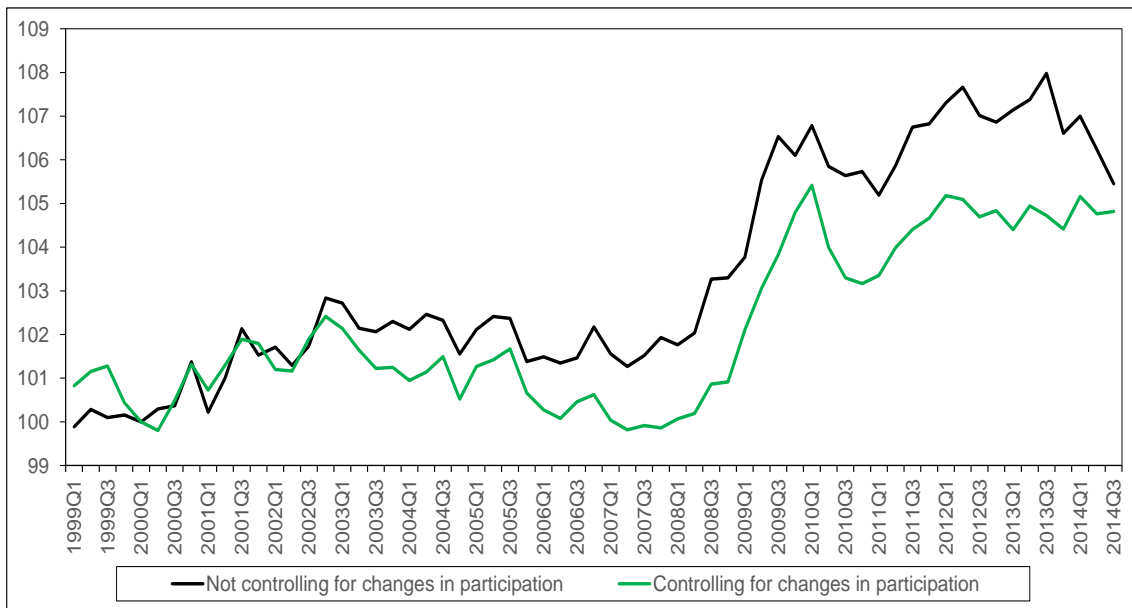
b) From density forecasts of inflation two years ahead (2002 Q1 = 100).
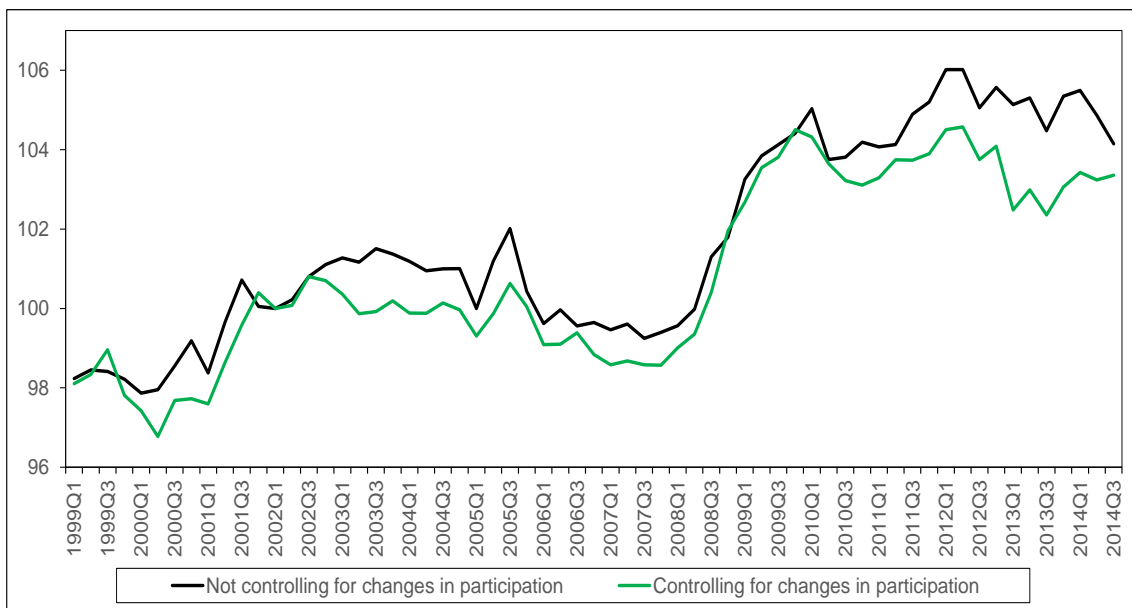
Fig. 4 Estimates of aggregate uncertainty for the ECB's SPF (cont.).

c) From density forecasts of GDP growth one year ahead (2001 Q2 = 100).
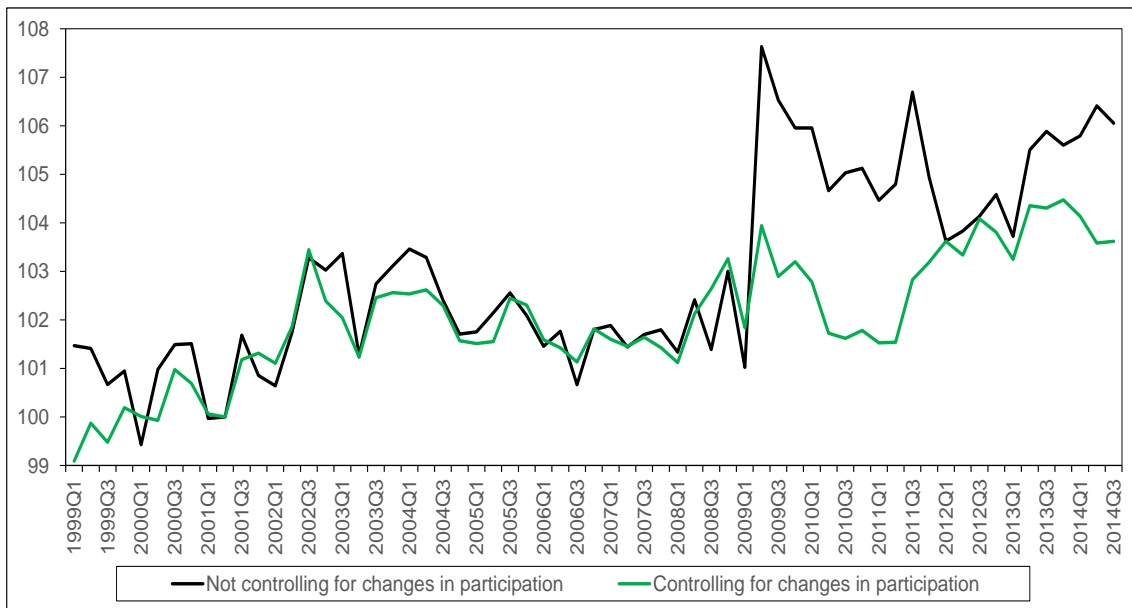
d) From density forecasts of GDP growth two years ahead (2002 Q1 = 100).
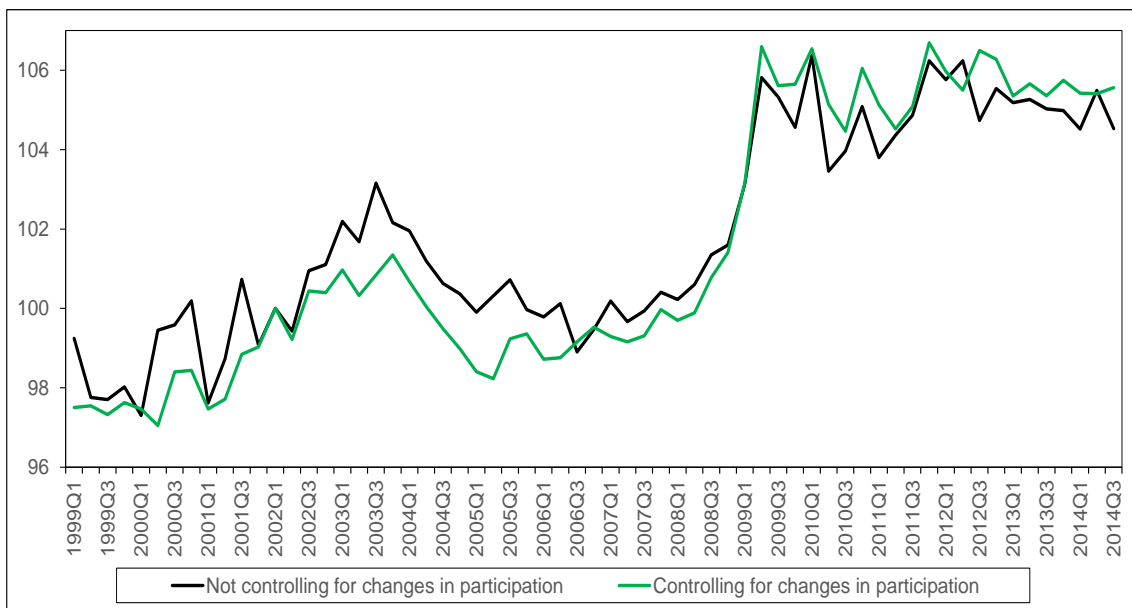
Fig. 4 Estimates of aggregate uncertainty for the ECB's SPF (cont.).

e) From density forecasts of unemployment one year ahead (2000 Q4 = 100).
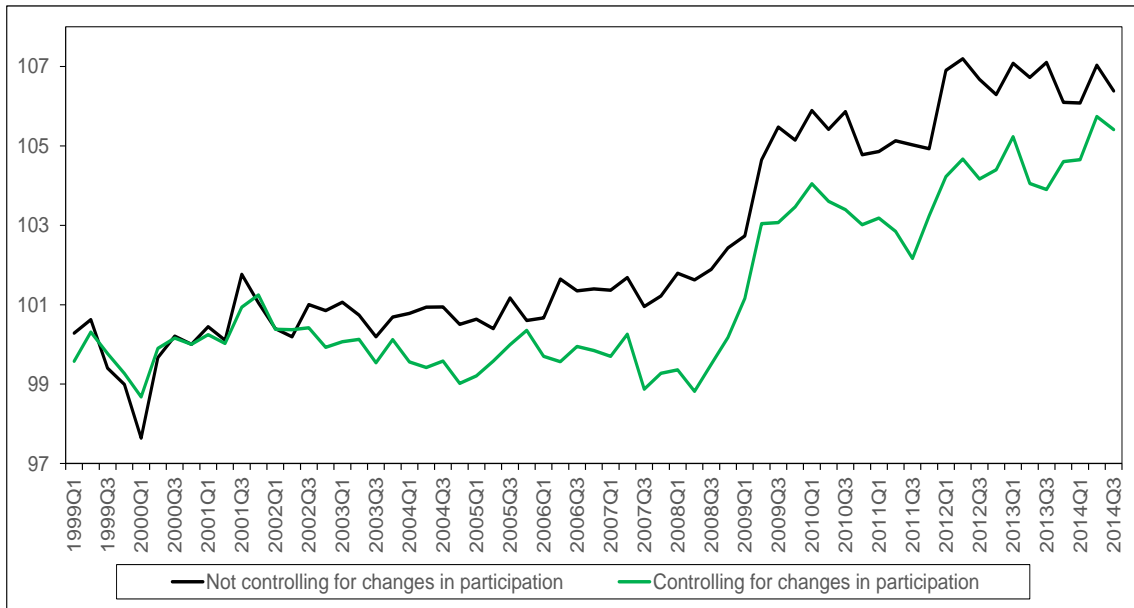
f) From density forecasts of unemployment two years ahead (2001 Q1 = 100).
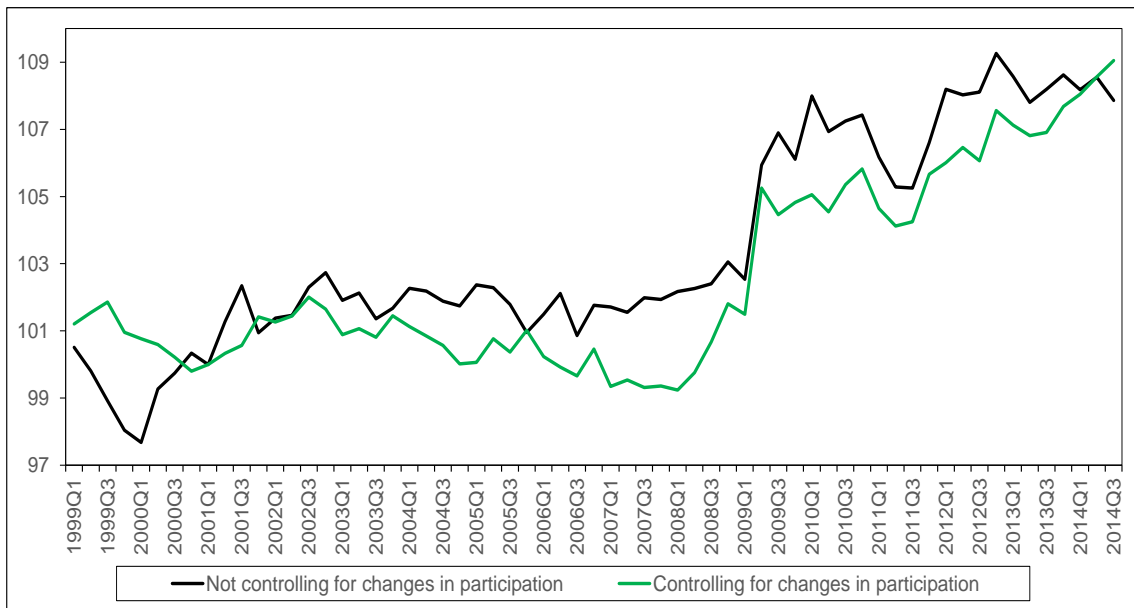
Fig. 5 Standardised estimates of financial-market volatility from European financial markets.
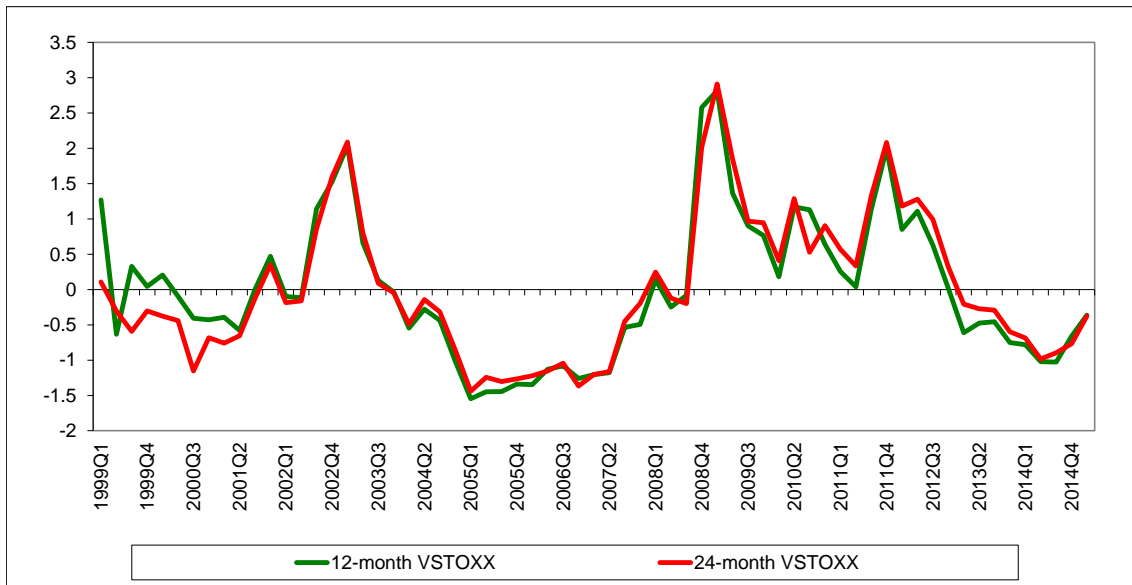
Source: stoxx.com and own calculations.

Fig. 6 Simple and weighted averages of individual measures of uncertainty.

a) From density forecasts of inflation one year ahead.
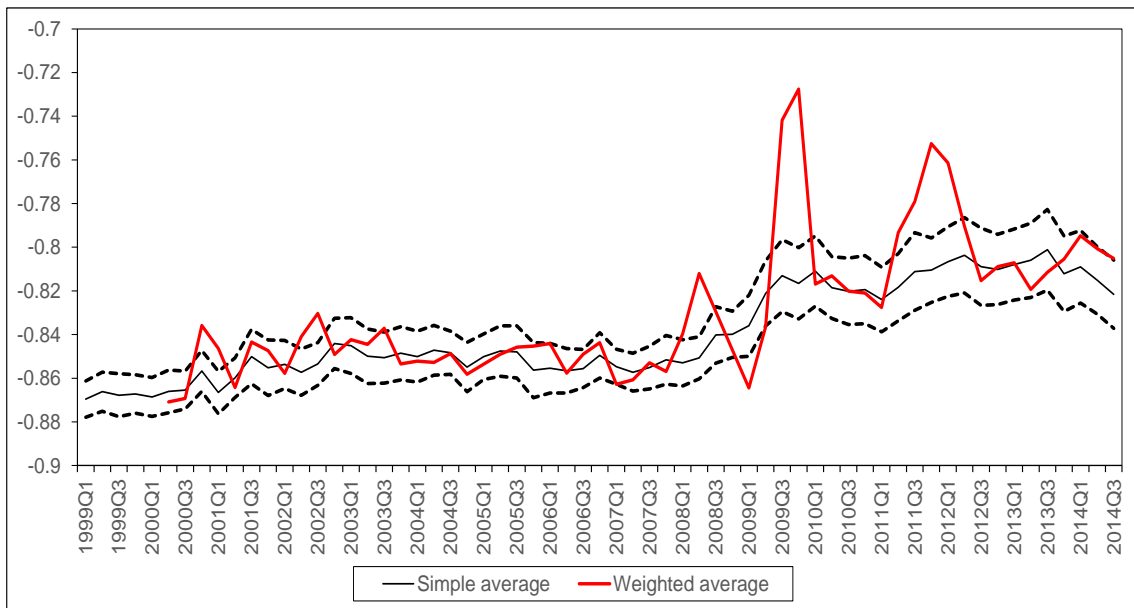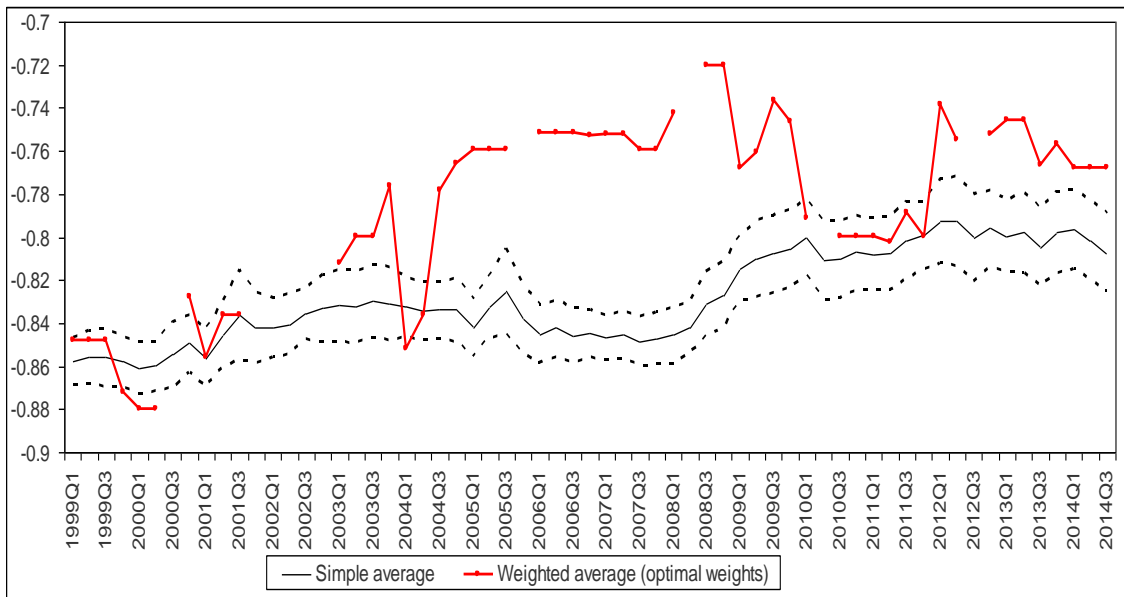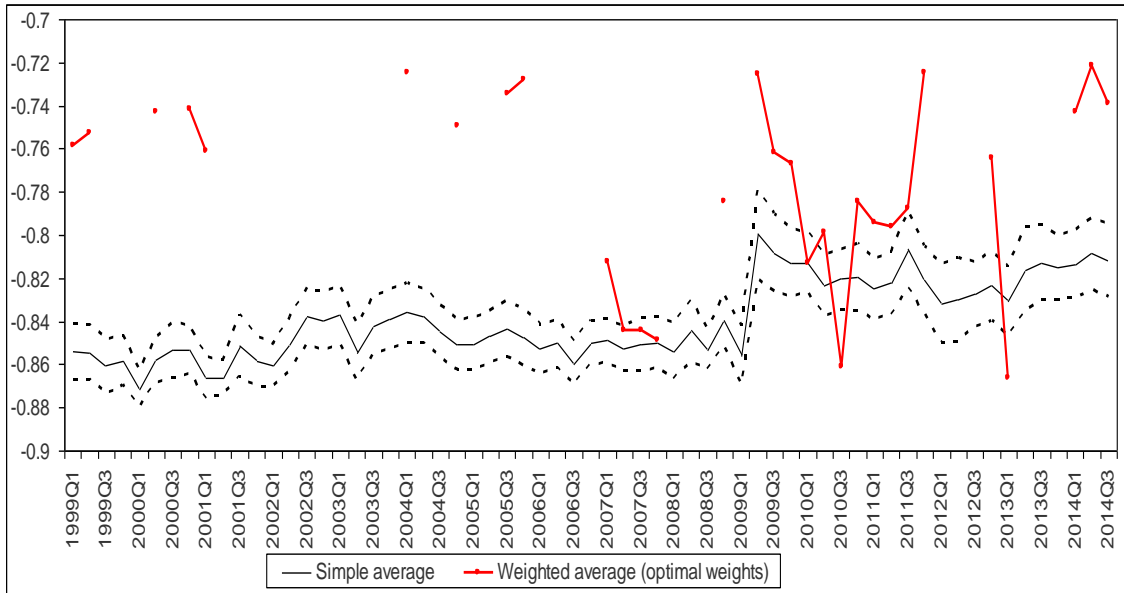

b) From density forecasts of inflation two years ahead.
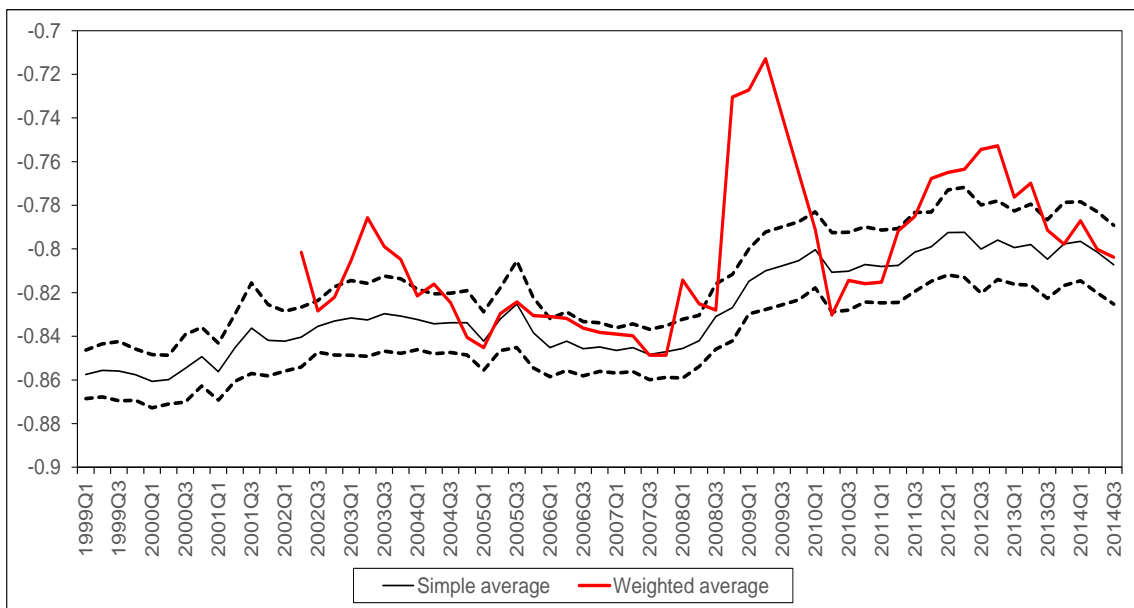
Fig. 6 Simple and weighted averages of individual measures of uncertainty (cont.).

c) From density forecasts of GDP growth one year ahead.
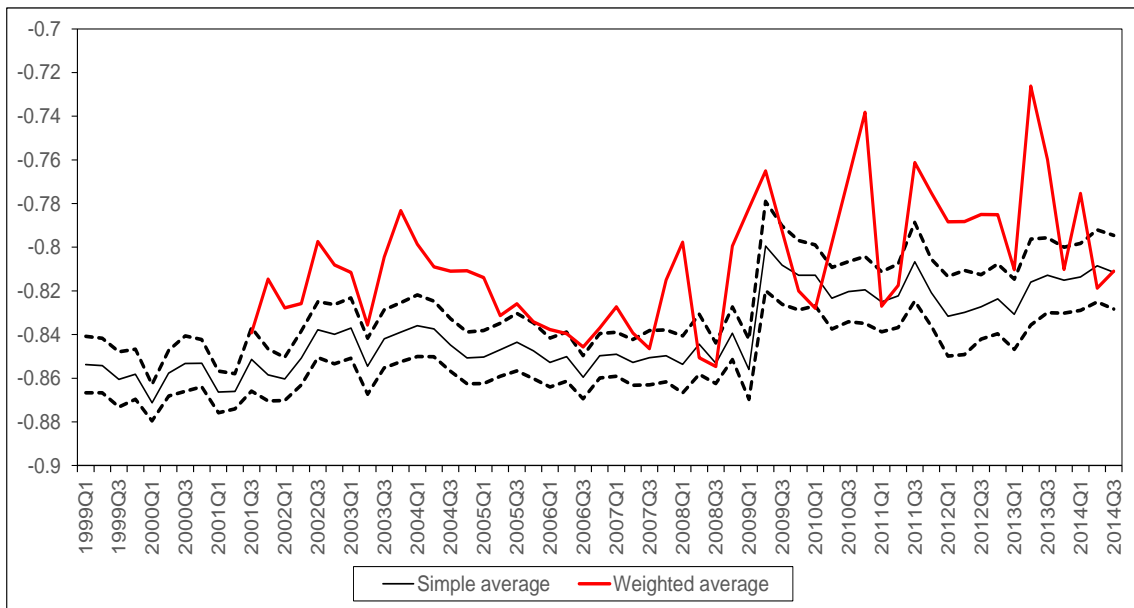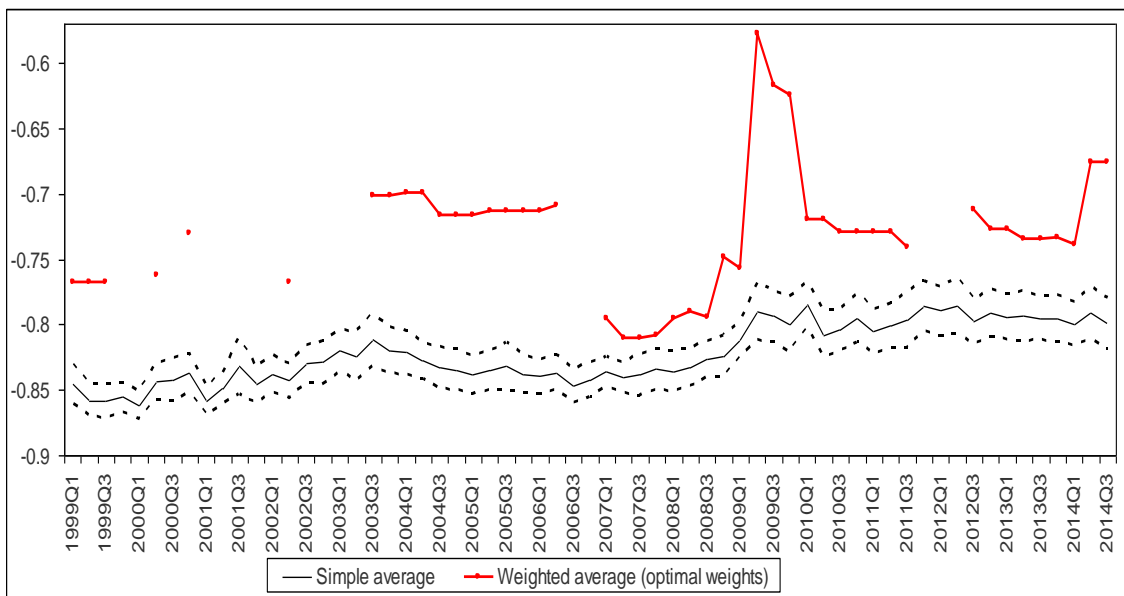

d) From density forecasts of GDP growth two years ahead.
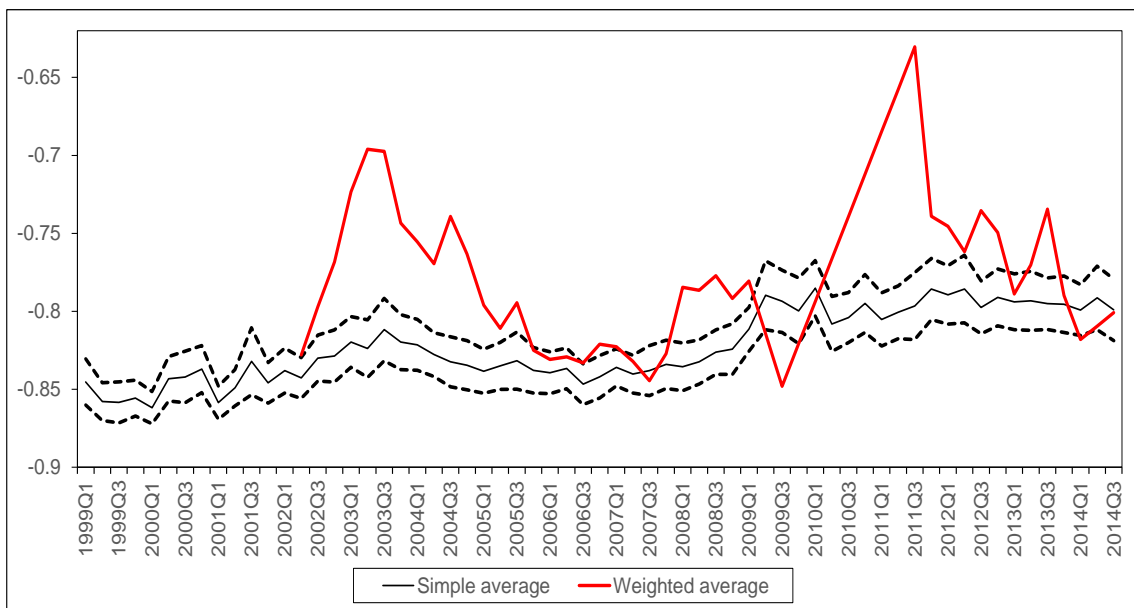
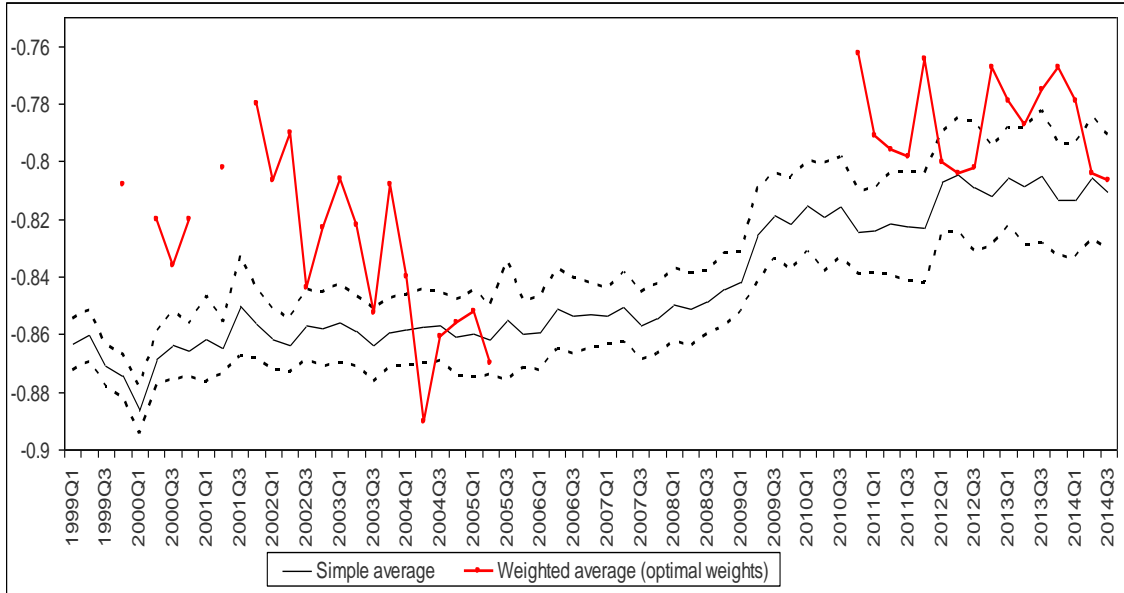**Fig. 6** Simple and weighted averages of individual measures of uncertainty (cont.).

e) From density forecasts of unemployment one year ahead.
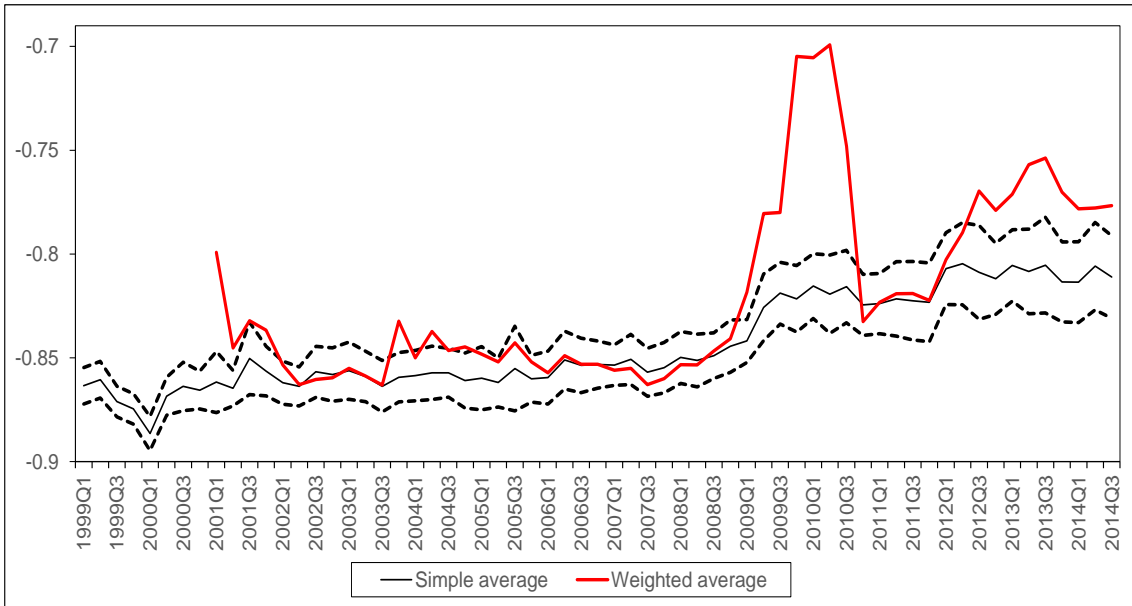
f) From density forecasts of unemployment two years ahead.
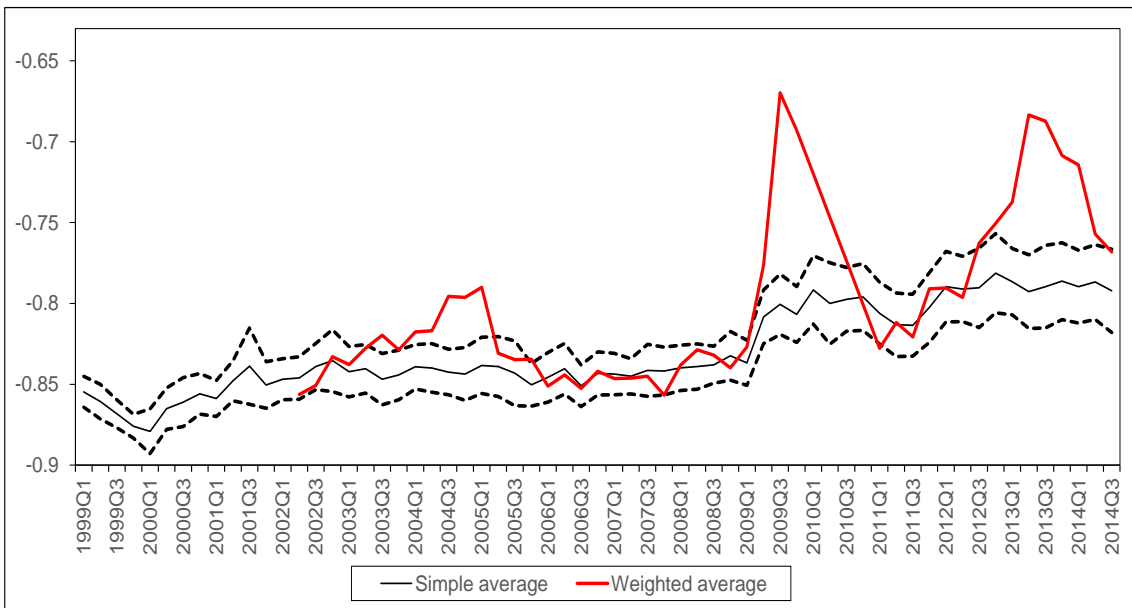
**Fig. 7** Weight assigned to the best forecaster.

a) Using density forecasts of inflation one year ahead.
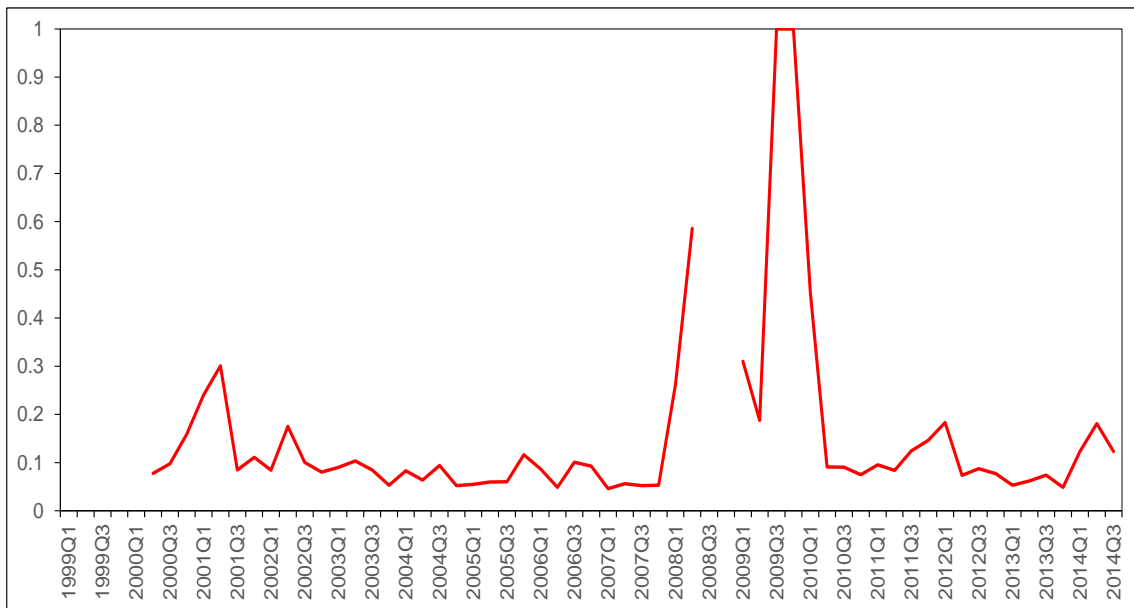
b) Using density forecasts of inflation two years ahead.
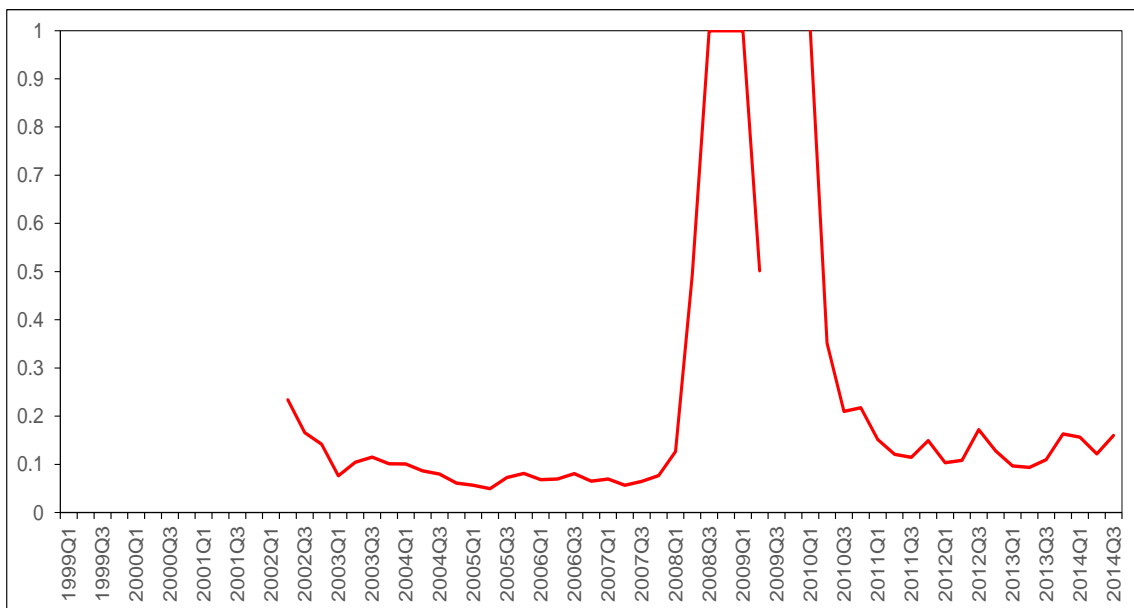
Fig. 7 Weight assigned to best forecaster (cont.).

c) Using density forecasts of GDP growth one year ahead.
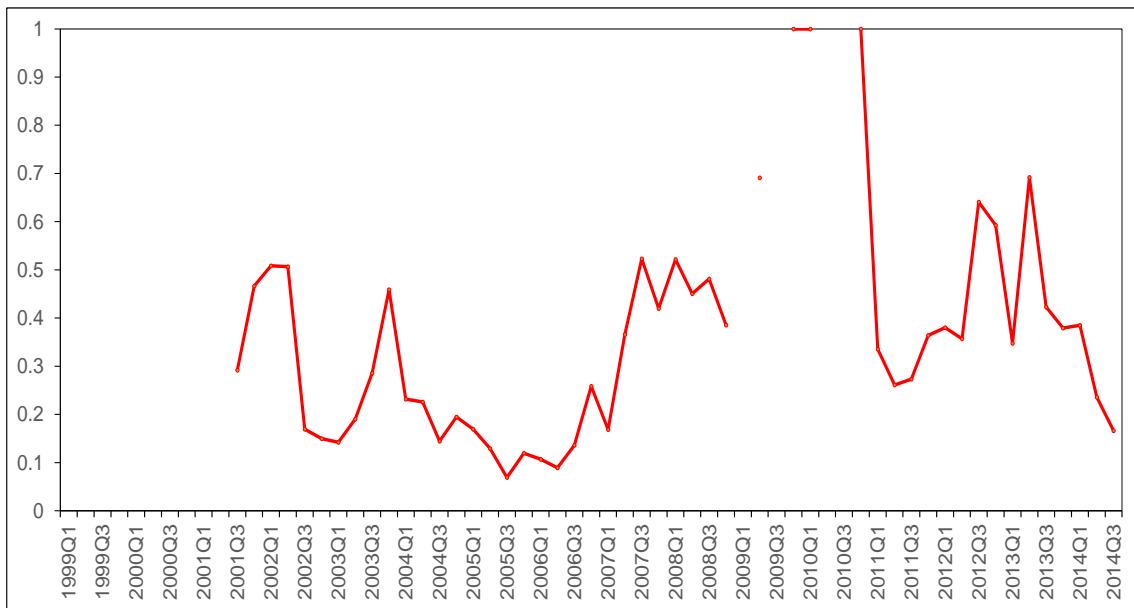
d) Using density forecasts of GDP growth two years ahead.
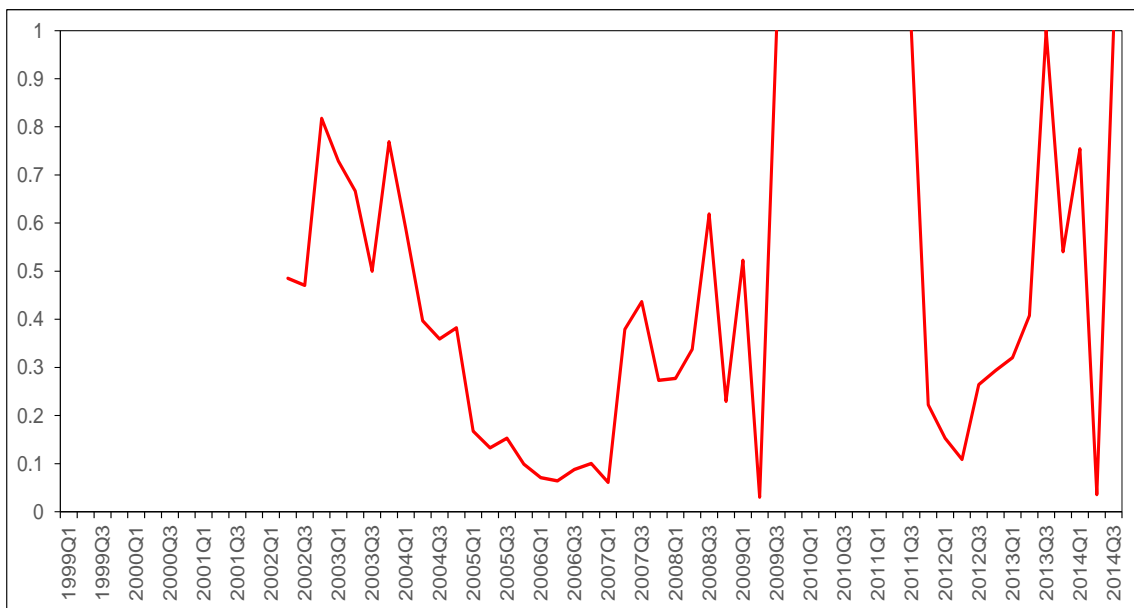
Fig. 7 Weight assigned to best forecaster (cont.).

e) Using density forecasts of unemployment one year ahead.
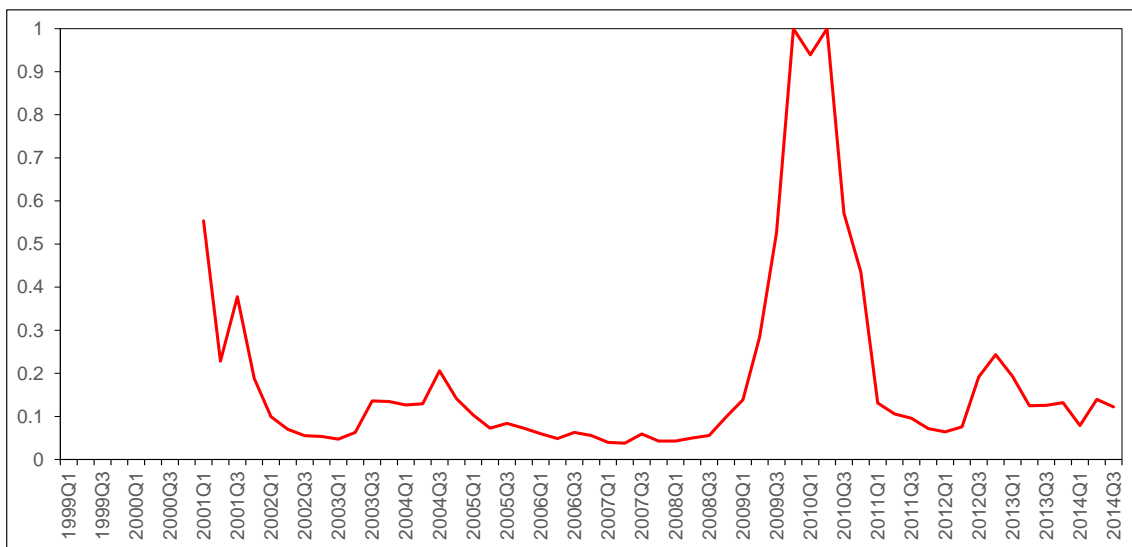
f) Using density forecasts of unemployment two years ahead.
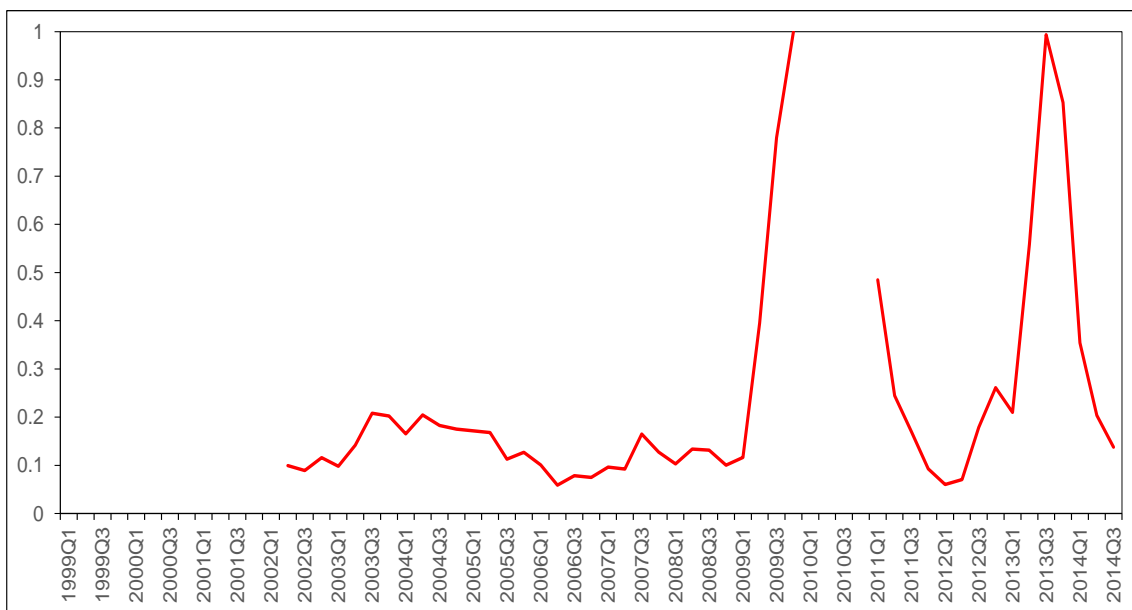
Fig. 8 Weighted averages of individual measures of uncertainty using the optimal weights.

a) From density forecasts of inflation one year ahead.
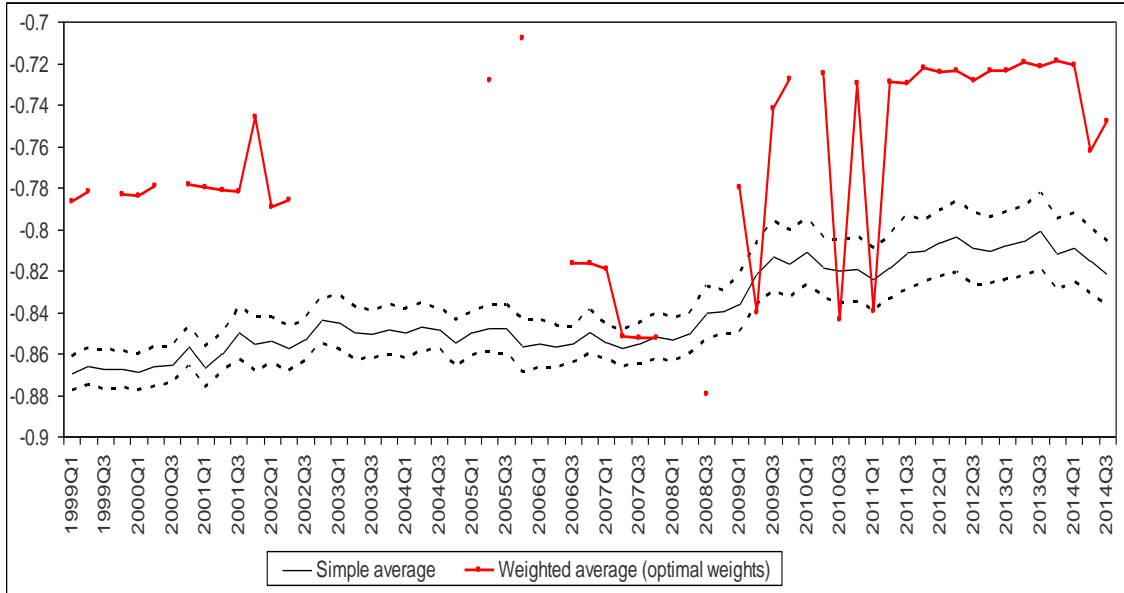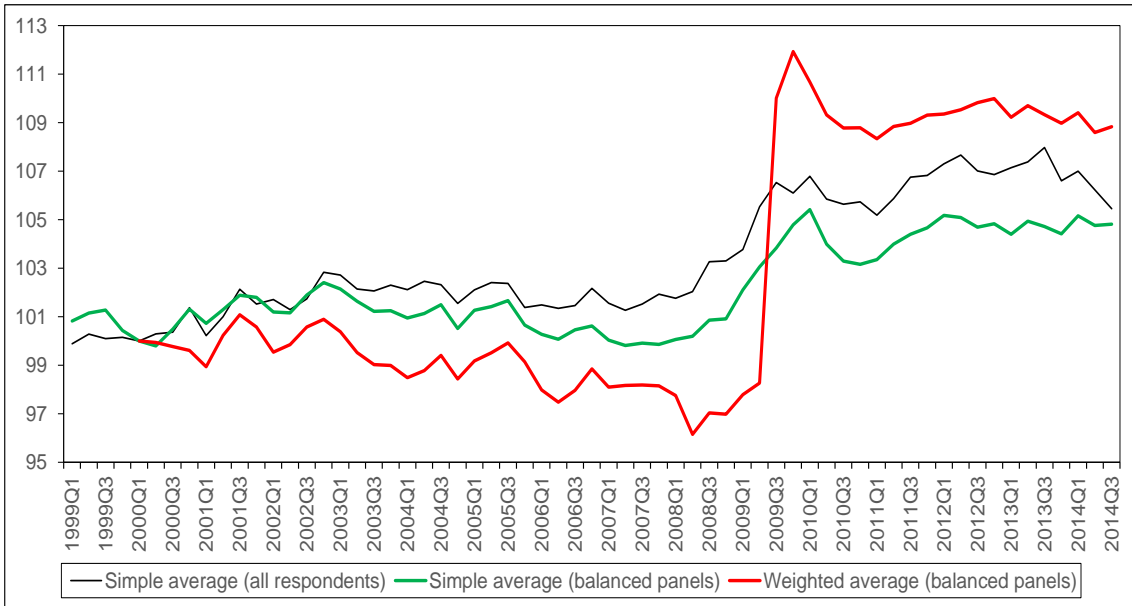

b) From density forecasts of inflation two years ahead.

Fig. 8 Weighted averages of individual measures of uncertainty using the optimal weights (cont.).

c) From density forecasts of GDP growth one year ahead.

d) From density forecasts of GDP growth two years ahead.

Fig. 8 Weighted averages of individual measures of uncertainty using the optimal weights (cont.).

e) From density forecasts of unemployment one year ahead.

f) From density forecasts of unemployment two years ahead.
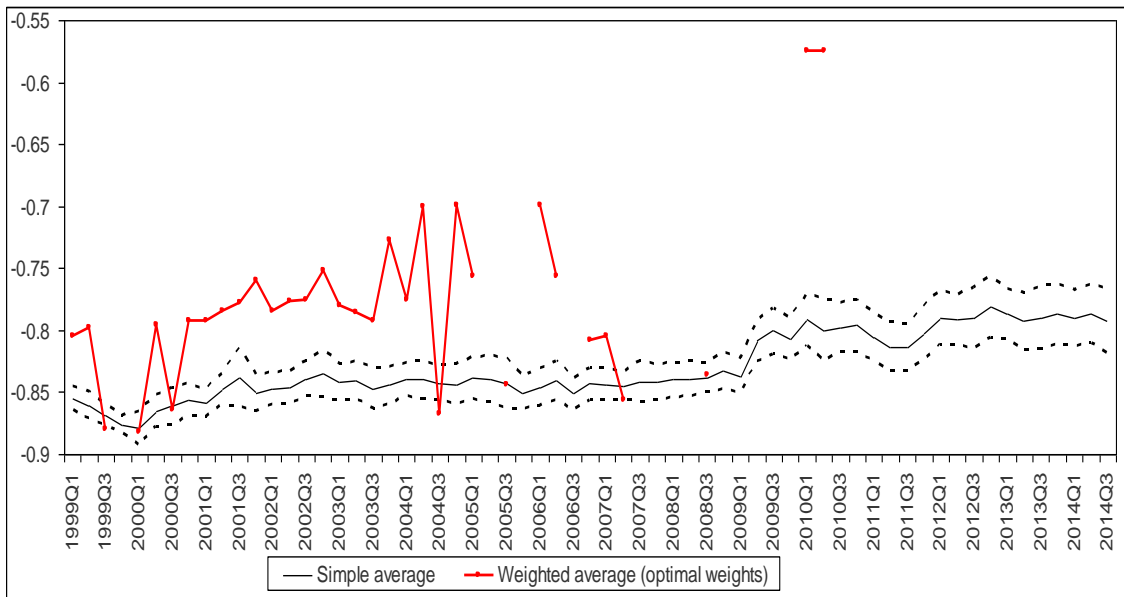
Fig. 9 Weighted averages of individual measures of uncertainty controlling for changes in participation.

a) From density forecasts of inflation one year ahead (2000 Q1 = 100).

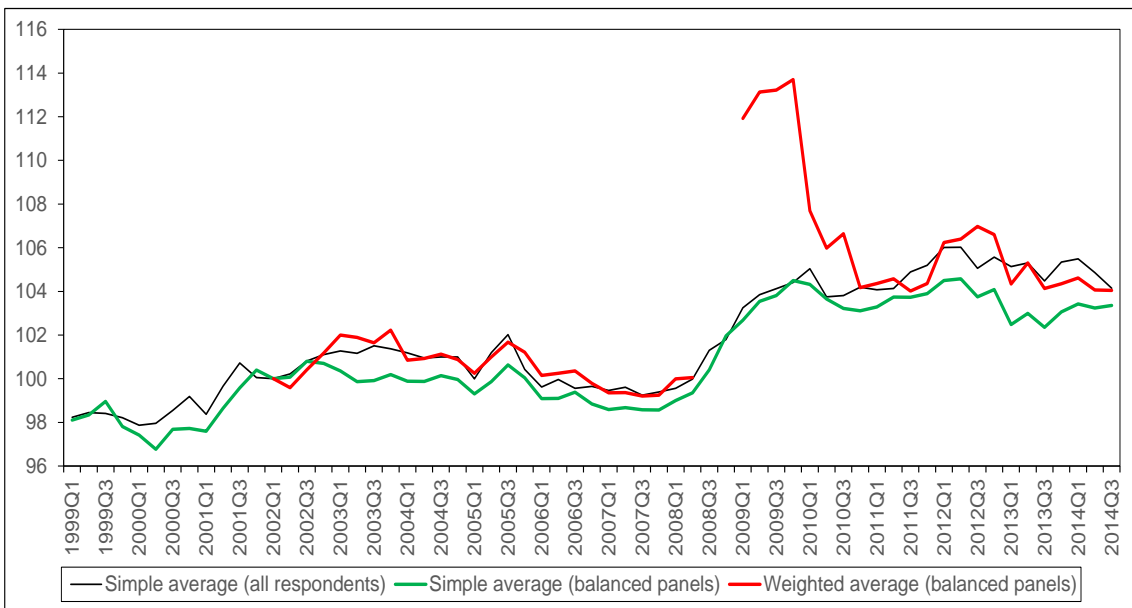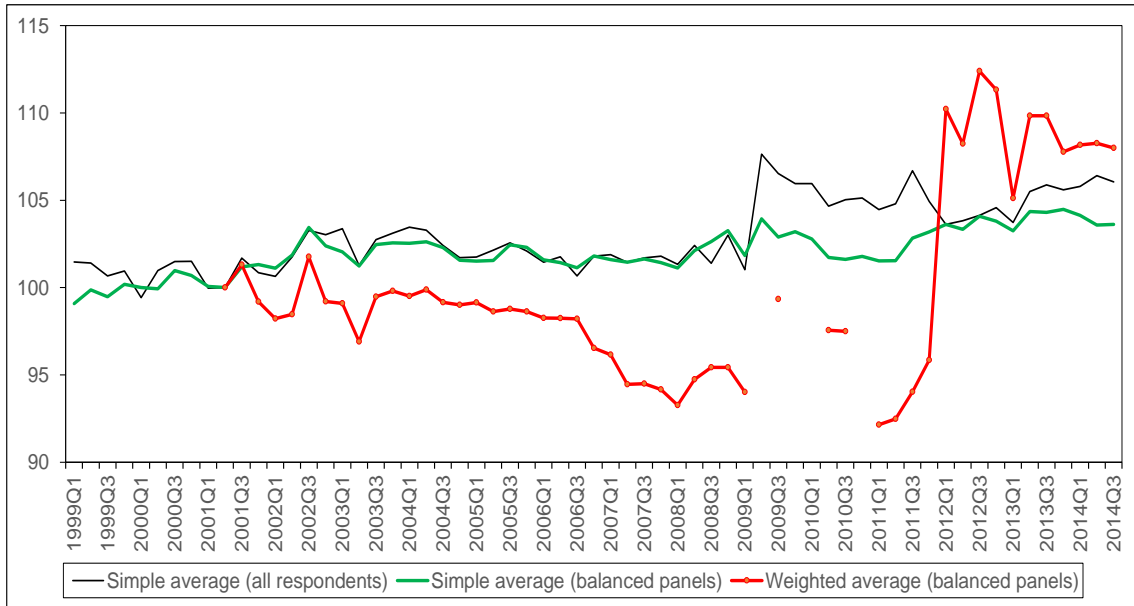b) From density forecasts of inflation two years ahead (2002 Q1 = 100).

Fig. 9 Weighted averages of individual measures of uncertainty controlling for changes in participation (cont.).

c) From density forecasts of GDP growth one year ahead (2001 Q2 = 100).

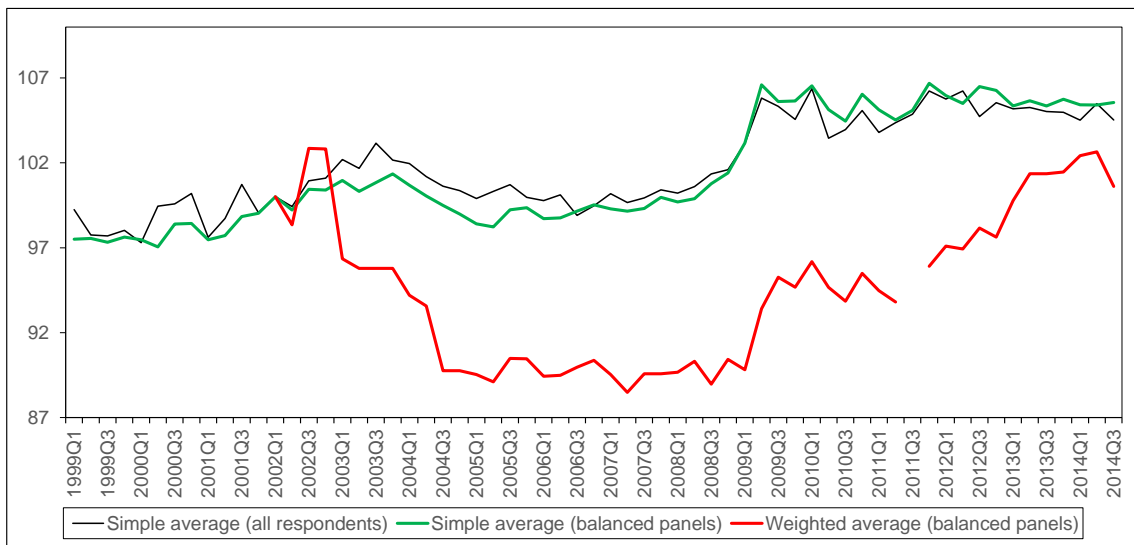d) From density forecasts of GDP growth two years ahead (2002 Q1 = 100).
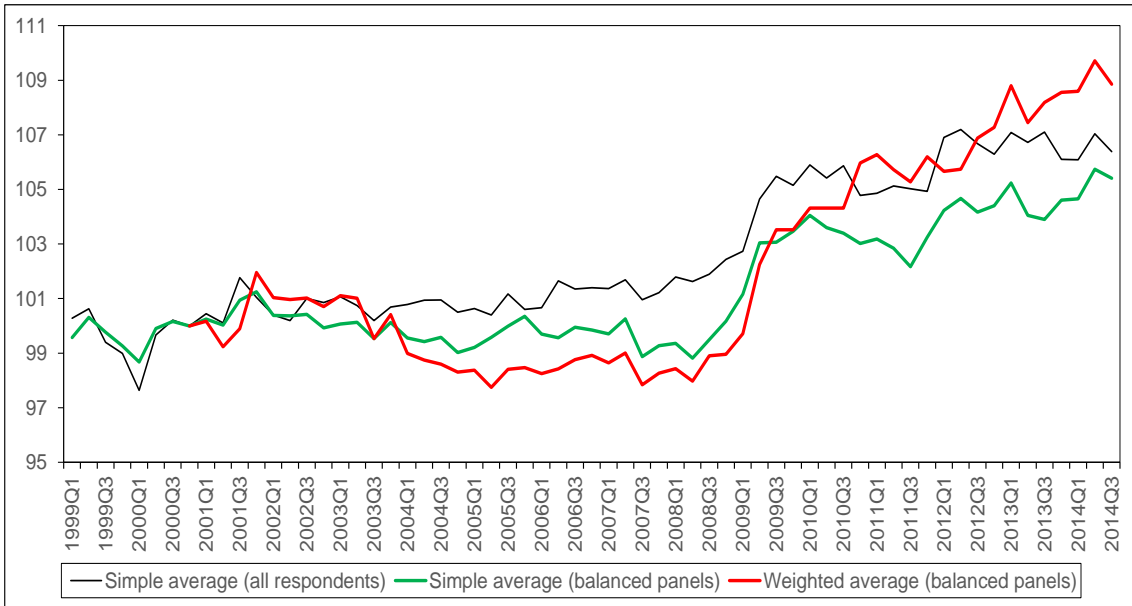
Fig. 9 Weighted averages of individual measures of uncertainty controlling for changes in participation (cont.).

e) From density forecasts of unemployment one year ahead (2000 Q4 = 100).

f) From density forecasts of unemployment two years ahead (2001 Q1 = 100).