Are experts’ probabilistic forecasts similar to the NBP projections?

Halina Kowalczyk, Ewa Stanisławska
Are experts’ probabilistic forecasts similar to the NBP projections?

Halina Kowalczyk, Ewa Stanisławska
Halina Kowalczyk – Economic Institute, Narodowy Bank Polski; halina.kowalczyk@nbp.pl
Ewa Stanisławska – Economic Institute, Narodowy Bank Polski; ewa.stanislawska@nbp.pl

Acknowledgements:
The authors would like to thank dr hab. Tomasz Łyziak and an anonymous referee for valuable comments. The paper expresses opinions of the authors and does not reflect a view of Narodowy Bank Polski.
# Contents

1 Introduction 5
2 Data: two kinds of forecasts 7
   2.1 NBP fan charts 7
   2.2 NBP Survey of Professional Forecasters 8
   2.3 Problems related to comparing forecasts 8
3 Comparing performance of probabilistic forecasts 12
   3.1 Comment on methods 12
   3.2 Results 13
4 Assessing similarity of distributions 18
   4.1 Comment on measurement 18
   4.2 Results 19
5 Conclusions 26
References 27
Appendix 1 Comparison of point forecasts 29
Abstract

We assess similarity of the Polish central bank’s forecasts published in Inflation Reports and economic experts’ forecasts (from NBP Survey of Professional Forecasters), an important issue in monetary policy. Contrary to other studies which use point forecasts, we are interested in comparing whole forecasts’ distributions. We are especially interested whether the SPF experts mirror the NBP projections. For this purpose, we propose employing methods based on distance between distributions. Unfortunately, substantial heterogeneity of forecasts, as well as short and atypical period analyzed, limit drawing firm conclusions with this respect.

JEL codes: D83, D84, E37

Keywords: survey data, fan charts, probabilistic forecasts, inflation forecasts, GDP growth forecasts, distribution similarity
1 Introduction

It is a common practice of central banks to regularly publish their forecasts of inflation and GDP growth, sometimes complemented with projection of interest rate path. This high degree of transparency is intended, inter alia, to support explaining actions of monetary authorities to the public and to influence its expectations about future performance of the economy. Ability to manage expectations of the public with various tools available to the central bank, especially with macroeconomic projections and the inflation target, in turn is believed to facilitate conducting monetary policy. Not surprisingly, the issue of influence of central bank forecasts on the private sector forecasters has drawn attention of researchers.

Fujiwara (2005) has shown that professional forecasters’ predictions of inflation and GDP growth are influenced by Japanese central bank projections, but not vice versa. Additionally, central bank projections reduce disagreement among forecasters. Similarly, Filacek and Saxa (2012) provide evidence that private sector forecasts get closer to the Czech central bank forecasts of inflation and interest rate after their announcements, and this effect is stronger during periods of higher uncertainty. Interestingly, central bank forecasts do not seem to affect GDP growth forecasts.

Empirical literature confirms that central banks affect inflation forecasts of the public sector also in Sweden, Canada, the UK, Switzerland and Japan (Hubert, 2013) and Chile (Pedersen, 2013). Ehrmann et al. (2010) takes wider perspective and considers the role of central bank transparency (publishing forecasts constitutes significant part of it) in influencing disagreement among professional forecasters (on future development of several economic indicators, such as inflation rate, real GDP growth rate, unemployment, interest rate). They find significant role of central bank transparency in reducing dispersion of professional forecasters in the euro area, but not among the general public (consumers). Similar study by Csávás et al. (2012), conducted for greater group of economies (including CEE countries), indicates that central bank transparency improves accuracy and lowers dispersion of short interest rate forecasts of professional forecasters, and to lesser degree of inflation forecasts. The influence on forecasts of real variables is rather weak. The most important components of transparency reducing uncertainty in the economy seems to be: publishing forecasts, assessing accuracy of past forecasts error and having qualitative target. The evidence for Poland suggests that publishing central bank forecasts reduces dispersion of private sector’s GDP growth predictions, but has not impact on inflation forecasts (Kotłowski, 2015). The author explains it by anchoring effect of the NBP inflation target and by greater uncertainty of economic activity than of inflation during the global economic crisis.

This paper contributes to this strand of literature by taking into account probabilistic character of central bank forecasts. Many central banks, including Narodowy Bank Polski (National Bank of Poland), publish their forecasts in the form of fan charts informing about uncertainty of future outcomes. Also the forecasts of the private sector
sometimes take probabilistic form. Therefore, instead of looking at the central path of the forecasts – as in the mentioned above papers – it is advisable to compare whole forecast distributions revealing uncertainty assessment of forecasters. Ignoring this additional dimension might lead to different conclusions.

In the paper we evaluate similarity between the Polish central bank’s forecasts and the professional forecasters’ predictions in terms of distributions. Due to data constraints, we do not compare forecasts made by the public before and after publication of the central bank projections, like in the papers mentioned above, but rather look whether central bank projections are reflected in the distributions of individual professional forecasters in general. If the two forecasts are similar, we are not able to prove with certainty the causality (that external forecasters follow the central bank) – similarity of distributions might result from various reasons – but nevertheless, likeness of expectations of the central bank and the public is a desirable feature. With this respect our goal differs from Kotłowski (2015), who directly tests whether central bank forecasts affect experts’ predictions.

As the analysis employs data from the Survey of Professional Forecasters, introduced in 2011 by Narodowy Bank Polski, it gives opportunity to learn about the way its’ participants formulate forecasts, especially when it comes to judgment of uncertainty. The evaluation of survey results, after 4 years of conducting the NBP SPF, constitutes the second goal of the paper. With this respect our paper is related to studies dealing with assessment of probabilistic forecasts surveys of professional forecasters (e.g. Kenny et al., 2012, 2014; Boero et al. 2011, 2014). Unfortunately, a short history of the NBP SPF limits the scope of the analysis and the conclusions should be treated as preliminary.

The paper is structured as follows. Section 2 presents briefly data employed in the analysis, i.e. the NBP inflation and GDP growth projections and the NBP Survey of Professional Forecasters. Comparing predictions from these two sources we take under consideration two dimensions: forecast performance, in section 3, and similarity of forecasts distributions, in section 4. In each section, description of results is preceded by presentation of tools applied in the analysis. The last section concludes. Additionally, in Appendix we present comparison of point forecasts.

---

1 See e.g. Survey of Professional Forecasters conducted by Philadelphia Fed, ECB Survey of Professional Forecasters or Bank of England Survey of External Forecasters.
2 The issue of uncertainty measurement/assessment has received increased attention since the beginning of the global financial crisis. Surveys, especially those of professional forecasters and formulated in probabilistic terms, constitute important source of information about uncertainty, as indicated by, inter alia, Rich et al. (2012), García and Mansanares (2007) or Andrade et al. (2014).
3 Kotłowski (2015) employs different data set (Reuters Thomson survey of financial sector analysts) than used in our study (Survey of Professional Forecasters). The advantage of Reuters Thomson survey is its long history, but it contains only point forecasts. The author analyzes level and dispersion of point forecasts of inflation and GDP growth. Sample covered by the SPF is much shorter, which limits choice of analytical tools, but probabilistic character of predictions gives richer information on experts’ opinions. As result, both studies ask different questions on links between experts’ forecasts and central bank projection and their results are not easily comparable.
2 Data: two kinds of forecasts

2.1 NBP fan charts

In the paper we compare two sets of probabilistic forecasts of inflation and GDP growth: one formulated by the professional forecasters (collected in the NBP SPF) and one published by Narodowy Bank Polski.

The NBP started to publish inflation projections in August 2004 and one year later complemented it with GDP growth forecast. Both forecasts are prepared under assumption of constant interest rate and cover up to three years (current year and next two years). The NBP projections appear three times a year as a part of Inflation Reports (in March, July and November), and are additionally presented during press conferences. From the beginning, both forecasts communicate not only the central path of the forecasted variable, but also the projection risk, portrayed in a fan chart. The model used for forecasting and the way of constructing fan charts underwent some changes over time. Until July 2011 fan charts were centered on the median and equal-probability ranges, each representing 15% probability, were used. Since then the central forecast has represented the most probable path of the forecasted variable, i.e. the mode of the forecast distribution, and the highest probability density ranges have been applied. The probability of the central range is equal to 30%; additionally extensions to 60% and 90% probability intervals are showed (Figure 1).

Figure 1 Exemplary projection fan chart published by the NBP

![Figure 1](image.png)

Source: NBP Inflation Report.

The distribution of forecast is described by a two-piece normal distribution⁴. Its parameters determining probability bands depend on past forecast errors, relation of uncertainty of exogenous variables in the current projection round to the mean values and assessment of skewness based on scenarios prepared by experts. The risks do not need to be distributed evenly on both sides of the central path.

⁴ The construction of fan chart is described in Inflation Reports published in October 2008 and July 2011.
2.2 NBP Survey of Professional Forecasters

The NBP SPF, introduced in the 3rd quarter 2011\(^{\text{1}}\), is the only probabilistic survey conducted in Poland. It collects, on quarterly basis, opinions of professional forecasters – mainly from financial institutions, but also from academic and research institutions as well as employee and employer organizations – on inflation and GDP growth in several horizons (ranging from one year to 5 years). Additionally, participating experts express their opinions about NBP reference rate and other relevant variables, like unemployment rate, exchange rate, wages growth, GDP growth in the euro area and oil price\(^{\text{2}}\).

Survey participants are asked to consider various scenarios, think about how probable they are, and provide the point central forecast together with the range of the possible values for a given variable. The central forecast is defined as the 50th percentile (median) of expert’s forecast subjective distribution, while the range of forecasted values corresponds to interval between 5\(^{\text{th}}\) and 95\(^{\text{th}}\) percentile. This formulation of the survey question, different from other SPF’s, like those conducted by the ECB or the Philadelphia Fed solves some problems related to probabilistic questions in the form of predefined intervals to which expert assigns probabilities, and facilitates expressing uncertainty by experts. Moreover, it allows straightforward measurement of this uncertainty by inter-quantile range – the difference between 95\(^{\text{th}}\) and 5\(^{\text{th}}\) percentiles given by the expert.

2.3 Problems related to comparing forecasts

2.3.1. Timing

In the study, we match NBP SPF forecasts with NBP projections for the same quarter in the period between 2011q3 and 2014q4. The forecast horizon equals +4 quarters. Relatively short period under analysis results from data availability. Table 1 shows matched pairs of forecasts – in the rest of the paper we will identify pairs of forecasts by referring to the survey date. NBP SPF is conducted in last two weeks of each quarter, extended by two first working days of next quarter. NBP projections are published three times a year, more or less in the middle of March, July and November. In all periods, except 2\(^{\text{nd}}\) quarter forecasts, the NBP projections are available prior to the survey. No NBP projection corresponds to the 3\(^{\text{rd}}\) quarter surveys.

---

\(^{\text{1}}\) Results from the first survey round, treated experimentally, are not published.

\(^{\text{2}}\) For additional information on NBP SPF see: Kowalczyk et al. (2013).
Data: two kinds of forecasts

<table>
<thead>
<tr>
<th>Survey round</th>
<th>Survey quarter</th>
<th>Survey date</th>
<th>Projection date</th>
<th>Forecast for</th>
<th>Projection before SPF?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2011q3</td>
<td>2011m09</td>
<td>-</td>
<td>2012q3</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>2011q4</td>
<td>2011m12</td>
<td>2011m11</td>
<td>2012q4</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>2012q1</td>
<td>2012m03</td>
<td>2012m03</td>
<td>2013q1</td>
<td>yes</td>
</tr>
<tr>
<td>4</td>
<td>2012q2</td>
<td>2012m06</td>
<td>2012m07</td>
<td>2013q2</td>
<td>no</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>2012m09</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>2012q4</td>
<td>2012m12</td>
<td>2012m11</td>
<td>2013q4</td>
<td>yes</td>
</tr>
<tr>
<td>7</td>
<td>2013q1</td>
<td>2013m03</td>
<td>2013m03</td>
<td>2014q1</td>
<td>yes</td>
</tr>
<tr>
<td>8</td>
<td>2013q2</td>
<td>2013m06</td>
<td>2013m07</td>
<td>2014q2</td>
<td>no</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>2013m09</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>2013q4</td>
<td>2013m12</td>
<td>2013m11</td>
<td>2014q4</td>
<td>yes</td>
</tr>
<tr>
<td>11</td>
<td>2014q1</td>
<td>2014m03</td>
<td>2014m03</td>
<td>2015q1</td>
<td>yes</td>
</tr>
<tr>
<td>12</td>
<td>2014q2</td>
<td>2014m06</td>
<td>2014m07</td>
<td>2015q2</td>
<td>no</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>2014m09</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>2014q4</td>
<td>2014m12</td>
<td>2014m11</td>
<td>2015q4</td>
<td>yes</td>
</tr>
</tbody>
</table>

Source: own elaboration.

Different timing of forecasts leads to potentially different information sets of forecasters. The time of preparing forecasts overlaps significantly in the case of forecasts made in the 2nd quarter, while the greatest discrepancies appear in 4th quarter. Comparing projection cut-off dates, survey dates and GUS publication calendar of GDP and inflation data (Table 2) leads to the following observations:

- In the case of three pairs of forecasts (prepared in 2nd quarters), we can assume that forecasters had the same information set.
- In the case of 4th quarter forecasts, SPF experts knew realization of GDP in the previous quarter – which wasn’t available for the NBP projection staff – and additionally two/three more realizations of inflation than NBP staff.
- In the case of forecasts prepared in 1st quarters, the SPF experts had advantage of knowing one or two more realizations of inflation, and in 2012, additionally, the latest GDP data.

If the realizations of macro variables were in line with expectations, the differences in information sets would not affect similarity/dissimilarity of forecasts. Therefore, in the Table 2 we indicated whether the realizations of GDP growth and inflation surprised the market or not, and in which direction. As a proxy for expected values of these variables we used data from Thomson Reuters survey among financial market analysts.
Additional complication is that in the case of two pairs of forecasts (2012q4 and 2013q1), the main NBP interest rate was lowered between the cut-off date and the time of survey.

Table 2 New data appearing between projection cut-off date and conducting survey (last day)

<table>
<thead>
<tr>
<th>Survey quarter</th>
<th>Projection cut-off date</th>
<th>End of survey</th>
<th>GDP data</th>
<th>Actual value relative to earlier forecasts:</th>
<th>CPI data</th>
<th>Actual value relative to earlier forecasts:</th>
<th>NBP interest rate change:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011q4</td>
<td>2011-09-28</td>
<td>2011-12-30</td>
<td>for 2011q3</td>
<td>higher</td>
<td>for 2011m09, m10, m11</td>
<td>higher</td>
<td></td>
</tr>
<tr>
<td>2012q1</td>
<td>2012-01-24</td>
<td>2012-03-31</td>
<td>for 2011q4</td>
<td>higher</td>
<td>for 2012m01, m02</td>
<td>about the same</td>
<td></td>
</tr>
<tr>
<td>2012q2</td>
<td>2012-06-15</td>
<td>2012-06-29</td>
<td>for 2012q3</td>
<td>about the same</td>
<td>for 2012m10, m11</td>
<td>lower</td>
<td>-50 bp</td>
</tr>
<tr>
<td>2013q1</td>
<td>2013-02-14</td>
<td>2013-04-03</td>
<td>for 2013q3</td>
<td>about the same</td>
<td>for 2013m01, m02</td>
<td>lower</td>
<td>-50 bp</td>
</tr>
<tr>
<td>2013q2</td>
<td>2013-06-13</td>
<td>2013-07-02</td>
<td>for 2013q3</td>
<td>about the same</td>
<td>for 2013m10, m11</td>
<td>lower</td>
<td></td>
</tr>
<tr>
<td>2013q4</td>
<td>2013-10-21</td>
<td>2014-01-07</td>
<td>for 2013q3</td>
<td>about the same</td>
<td>for 2014m02</td>
<td>about the same</td>
<td></td>
</tr>
<tr>
<td>2014q1</td>
<td>2014-02-14</td>
<td>2014-04-02</td>
<td>for 2014q3</td>
<td>higher</td>
<td>for 2014m10, m11</td>
<td>lower</td>
<td></td>
</tr>
<tr>
<td>2014q2</td>
<td>2014-06-13</td>
<td>2014-07-02</td>
<td>for 2014q3</td>
<td>higher</td>
<td>for 2014m10, m11</td>
<td>lower</td>
<td></td>
</tr>
<tr>
<td>2014q4</td>
<td>2014-10-20</td>
<td>2015-01-06</td>
<td>for 2014q3</td>
<td>higher</td>
<td>for 2014m10, m11</td>
<td>lower</td>
<td></td>
</tr>
</tbody>
</table>

Source: own calculations based on NBP, GUS and Reuters Thomson data.
Notes: Assessment of direction of forecast error is based on Reuters Thomson survey among financial market analysts (median of individual responses treated as consensus).

2.3.2. Interest rate assumption

Additional complication in comparing the NBP SPF forecasts with the NBP projections refers to interest rate path in the forecasted period, as the central bank projections are prepared under constant interest rate assumption. It is a technical assumption, chosen by many central banks mainly because it is easy to communicate (other interest rate paths might be misunderstood as a commitment) and solves some organizational problems (like reaching consensus on future interest rate development by all members of the monetary authority). The shortcoming of projections based on constant interest rate is that they might be not reliable and less accurate than forecasts based on most probable interest rate path7.

On the contrary, SPF experts present unconditional forecasts of inflation and GDP growth; they report also probabilistic forecasts of interest rate. As the level of interest rates affects inflation and economic activity (with lag), differences between expected interest rate path by experts and assumed in the NBP projections might lead to

7 Alternatively, market or monetary authorities expectations of future interest rate path might be employed. For arguments pro and against various interest rate assumptions in forecasting by central banks see, e.g. Goodhart (2009), Knuppel and Schulte-Frankenfeld (2013).
discrepancies between compared forecasts. Table 3 summarizes level of interest rate assumed in NBP projection and interest rate predicted by SPF experts. The greatest differences refer to forecasts in 2012q4 and 2013q1 and they amount to more than 100 bp and about 50 bp, respectively. In these two quarters we might expect differences in the assessed forecasts resulting from assumed different developments of interest rate. In other periods distinct assumptions about interest rate path should not disturb comparability of forecasts.

Additional argument that constant interest rate assumption should not substantially distort our analysis is provided by Knuppel and Schultefrankenfeld (2013), who show that the underlying interest rate assumption does not affect in practice accuracy of inflation and output forecasts.

<table>
<thead>
<tr>
<th>Survey quarter</th>
<th>Interest rate in NBP projection</th>
<th>Implied interest rate predicted by SPF experts</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011q4</td>
<td>4.50</td>
<td>4.27</td>
</tr>
<tr>
<td>2012q1</td>
<td>4.50</td>
<td>4.50</td>
</tr>
<tr>
<td>2012q2</td>
<td>4.75</td>
<td>4.83</td>
</tr>
<tr>
<td>2012q4</td>
<td>4.75</td>
<td>3.61</td>
</tr>
<tr>
<td>2013q1</td>
<td>3.75</td>
<td>3.22</td>
</tr>
<tr>
<td>2013q2</td>
<td>2.75</td>
<td>2.41</td>
</tr>
<tr>
<td>2013q4</td>
<td>2.50</td>
<td>2.56</td>
</tr>
<tr>
<td>2014q1</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>2014q2</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>2014q4</td>
<td>2.00</td>
<td>1.89</td>
</tr>
</tbody>
</table>

Source: own calculations based on NBP data.
Notes: In the case of forecasts prepared in the 4th quarter, interest rate predicted by SPF experts refers to median of aggregated distribution of forecasts for the next year, while in the case of forecasts prepared in the 1st and 2nd quarters it is implied predicted interest rate for the rest of the current year (based on median of aggregated distribution of forecasts for the current year).

2.3.4. Predictive distributions

Additional problems emerge while applying similarity measure for probability density functions in order to compare probabilistic predictions (see section 4 for description of this measure). These problems stem from the fact that the compared distributions belong to different families. In order to eliminate this effect on similarity/divergence measure, all analyzed probability distributions (pdfs) are transformed by a triangular approximation (Johnson, 2002; details are given in Kowalczyk et al., 2013). The parameters of triangular pdfs are determined so as not to change the medians and 5th and 10th percentiles of original probabilistic forecasts. Choosing triangular distributions seems quite natural – it is equivalent to the assumption that reading the published forecasts, recipients focus attention mainly on the most probable scenario and on the range of possible values and that probability decreases linearly with distance from the main scenario.
3 Comparing performance of probabilistic forecasts

3.1. Comment on methods

Comparing probabilistic forecasts is a very complex issue. It is difficult to point a method that fully takes into account all aspects of such forecasts and which could be deemed as entirely adequate and complete. Generally, there are two groups of methods for evaluating performance of probabilistic forecasts.

The first one applies so called scores which are functions of the form \( s(f(x), x_0) \), where \( f(x) \) is a probability distribution (pdf or cdf) and \( x_0 \) is the observed value. The score informs how far the observation is from the predicted distribution, or how the predicted distribution is spread around the observation. The widely used scores of this kind are the Brier score, the ranked probability score (RPS) and the logarithmic score.⁸

In our analysis we employ the RPS in a histogram context (Boero et al., 2011), which means that each probabilistic forecasts is represented by the vector of bin probabilities \( \mathbf{p}(x) = [p_1, p_2, ..., p_k] \), and realization by the vector \( \mathbf{r}(x0) = [r_1, r_2, ..., r_k] \), where \( r_i = 1 \) if the outcome \( x_0 \) falls in bin \( i \), and \( r_i = 0 \) otherwise.

The score for prediction \( \mathbf{p}(x) \) is defined as

\[
s(F_p, x_0) = \sum_{k=1}^{k} (F_p(k) - F_r(k))^2 ,
\]

where \( F_p(k) \) and \( F_r(k) \) are cdfs for bin \( k \) calculated on the basis of bin probabilities:

\[
F_p(k) = \sum_{i=1}^{k} p_i \quad \text{and} \quad F_r(k) = \sum_{i=1}^{k} r_i .
\]

The RPS is used to evaluate a set of \( N \) forecasts \( \{p^j\}_{j=1...N} \) (delivered by one forecaster) on the basis of forecast-outcome pairs \( [(F_{p^j}^i, x_{0i})] \) and it is equal to average score over these pairs:

\[
RPS = \frac{1}{N}\sum_{j=1}^{N} s(F_{p^j}, x_{0j}) .
\]

The second group of methods applied in evaluating forecasts performance refers to their calibration. Assessment of calibration involves checking the statistical consistency of the predicted probabilities with observed frequencies. A very popular probability integral transform (PIT) belongs to this category of methods (Dawid, 1984; Diebold et al., 1998).

However, calculating calibration does not suffice to evaluate forecasts, as pointed out by Gneiting et al. (2007), who formulated the “paradigm of maximizing the sharpness of the predictive distribution subject to calibration”. The

---

⁸ See Boero et al. (2011) for discussion of these scores in application to the Bank of England’s Survey of External Forecasters.
Comparing performance of probabilistic forecasts

second criterion is sharpness, i.e. concentration of predictive distributions. Good forecasting is characterized by good calibration and high concentration of predictive distributions. More concentrated distribution is more informative. Having two forecasts with the same calibration, the sharper one is better. Cooke (1991) introduced a combined score – the product of calibration score and information score, under condition that the calibration satisfies certain minimum. These scores in the context of application to the NBP SPF were discussed in Kowalcyzk (2010) and Kowalcyzk et al. (2013). Unfortunately, our sample is not sufficient to calculate such combined scores, but we will use the ideas underlying the method of evaluating calibration proposed by Cooke.

Recall that the SPF participants provide the values of three percentiles – the 5th, 50th and 95th. These percentiles divide a range of possible values into four intervals with probabilities 0.05, 0.45, 0.45, 0.05. To assess expert’s calibration, we inspect how often the observed values fall into each inter-quantile interval and create empirical probability distribution \( f_i = [r_1, r_2, r_3, r_4] \), where \( r_i \) is the number of outcomes in the i-th interval. The expert is perfectly calibrated if 5% of the outcomes fall into his first inter-quantile intervals (below the 5th percentile), 45% into the second (between the 5th and the 50th percentile), 45% into the third (between the 50th and the 95th percentile) and 5% to the fourth (above the 95th percentile). Such a perfect calibration corresponds to the theoretical distribution \( f_i = [0.05, 0.45, 0.45, 0.05] \), shown as a benchmark in Figure 9. The lower is the distance, between the expert’s empirical distribution and the theoretical distribution, the better is calibration. As a measure of calibration we use the Kullback-Leibler divergence:

\[
l(f_e, f_i) = \sum_{i=1}^{4} r_i \ln \left( \frac{r_i}{f_i} \right), \quad \text{where} \quad p_1 = 0.05, \quad p_2 = 0.45, p_3 = 0.45, p_4 = 0.05 ,
\]

As mentioned earlier, calibration without taking into account informativeness is not a sufficient to infer about forecast performance. To measure informativeness of a set of expert’s forecasts we use average inter-quantile range of expert’s forecasts (i.e. the difference between the 5th and the 95th percentiles).

3.2 Results

We confront sets of forecasts produced by individual experts, the aggregate SPF forecasts\(^9\) and projections of the NBP, all referring to inflation and GDP growth in +4q. Describing results of our analysis, we focus in the first place on differences between forecasters, while the time dimension of our data is rarely exploited. If not specified differently, the analysis is limited to 18 experts, who participated in at least half of the surveys paired with the NBP projection and joined the SPF before 2013. Additionally, for the purpose of assessing accuracy of forecasts, only survey rounds when the NBP projections were available were employed.

\(^9\) The aggregated forecast is equal to a mixture of individual probabilistic distributions with equal weights. Method of aggregating probabilistic forecasts is described in Kowalcyzk et al. (2013).
The aggregated probabilistic SPF inflation forecasts perform slightly worse, in terms of RPS, than the NBP projections. RPS of the aggregated SPF forecast is equal to 1.89, while of the NBP projection – to 1.54 (Figure 2). However, it is characterized by lower RPS than any individual expert. The accuracy of individual experts’ forecasts is very diverse.

Also in the case of GDP growth the NBP projections outperform slightly the aggregated SPF. Performance of some individual experts is also good: their scores are very similar to or even lower than for the NBP projections\(^6\).

![Figure 2: Ranked probability score of probabilistic forecasts](image)

Source: own calculations based on NBP and GUS data.
Notes: Sample limited to quarters with available NBP projections.

We could see in Figure 2 that the aggregated SPF forecasts of 4-quarter-ahead inflation are characterized by similar RPS as forecasts of 4-quarter-ahead GDP growth. The same applies to the NBP projections. However, majority of individual experts make smaller mistakes (in probabilistic terms) in the case of GDP growth than inflation (Figure 3).

It suggests that experts have found it more difficult to forecast inflation than economic activity. In order to check whether it holds also for different forecast horizons, we have calculated the RPS also for current year forecasts (Figure 3). It turns out that in the case of this horizon, the RPS of inflation forecasts is smaller than the RPS of GDP growth forecast. So, experts find it difficult to forecast inflation in the 4-quarter horizon, but not in the shorter one. We will return to this problem while discussing calibration.

\(^6\) The very good performance of two experts results partially from the smaller number of surveys they participated, but even if we accounted for this effect by adjusting the sample, their RPS would be smaller than RPS of NBP projection.
Comparing performance of probabilistic forecasts

Figure 3 Ranked probability scores of individual experts’ forecasts of inflation and GDP growth

![Graph showing ranked probability scores for inflation and GDP growth forecasts.]

Source: own calculations based on NBP and GUS data.
Notes: All survey rounds.

Figure 4 (upper left panel) reveals poor calibration of the 4-quarter-ahead inflation forecasts. Very large proportion of realizations of inflation fell into the first inter-quantile intervals, below the 5th percentiles indicated by experts, and very small proportion of realizations fell inside the 90% probability interval (inside the second or the third inter-quantile intervals). This indicates that all forecasters declared biased forecasts. The empirical distribution for the best calibration (expert number 16) – when it comes to +4q inflation forecasts – is the following: [0.29, 0.57, 0.14, 0], meaning that: in 29% of the cases the realized inflation was below 5th percentiles given by the expert, in 57% of cases between 5th and 45th percentiles, in 14% of cases they fell between 45th and 95th percentiles, while no inflation outcome exceeded the 95th percentiles of his/her probabilistic forecasts. The performance of the aggregated SPF forecast was not superior to individual forecasts as well. Calibration of the NBP projection over the analyzed period was slightly better than in the case of individual or aggregated SPF forecasts, but even these forecasts overestimated future CPI inflation. These common errors of all forecasts were associated mainly with economic shocks to food and energy prices in years 2013-2014.

GDP growth forecasts 4-quarters-ahead are characterized by much better calibration than inflation forecast (Figure 4, right panel). In the case of about half of experts, a significant share of realizations of the forecasted variable were placed inside the declared 90% probability intervals. For the others, the declared intervals were too narrow. The statistical consistency of the aggregated SPF forecasts with observations was much better. In case of NBP projections all realizations fell inside 90% probability interval which might indicate overestimation of uncertainty.

It is interesting to notice that experts’ short-term inflation forecasts (current year), are much better calibrated than in the 4-quarter horizon (Figure 4, lower left panel). This might be linked to forecasts anchoring to the target in +4q horizon or wrong assessment of durability of shocks. In the case of GDP growth, calibration for 4-quarter and one-year horizon forecasts are more similar.
Figure 4 Empirical distributions for inflation and GDP growth forecasts

![Graphs showing empirical distributions for inflation and GDP growth forecasts.](image)

Source: own calculations based on NBP and GUS data.
Notes: Only projection quarters.

Figure 5 shows the calibration (Kullback-Leibler distance between empirical distribution and the theoretical distribution corresponding to the ideal calibration) together with informativeness, measured by mean inter-quantile range (between 95th and 5th percentiles) of individual forecasts, aggregated forecast and the NBP projection. It confirms poor calibration of inflation forecasts declared by individual experts, which can’t be explained only by their high informativeness – poor calibration characterizes forecasts with both very low and relatively large inter-quantile range. Also the aggregated SPF forecasts, despite higher inter-quantile range than all but one individual expert’s forecasts, can’t be considered as well calibrated. The problem is an overestimation of future inflation in this forecast horizon.

When it comes to the GDP growth forecasts (4 quarters ahead), there are two distinct groups of experts: one characterized by good calibration, but very low informativeness (NBP projections and aggregated SPF forecasts are among them), and the second group with much better informativeness but worse calibration. Poorer calibration in the latter group is caused by too narrow 90% probability intervals and not by the bias as in the case of inflation. Only few experts fulfill both criteria regarding calibration and informativeness.
Comparing performance of probabilistic forecasts

Figure 5 Calibration and inter-quantile range (informativeness) of inflation and GDP forecasts

Source: own calculations based on NBP and GUS data.
Notes: sample limited to projection quarters.
4 Assessing similarity of distributions

4.1 Comment on measurement

Our goal is to assess whether the NBP projections are reflected in forecasts provided by the participants of the NBP SPF – in other words, how far experts’ opinions are from the view of the NBP staff. Since both, the projections and experts’ forecasts, fall into the category of density forecasts, limiting the analysis to comparison of the expected values, variances, etc. seems insufficient.

We were looking for measures which allow comparing the probability distribution functions (pdfs) as a whole, not only parameters of the distributions, and therefore answering the question how similar the pdf of experts’ forecasts and pdf corresponding to the NBP fan chart are. Similarly as in the previous section, we represent pdfs as histograms, which allows applying numerous similarity/distance measures for discrete distributions (Cha, 2007).

If we consider histogram as a vector, several metrics for Euclidian space might be employed. For instance, for two probability distributions $p = [p_1,\ldots,p_n]$ and $q = [q_1,\ldots,q_n]$, where $p_i$ and $q_i$ are probabilities of the $i$-th bin, we have:

- **Euclidean L2 distance:**
  $$L_2(p,q) = \sqrt{\sum (p_i - q_i)^2}$$

- **Minkowski distance:**
  $$L_{\text{Mk}}(p,q) = \sqrt[k]{\sum |p_i - q_i|^k}$$

- **City block distance:**
  $$L_{\text{CB}}(p,q) = \sum |p_i - q_i|$$

- **Chebyshev distance:**
  $$L_{\text{CN}}(p,q) = \max |p_i - q_i|$$

The other possible approach, when pdfs are represented as histograms, is applying discrete versions of information-theoretic divergences, such as the Kullback-Leibler (KL) divergence, called also a relative entropy or information deviation, the Jeffreys divergence ($L$-divergence), K-divergence or the Jensen-Shanon (JS) divergence:

- **Kullback-Leibler divergence:**
  $$KL(p,q) = \sum p_i \log \frac{p_i}{q_i}$$

- **Jeffreys divergence:**
  $$J(p,q) = \sum (p_i - q_i) \log \frac{p_i}{q_i}$$

- **K-divergence:**
  $$K(p,q) = \sum p_i \log \frac{2p_i}{p_i + q_i}$$

- **Jensen-Shanon divergence:**
  $$JS(p,q) = \frac{1}{2} \left( \sum p_i \log \frac{2p_i}{p_i + q_i} + \sum q_i \log \frac{2q_i}{p_i + q_i} \right)$$
Assessing similarity of distributions

The Jensen-Shannon divergence is particularly attractive for our application for several reasons. Firstly, it is closely related to the KL divergence – the best-known information-theoretic distance – as it is the arithmetic average of two KL divergences:

\[ JS(p, q) = \frac{1}{2}KL(p, \frac{1}{2}(p + q)) + \frac{1}{2}KL(q, \frac{1}{2}(p + q)) \]

where \( \frac{1}{2}(p + q) \) is the mixture of distributions \( p \) and \( q \).

Suppose that the distribution \( p \) results from the fan chart and \( q \) is the subjective distribution provided by the NBP SPF expert. If the expert has adopted the NBP view entirely, then his/her distribution \( q \) is equal to \( p \) and both are equal to the mixture \( \frac{1}{2}(p + q) \), therefore the \( JS(p, q) = 0 \). In a more realistic situation, knowing how much the pooled distribution differs from experts’ forecast distribution and the NBP projection distribution, allows inferring about a potential impact of the projection on the forecast provided by the expert.

Secondly, in contrast to the Kullback-Leibler, the Jensen-Shannon distance is symmetric and, what is very important for our application, it does not require that for all bins (denoted by \( i \)) \( q_i \neq 0 \) when \( p_i \neq 0 \).\(^{11}\)

4.2 Results

4.2.1 Overview

Figure 6 and Figure 7 (panel A) show great heterogeneity of average distances of experts’ probabilistic forecasts from the NBP inflation and GDP growth projections. Two experts stand out from the rest of participants by formulating forecasts that are very close to NBP projections in all survey rounds. These two experts declare point forecasts closer to NBP projections (of inflation and GDP growth) than an average expert and declare the widest range of possible values (IQRs). At the same time, their inflation forecasts are one of the most accurate (in terms of mean absolute error and RPS), while the accuracy of their GDP growth forecasts is close to an average expert. Also the aggregated probabilistic forecast is quite close to the NBP projections.

\(^{11}\) The is no problem of division by zero: 0/0 is treated as 0 and similarly 0*log0=0.
Figure 6 Characteristics of probabilistic inflation forecasts

Source: own calculations based on NBP and GUS data.
Notes: SPF denotes aggregated forecast, NBP denotes central bank projection, consecutive numbers denote individual experts.

Figure 7 Characteristics of probabilistic GDP growth forecasts

Source: own calculations based on NBP and GUS data.
Notes: SPF denotes aggregated forecast, NBP denotes central bank projection, consecutive numbers denote individual experts.
4.2.2 Role of uncertainty assessment for similarity

Differences in forecast distributions between experts and NBP projections, measured by the Jensen-Shannon distance (JS hereafter), stem mainly from different assessment of uncertainty. In Table 4 we present correlation coefficients between the JS and differences in central forecasts, and between the JS and differences in IQRs for the whole sample (on individual data) and for three subgroups of experts’ forecasts: with low IQRs (below 0.33 centile), medium level of IQR (between 0.33 and 0.66 centile) and with high IQRs (above 0.66 centile). Differences in JS go in line with differences in IQRs, while the impact of differences in central forecasts gain in importance while comparing while comparing forecasts with similar uncertainty assessment.

Table 4. Spearman rank correlation coefficients between JS distance and differences in central forecasts and difference in IQRs (between forecaster and NBP projection)

<table>
<thead>
<tr>
<th></th>
<th>Whole sample</th>
<th>IQR&lt;p(0.33)</th>
<th>p(0.33)</th>
<th>IQR&gt;p(0.66)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inflation forecasts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference in central</td>
<td>0.13</td>
<td>0.21*</td>
<td>0.38***</td>
<td>0.73***</td>
</tr>
<tr>
<td>forecasts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference in IQRs</td>
<td>0.94***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>GDP growth forecasts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference in central</td>
<td>0.01</td>
<td>0.49***</td>
<td>0.15</td>
<td>0.46***</td>
</tr>
<tr>
<td>forecasts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference in IQRs</td>
<td>0.96***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Source: own calculations based on NBP data. Notes: Individual data. */**/*** denotes statistical significance at 10%/5%/1% level.

Differences in uncertainty assessment between experts’ forecasts and the NBP projections are quite strong. Among survey participants included in this analysis, only one expert declared 90% probability intervals (for both forecasted variables) of comparable width to NBP fan charts – (panels B in Figure 6 and Figure 7). IQRs of other experts’ forecasts were much narrower than IQRs of the NBP projections in virtually all cases. The dissimilarity of views regarding uncertainty might be explained by the fact that the inter-quantile ranges of these forecasts describe different kinds of risks. The experts express their subjective uncertainty, based on scenario analysis, while the large part of the NBP projection fan chart corresponds to uncertainty related to past forecast errors. This explains also why forecast uncertainty of the NBP projections are quite stable over time.

The 90% probability interval of the aggregated probabilistic forecast was closer to the bands of NBP fan charts, because it is affected not only by uncertainty assessment by individual experts, but also by disagreement among them. However, even in this case, it was narrower than 90% probability interval of the NBP projection.
NBP projections suggest greater uncertainty of GDP forecasts than inflation forecasts. This relation is preserved only in the case of forecasts of few individual experts – other experts indicate ranges of possible values of similar width for both variables under consideration (Figure 9).

**Figure 8** Uncertainty of inflation and GDP growth forecasts in subsequent NBP SPF rounds

<table>
<thead>
<tr>
<th>Source: own calculations based on NBP data.</th>
</tr>
</thead>
</table>

Similarity of experts point forecasts to the NBP projections does not correspond to similarity of uncertainty assessment. Some experts whose views on central paths of inflation and GDP are similar to NBP staff’s views, do not necessarily share the bank’s opinion about uncertainty, and the other way round, those experts whose IQRs are closer to the fan chart intervals, might differ more than the other experts in terms of central projections. This outcome motivates the need for comparing the whole forecast distributions instead of comparing point forecasts, as the NBP projections might be incorporated into experts forecasts in various ways.

### 4.2.3. Similarity of inflation forecasts vs. similarity of GDP growth forecasts

The other observation related to a comparison of forecast distributions, is that similarity of NBP SPF experts’ inflation forecasts to the NBP projections goes in line with similarity of experts’ GDP growth forecasts to the NBP projections. Figure 9 plots experts’ average JS distances from the NBP inflation and GDP growth forecasts. They are highly correlated (Spearman rank correlation coefficient equals 0.95 and is highly statistically significant). Interestingly, for almost all experts inflation forecasts are on average closer to the NBP projections than GDP growth forecasts.
Assessing similarity of distributions

Figure 9 Similarity of inflation forecasts vs similarity of GDP growth forecasts

Source: own calculations based on NBP and GUS data.
Notes: Average JS distances across experts. Filtered panel.

4.2.4. Similarity to NBP projections and forecast accuracy

Some researches argue that one reason for taking into consideration central bank forecasts by professional forecasters is that central bank forecasts might be more accurate (see eg. Hubert, 2013). Therefore, it is interesting to find out whether incorporating the central bank view pays off in terms of smaller RPS. It turns out that the relationship between similarity of experts’ inflation forecasts to NBP projections and accuracy of inflation forecasts is not statistically significant. The opposite is true for the GDP growth forecasts: there is a positive and statistically significant correlation between accuracy (measured by RPS) and similarity\textsuperscript{12}.

Table 5 Correlation coefficients (Spearman) between ranked probability score and average distance of experts’ forecast from NBP projection

<table>
<thead>
<tr>
<th>RPS of inflation/GDP growth forecasts</th>
<th>Average JS distance of experts’ forecasts from NBP projection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inflation</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
</tr>
</tbody>
</table>

Source: own calculations based on NBP SPF.

\textsuperscript{12} This result holds even if we remove from the sample the 2 forecasters with the exceptionally low RPS.
4.2.5. Development of JS divergence measure over time

So far we have discussed differences among experts in the context of similarity of their forecasts to the NBP projections. Now we turn to analysis of developments of this distance over time. Figure 11 shows average JS distance between experts’ forecasts and the NBP projection.

In the case of inflation, the average distance of experts’ forecasts from the central bank projections was quite stable over time (about 0.2), except 2013q2, 2014q2 and 2014q4, when it took the highest values. The average distance for GDP growth forecasts was about 0.25, with the lowest values in 2011q4 and 2012q2, and the highest value in 2014q4.

We might think about three sources of divergence between these forecasts: differences in information sets of SPF experts and NBP staff resulting from different timing of forecasts, availability of NBP projection prior to SPF round and differences in mechanism of expectations formation.

Taking under consideration differences in timing of SPF and NBP projections, we would expect the greatest differences in forecasts made in the 4th quarters and the smallest differences in forecasts made in the 2nd quarters (see section 2.3). However, this intuition is not confirmed. Additionally, greater differences should be visible in 2012q4 and 2013q1 when the differences in actual and predicted interest rates were the highest. Indeed, the distance between forecasts in these quarters is slightly higher than in the remaining quarters under consideration.

15 Apart from differences in interest assumption, which should not play a significant role as argued in section 2. The effect of interest rate change between projection cut-off date and survey date is taken under consideration while comparing information sets.
NBP forecasts also constitute part of experts’ information set. Assuming that the NBP forecasts are influential, we would expect higher JS distance in 2nd quarters. Indeed, aggregated inflation forecast are relatively further from the central bank projection in 2nd quarters, but it’s not true for GDP growth forecasts. Of course this observation is not sufficient to claim that SPF experts are under influence of the central bank inflation forecasts, but at least is not contradictory with this statement.

The last explanation of differences between forecasts refers to the process of expectations formation. Interestingly, large differences in the inflation forecasts (2013q2 and 2014q2) correspond to periods when inflation projections were very low – they dropped below the lower limit of the inflation target range. As showed in Appendix, the experts’ forecasts are affected by the NBP inflation target and usually are closer to it than the NBP projections. Also probability of inflation in target range according to experts’ forecasts is higher than for NBP predictions (Figure 6, panel D). So part of the differences in inflation predictions of NBP staff and SPF experts might be contributed to anchoring property of the latter. The largest difference between forecasts (of inflation and GDP growth), measured by JS distance, appeared in the 2014q4. In this quarter experts’ forecasts moved down below the NBP projection and further from the NBP inflation target, which might be a first sign of de-anchoring of expectations.
5 Conclusions

In this study we attempted to compare forecasts produced by the staff of the Polish central bank, published in the Inflation Reports, with experts’ probabilistic forecasts collected in the NBP Survey of Professional Forecasters. This task turned out difficult due to several problems mentioned in the text.

We were interested not only in accuracy of probabilistic forecasts and their calibration, but mainly in finding out whether experts participating in the SPF mirror the NBP projections, an important issue in studying expectations. In the paper, we proposed employing methods based on distance between probability density functions, which seem to be useful in analyzing coincidence of probabilistic forecasts. Among many measures of this kind, these originating from the information theory seem to us the most adequate for this purpose. However, the most popular distribution similarity measure, the Kullback-Leibler distance, turned out to be impossible to apply, due to violation of assumption about domains of the distributions under consideration. Therefore, we have chosen the Jensen-Shannon measure, which is free of this limitation, and, additionally, has advantage of symmetry. We have applied the Jensen-Shannon distance to compare the NBP projections with NBP SPF forecasts (on individual and aggregated level).

Due to substantial heterogeneity of forecasts, as well as short (3.5 years) and atypical period taken under consideration (changes in inflation were dynamic and difficult to predict) it is not possible to draw firm conclusions based on statistical analysis. Operating on averages across experts might lead to mistakes, similarly as treating the aggregated distributions as consensus. On the other hand, analyzing individual data does not make it possible to draw general conclusions.

Having in mind these limitations, the following observations might be made. In the period under consideration, the NBP inflation projections for 4 quarters ahead outperform forecasts of SPF experts (on both individual and aggregated level) in terms of RPS and calibration which measure how good predicted distribution are in matching the observed values. In the case of GDP, about 20% of experts had better statistical performance. It seems that experts’ difficulties in forecasting inflation refer only to 4-quarter horizon - forecasts for the current year are much better. Two explanations are possible: NBP SPF experts underestimated persistence of shocks lowering inflation or their forecasts (for +4 quarters) are under strong influence of the central bank inflation target. Substantial heterogeneity of experts’ forecasts is revealed also when we look at their similarity to NBP projections. The distance between SPF forecasts and NBP projections, in terms of distributions, is driven mainly by different assessment of uncertainty. Similarity of experts’ inflation forecasts to the NBP projections goes in line with similarity of experts’ GDP growth forecasts to the NBP projections.


References


Pedersen M., 2013, What Affects the Predictions of Private Forecasters? The Role of Central Bank Forecasts, Documentos de Trabajo, No 686, Banco Central De Chile.

Appendix 1 Comparison of point forecasts

Similarity of central forecasts

The median of SPF forecasts is quite close to the NBP projection: the mean absolute difference amounts to 0.35 pp. for inflation (Figure 12, panel A) and 0.27 pp. for GDP growth (Figure 13, panel A). However, if we look at individual forecasters’ performance, the figures reveal significant heterogeneity with this respect. The mean absolute difference statistics calculated for individual experts ranges from 0.30 to 0.85 pp. for inflation (with average equal to 0.50 pp.) and from 0.22 to 0.81 pp. for GDP growth (with average equal to 0.47 pp.). The distance from NBP inflation forecasts does not go in with distance from NBP output forecast: the correlation coefficient is positive (0.4), but statistically significant only at 10% significant level. This could mean that even if SPF experts taken into account the NBP projections, they treated inflation and GDP forecasts differently.

Figure 12 Characteristics of experts’ central forecasts – inflation

![Graphs A, B, C, and D](image)

Source: own calculations based on NBP data.
Notes: SPF denotes median of individual experts’ central forecasts, NBP denotes central bank projections (central path), consecutive numbers denote individual experts. In all graphs, except B, sample covered only quarters with available NBF projection.

---

14 Annex 1 includes figures showing development of differences in these forecasts over time for each expert. Also tables with detailed descriptive statistics for individual experts are relegated outside the main text (Appendix 2).
As far as the direction of deviations of experts’ central forecasts from the NBP projections is concerned, only 4 experts (out of 18) formulated inflation forecasts on average below the NBP projection, while the rest of them tended to deliver higher forecasts. The pattern of deviations has changed over time. At the beginning of the analyzed period – when the current inflation and the NBP projection were above the NBP inflation target – experts’ forecasts were on average equal to the NBP central projections, while from 2013q1 to 2014q2 – when the current inflation and the NBP inflation projection deviated from the target downwards – experts’ forecasts tended to be higher than NBP projections Figure 14. It suggests that forecasts of SPF experts are, at least to some extent, anchored to the NBP inflation target. On this background, forecasts from the last quarter under consideration (2014q4) are distinctive, as virtually all experts’ forecasts placed below the NBP central path and further from the 2.5% the central bank projection. It might be interpreted as a first sign of deanchoring of inflation expectations in this horizon. Another explanation is that in October and November 2014 CPI inflation figures published after the projection cut-off date, but before SPF deadline were surprisingly low, which could also contribute to such low experts’ forecasts in this quarter.

When it comes to the GDP growth projections, the largest differences between the experts and the central bank projections appeared in 2011-2012, when in the environment of slowing down (but still relatively high) economic
growth, the SPF experts expected stronger GDP growth in +4q than the central bank (Figure 14). In the remaining part of the sample period, average forecasts of the experts were similar to the central bank predictions.

Figure 14 NBP forecast and SPF experts’ forecasts – development over time

Source: own calculations based on NBP data.
Notes: Full panel.

Anchoring of inflation expectations

Łyziak (2012), using a different survey data and a consensus measure, showed that inflation forecasts of economists are under strong influence of the NBP inflation target. This observation seems to be confirmed also if individual forecasts from the NBP SPF are employed. Central forecasts of individual experts in +4 quarters horizon typically are closer to the targeted 2.5% value than the NBP central forecasts: all but 3 forecasters from the filtered panel declared inflation forecasts closer on average to the NBP target than central bank projections (in absolute terms) – see Figure 12,
panel D. However, there is large heterogeneity of experts in terms of an average absolute deviation from the target (it ranges from 0.24 pp. to 0.87 pp.)\textsuperscript{13}.

The anchoring of expectations is also visible if we look at deviations from the target in each survey round. Figure 15 shows deviations of forecasts from the inflation target for individual experts against deviations of NBP projections from this target, together with 45 degrees line. In majority of cases, experts’ forecasts were closer to the 2.5% value than NBP projections – especially when magnitude of projection’s deviation was large. As mentioned previously, the last quarter (2014q4) is distinctive, as virtually all experts’ forecasts were further from the 2.5% than the NBP projection.

**Figure 15 Distance from inflation target**

![Figure 15](image)

Source: own calculations based on NBP data.

**Point forecasts accuracy**

One reason for taking into consideration central bank forecasts by professional forecasters is that central bank forecasts are more accurate\textsuperscript{14} (see eg. Hubert, 2013). Therefore we might ask whether formulating central forecasts similar to the NBP central projection pays off in terms of smaller forecast errors. In the case of inflation, mean absolute forecast error of the central bank projection was lower than the error of average expert by 0.27 pp. (MAEs were equal to, respectively, 1.71 and 1.44 pp.) and lower than all individual experts forecasts errors but two. When it comes to GDP growth forecasts, MAE of average forecaster is quite close to the MEA of the NBP projection (equal, respectively,\textsuperscript{15}.

\textsuperscript{13} See Figure 18 in Appendix 1 for panel plots of deviations from the target and Appendix 2 for table with statistics.

\textsuperscript{14} The other reason is that central bank forecast might reveal information about future decisions and/or preferences of monetary authorities.
to 1.18 and 1.09 pp.). If we look at MAE of individual forecasts, 5 experts formulated more accurate point forecasts than NBP projection7 (Figure 12 and Figure 13, panels C).

Analysis of correlations shows that greater deviations of individual experts’ forecasts from the NBP projections are associated with greater forecast errors, for both variables: inflation and GDP growth (Table 7). The positive correlation between the absolute difference of experts’ inflation forecast from the NBP projection and size of forecast error is confirmed if, instead of individual data, we consider mean characteristics of experts. It means that those experts, who tend to declare forecasts closer to the NBP projection, also are inclined to formulate more accurate forecasts. In the case of the GDP growth forecast, there is no such relationship on the level of these mean characteristics. This result is driven by two forecasters (number 17 and 18), who joined the SPF later than the others, in the period when forecast errors were smaller for all forecasters – if we exclude them from the sample the correlation coefficient equals 0.62 (p-val.<0.01).

Table 6 Correlation coefficients (Pearson) between mean absolute forecast errors and deviations of central forecasts from NBP projections

<table>
<thead>
<tr>
<th>Absolute error of experts’ forecasts</th>
<th>Average across experts, filtered panel</th>
<th>Individual data, filtered panel</th>
<th>Individual data, full panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute difference from NBP projection</td>
<td>inflation</td>
<td>0.71***</td>
<td>0.25***</td>
</tr>
<tr>
<td>GDP growth</td>
<td>-0.07</td>
<td>0.51***</td>
<td>0.52***</td>
</tr>
</tbody>
</table>

Source: own calculations based on NBP SPF.
Notes: *** denotes significance at 10%/5%/1%.

7 This result is robust even if we compare experts’ predictions with NBP projections pairwise to take into account the fact that some experts joined the survey later.