NATIONAL BANK OF POLAND
WORKING PAPER
No. 142

A new approach to probabilistic surveys of professional forecasters and its application in the monetary policy context

Halina Kowalczyk, Tomasz Łyziak, Ewa Stanisławska

Warsaw 2013

Halina Kowalczyk – National Bank of Poland, Economic Institute, e-mail: Halina.Kowalczyk@nbp.pl.

Tomasz Łyziak – National Bank of Poland, Economic Institute, e-mail: Tomasz.Lyziak@nbp.pl

Ewa Stanisławska – National Bank of Poland, Economic Institute, e-mail: Ewa.Stanislawska@nbp.pl

Design:

Oliwka s.c.

Layout and print:

NBP Printshop

Published by:

National Bank of Poland
Education and Publishing Department
00-919 Warszawa, 11/21 Świętokrzyska Street
phone: +48 22 653 23 35, fax +48 22 653 13 21

ISSN 2084–624X

http://www.nbp.pl

Contents

Abstract 2
1. Introduction 3
2. Methodology of the NBP Survey of Professional Forecasters 6
2.1. Drawbacks of standard surveys of professional forecasters 6
2.2. The design of the NBP SPF 8
2.3. Elicitation of probability and form of presenting the results 10
2.4. Deriving the aggregated probability function 11
2.4.1. Aggregation with equal weights 12
2.4.2. Robustness check 14
2.4.3. Aggregation with performance-based weights 16
3. Interpretation of the results of the NBP SPF 20
3.1. Overview of data 20
3.2. Analysis of expectations - one-dimensional vs. two-dimensional approach 24
3.2.1. Development of expectations 24
3.2.2. Assessing central bank credibility 29
4. Conclusions 36
References 38
Annex: NBP SPF questionnaire (of the survey taken in II quarter 2012) 40

Abstract

In this paper we present the NBP Survey of Professional Forecasters introduced in 2011 by the National Bank of Poland. It is a new survey that allows analysis of macroeconomic forecasts of professional economists, including their probabilistic forecasts of CPI inflation, GDP growth and the NBP reference rate. In the paper we discuss in detail survey methodology, whose some elements are novel. It refers especially to the construction of probabilistic survey questions. Instead of declaring probabilities that in a certain horizon a given variable will be in pre-defined intervals, NBP SPF experts declare median and the limits of a 90-percent probability range between the 5^{th} and 95^{th} percentile of their subjective probability distributions. To present the benefits from the applied design of the NBP SPF, we describe the first results obtained from the NBP SPF.

JEL: C82, D84, E52

1. Introduction

There are different reasons why central banks are interested in macroeconomic forecasts of professional economists and conduct their own surveys among them (so-called Surveys of Professional Forecasters). First of all, professional economists are capable to make informed, forward-looking forecasts. Observing them provides central banks with a cross-check of their own macroeconomic projections. Interactions between central bank economists and outside forecasters can improve the understanding of macroeconomic prospects and ability to predict them by both groups of economists. Another benefit for central banks is that financial sector agents are capable to make long-term forecasts, with the horizon consistent with the lags in the monetary transmission mechanisms or even longer. Especially direct measures of long-term inflation expectations are needed for central banks since they are helpful in assessing central bank credibility. Finally, forecasts produced by professional economists can exert a strong influence on expectations of economic agents less specialized in macroeconomic forecasting (consumers, producers).

The results of empirical studies that exploit inflation forecasts of professional economists confirm their usefulness in forecasting inflation. E.g., Mehra (2002) shows that although US professional forecasters make biased inflation forecasts (i.e. inflation expectations are statistically different from actual inflation on average), they adequately process available information on inflation, output gap, money growth and oil prices. Trehan (2010) demonstrates that inflation forecasts from the US Survey of Professional Forecasters are more accurate than forecasts from statistical models (random walk forecast, UC-SV models) and forecasts based on lagged headline and core inflation. Mixed evidence concerns their superiority over forecasts based on the Phillips curve. It seems that professional forecasters rely too much on recent inflation figures while forming their inflation expectations, which leads to deterioration of forecasting accuracy of those measures. Scheufele (2011) shows that a model using economic experts' survey expectations in Germany outperforms most of the competing models, such as: AR, ARMA, random walk or Phillips curve models.

Another factor making direct measures of inflation expectations important for central banks is their role in actual price formation, as confirmed in empirical studies estimating different versions of the Phillips curve using those measures as proxies for inflation

expectations in the economy. E.g., Henzel and Wollmershaeuser (2006) present an overview of literature, providing estimates of the hybrid Phillips curve for the euro area, Germany and US. In those studies different measures of inflation expectations are used (based on consumer or professional economists' surveys). In all the cases under consideration direct measures of expectations, including those formed by professional forecasters, appear statistically significant.

From the theoretical point of view the Phillips curve should contain inflation expectations of price setters. Statistical significance of professional forecasters' expectations confirmed in the literature is usually justified with the fact that inflation forecasts of professional economists have impact on inflation expectations of other agents in the economy. E.g. Carroll (2003) using epidemiological model shows that inflation forecasts of professional economists in the US have significant impact on consumer inflation expectations. Döepke et al. (2008) present similar analysis for Germany, France, Italy and UK, while Łyziak (2012) – for Poland.

In 2011 the National Bank of Poland introduced its own Survey of Professional Forecasters (NBP SPF). This tool allows collecting macroeconomic forecasts of a broad group of professional economists – broader than in surveys existing in Poland before (e.g. monthly surveys by Reuters). Moreover, it facilitates interactions between forecasters from the central bank and from the outside. Conducting the survey by the central bank allows tailoring the list of variables, time horizons and survey questions to specific needs of monetary authorities, e.g. by lengthening of the forecast horizons, focus on probability forecasts instead of point forecasts etc.

Design of the NBP Survey of Professional Forecasters refers to examples of similar surveys carried out by other central banks with long traditions in this respect, especially by the Federal Reserve, the European Central Bank and the Bank of England. Many of the features of the NBP SPF, for example, the composition of the panel of professional forecasters, the set of variables and the time horizon of the forecasts, are consistent with the solutions adopted in those banks. However, the way, in which the NBP SPF experts are asked to reveal the degree of uncertainty faced in analysing macroeconomic prospects, is different. Moreover, the tools used to present survey results and interpret them from the point of view of expectations' analysis and macroeconomic forecasting are also different.

Methodological novelties introduced in the design of the NBP SPF result from the conclusions of some recent empirical studies that compare macroeconomic point and probabilistic forecasts in the surveys carried out by the Fed (Engelberg et al. 2009a), the Bank of England (Boero, Smith, Wallis 2008a; Boero, Smith, Wallis 2008b) and the European Central Bank (Bowles et al. 2007) as well as from studies analysing different aspects of assessment of subjective probability (e.g. Tversky and Kahneman 1974; Savage 1971; Hogarth 1975; Cooke 1991).

The paper is organized in the following way. Section 2 describes methodological foundations of the NBP SPF as well as details of its design. We attempt to show how the drawbacks of existing surveys of this kind, as indicated in the literature, are solved in the construction of the NBP SPF. They should make the questionnaire more user-friendly to potential participants, and – most importantly – increase reliability of the results. Section 3 is devoted to preliminary interpretations of the results of the NBP SPF. After an overview of the data, we analyse how macroeconomic forecasts developed and illustrate the usefulness of the NBP SPF in analysing central bank credibility. It should be stressed that due to the short period covered by the NBP SPF (only 4 surveys has been carried out so far) the analysis is rather descriptive and its results are tentative. The last section concludes the study.

2. Methodology of the NBP Survey of Professional Forecasters

2.1. Drawbacks of standard surveys of professional forecasters

Methodology of existing surveys of professional forecasters displays some weaknesses, both in the case of point and probabilistic surveys. As far as the former ones are concerned, empirical testing of a consistency between subjective probabilistic forecasts and point forecasts declared by experts in surveys shows that the interpretation of the point values provided by respondents is problematic (Engelberg et al. 2009a, Boero et al. 2008). Experts asked for point forecasts synthesize their subjective probability distributions in different ways (ECB 2009). Their point forecasts can thus be incomparable with each other and the calculation of the median or averaging the values in the group of forecasters may lead to invalid interpretations.

Also the form of probabilistic questions in surveys of professional forecasters seems debatable. The best-known macroeconomic surveys, i.e.: Surveys of Professional Forecasters conducted by the Federal Reserve and the European Central Bank and the Survey of External Forecasters carried out by the Bank of England, use pre-defined intervals in the probability survey questions. The survey respondents provide their estimates of probabilities that in a certain horizon a given variable will be between upper and lower limits of those intervals. Such design of the survey questions imposes some problems. Firstly, the number of intervals is usually large[1] and the intervals are relatively narrow, which makes the assessment of respective probabilities difficult (it is difficult to make comparisons and express probabilities), therefore survey responses can imperfectly reflect experts' opinions. Secondly, defining intervals by institutions conducting surveys can amplify the natural (identified by psychologists) tendency to 'anchor' respondents' opinions (e.g. in the central interval). Thirdly, with a pre-defined range of possible values of a variable, it is inevitable need to change that range from time to time.[2]

[1] E.g. the ECB SPF in 2009Q4 used as many as 14 intervals in the case of inflation forecasts.

[2] E.g. see the changes in the intervals of inflation forecasts of the ECB SPF between 2007Q4 and 2009Q4:
- 2007Q4: <0%, 0 to 0.4%, 0.5 to 0.9%, 1.0 to 1.4%, 15 to 1.9%, 2 to 2.4%, 2.5 to 2.9%, 3 to 3.4%, >3.5%;
- 2009Q4: <-2%, -2% to -1.6%, -1.5% to -1.1%, -1% to -0.6%, -0.5% to -0.1%, 0-0.4%, 0.5-0.9%, 1.0-1.4%, 15-1.9%, 2-2.4%, 2.5-2.9%, 3-3.4%, 3.5% to 3.9%, >4.0%.

Another problem, for a long time neglected in studies on expectations and uncertainty, is the need to distinguish between intrapersonal (internal, personal) uncertainty from interpersonal uncertainty, resulting from the differences of opinions (Zarnowitz and Lambros 1983). The high degree of compliance of individual forecasts does not necessarily reflect a low level of uncertainty faced by individual experts while formulating their opinions, and vice versa. Moreover, the standard way of presenting group forecast, i.e. aggregated histogram, leads to mixing these different types of uncertainty and makes it impossible to identify them (for discussion see e.g. Boero et al. 2008a, Giordani and Soderlind 2003).

To avoid the problems outlined above, in the NBP Survey of Professional Forecasters each expert is asked to consider various scenarios of macroeconomic developments and assess a possible range of values of a given variable, indicating the limits of a 90-percent probability range laying between the 5^{th} and 95^{th} percentile of her/his subjective (personal) probability distribution as well as the median of this distribution. It should be underlined that meaning of this distribution is strictly defined as the reflection of experts' beliefs on different macroeconomic scenarios they are asked to consider. It is neither an estimate of the objective probability distribution of a given variable (we treat macroeconomic variables as unknowns, not stochastic), nor a distribution describing experts' past forecast errors or the forecast errors of the models they use. As pointed out by Kowalczyk (2010), asking experts about the probability distribution of inflation would be justified if we could assume that future inflation (in a given horizon), is a random phenomenon, which is subject to a certain unknown law of probability, that expert is able to identify. Focusing probabilistic questions of the survey on the distribution resulting from past forecast errors would depend on experts' self-assessment of their forecasts' errors. So it seems adequate to accept the interpretation, according to which experts should determine their uncertainty about various scenarios concerning macroeconomic developments in the future.[3]

[3] Such an approach seems also to be consistent with the intentions of the creators of the ASA-NBER survey (Zarnowitz and Lambros 1983), which became a benchmark for all the surveys of professional forecasters.

2.2. The design of the NBP SPF

The NBP Survey of Professional Forecasters is conducted on quarterly basis among professional economists nominated by different institutions, mainly by commercial banks and financial institutions, but also by research institutes preparing macroeconomic forecasts as well as labour unions and employer organizations.

The main questions of the survey concern forecasts of CPI inflation and GDP growth for different time horizons (+4 quarters, +8 quarters, average in the current year, annual average in the next 2 years, average during the nearest 5 years). Experts participating in the survey are asked to consider various scenarios of economic developments and to provide – on the basis of the conducted analysis – the range of possible values and a central point forecast for those variables. Additional survey questions concern forecasts of the NBP reference rate, exchange rate, unemployment rate, average wage growth, oil prices and GDP growth in the euro area. Except for the first of those variables, additional questions concern point forecasts. Table 1 describes the main features of the NBP Survey of Professional Forecasters as compared with similar surveys conducted by the Bank of England, ECB and Fed. Annex 1 presents the questionnaire of the NBP SPF.

Table 1. Features of surveys of professional forecasters conducted by selected central banks and the NBP

Survey 1) central bank 2) year of introducing probability questions 3) frequency	**Panel of experts**	**Variables forecasted**	**Forecast horizon**
Survey of External Forecasters 1) Bank of England 2) 1996 3) quarterly	• approx. 30 experts (approx. 20 in a regular manner) • experts of financial institutions from the London City, research institutes, private consultants	Probability forecasts (2): • CPI (till 2004 – RPIX) • GDP (since 1998) Point forecasts (2): • official interest rate • exchange rate	Quarterly indices Q/Q_{-4}: +1Y, +2Y, +3Y
Quarterly Survey of Professional Forecasters 1) European Central Bank 2) 1999 3) quarterly	• approx. 70 experts • experts experienced in macroeconomic forecasting nominated by national central banks, represent their institutions (financial sector, research institutes, employers' associations and trade unions/labour institutes)	Probability forecasts (3): • HICP • GDP • unemployment rate Point forecasts (4): • ECB's interest rate (MRO) • oil price • exchange rate • ULC	Probability forecasts: annual averages in the current year, +1Y, +2Y,+5Y HICP and unemployment rate: monthly indices M/M_{-12}: +1Y,+2Y GDP: quarterly indices Q/Q_{-4}: +1Y, +2Y
Survey of Professional Forecasters 1) Federal Reserve Bank of Philadelphia 2) 1990 (earlier: ASA/NBER) 3) quarterly	• more than 50 experts (not all provide forecasts for all variables; average number of forecasts: 30-40) • the panel is dominated by representatives of financial institutions, there are some experts from research institutes and consulting firms	Probability forecasts (5): • GDP • GDP deflator • CPI core • PCE core • probability of recession Point forecasts (more than 20 variables)	GDP: current year, +1Y Inflation: quarterly indices $Q4/Q4_{-4}$: current year, +1Y; annual averages (point forecasts): next 5 and 10 years Probability of recession: next 5 quarters
NBP Survey of Professional Forecasters 1) National Bank of Poland 2) 2011 3) quarterly	• approx. 30 experts (approx. 20 in a regular manner) • professional economists nominated by commercial banks and financial institutions; research institutes, labour unions and employer organizations	Probability forecasts (3): • GDP • CPI inflation • NBP refinancing rate Point forecasts (5): • exchange rate • unemployment rate • average wage growth • oil prices • GDP growth in the euro area	Probability forecasts: +4Q, +8Q, current year, +1Y, +2Y, annual averages next 5 years Point forecasts: current year, +1Y, +2Y

Source: websites of the central banks under consideration.

2.3. Elicitation of probability and form of presenting the results

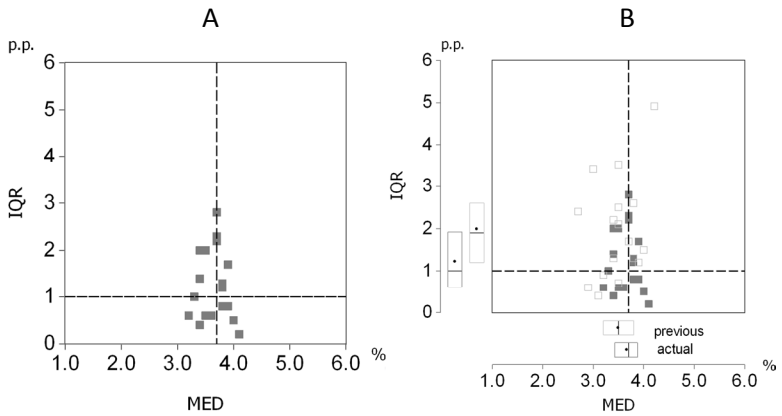
As outlined above, experts of the NBP SPF are asked to consider possible macroeconomic scenarios and state 5th, 50th and 95th percentile of their subjective distributions[4] for the CPI inflation, GDP growth and the NBP reference rate. Such a construction of probabilistic survey questions, provides not only unambiguous measure of central tendency of a forecasted variable (median), but also direct measure of individual uncertainty attached to the forecast, calculated as an interquantile range, i.e. difference between 95th and 5th percentiles of the subjective distribution. It should be stressed that using other probabilistic surveys, like Philadelphia Fed SPF, requires estimation of these values from individual histograms provided by forecasters (see e.g. Engelber et al. 2009b).

Results of the NBP surveys are presented in the form of the scatter graphs, in which each individual forecast is considered separately, as well as in the form of aggregated distributions, formed as an equal-weight-mixture of individual distributions estimated on the basis of percentiles provided by each forecaster (see section 2.4).

Engelberg et al. (2009b) stress the need of analysing changes in individual forecasts instead of aggregated measures, as averaging individual forecasts blurs information about disagreement among survey participants and level of uncertainty. Moreover, due to the existence of individual characteristics of forecasters (e.g. optimists and pessimists, overconfident and cautious) combined with unknown pattern of changes in the panel of survey participants, aggregated measures of their forecasts might not reflect true changes in expectations. As a simple way to summarize the cross-sectional distribution of forecaster's beliefs, Engelberg et al. (2009b) recommend scatter-graphs, in which each point corresponds to an individual probabilistic forecast (see Figure 1A). The median (MED) of forecaster's subjective distribution measures the central tendency of his/her beliefs. The interquantile range (IQR) measures the uncertainty that forecaster perceives. Dispersion of the points across the horizontal axis indicates about disagreement in the central forecasts. The shift up/down of the points informs about increase/decrease of experts' uncertainty while making predictions (Figure 1B).

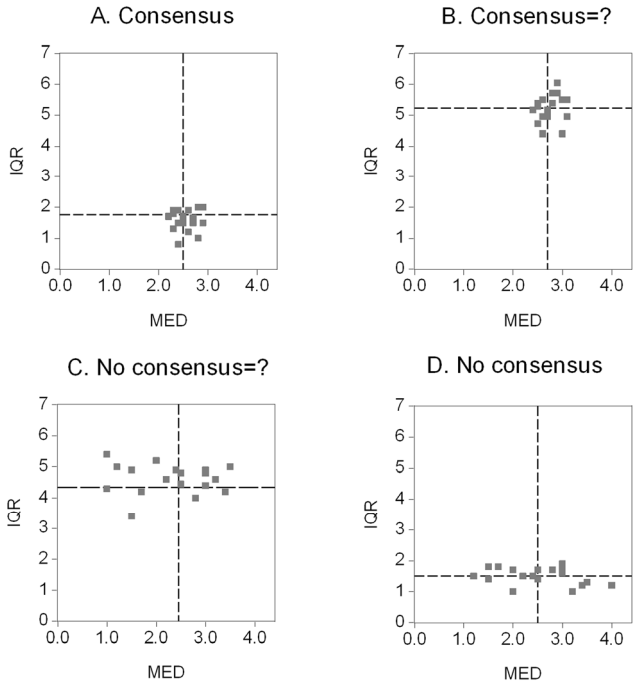
[4] By subjective probability we mean probability which reflects personal beliefs about specific outcomes. It refers to uncertainty due to imperfection of knowledge and is an attribute of a person, not phenomenon he/she describes. For discussion see e.g. Kowalczyk (2010). As a consequence, expert subjective distribution should be assessed as good, if properly reflects his/her beliefs.

Figure 1. Scatter graphs



Scatter graphs are more informative about consensus and disagreement than analyses of point predictions. The latter are difficult to interpret if nothing is known about uncertainty the forecasters face. In line with the interpretation proposed by Kowalczyk (2010), scatter graphs show a consensus, when a low disagreement between forecasters in terms of their median forecasts is accompanied by a low degree of uncertainty (see Figure 2).

Figure 2. Scatter graphs and analysis of consensus between forecasters



2.4. Deriving the aggregated probability function

As mentioned above, aggregated distributions should not be regarded as consensual forecasts and therefore they should rather play a complimentary role in analysing

expectations. However, their usefulness in forecasting, i.e. obtaining objective probability of some uncertain events, is beyond doubt. In our opinion, employing survey data in these two distinct areas, i.e. in analysing expectations and macroeconomic forecasting, requires different treatments of the data. Analysis of expectations requires experts' opinions to be reflected in the most accurate way, i.e. they should not be subject to many additional transformations. In the case of macroeconomic forecasting, survey outcomes can be processed with the aim of achieving the best forecasts (e.g. it seems reasonable to correct possible biases and differences in quantifying probability by experts – see section 2.4.2).

2.4.1. Aggregation with equal weights

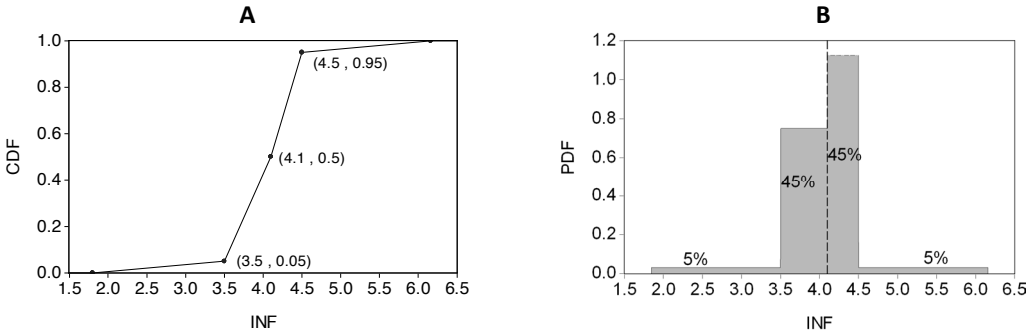
Deriving aggregated probability distributions in the case of the NBP Survey of Professional Forecasters is slightly more difficult than averaging histograms, since it requires fitting probability densities to each expert's assessment. The method proposed by Cooke (1991) is applied and the distribution with the minimum information (maximum entropy) is fitted.

The values of 5th, 50th and 95th percentiles provided by the expert for a variable X divide a range of possible values into four intervals with probabilities 0.05, 0.45, 0.45, 0.05, i.e.:

$$P(X \le x_{05}) = 0.05 \qquad P(x_{05} < X \le x_{50}) = 0.45 \qquad P(x_{50} < X \le x_{95}) = 0.45 \qquad P(X > x_{95}) = 0.05$$

The distribution with minimum information, which satisfies the expert's quantiles is uniform between these quantiles. Figure 3A presents an example of the cumulative distribution and Figure 3B – the density function for the expert, who stated the following percentiles in declaring forecasts of the CPI inflation ($X = INF$): x_{05} =3.5, x_{50} =4.1 and x_{95} =4.5. The half-bounded interquantile intervals $(-\infty, x_{05}]$ and (x_{95}, ∞) are replaced by the bounded ones: $[a, x_{05}]$ and $(x_{95}, b]$ and the piece-wise uniform density on $[a, b]$ is constructed. The boundaries *a* and *b* are determined based on the minimum and maximum values declared by all the experts (see section 2.4.3) for the forecasted variable

Figure 3. Example of an interpolated cumulative distribution (A) and the corresponding probability density function (B)



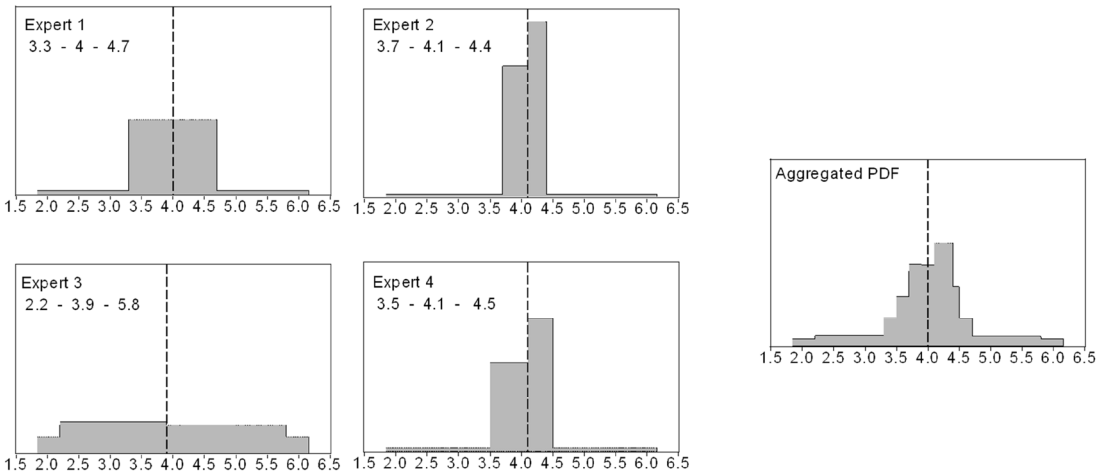
Source: calculations on the basis of randomly selected expert participating in the NBP SPF in 2011Q3.

In the second step individual probability distributions are aggregated (Figure 4). Formally, if $g_i(x)$ denotes the probability density function for the forecasted variable (CPI, GDP growth or the NBP reference rate) provided by expert e_i, then the aggregated distribution that results from combining forecasts of *N* experts, takes the following form:

$$g_A(x) = \frac{1}{N}\sum_{i=1}^{N} g_i(x)$$

This simple arithmetic average of expert's probability distributions has many useful properties (see: Clemen and Winkler, 1999) and describes the following hierarchical stochastic model of making decisions: in the first step an expert is randomly selected with probability $\frac{1}{N}$, while in the second step, the value of the variable of interest is randomly drawn according this expert's distribution.

Figure 4. Example of aggregation of individual distributions (4 randomly selected experts participating in NBP SPF)



2.4.2. Robustness check

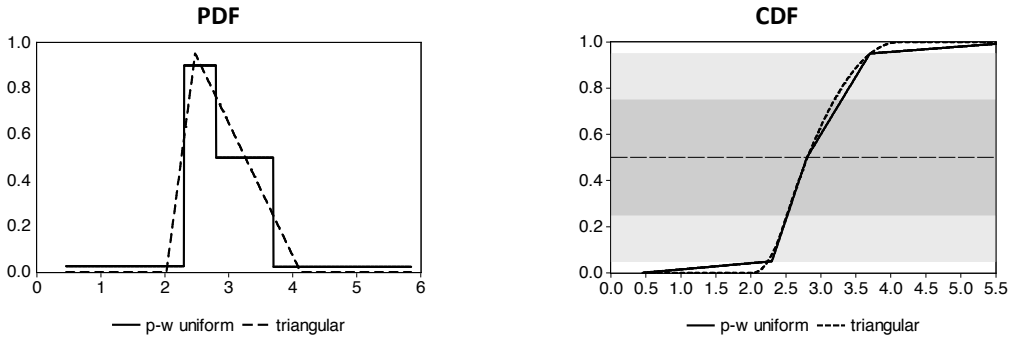
Deriving aggregate probabilistic forecast we assume that the individual distributions are piece-wise uniform ones. It is debatable whether such type of the distribution is the most appropriate if the true individual subjective distribution has one distinct mode. NBP SPF does not provide information whether the expert's distribution is one- or multimodal. It can be treated as its disadvantage, since it provides less information about modes compared to histograms used in traditional SPFs. Therefore we analyse how the median and interquantile range of the aggregated distribution would change if instead of the piece-wise uniform, triangular or beta (approximated by triangular) distributions were applied. Those types of the distributions are commonly used in risk analyses and could be more appropriate in the case of unimodal distributions.

Assessing the impact of the assumed individual subjective distribution on the aggregated one, we refer to the results of the survey conducted in the 1st quarter 2012. We apply triangular approximation proposed by Johnson (2002) and estimate parameters of the triangular distribution: its lower limit *a*, upper limit *b* and the mode *m* by linear combination of the 5th, 50th and 95th percentiles:

$$\hat{a} = (23x_{05} - 6x_{50} - x_{95})/16$$
$$\hat{b} = (-x_{05} - 6x_{50} + 23x_{95})/16$$
$$\hat{m} = (-13x_{05} + 42x_{50} - 13x_{95})/16$$

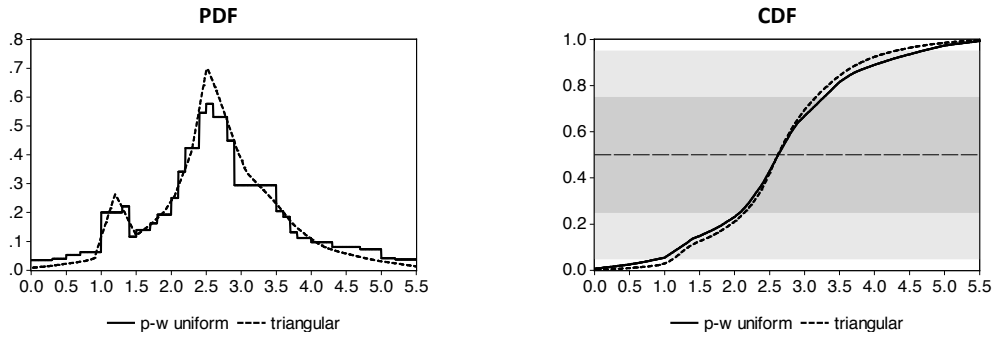
Johnson (2002) shows that above expressions provide a good fit to the wide range of distribution functions (beta, gamma, lognormal, Golenko-Ginsburg), while used for approximation based on median and the 5th and 95th percentiles. Figure 5 shows the triangular distribution obtained in this way, together with the corresponding piece-wise uniform.

Figure 5. The triangular versus piece-wise uniform density based on 5th, 50th and 95th percentiles, and the corresponding cumulative distributions (x_{05}=2.3, x_{50}=2.7, x_{95}=3.7).



We apply triangular approximation for each individual inflation forecasts presented by experts in the 1st quarter 2012 and aggregate these densities by pooling. Figure 6 illustrates differences between resulting aggregated distributions (triangular versus piece-wise uniform) for the horizon of 8 quarters.

Figure 6. Aggregated distributions obtained for different type of individual pdfs (piece-wise uniform and triangular)



The choice of type of individual distribution had no noticeable effect neither on median nor on mean. The 50% probability interval and the probability of future inflation being within the range of permitted deviations from the NBP inflation target, i.e. [1.5%, 3.5%], are influenced by alternative assumptions concerning the distribution type to a little extent. Greater differences are observed in the case of 90% probability intervals – for

triangular distribution they are narrower. The same conclusion applies in the case of aggregated results for other forecast horizons (Table 2).

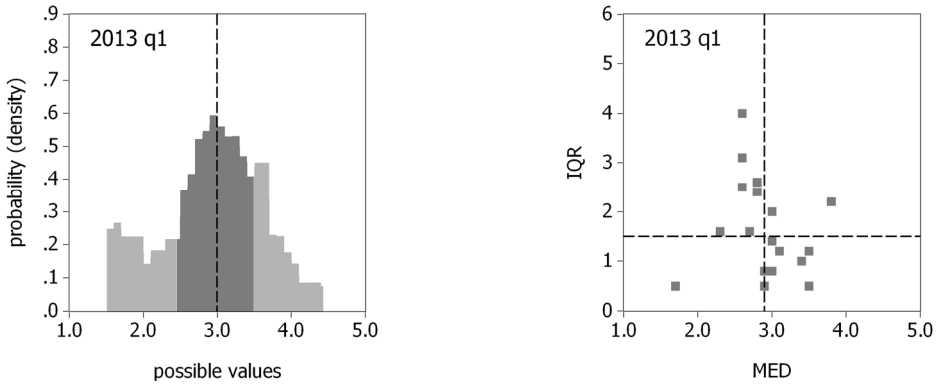
Table 2. Medians and probability intervals of aggregated distribution obtained under different assumption for the shape of the individual subjective pdf's

Horizon	Assumed individual pdf	Result of aggregation				
		median	mean	50% probability	90% probability	P([1.5 , 3.5])
2013Q1	p-w uniform	3.0	2.9	2.5 - 3.5	1.5 - 4.4	0.70
	triangular	3.0	2.9	2.5 - 3.4	1.6 - 4.1	0.75
2014Q1	p-w uniform	2.6	2.7	2.1- 3.3	0.9 - 4.7	0.67
	triangular	2.6	2.6	2.1 - 3.2	1.1 - 4.3	0.72
2012	p-w uniform	3.7	3.7	3.3 - 4.1	2.7 - 4.7	0.37
	triangular	3.7	3.7	3.4 - 4.0	2.8 - 4.5	0.37
2013	p-w uniform	2.9	2.9	2.4 - 3.5	1.4 – 4.5	0.70
	triangular	2.9	2.9	2.4 - 3.4	1.6 - 4.1	0.77
2014	p-w uniform	2.6	2.7	2.0 – 3.3	1.0 – 4.9	0.69
	triangular	2.6	2.7	2.1 - 3.1	1.3 - 4.5	0.75
2012-2016	p-w uniform	3.0	3.0	2.5 - 3.5	1.5 - 4.8	0.70
	triangular	3.0	3.0	2.6 – 3.4	1.9 – 4.4	0.76

2.4.3. Aggregation with performance-based weights

It remains an open question, whether equal "weighting" is optimal in the case of heterogeneity of experts. Experts have to quantify their uncertainty, and might to do that in different ways. Some of the experts might be overconfident, which would lead to underestimation of macroeconomic uncertainty, while the others, excessively cautious, can overestimate uncertainty. Significantly different variances of individual distributions might also lead to the interpretation problems. For instance, multimodality of the aggregated distribution shown on Figure 7 results not from the fact that there are two groups of experts with different views, as might be expected, but it is a consequence of a small IQR accompanying the lowest central forecast.

Figure 7. Example of an influence of a single individual forecast on the aggregated distribution



Source: NBP SPF conducted in 1Q2012.

As in traditional SPFs, the form of the aggregated probability distribution obtained by equal weighting scheme is influenced both by intrapersonal and interpersonal uncertainty, but contrary to the former ones we are able to assess the impact of these components by analysing the scatter graphs. After the NBP SPF covers longer period (so far only 4 quarterly surveys has been conducted), we plan to apply the Cooke's "classical model" (1991) for combining individual probability distributions. It will solve, at least to some extent, problems resulting from heterogeneity of experts. The Cooke's "classical model", which we describe below in a simplified manner, is widely used in engineering and natural sciences for combining expert judgements.

In the Cooke's method, the aggregated distribution is a mixture of individual distributions with weights, which are a product of two indicators: *calibration score* and *information score*. Those scores are calculated using experts' responses to questions about a set of so-called *seed variables* (unknown for experts and known for persons conducting elicitation). In the case of the NBP SPF the historical forecasts with known outcomes can play a role of seed variables.

The calibration score $Cal(e_i)$ is a statistical measure of the average compatibility of the expert's predictions with realizations. In the case of a forecast characterized by three quantiles: x_{05}, x_{50}, x_{95}, the expert is perfectly calibrated (gets the maximum calibration score) if 5% of realizations of the forecasted variables used for calibration process, fall into the first inter quantile interval, 45% into the second, 45% into the third and 5% to the fourth. This ideal situation is described by the theoretical distribution $f_p = \{p_1, p_2, p_3, p_4\}$, where $p_1 = 0.5$, $p_2 = 0.45$, $p_3 = 0.45$, $p_4 = 0.5$. It's common for all the experts.

If we denote the real empirical distribution characterizing past performance of expert e_i by $f_{ri} = \{r_{i1}, r_{i2}, r_{i3}, r_{i4}\}$, the divergence between the distributions f_{ri} and f_p could be measured by the Kullback-Leibler distance:

$$I(f_{ri}, f_p) = \sum_{j=1}^{4} r_{ij} \ln(\frac{r_{ij}}{p_j}) \ .$$

The lower is the distance, the higher is the calibration score.

The calibration score is a p-value of a statistical test of the hypothesis that the expert is calibrated. To test this hypothesis the fact is used that, if the realizations are drawn from expert's distributions corresponding to his quantiles given for K seed variables, the statistic:

$$2KI(f_{ri}, f_p) = 2K\sum_{j=1}^{4} r_{ij} \ln(\frac{r_{ij}}{p_j})$$

has asymptotic χ^2 distribution with 3 degrees of freedom.

The information contained in a distribution depends on degree to which the distribution is concentrated. In terms of the NBP SPF: the narrower the range of possible values of the forecasted variable declared by an expert is, the more informative is her/his interval prediction. Calculating the information score $Inf(e_i)$ requires assigning a density g_k to each quantile assessment of an unknown variable X_k, and defining the background density u_k for comparison. It is assumed that g_k is the piece-wise uniform (see: section 2.4.1), while u_k is the uniform distribution concentrated on the so-called *intrinsic range*, which is an interval containing all the assessed quantiles of all experts and the outcome (the smallest interval of this property for each X_k is determined and next its boundaries are extended by l% according to the so-called l% *overshoot rule*). Information of g_k is measured by the relative entropy , i.e. the Kullback-Leibler distance $I(g_k, u_k)$. Information score is obtained by averaging over a set of K variables used for calibration:

$$Inf(e_i) = \frac{1}{K}\sum_{k=1}^{K} I(g_k, u_k) \ .$$

The combined score, i.e. the weight of expert e_i, is calculated as:

$$w(e_i) = Inf(e_i) * Cal(e_i) * 1_{\{Cal(e_i) \geq \alpha)}$$

where α denotes the significance level for calibration.

This score rule is asymptotically *strictly proper*, i.e. the expert maximize his long-run expected scores only by stating percentiles according to his own beliefs.

3. Interpretation of the results of the NBP SPF

3.1. Overview of data

The first survey was conducted in September 2011, which gives 4 rounds till now, however the initial survey is treated as experimental and not analysed (section 3.1 and 3.2.1) or analysed with caution (section 3.2.2). In each survey round participated 18-19 experts, out of 25 registered. As far as turnover in panel members is considered, the number of dropped experts from survey to survey and those who re-entered, varied from 2 to 4.

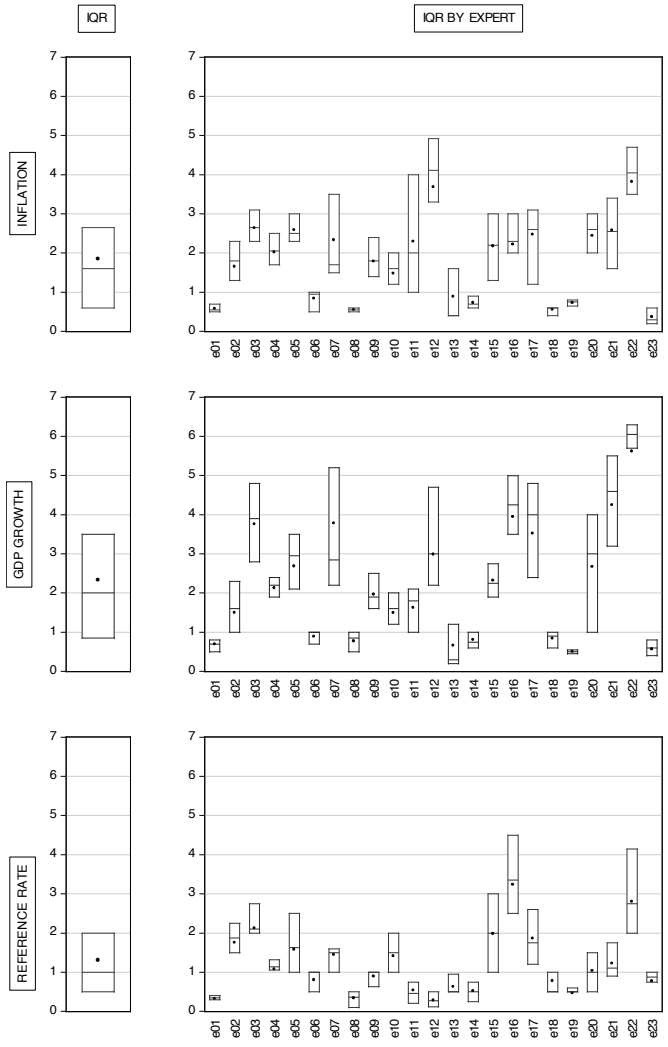
Before moving to economic analysis of professional forecasters' expectations, it's worth to look at some statistical characteristics and patterns observed in the collected data. We focus on three questions: how wide ranges of possible values survey participants declared? are the experts heterogeneous, i.e. whether they have permanent tendency to attach high/low uncertainty to the forecasts[5]? and are the individual distributions of variables symmetric? If the intervals determined by 5^{th} and 95^{th} percentiles of experts' subjective distributions were very extensive, they would have low informative value. As discussed in the previous section, heterogeneity of experts is important, from the point of view of analysing development of expectations as well as of forecasting. Finally, asymmetry of subjective distributions not only justifies precise defining central forecast in the survey question (due to the fact that in asymmetric distributions median, mean and mode are not equal to each other), but also communicates how the experts asses possible risks to the economy.

Boxplots in Figure 5 show distribution of interquantile ranges, based on percentiles declared by experts, for each forecasted variable (for all forecast horizons in surveys from 2011Q4 to 2012Q2). As seen on the left panel, presenting IQRs for all experts together, the narrowest ranges of possible realizations were stated for the NBP reference rate – median IQR amounted to 1 p.p., while the widest for GDP growth with median IQR of 2 p.p. Moreover, the IQRs for interest rate were the least, and for GDP growth the most diversified: the 50% middle values of IQRs, confine between 0.6 and 2.6 p.p. for inflation, 0.9 and 3.5 p.p. for GDP growth, 0.5 and 2.5 p.p. for interest rate.

[5] As we have hardly any realizations of the forecasted variables, we cannot assess experts' heterogeneity with respect to outcomes.

Turning to the distribution of IQRs by single expert (right panel), there is a pronounced heterogeneity of forecasters with this respect. For example, expert e22 in all surveys declared very high level of uncertainty for all variables, and expert e01 – very low uncertainty.[6] Significant heterogeneity of forecasters, both point and probabilistic, has been documented in other surveys, like Survey of External forecasters of Bank of England, NBER SPF or ECB SPF (see: Boreo et al. 2008a, 2008b; Bowles et al. 2007; Lahiri and Liu 2006).

Figure 5. Distribution of inter quantile ranges (IQR) for various variables in surveys from 2011Q4-2012Q2



We measure skewness of individual subjective distributions in terms of quantiles, using the fact that in case of symmetry the 95th and 5th percentiles are equally distant from the

[6] The caution is that some experts participated only in one or two surveys (see Table 3). However, this observation is also true if we consider only these forecasters who took part in all survey rounds.

median. The formula is the following: $A = \frac{(Q_{95}-Q_{50})-(Q_{50}-Q_{05})}{Q_{95}-Q_{05}}$, where Q denotes proper percentile declared by a forecaster.

As presented in Table 3, more than half of individual subjective distributions are asymmetric, with the greatest share, equal to 59%, for GDP growth. Also in this case there is a vast heterogeneity among forecasters: some of them almost always indicated symmetric intervals (e.g. e010, e018), others almost always asymmetric (e.g. e04, e21). Indication of symmetric intervals surrounding central forecast by some of respondents can reflect scenarios of macroeconomic developments considered by experts, but it can also result from assessing uncertainty on the basis of past forecast errors that would be incompatible with the concept of uncertainty adopted in the survey design.

Distributions of non-zero asymmetry coefficients for inflation, GDP growth and reference rate are presented in Figure 6. The asymmetry of individual subjective distribution of future inflation and GDP growth is not strong – the quantile coefficient of skewness rarely exceeds +/-0.4. When it comes to direction of skewness, inflation forecasts are rather positively skewed, while of GDP growth – negatively.[7] This would mean that forecasters expect greater risks of higher inflation and of lower GDP growth. On the contrary, for reference rate the number of positively and negatively skewed distributions is about equal. Also for this variable, there are more strongly skewed individual subjective distributions, with some cases with central forecast being equal to the 95th or 5th percentile.[8]

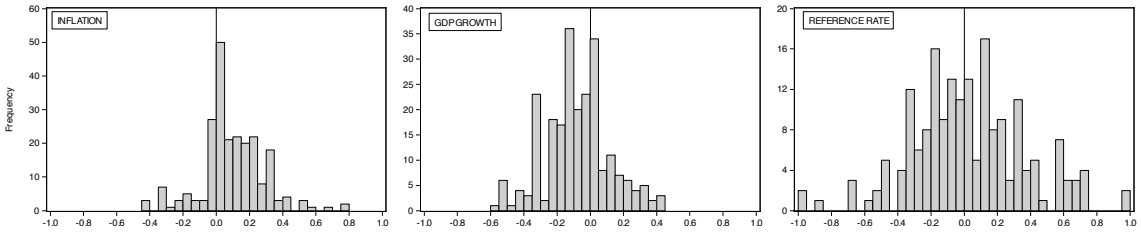
[7] Bowles et al. (2007) observed also positive skewness of inflation and negative skewness of GDP growth forecasts, but in the aggregated distributions.

[8] Such a situation happened in the two last surveys in the case of interest rate forecasts. Due to the fact that such a distribution do not exist, we interpret this result as a mistake of an expert and exclude from the sample in the further analysis.

Table 3. Share of asymmetric forecasts in all forecasts in total and for particular forecasters, surveys from 2011Q4 to 2012Q2

Forecaster	Inflation		GDP growth		Reference rate	
	share	n	share	n	share	n
e01	0.61	18	0.94	18	0.61	18
e02	0.50	6	0.67	6	0.33	6
e03	0.89	18	0.89	18	0.72	18
e04	0.94	18	0.72	18	0.94	18
e05	0.94	18	0.94	18	0.67	18
e06	0.67	18	0.33	18	0.17	18
e07	0.78	18	0.72	18	0.83	18
e08	0.67	6	0.67	6	0.67	6
e09	0.00	12	0.08	12	0.00	12
e10	0.00	18	0.06	18	0.22	18
e11	0.22	18	0.56	18	0.61	18
e12	0.50	6	0.33	6	0.17	6
e13	0.50	18	0.50	18	0.33	12
e14	0.56	18	0.83	18	0.67	18
e15	0.42	12	0.83	12	0.82	11
e16	0.67	18	0.94	18	0.72	18
e17	0.33	18	0.33	18	0.50	18
e18	0.00	18	0.06	18	0.00	18
e19	0.50	4	0.75	4	0.33	6
e20	0.00	6	0.00	6	0.00	6
e21	0.78	18	0.89	18	0.78	18
e22	0.56	18	0.78	18	0.83	18
e23	0.00	6	0.00	6	0.50	6
All	**0.52**	328	**0.59**	328	**0.54**	323

Figure 6. Histograms of non-zero asymmetry coefficients for various variables, surveys from 2011Q4 to 2012Q2



3.2. Analysis of expectations - one-dimensional vs. two-dimensional approach

In this section, analysing results from the first three NBP surveys, we aim to present how the design of survey, by including information about two dimensions of expectations, enhances possibilities of interpretation of expectations. In the first place, we describe development of forecasts of inflation, GDP growth and reference interest rate during last year, with a special interest in uncertainty and disagreement measures. Secondly, we show how this data might be used to asses central bank credibility.

3.2.1. Development of expectations

First row of Figure 7 presents predictions of inflation, GDP growth and NBP's reference rate in 2012 made by professional forecasters in December 2011. At that time Poland experienced period of high inflation, staying since the beginning of the year significantly above the upper limit of deviations from NBP inflation target (2.5% +/-1 p.p.), reaching 4.8% in November. It was driven mainly by high energy and food prices in the first half of the year, and significant depreciation of Polish zloty in the second half. Additionally inflation was affected by increases in administered and regulated prices. The elevated level of inflation was accompanied by relatively high economic growth, amounting to 4.2% (y/y) in the 3rd quarter 2011, but with weakening domestic demand. The NBP reference rate, after a series of increases in the first half of the year, was equal to 4.5%.

As showed on scatter-graphs, forecasts for all variables were characterised by quite large degree of disagreement among experts and high level of uncertainty. Forecasters expected inflation to decrease next year below the current level – median of central forecasts amounted to 3.5% – but their opinions varied with respect to the scale of decline of price dynamics. 50% of middle forecasts were placed between 3.25 and 3.8%. The uncertainty indicated by experts was also very diversified. The most confident expert indicated as narrow range of possible realizations as 0.4 p.p., while the most cautious as wide as 4.92 p.p. The median IQR amounted to 1.9 p.p.

Asked about the future economic activity, the majority of experts declared values of about 3.2% with quite tight ranges of possible realizations – not exceeding 2 p.p. – what is visible as a cluster in the middle-low part of the scatter graph. However, at the same time a group

of forecasters expecting weaker GDP growth (about 2.3%), but with higher uncertainty (IQRs between 4 and 6 p.p.), appeared.

There was no consensus among forecasters also with respect to the NBP reference rate next year. Four experts out of 17 expected it to remain on the current level (4.5%), while the rest predicted monetary easing, even to 3.75%. As in the case of inflation and GDP growth, the level of uncertainty was very diverse, but generally high.

High level of uncertainty and disagreement among forecasters about future development of Polish economy during this survey probably steamed from intensifying of tensions on international financial markets associated with sovereign debt crisis in the euro area at that time, which affected Poland through the exchange rate. Moreover there were some unfavourable signals about future domestic economic activity, as well as worsening of economic outlook in the euro area.

Looking on the first column of Figure 7, presenting inflation forecasts for 2012 in three consecutive surveys, one notice, that after initial variation of views described above a consensus was formed.[9] On both scatter-graphs, showing results of surveys conducted in 1st and 2nd quarter 2012, the cloud of points moved toward the right-down corner and became more compact. It resulted from upward revision of central forecasts, together with decrease of uncertainty and disagreement about central forecasts. The median IQR decreased from 1.9 p.p. in the first survey to 0.7 p.p. in the third, while the dispersion of central forecasts, measured by inter quartile range of medians, from 0.6 p.p. to 0.2 p.p. These changes were reflected also in the aggregated forecast: median of the aggregated distribution in 2nd quarter 2012 amounted 3.8% comparing to 3.4% in 4th quarter 2011, and the range of 50% probability narrowed to 3.6 – 4.1% from 3.0 – 4.1% (see Table 4).

If one consider only the medians of central forecasts of GDP growth, presented as dotted vertical lines in the second column of scatter graphs, he would conclude that in the analysed period expectations of future economic activity remained constant at 3.0%. However, if one takes under consideration disagreement and, especially, the uncertainty measure, he will get much richer information. In the second survey experts became more unanimous about future economic activity and more certain of their predictions. These

[9] Apart from economic factors, like some mitigation of the turmoil in the global financial markets at the beginning of 2012, affecting forecasts in this period, forming a consensus was facilitated by shortening of the length of forecasting horizon.

experts who at the previous survey round indicated the lowest central values and the widest intervals of possible realizations, revised strongly their forecast toward higher GDP growth and indicated narrower intervals of possible realizations. In the last survey another decrease of uncertainty was observed, with slight movement of medians toward lower GDP growth. In terms of the aggregated distribution, in the analysed period the median dynamics of GDP forecasted for 2012 decreased from 3.1% to 2.9%, and 50% probability interval shifted from 2.4-3.6% to 2.7-3.3% (Table 4).

Similar pattern of forming consensus, i.e. gradual decrease of uncertainty and disagreement, is visible in the third column of scatter-graphs in Figure 7, showing NBP reference rate forecasts. In the 1st quarter 2012 experts shifted their central forecasts towards higher rates, expecting the interest rate staying on the current level, rather than decreasing as in the previous survey. Short after conducting this survey, the Policy Council, concerned about risk of fixing inflation on elevated level due to smaller than expected slowdown of economic activity and high inflationary expectations, started to communicate possible tightening of monetary policy[10], what took place in May this year (rise to 4.75% from 4.5%). In the next survey forecasters again adjusted the level of predicted reference rate to the current level of this rate. As there was almost no disagreement between forecasters (inter quantile range of medians amounts to 0.03 p.p.), and the level of uncertainty was low (median IQR equal to 0.3 p.p.), it might be said that there was a consensus about the level of reference rate in 2012.

[10] See: *Information from the meeting of the Monetary Policy Council held on 3-4 April 2012* (http://www.nbp.pl/homen.aspx?f=/en/onbp/organizacja/rada_archiwum.html).

Figure 7. Individual forecasts of inflation, GDP growth and reference interest rate in 2012

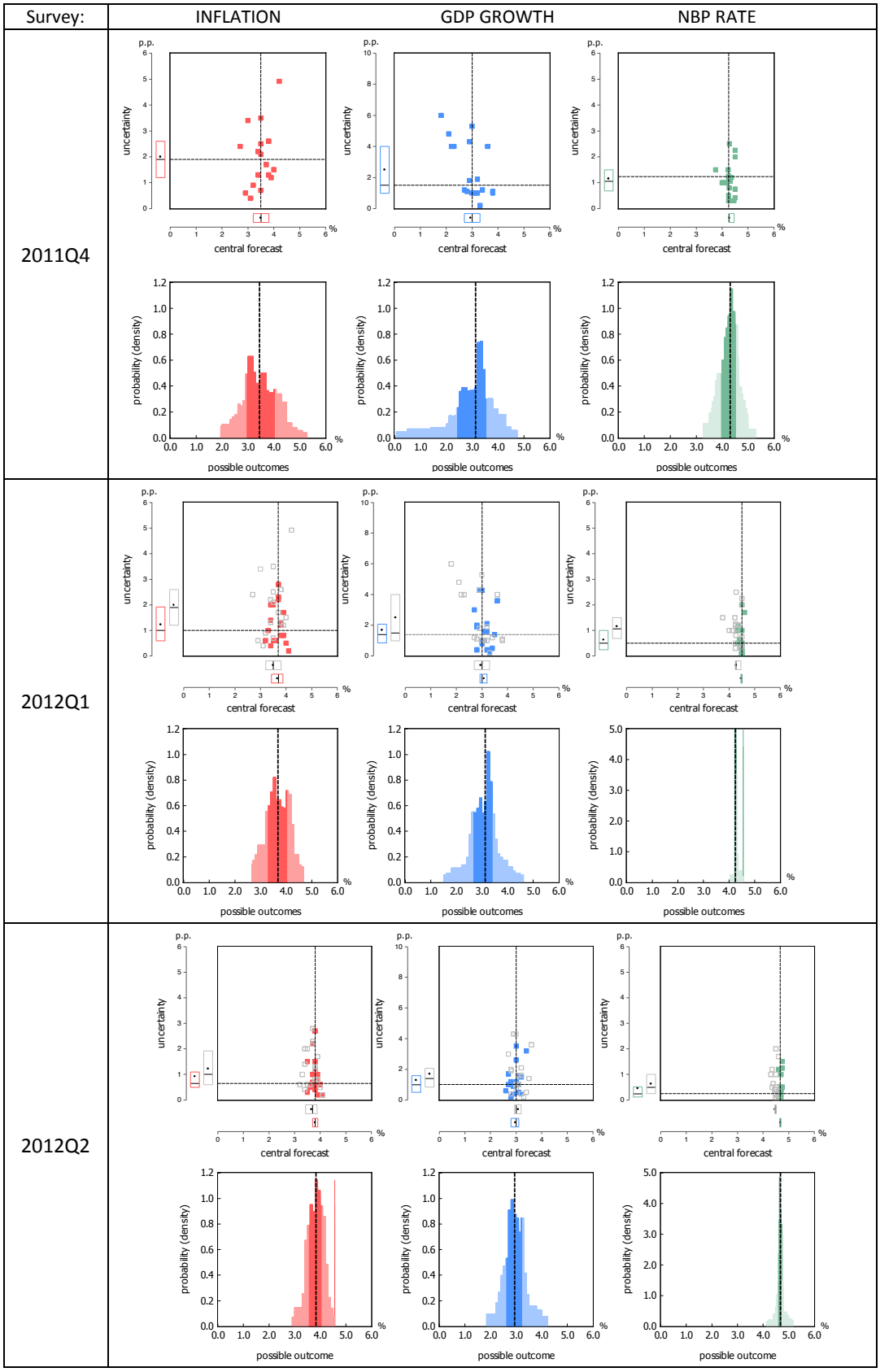


Table 4. Characteristics of the aggregated distributions of forecasts for 2012

	INFLATION			GDP GROWTH			NBP REFERENCE RATE		
survey	2011Q4	2012Q1	2012Q2	2011Q4	2012Q1	2012Q2	2011Q4	2012Q1	2012Q2
median	3.4	3.7	3.8	3.1	3.1	2.9	4.3	4.5	4.7
50% probability	3.0 - 4.1	3.3 - 4.1	3.6 - 4.1	2.4 - 3.6	2.7 - 3.4	2.7 - 3.3	4.0 - 4.6	4.3 - 4.6	4.6 - 4.8
90% probability	1.9 - 5.3	2.7 - 4.1	2.9 - 4.6	0.1 - 4.7	1.5 - 4.6	1.8 - 4.2	3.3 - 5.3	3.8 - 5.1	4.1 - 5.2

Source: NBP SPF.

Table 5 presents development of individual uncertainty, measured by median IQR, and disagreement among forecasters over time, assessed using inter quartile range of central forecasts[11]. Despite the fact that these two characteristics represent different theoretical concepts, as pointed out by Zarnowitz and Lambros (1983) and discussed in Section 2.1, many empirical studies use dispersion of point forecasts as a proxy of macroeconomic uncertainty, arguing that both measures are positively correlated (for example: Giordani and Soderlind 2003; Boreo et al. 2008a). Very short history of the NBP SPF does not allow to conduct any formal analysis of the relationship between these two characteristics, but below we describe some patterns visible in our data.

As expected, for all variables and periods the lowest uncertainty is attached to forecast with the shortest horizon. It monotonically increases with lengthening of the horizon from the current year to two years. Interestingly, experts are less certain what happens in two years than during next 5 years. It steams probably from the presumption that during such a long period – one should keep in mind that survey question asks about 5-year average and not value of a given variable in 5 years – influence of possible shocks will cancel out.

With the exception of interest rate forecasts, the relationship between the length of horizon and level of disagreement among forecasters is not so evident. For example, in the 4th quarter 2011 experts had more divergent views on GDP growth in 2012 than in 2013 (inter quartile range of medians amounted to 0.60 p.p. and 0.50 p.p., respectively), and in the next survey they were more unanimous about level of economic activity in 2014 (0.30 p.p.) than in 2013 (0.65 p.p.). Similarly as for the uncertainty, disagreement on forecasts of 5-year average was lower than on predictions in two-year horizon. As pointed

[11] Other measures useful in assessing disagreement among forecasters, employed in the literature are: standard deviation of central forecasts, quasi standard deviation or mean absolute difference in medians (see: e.g. Giordani and Soderlind 2003, Boero et. al 2008b, Engelberg et al. 2006). We decided to use interquartile range of medians, as this characteristic is readily read from the scatter-graphs and gives similar results as other measures.

out Giordani and Soderlind (2003), disagreement among forecasters results from different information set and methods of processing data. Since evidence relevant for forecasting in such a long horizon is very limited, it seems reasonable that experts views are not very diversified.

Table 5. Uncertainty and disagreement measures for various forecast horizons

	INFLATION			GDP GROWTH			REFERENCE RATE		
survey	2011Q4	2012Q1	2012Q2	2011Q4	2012Q1	2012Q2	2011Q4	2012Q1	2012Q2
horizon	**Uncertainty – median IQR**								
2011	0.20	-	-	0.40	-	-	0.02	-	-
2012	1.90	1.00	0.65	1.50	1.40	1.00	1.06	0.50	0.29
2013	2.50	1.80	1.80	2.35	2.20	2.15	1.75	1.00	1.00
2014	-	2.90	2.95	-	2.70	2.30	-	1.50	1.16
5 years	2.20	1.60	2.00	2.25	2.20	2.20	1.36	1.33	1.05
	Disagreement – interquartile range of central forecasts								
2011	0.00	-	-	0.10	-	-	0.00	-	-
2012	0.60	0.45	0.20	0.60	0.27	0.30	0.22	0.04	0.04
2013	0.50	0.57	0.60	0.80	0.65	0.70	0.50	0.46	0.33
2014	-	0.50	0.20	-	0.30	0.70	-	0.66	0.50
5 years	0.22	0.45	0.30	0.44	0.47	0.40	0.44	0.40	0.33

3.2.2. Assessing central bank credibility

The need to monitor long-term inflation expectations of professional forecasters that are treated as measures of central bank credibility was one of the reasons of introducing the NBP Survey of Professional Forecasters. There were attempts to analyse central bank credibility in Poland using short-term inflation expectations of financial sector analysts as declared in the Reuters survey (e.g. Łyziak, Mackiewicz, Stanisławska 2007), also with some adjustments reflecting a short-term nature of those expectations (Łyziak 2012). In general the conclusions from the above studies suggested a relatively high degree of central bank credibility among financial sector analysts, i.e. a significant role of inflation target in the formation of those expectations.

Assessment of central bank credibility involves two steps (Demertzis et al. 2008): verification, to what extent long-run inflation expectations are anchored to a constant and whether the anchor matches the objective of the central bank. NBP Survey of Professional Forecasters allows measurement of different aspects of central bank credibility and its probabilistic form enriches analysis in this area.

In this part of the paper we are rather going to present possible ways of analysing central bank credibility in Poland on the basis of the NBP SPF data than making any formal tests of this important feature of expectations. It is due to the fact that the NBP SPF is a relatively new tool and the number of available observations is small, which makes conducting such tests impossible. It should be also noted that since the introduction of this survey, its forecasting rounds have been conducted in the environment of a relatively high inflation, that exceeded significantly the NBP inflation target (2.5%) and even its higher bound of tolerable deviations (3.5%). Moreover, current CPI inflation increased only at the end of 2011 (i.e. between the 1st and the 2nd round of the NBP SPF), which limits the possibility of analysing the degree to which inflation forecasts are anchored.

In our analytical framework for analysing central bank credibility we rely on long-term inflation forecasts, exceeding the lags of the monetary transmission mechanism in Poland (according to the recent assessment described in Demchuk et al. 2012, section III.2), i.e. forecasts for inflation in 8 quarters and implied average annual inflation in the 4th and 5th year ahead.[12] We focus both on the anchoring property, according to which long-term inflation expectations should be relatively stable, insensitive to movements in current inflation and short-term inflation expectations (in this role we use forecasts for inflation in 4 quarters), as well as on the consistency between long-term inflation forecasts and the NBP inflation target.[13] Analysis is conducted separately using point forecasts (one-dimensional approach) and probabilistic forecasts with the aim to show how the analysis of central bank credibility can be enriched while having probabilistic forecasts and how the conclusions can change in this case.

Except standard measures of central bank credibility, i.e. the deviations of the medians of point inflation forecasts from the NBP inflation target, the design of the NBP SPF allows developing more sophisticated measures of central bank credibility, that refer to different aspects of central bank credibility. Those measures – including the deviation of the median of aggregated distribution of forecasts from the NBP inflation target, the order of the

[12] Implied average annual inflation in the 4th and 5th year was calculated on the basis of NBP SPF inflation forecasts for the next 5 years and for the first 3 subsequent years. It seems that this measure outperforms the forecast of average annual inflation in the next 5 years due to the fact that the latter one is influenced – by its construction – by changes in short-term inflation expectations and can be used as a measure of central bank credibility only in a long-term perspective, with observations available for different inflation episodes.

[13] Since February 2003 there has been a continuous target 2.5%±1p.p.

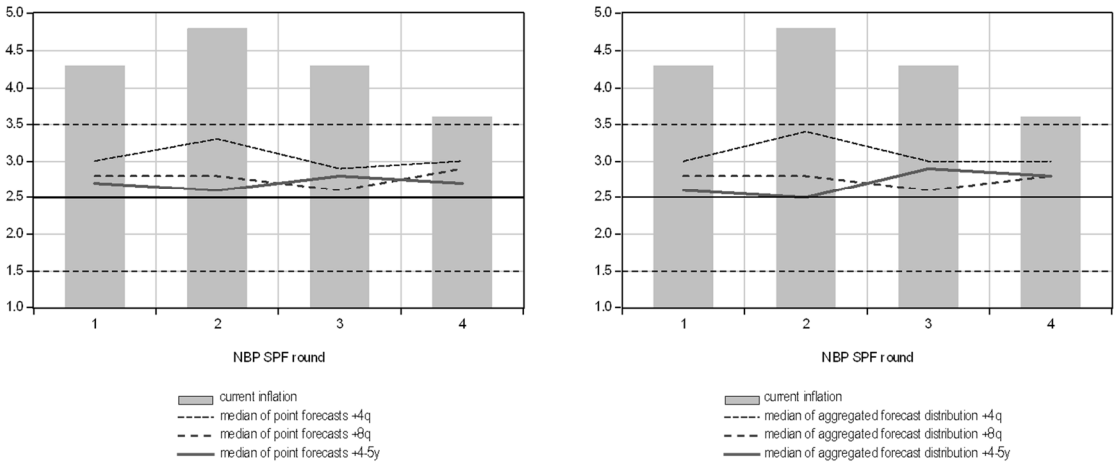
quantiles of the aggregated distribution of individual inflation forecasts corresponding to the NBP inflation target (2.5%) or the probability of future inflation being within range of permitted deviations from the NBP target (1.5%, 3.5%) – make use of both dimensions of the forecasts declared in the NBP SPF, i.e. medians and ranges of possible outcomes. In addition we analyse the characteristics of individual forecasts, in particular medians of short- and long-term individual forecasts.[14]

Taking into consideration measures of central bank credibility described above there are the following observations that can be made on the basis of the results of the NBP Survey of Professional Forecasters:

1) Medians of central (point) forecasts for all the horizons under consideration were above the NBP inflation target, but below the upper bound of permitted fluctuations. Deviations of the NBP SPF experts' inflation forecasts from the NBP inflation target were significantly lower than the deviations of current CPI inflation from the target, especially in the case of longer-term forecasts (Figure 8).
2) There are some differences between the medians of central (point) forecasts and the medians of aggregate forecast distribution. E.g. in the 2nd round of the NBP SPF the latter measure, reflecting both the individual point forecasts and uncertainty, was consistent with the NBP inflation target even if the former stayed above the target (Figure 8). Such differences can appear especially during the episodes of high dispersion of individual assessments of forecast uncertainty. In such circumstances analysing point forecasts only can lead to serious misinterpretations.
3) Fluctuations of inflation forecasts for the horizons of 4 quarters, 8 quarters and 4-5 years were significantly lower than changes in current CPI inflation. In line with the concept of anchored inflation expectations, long-term inflation forecasts of the NBP SPF experts seem to be less affected by changes in current inflation than their short-term inflation forecasts. For example, the increase of current CPI inflation between the 1st and the 2nd round of the NBP SPF had no effect on 8-quarter inflation forecast and the average predicted inflation in 4-5 years even decreased (Figure 8).

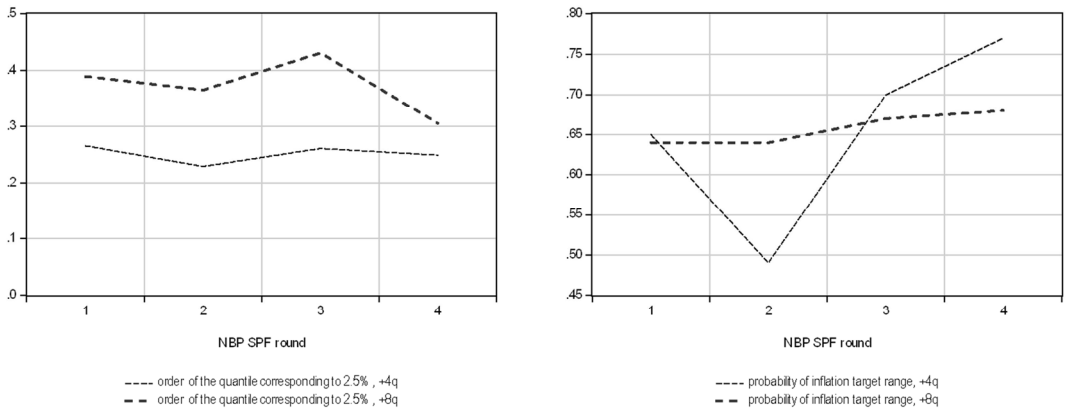
[14] There are also other measures that can be calculated on the basis of the NBP SPF, such as the quantiles of implied individual forecast distributions corresponding to the NBP inflation target (2.5%) or probabilities of future inflation being within the range of permitted deviations from the NBP target (1.5%, 3.5%) based on implied individual forecast distributions.

Figure 8. Medians of point inflation forecasts (left panel, %) / medians of aggregate forecast distribution (right panel, %) vs. current CPI inflation



4) The orders of quantiles corresponding to the NBP inflation target based on the aggregated forecast distribution are closer to the 0.5 (i.e. the order of the median) in the case of long-term forecasts (horizon: 8 quarters) than short-term forecasts (horizon: 4 quarters). Probability of the inflation target range is significantly less volatile in the long-term horizon, insensitive to changes in respective probability displayed by short-term forecasts. E.g., a decrease in the probability inflation in 4 quarters being within the inflation target range observed between the 1st and the 2nd round of the NBP SPF had no influence on the long-term forecasts (Figure 9).

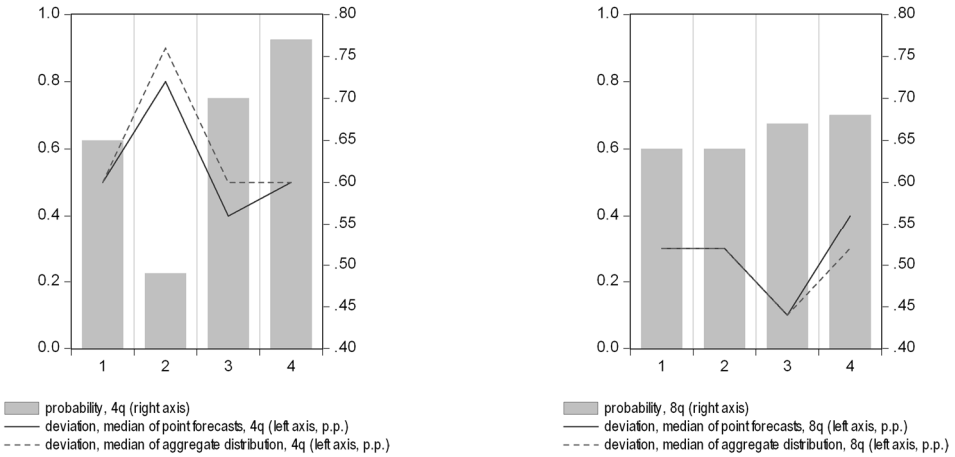
Figure 9. Quantile ranks corresponding to the NBP inflation target (left panel) and probability of future inflation being within the inflation target range based on the aggregated distribution of inflation forecasts



5) Assessment of central bank credibility based on deviations of point inflation forecasts from the NBP inflation target and probabilities of future inflation being within the NBP inflation target range, are broadly consistent with each other, which means that larger

deviations of point forecasts from the target coincide with lower probabilities of achieving the target. There was, however, an exception from this regularity, in the 4th round of the NBP SPF, when the deviation of long-term inflation forecasts from the NBP inflation target increased with respect to the 3rd NBP SPF round, while probability of the inflation target range increased (Figure 10).

Figure 10. Deviations of inflation forecasts for +4 quarters (left panel) and +8 quarters (right panel) from the NBP inflation target and probability of future inflation being within the inflation target range



6) Analyses of individual inflation forecasts confirms above conclusions. Taking into consideration the experts of the NBP Survey of Professional Forecasters active in all the forecasting rounds, it seems that in the majority of cases the increase of the medians of their individual forecasts distributions for 4 quarters ahead was accompanied by either a smaller increase of the median of 8-quarter or 4-to-5-year forecast distribution, its stabilization or decrease (Figure 11.A). Analysing scatter graphs showing changes in short- and long-term forecasts it can be observed than in majority of cases long-term forecasts were less volatile than the short-term ones (Figure 11.B).

Figure 11.A. Changes in individual central inflation forecasts , all the NBP SPF forecasting rounds, sub-sample of experts active in all the rounds

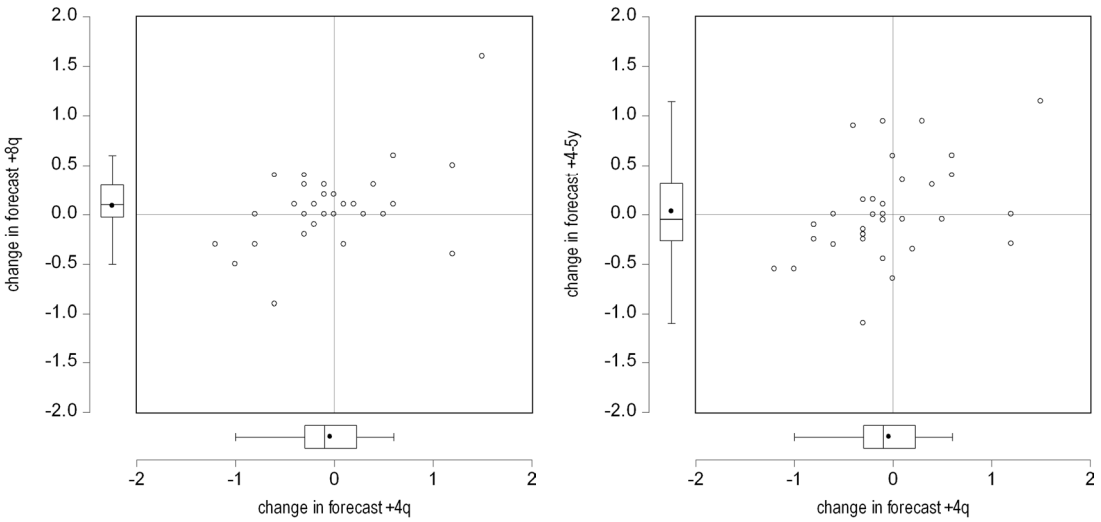
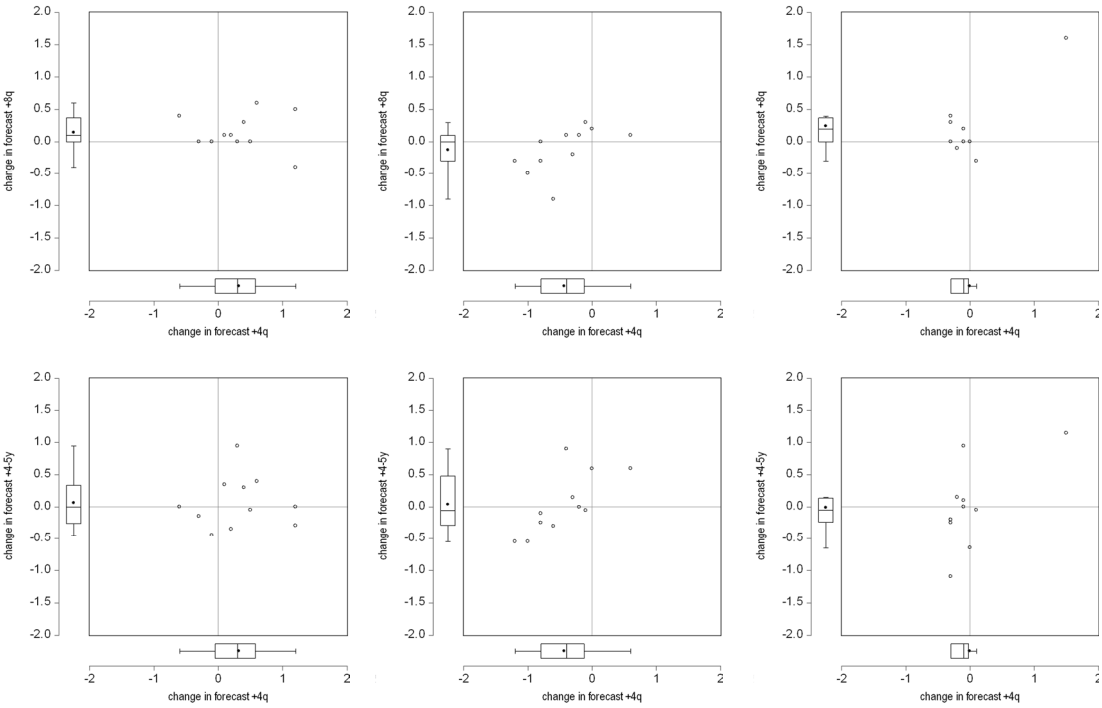


Figure 11.B. Changes in individual inflation central forecasts (in subsequent forecasting rounds, sub-sample of experts active in all the rounds



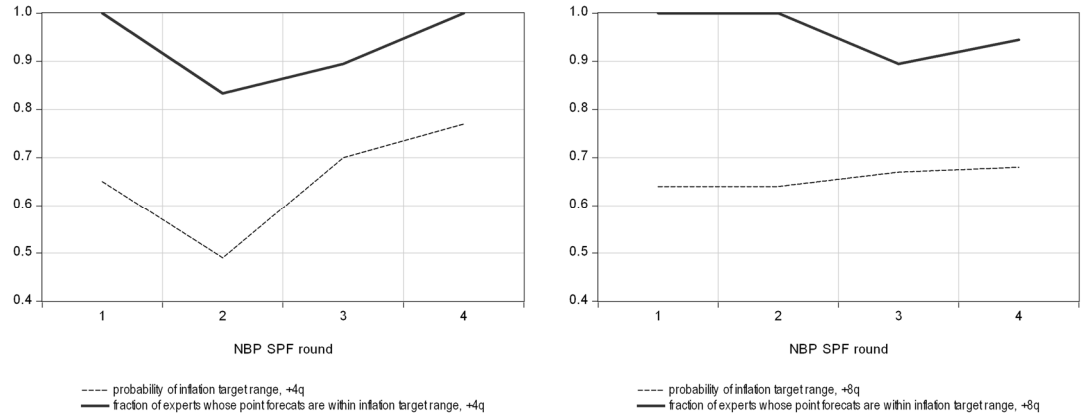
2nd NBP SPF round 3rd NBP SPF round 4th NBP SPF round

The measures of central bank credibility proposed above, indicating different aspects of credibility, should be treated as complementary to each other. It can happen that in some cases conclusions based on a single measure are not consistent with each other. For example, stable deviations of inflation forecasts from the NBP inflation target can coexist

with the quantile of the aggregated distribution corresponding to the NBP inflation target getting closer to median or the probability of future inflation being within the inflation target range can increase in spite of the fact that other measures indicate worsening of central bank credibility. Therefore all the measures of this kind should be monitored and cross-checked with each other.

Finally, it should be noted that the assessment of central bank credibility with the probability of the inflation target range based on aggregate forecast distribution and with the fraction of NBP SPF experts whose point forecasts are within the inflation target range does not give unique conclusions. The further indicator was used by Łyziak, Mackiewicz, Stanisławska (2007) while analysing financial sector agents' inflation expectations (due to lack of probability forecast at the disposal). Even if tendencies of both measures are similar in the case of the 4-quarter horizon, they differ significantly in the case of longer-term forecasts (Figure 12).

Figure 12. Probability of the inflation target range based on aggregate forecast distributions vs. the fraction of experts whose point forecasts are within the inflation target range



Interpreting the results of the NBP SPF we can observe that in the analysed period – too short to make any final conclusions – inflation forecasts of the NBP SPF experts seemed to be in line with the concept of anchored expectations, however, the anchor was slightly above the NBP inflation target.

4. Conclusions

The literature referred to in the paper suggests that relying on point forecasts may lead to invalid interpretations, because of different meanings of point forecasts declared by individual forecasters. In analysing macroeconomic forecasts there is a need to observe not only point forecasts, but also uncertainty faced by experts. It seems especially important in the present time of huge uncertainty in the global macroeconomic environment. For those reasons, the use of probability forecasts should be preferable. At the same time it is necessary to take into consideration difficulties experts can face in the quantification of uncertainty and distinguishing between objective and subjective probability. For this reason, the NBP Survey of Professional Forecasters introduced in 2011, has other methodological foundations than traditional SPFs. The NBP SPF experts are asked to consider various scenarios of macroeconomic developments and assess a possible range of values of a given variable in a certain horizon, indicating 5^{th}, 50^{th} and 95^{th} percentiles of their subjective (personal) probability distributions. Such a form of survey questions related to the CPI inflation, GDP growth and the NBP reference rate, makes it possible to analyse separately intrapersonal uncertainty and interpersonal uncertainty (i.e. disagreement).

In terms of using the NBP SPF outcomes we follow the principle of a distinct treatment of data in two different areas of their use. In the analysis of expectations experts' forecasts should be presented in the way fully consistent with subjects' responses. Therefore scatter graphs are perceived as the most accurate way of analysing expectations, while aggregated distributions, and in particular their point characteristic, should be treated cautiously while analysing expectations.

In the case of macroeconomic forecasting survey data can be processed with the aim of achieving the best forecasts. It can be reasonable to construct aggregate forecast distributions based on percentiles provided by individual forecasters. Aggregated distributions can be derived either with equal weights attached to approximated individual forecast distributions or – in order to address the problem of heterogeneity of experts and personal biases – with differentiated weights, based on calibration and information scores (Cooke 1991). As far as individual forecasts distributions are concerned, we approximate them with the use of the piece-wise uniform distributions. As shown in the paper, applying

other types of distributions (beta, triangular) in the case of the results obtained from the NBP SPF had a little effect on the features of aggregate distributions important from the point of view of monetary policy makers.

Discussing the results of the first four rounds of the NBP SPF we attempt to present the usefulness of this analytical tool and rich interpretation possibilities it offers. Two-dimensional analysis, taking into account not only point (individual or aggregate) forecasts, but also subjective assessments of uncertainty, makes the analysis of macroeconomic forecast and their changes richer. E.g. the results of the NBP SPF allow developing various measures of central bank credibility, reflecting different aspects of this qualitative feature. Constrained by a small number of observations we do not pretend to form decisive conclusions concerning the formation of macroeconomic forecasts by professional economists in Poland. Some deliberations of this kind provided in the paper – e.g. the observation that inflation forecasts of the NBP SPF experts seem in line with the concept of anchored expectations, but the anchor exceeds the NBP inflation target – are strongly influenced by the specific features of the time period under consideration (huge uncertainty, relatively high inflation) and should be treated as hypotheses to be verified in the future.

In general macroeconomic forecasts collected in the previous rounds of the NBP SPF are characterized by high level of uncertainty and disagreement among forecasters, which is not surprising given the turbulences in the global economy. Macroeconomic forecasts collected in the NBP SPF – similarly as in other surveys of this kind – seem to display a significant heterogeneity, both point and probabilistic. More than a half of individual subjective distributions are asymmetric with their highest share in the case of the forecasts of the GDP growth. Inflation forecasts seem positively skewed, while GDP growth forecasts – negatively skewed, however the asymmetry is not strong.

References

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

Boero G., Smith J., Wallis, K. (2008b), Evaluating a three-dimensional panel of point forecasts: the Bank of England Survey of External Forecasters, *International Journal of Forecasting*, 24, 354-367.

Bowles C., Friz R., Genre V., Kenny G., Meyler. A., Rautanen. T. (2007), The ECB Survey of professional forecasters (SPF). A review after eight years' experience, *ECB Occasional Paper Series*, 59.

Carroll C. D. (2003), Macroeconomic expectations of households and professional forecasters, *Quarterly Journal of Economics*, 118(1), 269-298.

Clemen R. T., Winkler L. R. (1999), Combining probability distributions from experts in risk analysis. *Risk Analysis*, 19 , 187-203.

Cooke R. (1991), *Experts in Uncertainty: Expert Opinion and Subjective Probability in Science,* Oxford University Press.

Demchuk O., Łyziak T., Przystupa J., Sznajderska A., Wróbel E. (2012), Monetary policy transmission mechanism in Poland. What do we know in 2011?, *National Bank of Poland Working Papers*, 116, National Bank of Poland.

Demertzis M., Marcellino M., Viegi N. (2008), A measure for credibility: tracking US monetary developments, *DNB Working Papers*, 187, Netherlands Central Bank.

Döepke J., Dovern J., Fritsche U., Slacalek J. (2008), The dynamics of European inflation expectations, *The B.E. Journal of Macroeconomics*, 8(1), article 12.

ECB (2009), Results of special questionnaire for participants in the ECB Survey of Professional Forecasters, http://www.ecb.int/stats/prices/indic/forecast/html/index.en.html.

Engelberg J., Manski Ch., Williams J. (2006), Comparing the point predictions and subjective probability distributions of professional forecasters, *National Bureau of Economic Research Working Paper*, 11978.

Engelberg J., Manski Ch., Williams J. (2009a), Comparing the point predictions and subjective probability distributions of professional forecasters, *Journal of Business & Economic Statistics*, 27, 30-41.

Engelberg J., Manski Ch., Williams J. (2009b), Assessing the temporal variation of macroeconomic forecasts by a panel of changing composition, http://faculty.wcas.northwestern.edu.

Giordani P., Soderlind P. (2003), Inflation forecast uncertainty, *European Economic Review*, 47, 1037-1059.

Henzel S., Wollmershäuser T. (2006), The New Keynesian Phillips Curve and the role of expectations: evidence from the Ifo World Economic Survey, *CESIfo Working Paper*, 1694.

Hogarth R. (1975), Cognitive processes and the assessment of subjective probability distributions, *Journal of the American Statistical Association*, 70 (350), 271-289.

Johnson D. (2002), Triangular approximations for continuous random variables in risk analysis, *The Journal of the Operational Research Society*, 53(4), 457-467

Kowalczyk H. (2010), O eksperckich ocenach niepewności w ankietach makroekonomicznych *[Expert assessment of uncertainty in macroeconomic surveys]*, *Bank i Kredyt*, 41(5), 101-122.

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

Łyziak T., Mackiewicz J., Stanisławska E. (2007), Central bank transparency and credibility: the case of Poland, 1998-2004, *European Journal of Political Economy*, 23(1), 67-87.

Łyziak T. (2012), Inflation expectations in Poland, *National Bank of Poland Working Papers*, 115, National Bank of Poland.

Mehra Y. (2002), Survey measures of expected inflation: revisiting the issues of predictive content and rationality, *Economic Quarterly*, 88(3), 17–36.

Savage L. (1971), Elicitation of personal probabilities and expectations, *Journal of the American Statistical Association*, 66(336), 783-801.

Scheufele R. (2011), Are qualitative inflation expectations useful to predict inflation?, *Journal of Business Cycle Measurement and Analysis*, 2011(1), 29-53.

Trehan B. (2010), Survey measures of expected inflation and the inflation process, *Working Paper*, 2009-10, Federal Reserve Bank of San Francisco.

Tversky A., Kahneman D. (1974), Judgement under uncertainty: heuristics and biases, *Science*, 185, 1124-1131.

Zarnowitz A., Lambros L. (1983), Consensus and uncertainty in economic prediction, *NBER Working Paper*, 1171.

Annex: NBP SPF questionnaire (of the survey taken in II quarter 2012)

NBP Survey of Professional Forecasters Q2 2012

Guidelines on completing the form

The survey concerns projections related to selected macroeconomic variables in different time horizons. *For quarterly year-on-year indices these are the next year and the year following it.* For annual indices, we ask about the current year and the following two years, and for some variables - additionally about the nearest five years. The main questions of the survey are provided in the part A of the questionnaire. They concern CPI inflation and GDP growth. In the part B, we ask for complementary forecasts.

Questions related to CPI inflation, GDP growth and the NBP reference rate are probabilistic.

We would like you to consider possible macroeconomic scenarios, think about how probable they are according to your subjective assessment and provide the central forecast and the range of possible values on the basis thereof.

The central forecast	is the value for which the occurrence of lower and higher values is equally probable. Put it another way, it is the median (50th percentile) of your subjective probability distribution. It is marked with ”50” index (INF_{50} for the median of the CPI inflation forecast, GDP_{50} for the median of the GDP forecast, REF_{50} for the median of the NBP reference rate forecast).
The range of forecasted values	is an interval to which you assign the probability of 0.9 and the probabilities of the occurrence of values below its lower and upper endpoints are each 0.05. Therefore, the lower endpoint of the interval is the 5th percentile and the upper endpoint - the 95th percentile of your subjective probability distribution. The lower endpoint is marked with ”05” index (INF_{05}, GDP_{05}, REF_{05}) the upper endpoint - with ”95” index (INF_{95}, GDP_{95}, REF_{95}).

For complementary questions about CPI inflation and GDP growth, please provide point forecasts.

We would like you to provide values of the variables consistent with your central forecasts for inflation and growth.

Forecast horizons in the current edition of the NBP Survey are as follows:

- quarters: Q2 2013, Q2 2014;
- years: 2012, 2013, 2014 and the period from: 2012 - 2016.

Data available upon the current survey:

- CPI inflation (%) ..... (*year-on-year)*
- GDP growth(%) ..... (*year-on-year)*

A. CPI inflation and GDP growth forecasts

A1. CPI inflation (%)

In a quarter (current year quarter on preceding year corresponding quarter) :

2013 Q2: INF_{05} = INF_{50} = INF_{95} =

2013 Q3: INF_{05} = INF_{50} = INF_{95} =

In a year (annual average) :

2012: INF_{05} = INF_{50} = INF_{95} =

2013: INF_{05} = INF_{50} = INF_{95} =

2014: INF_{05} = INF_{50} = INF_{95} =

During the period of the nearest 5 years (annual average) :

2011 - 2015: INF_{05} = INF_{50} = INF_{95} =

A2. GDP growth (%)

In a quarter (current year quarter on preceding year corresponding quarter) :

2013 Q2: GDP_{05} = GDP_{50} = GDP_{95} =

2013 Q3: GDP_{05} = GDP_{50} = GDP_{95} =

In a year (annual average) :

2012: GDP_{05} = GDP_{50} = GDP_{95} =

2013: GDP_{05} = GDP_{50} = GDP_{95} =

2014: GDP_{05} = GDP_{50} = GDP_{95} =

During the period of the nearest 5 years (annual average) :

2011 - 2015: GDP_{05} = GDP_{50} = GDP_{95} =

B. Complementary forecasts

B1. NBP reference rate

In a quarter (average) :

2013 Q2: REF_{05} = REF_{50} = REF_{95} =

2013 Q3: REF_{05} = REF_{50} = REF_{95} =

In a year (average) :

In a year (average) :

2012: REF $_{05}$ = REF $_{50}$ = REF $_{95}$ =

2013: REF $_{05}$ = REF $_{50}$ = REF $_{95}$ =

2014: REF $_{05}$ = REF $_{50}$ = REF $_{95}$ =

During the period of the nearest 5 years (average) :

2011 - 2015: REF $_{05}$ = REF $_{50}$ = REF $_{95}$ =

B2. CPI and GDP determinants

If the following variables were important while formulating the forecasts from the part A, we would like you to provide the values assumed while estimating the central forecasts for inflation and growth.

Average values in the year	2012	2013	2014
EUR to PLN exchange rate			
Registered unemployment rate (%)			
Nominal annual growth of the total average gross wages and salaries (%)			
Brent oil prices in USD/barrel			
Annual GDP growth in the euro area (%)			

B3. Other inflation and GDP determinants which are especially important in the current forecasting cycle (please comment):