RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS No. 91
Econometrics 28 2010

**Maria M. Kaźmierska-Zatoń, Wojciech Zatoń**
University of Łódź

# MULTI-CRITERIA COMBINED FORECASTS

**Abstract:** Using combination of forecasts results from the underlying belief that there is no such a single model that could comprise the whole actual and often complex economic process. The aim of this paper is an attempt to look at combining forecasts as a multi-criteria decision-making process. It is a multi-criteria process as we need to select: (a) individual component forecasts (quantity, quality, type), (b) values of their weights (objective determination of weights or their arbitrary choice), (c) a mathematical representation of combination and (d) a measure to assess the quality of a combined forecast. In the paper we provide an overview of methods applied for combining forecasts and show the results of our study of a combined forecast for inflation in Poland. The analysis indicates that combining forecasts is fully justified, especially if weights for individual forecasts are determined on the basis of multi-criteria optimization. The results also show that combining forecasts is useful even if a dominant individual forecast exists.

**Key words:** forecasts, combining, weights, criteria, optimization.

## 1. Introduction

It is not an easy task to forecast various types of phenomena, not only economic ones, even when we can choose from a broad scope of methods. These methods often generate different forecasts and it is difficult to assess them and select the best one. Therefore it is often advisable to apply combined forecasts which are usually an average of component forecasts derived from different methods or developed by different authors. Using combination of forecasts results from the underlying belief that there is no such a single model that could comprise the whole actual and often complex economic process. Moreover, a number of various sets of data can be applied to describe the process of a given phenomenon or different research institutions can forecast it from different points of view. The aim of this paper is an attempt to look at combining forecasts as a multi-criteria decision-making process. It is a multi-criteria process as we need to select: (a) individual component forecasts (quantity, quality, type), (b) values of their weights (objective determination of weights or their arbitrary choice), (c) a mathematical representation of combination and (d) a measure to assess the quality of a combined forecast. In parts 2-4 we provide an overview of methods applied for combining forecasts taking account of problems and recommendations concerning development of such forecasts. In order to illustrate this, part 5 includes the results of our study of a combined forecast for inflation in Poland.

## 2. Selecting individual forecasts for combination

A combined forecast for a given phenomenon over a specific period of time is defined as a combination of individual forecasts, each of them being developed for a given period of time with the use of a specific forecasting method. In our paper we do not describe the ways (methods) how individual forecasts are developed since there are many of them and it depends on a particular phenomenon and subjective approach of a researcher which one to select. As far as combining forecasts is concerned it is worth stressing that different research institutes can develop individual forecasts for the same variable (the same phenomenon) with the use of different data and different models. Moreover, different mathematical functions can be applied for the purpose of combining forecasts. Generally speaking, the more independent and varied in their construction individual forecasts are, the higher a chance to get a good quality combined forecast they constitute.

A good starting point for an overview of problems related with combining forecasts is a general scheme developed by Flores and White [Flores, White 1988, pp. 95-103]. The authors divide methods applied for developing individual forecasts into objective and subjective ones. Objective methods are quantitative methods (e.g. trend models, exponential smoothing, moving averages) whereas subjective ones include expert methods or surveys. Methods of combining forecasts, on the other hand, fall into two categories: systematic and intuitive ones. The former are strictly mathematical methods (mainly based on a simple or weighted average derived from individual forecasts) which are programmable and repeatable without a researcher's intervention. Such methods can be referred to as mechanical ones. Intuitive methods cannot be automatically repeated and require a researcher's consultation any time they are applied. The most formalized approach to combining forecasts, i.e., systematic methods for combining objective forecasts is the most widely discussed in the literature. Bates and Granger's article [Bates, Granger 1969] was one of the first to address this subject.

Given a large number and broad variety of methods for developing individual forecasts selection of forecast components is not an easy task. The most frequently applied criterion for this purpose is accuracy of individual ex post forecasts. Its statistical measures include for instance mean measures such as MAE, RMSE, MAPE and their modifications (e.g. a symmetric error MAPE – sMAPE), the Theil coefficient, turning points accuracy rate (e.g. [Zeliaś, Pawełek, Wanat 2003, p. 45; Cieślak (ed.) 2000, p. 49; Chen, Yang 2004; Makridakis et al. 1982, pp. 111-153; Greszta, Maciejewski 2005, pp. 49-61; Milo (ed.) 2002, p. 37]). Significant elements of the analysis of forecasting accuracy should include assessment of serial correlation of errors and characteristics of error distribution [De Menezes, Bunn, Taylor 2000, pp. 190-204].

The above mentioned criteria to select individual forecasts for forecast combination can be complemented by the analysis of so called forecast encompassing. A forecast $f_1$ encompasses forecast $f_2$ if $f_2$ does not provide any additional useful forecasting information other than $f_1$. Encompassing of individual forecasts which are potential components of a combined forecast can be verified by studying differences in their error values applying adequate statistical tests, e.g. $t$-Student and $F$ significance tests (for instance [Harvey, Newbold 2000, pp. 471-482; Clements, Hendry 2002, pp. 268-274]), the HLN test [Harvey, Leybourne, Newbold 1998, pp. 254-259; Costantini, Pappalardo 2008], the DM test [Diebold, Mariano 1995, pp. 253-263], the Hansen test [Hansen 2005, pp. 365-380].

Testing encompassing shall contribute to selecting individual forecasts of various prognostic values. In our opinion this will be more probable if potential component forecasts are developed on the basis of different data and methods in particular. This seems to be in line with recommendations resulting from research of combined forecasts described in literature (for instance [Armstrong 2001, p. 435; Makridakis, Hibon 2000, pp. 451-476]).

In case of a large number of individual forecasts a multi-tier and multi-level approach to combination can be applied. In this case a final combined forecast is a result of combining forecasts which are clusters of individual forecasts (similar in terms of their accuracy). The best forecast is selected with a support of a genetic algorithm [Lemke, Gabrys 2008, pp. 231-247].

The above mentioned criteria of assessment of accuracy of individual forecasts used to select component forecasts for combination can also be applied for assessment of the quality of a combined forecast.

## 3. Determining mathematical representation of combination and weight values for individual forecasts

The simplest way to combine forecasts is to apply an arithmetic mean or a median of individual forecasts. It is not obvious which method to choose as the literature includes research results which favour both a mean (e.g. [Clemen 1989, pp. 539-583]) and a median (e.g. [McNees 1992, pp. 703-710; Agnew 1985, pp. 363-376]). In case of applying an arithmetic mean the fact that weights for component forecasts are the same and do not take account of any errors of individual component forecasts constitutes a sort of a disadvantage. For this reason many a researcher attempted to test different methods that use various weights and additionally tried to add to their variability in time by using errors generated by individual forecasts.

Let us study a case of combining two forecasts and assume that for time $t$ two individual forecasts, $f_{1t}$ and $f_{2t}$, for $y_t$ variable are developed. In a general case a combined forecast can be calculated as a weighted average of the two forecasts:

$$f_{ct} = \beta f_{1t} + (1 - \beta) f_{2t}, \tag{1}$$

where $f_{ct}$ is a combined forecast, whereas $\beta$ and $(1 - \beta)$ are weights for individual forecasts.

$\sigma_1^2$ and $\sigma_2^2$ denote error variances of individual forecasts and $\rho$ is a correlation coefficient between error series of these forecasts. In order to minimize error variance of a combined forecast an optimal weight can be calculated using the following formula (variance – covariance method, (e.g. [Clements, Hendry 2002, p. 271]):

$$\beta = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \tag{2}$$

This paper includes proposals for determining weights for individual forecasts by optimizing also other accuracy measures for a combined forecast. The results are presented in section 5.

Weights obtained according to this formula (2) can exceed the range (0;1) which is hard to justify [Clements, Hendry 2002, p. 272]. In order to avoid such situations a restricting condition $0 \leq \beta \leq 1$ can be applied. In practice it is often assumed that errors of individual forecasts are not correlated and the weight $\beta$ is determined in the following way:

$$\beta = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}. \tag{3}$$

For instance errors from the last period can be applied instead of error variance of individual forecasts [Granger, Newbold 1986, p. 269]:

$$\beta = \frac{e_{t-1,1}^2}{e_{t-1,1}^2 + e_{t-1,2}^2}. \tag{4}$$

The formula (1) can be made more general for the case $k$ of individual forecasts:

$$f_{ct} = \beta_1 f_{1t} + \beta_2 f_{2t} + \ldots + \beta_k f_{kt}. \tag{5}$$

Weights $\beta_i$ ($i = 1, \ldots, k$) can also be determined as an estimation of regression coefficients:

$$y_t + \beta_0 + \beta_1 f_{1t} + \beta_2 f_{2t} + \ldots + \beta_k f_{kt} + \varepsilon_t, \tag{6}$$

where $y_t$ is a series of actual values given forecast phenomenon.

Predicted values of a variable provided for in formula (6) constitute a combined forecast.

The parameters in the formula (6) can be estimated for instance with the use of an ordinary least squares method, restricted least squares method as regards parameters ($\beta_0 = 0$, $\beta_1 + \beta_2 + \ldots + \beta_k = 1$), weighted least squares method or quantile regression. The results of such research can be found in works for instance by [Diebold, Pauly 1987, pp. 21-40; Taylor, Bunn 1998, pp. 193-206, 1999, pp. 325-339]. Some researchers [Li, Tkacz 2004] apply methods of nonparametric estimation (kernel estimators) thus pointing out a possibility to obtain better results than those obtained with the use of other methods of linear and nonlinear estimation. Yet a different approach to determine weights is based on Bayes methods. It was first shown in 1975 by Bunn [Bunn 1975, pp. 325-329] who proposed to determine weight $\beta$ in formula (1) as a frequency of an event so that forecast $f_{1t}$ in the past was better (had a smaller error) than forecast $f_{2t}$. Determination of a combined forecast as a linear combination of individual forecasts is often used due to a wider selection of weight estimation methods. In case of choosing a nonlinear relation between forecasts nonlinear estimation methods or neural networks are applied [Shi, Da Xu, Liu 1999, pp. 49-54; Donaldson, Kamstra 1996, pp. 49-61].

## 4. Selected comparative studies of forecasting methods. Conclusions

In practice forecasting can support decision-making processes. Therefore the quality of forecasts directly translates into real profits or losses resulting from decisions made on the basis of a forecast. Thus, the analysis of the quality constitutes an important element of forecasting. In 1982-2000 comparative studies of forecasting methods initiated by S. Makridakis gained a lot of interest. In a series of studies (M-competition [Makridakis et al. 1982, pp. 111-153], M2-competition [Makridakis et al. 1993, pp. 5-22], M3-competition [Makridakis, Hibon 2000, pp. 451-476; Armstrong, Lusk 1983, pp. 259-311]) a lot of economic categories (from 1000 to 3000 series) of US economy both on macro-level and on the level of selected companies were subject to forecasting. The studies used 21 forecasting methods, mainly different variants of extrapolation methods such as moving average, exponential smoothing, Box-Jenkins model as well as regression, Bayes methods, some expert systems, a method of artificial neural networks and also methods of forecast combination. Data series of different frequency were analyzed and forecasted. A forecast horizon was determined in a multi-variant way.

General conclusions were as follows:

a) simpler methods (e.g. a simple method of exponential smoothing) resulted in most cases in better forecasts for one period ahead compared to more complex methods;

b) ranking of methods depended on applied accuracy criterion;

c) accuracy of methods depended on a forecast horizon;

d) in case of a longer forecast horizon (i.e. with big changes in a forecast series) more accurate forecasts are obtained in case of:

- using large number of historic data;
- attenuating a trend component in a forecasting method (a profound analysis of a trend, an attempt to differentiate between a constant trend change and a change resulting from a random walk process);
- application of combined forecasts.

Recommendations based on the studies can be formulated in the following way:

1. A special attention should be paid to a trend included in a variability of a forecast category which can often change due to different reasons (e.g. structural changes, cycle changes, a part of a random walk process). Therefore a more practical (not purely mechanical) approach to extrapolation of a trend in a series is recommended taking into account economic evaluation of current situation as far as a given category is concerned.

2. Since accuracy of methods depends on a forecasting horizon it is advisable to look for ways of combining short and long-term forecasts;

3. Individual forecasts used for combining should be derived by different methods, more than five if possible.

4. Weights applied in combination should be based on profound characteristics of a forecast phenomenon.

5. In case of high level of uncertainty concerning the future, equal weights should be applied.

Recommendations resulting from the above-mentioned conclusions concerned higher flexibility of applied methods and use of knowledge and rules proposed by experts. All this led to development of expert forecasting systems such as RBF [Collopy, Armstrong 1992, pp. 1394-1414]. It was (and still is, see www.forecasting-principles.com) a kind of an expert system based on a set of 99 rules (Rule-Based Forecasting) used in forecasting depending on observed changes in an analysed series and on a forecasting horizon. In this system the final forecast is a combination of four simple extrapolation methods (random walk, linear regression for estimation of long-term trend, Brown and Holt methods) corrected according to a coded system of rules. The set of rules referred to [Armstrong, Collopy 1993, pp. 103-115]:

1. Identification and analysis of components of a series of a forecast phenomenon (it also included expert knowledge concerning reasons for change of a trend, e.g. a change of a cycle phase, launching price promotions in product sales, promotion campaigns, etc.).

2. Determination of parameters for exponential smoothing.

3. Development of short- and long-term models (they were estimated separately due to different factors determining a trend in a short- and long-term perspective).

4. Determination of weights for combined forecasts for a given series (mainly due to forecast horizon and compatibility of trend direction in a short- and long-term model).

Results obtained by Armstrong and Callopy provided evidence of effectiveness of the developed system. In majority of cases the forecasts they obtained were more accurate compared to forecasts based on particular methods, i.e. forecasts based on random walk model and forecasts resulting from a combination of individual forecasts using equal weights. The longer a forecast horizon the bigger the differences.

At present commercial expert systems constitute an important forecasting and management tool for companies and concerns in particular. They provide conditions for effective forecasts by integration of quantitative methods and expert evaluation [Dittmann 2008].

## 5. Example

In order to illustrate the problems related with combining forecasts discussed in the previous sections we have developed forecasts for inflation in Poland. We used statistical data for monthly price index of consumer goods and services in Poland in the period from January 2005 to August 2009 (source: www.stat.gov.pl, aggregated studies, price index of consumer goods and services, analogues period of the previous year = 100). For the purpose of calculations index values reduced by 100 were employed. A sample of data from March 2005 to December 2008 was used to develop the following models: a trend model (third order polynominal), AR(2) autoregressive model and Holt model (adopting optimal values of smoothing parameters due to mean-squared error of ex post MSE forecast). These models were applied to develop individual ex post forecasts for the period from March 2005 to December 2008 (see Figure 6) and expired forecasts for the period from January 2009 to August 2009 (see Figure 7). The forecasts, complemented with an expert forecast for the period from January 2009 to August 2009 (developed within a contest for the best team of macroeconomic analysts organized by the National Bank of Poland, *Rzeczpospolita* and *Parkiet*) were included in a linear combination which determines values of combined inflation forecasts. The forecasts were developed in a multi-variant way determining weights for individual forecasts and assuming different criteria for assessing the quality of a combined forecast. For the purpose of evaluation of individual forecasts accuracy of forecasting based on MSE, MAPE, sMAPE values and the Theil coefficient and error autocorrelation coefficient as well as encompassing (Diebold-Mariano test with the square loss function to test validity of the null hypothesis of equal accuracy of two forecasts) were applied. The recorded results are presented in Tables 1-5 in Appendix to this paper and in Figures 1-9. The forecasts are designated in the following way:

P1 – a forecast based on a trend model;

P2 – a forecast based on Holt model;

P3 – a forecast based on an autoregressive model;

P4 – an expert forecast;

P5 – a combined forecast, arbitrarily adopted equal weights as 0.33 for every ex post component forecast and 0.25 for an expired forecast;

P6 – a combined forecast, weights from estimation of regression parameters (see formula (6), $k = 4$) least squares method with collateral conditions, $\beta_0 = 0, \beta_1 + \beta_2 + ... + \beta_k = 1$;

P7 – a combined forecast, weights for minimizing MAPE value;

P8 – a combined forecast, weights for minimizing the absolute value of error autocorrelation coefficient (of the first order);

P9 – a combined forecast, weights for minimizing sMAPE value;

P10 – a combined forecast, weights for minimizing the value of the Theil coefficient;

P11 – a combined forecast, weights for multi-criteria optimalization and simultaneous minimalization of sMAPE/100 value, the value of error autocorrelation coefficient (absolute) and the value of the Theil coefficient, criteria are of equal importance (weight of every criterion equals 1);

P12 – a combined forecast, weights for multi-criteria optimization and simultaneous minimization of sMAPE/100 value, the value of error autocorrelation coefficient (absolute) and the value of the Theil coefficient. The priority given to sMAPE/100 measure, it is 50 times more important than other criteria.

The objective function in case of P11 and P12 variants was a weighted average of the above-mentioned measures. It is labelled as sMTA.

As far as the assessment criteria are concerned the worst individual forecast is the one based on a trend model (big mean errors, serial correlation of errors, asymmetric distribution of ex post errors) whereas the best included the forecast based on the Holt model (in the group of ex post forecasts) and the expert forecast (in the group of expired forecasts) (see Tables 1 and 2 in the Appendix and Figures 1, 2 and 3). In our opinion the poor quality of forecasts based on the trend model results from the fact

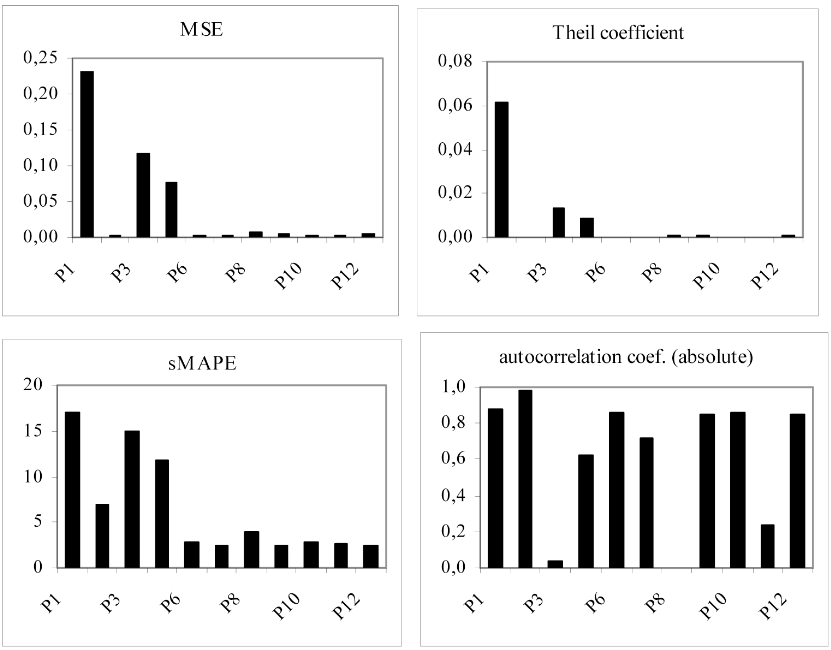



**Figure 1.** Properties of ex post individual and combined forecasts.
Notice: first three columns indicate values for individual forecasts

Source: own study.

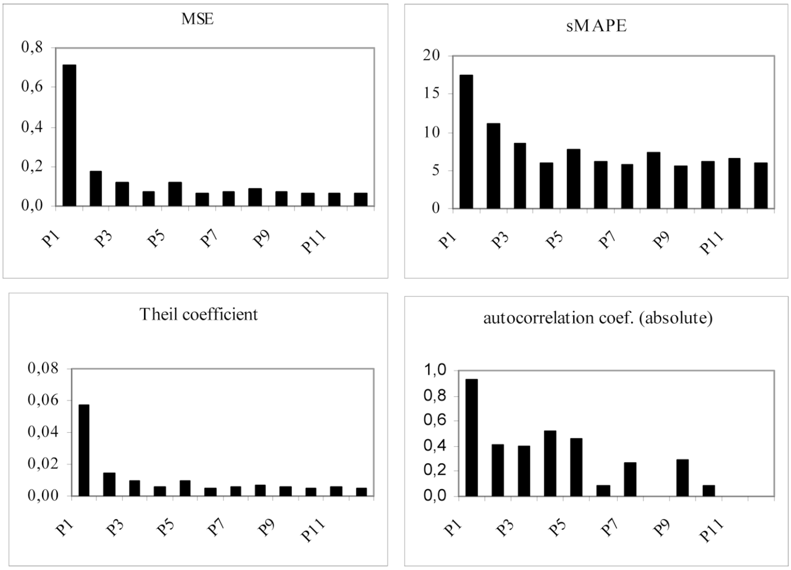
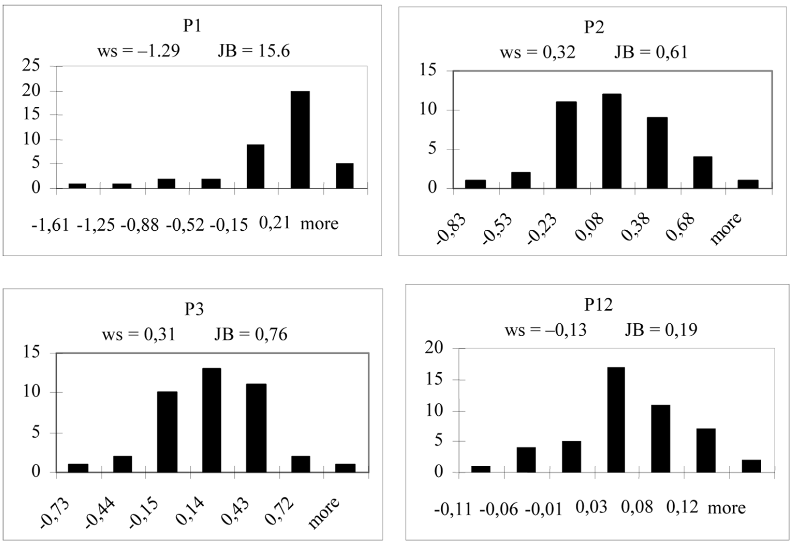



**Figure 2.** Properties of expired individual and combined forecasts.
Notice: first four columns indicate values for individual forecasts

Source: own study.




Notice: ws – skewness coefficient, JB – Jarque-Bera statistics (critical value for the significance level of 0.05 amounts to 5.99).

**Figure 3.** Distribution of errors of individual ex post forecasts and the best combined forecast

Source: own study.

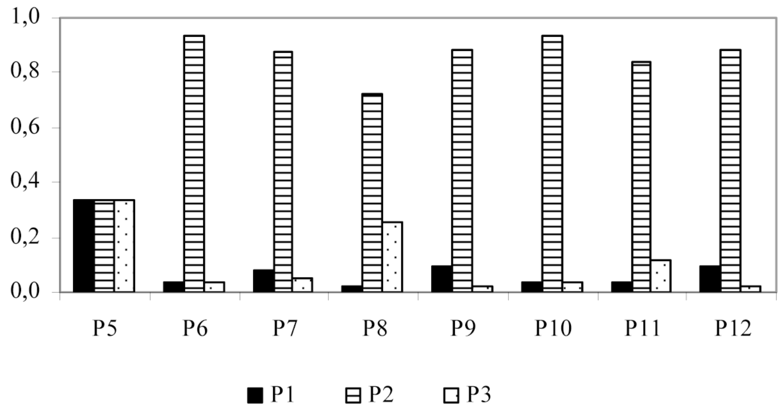



**Figure 4.** Weights for individual ex post forecasts

Source: own study.

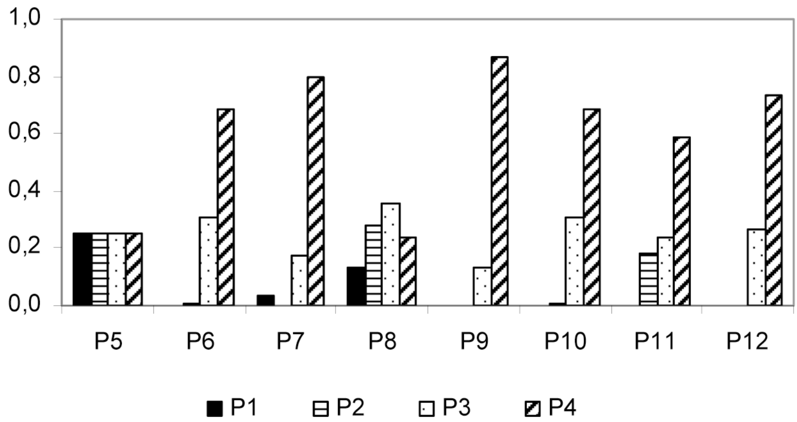



**Figure 5.** Weights for individual expired forecasts

Source: own study.

that this model takes account of long-term changes of a series in focus whereas the autoregressive and Holt models provide for volatility in an adaptive way in short periods (see Figure 9). In the study no modification of the trend due to short-term impact factors was made.

Values of a DM encompassing test (see Table 3 in the Appendix) indicate significant differences in forecasting accuracy with the use of the trend model compared to other forecasts. There is no evidence indicating significant differences in case of forecasts based on Holt, autoregressive and expert models. Taking into account the scale of accuracy and an encompassing test the combined forecast should be developed on the basis of the values of the trend and expert model forecasts. Nevertheless, all the forecasts were left as combination components in order to meet the requirement of diversity and the number of methods applied for component forecasts.

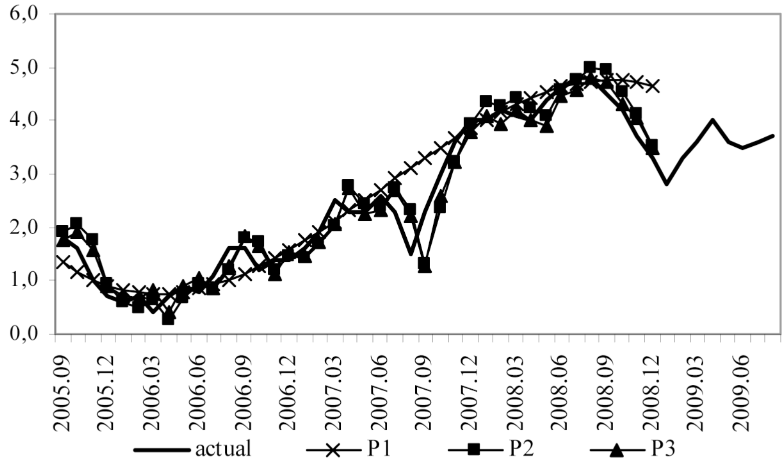



**Figure 6.** Inflation in Poland, actual results and individual ex post forecasts

Source: own study.

The obtained results (see Figures 1-8 and Tables 1 and 2 in the Appendix) indicate that in majority of combining variants individual forecasts are worse than combined forecasts as far as different criteria are concerned. In particular it refers to ex post forecasts. In case of combined forecasts, the forecast developed as a simple average (P5) has the worst properties whereas multi-criteria forecasts P11 and P12 have the best ones (see for instance sMTA values in Tables 1 and 2 in the Appendix). The latter forecasts exhibit sMAPE mean errors a little bit bigger than the best variant for this criterion namely P9, but they contain no error autocorrelation. The best combination P12 also has smaller asymmetry in ex post error distribution than individual forecasts and on the significance level of 0.05 there are no grounds for rejecting the hypothesis that it is a normal distribution (see Figure 3).

It is worth stressing that in case of discussed variants weights for individual forecasts are volatile (see Figures 4, 5). In both analysed periods there are dominating forecasts (of the highest accuracy) and in case of ex post forecasts (from March 2005 to December 2008) it is the forecast based on the Holt model and in case of expired forecasts (from January 2009 to August 2009) it is the expert forecast. Nevertheless, their weight diversity for ex post period is much smaller than in case of expired forecasts.

Combined expired forecasts for inflation simulate its actual volatility very well, much better than individual forecasts (see Figures 7 and 8) with the exception of two last analysed months which can be justified by rather distant short-term forecast horizon.

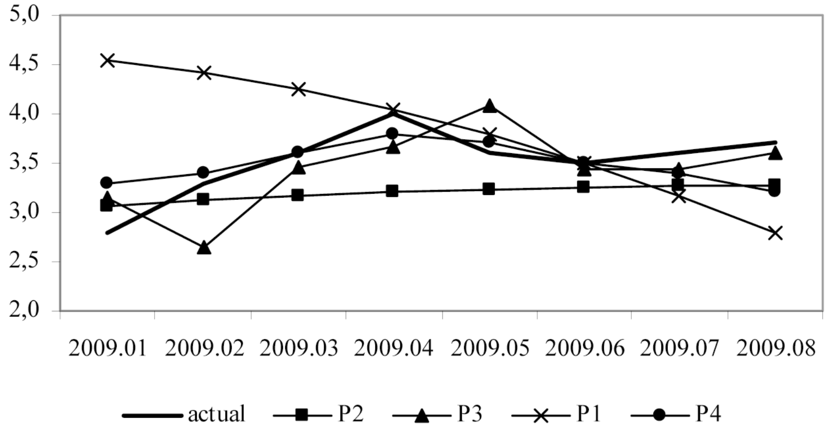



**Figure 7.** Inflation in Poland, actual results and individual expired forecasts

Source: own study.

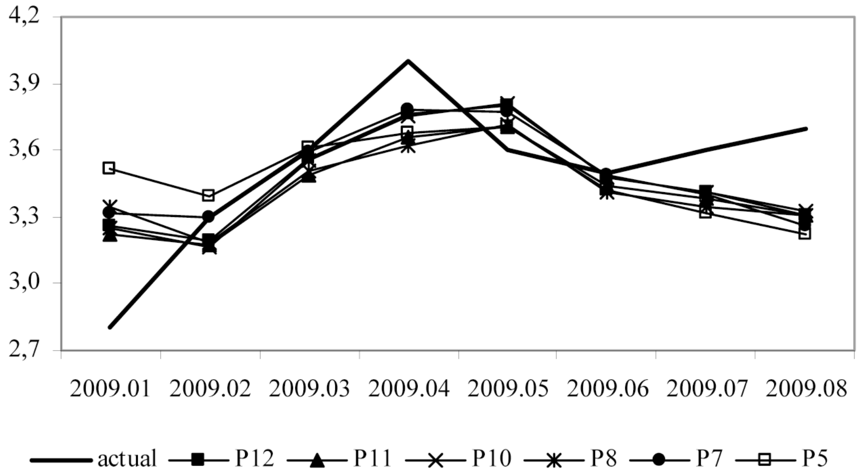



**Figure 8.** Inflation in Poland, actual results and expired combined forecasts

Source: own study.

The analysis of error values of individual and the best combined expired forecasts in subsequent periods (see Figure 9) indicates that combining forecasts is fully justified (providing for possibility to reduce mean errors), especially if weights for individual forecasts are determined on the basis of multi-criteria optimization.

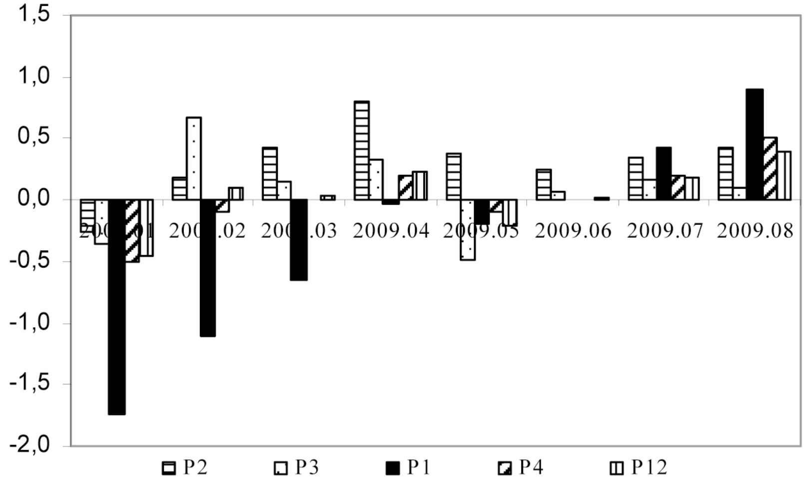



**Figure 9.** Errors of expired individual forecasts and the best combined forecast (P12)

Source: own study.

## 6. Conclusions

Forecasts bear significant practical implications. On the one hand, it costs to develop them, on the other, the quality of forecasts translates into actual profits or losses resulting from the decision made on the basis of a forecast. Although combined forecasts can be more expensive compared to individual ones they allow for predicting the value of a phenomenon on the basis of a larger scope of information which means reduction of a forecast error which is proved by many studies discussed in the literature and in this paper. In order to achieve this objective it is advisable to apply different assessment measures for forecast quality which allow for comprehensive evaluation (concerning accuracy, error autocorrelation, properties of error distribution) or an attempt to include several criteria simultaneously. However, it is difficult to draw unambiguous conclusions and provide recommendations since as Armstrong [Armstrong 2001, pp. 425-427] points out the main prerequisites for combining forecasts are lack of a dominant forecast, uncertainty of changes in a forecast phenomenon, high cost of a forecast-based wrong decision. Our study indicates that combining forecasts is also useful when a dominant individual forecast exists.

# Appendix

**Table 1.** Properties of ex post individual and combined forecasts

| Forecast variant | MSE | MAPE | sMAPE | Theil coefficient | Error autocorrelation coefficient | sMTA criteria weight = 1 |
|---|---|---|---|---|---|---|
| P1 | 0.51 | 34.56 | 30.21 | 0.0612 | 0.87 | |
| P2 | 0.00 | 3.88 | 7.00 | 0.0004 | 0.98 | |
| P3 | 0.12 | 16.17 | 14.98 | 0.0135 | 0.04 | |
| P5 | 0.08 | 12.95 | 11.71 | 0.0089 | 0.62 | 0.75 |
| P6 | 0.00 | 2.78 | 2.88 | 0.0002 | 0.86 | 0.89 |
| P7 | 0.00 | 2.36 | 2.40 | 0.0004 | 0.72 | 0.75 |
| P8 | 0.01 | 3.94 | 3.99 | 0.0009 | 0.00 | 0.04 |
| P9 | 0.00 | 2.35 | 2.39 | 0.0004 | 0.84 | 0.87 |
| P10 | 0.00 | 2.78 | 2.88 | 0.0002 | 0.86 | 0.89 |
| P11 | 0.00 | 2.59 | 2.65 | 0.0003 | 0.24 | 0.26 |
| P12 | 0.00 | 2.35 | 2.39 | 0.0004 | 0.84 | 0.87 |

Source: own study.

**Table 2.** Properties of expired individual and combined forecasts

| Forecast variant | MSE | MAPE | sMAPE | Theil coefficient | Error autocorrelation coefficient | sMTA criteria weight = 1 |
|---|---|---|---|---|---|---|
| P1 | 0.72 | 19.54 | 17.40 | 0.0575 | 0.93 | |
| P2 | 0.18 | 10.60 | 11.18 | 0.0142 | 0.40 | |
| P3 | 0.12 | 8.42 | 8.58 | 0.0097 | −0.40 | |
| P4 | 0.07 | 5.97 | 5.93 | 0.0060 | 0.52 | |
| P5 | 0.12 | 7.83 | 7.66 | 0.0095 | 0.45 | 0.54 |
| P6 | 0.05 | 6.22 | 6.17 | 0.0051 | −0.08 | 0.15 |
| P7 | 0.07 | 5.81 | 5.73 | 0.0058 | 0.26 | 0.33 |
| P8 | 0.09 | 7.29 | 7.25 | 0.0070 | 0.00 | 0.08 |
| P9 | 0.07 | 5.63 | 5.59 | 0.0054 | 0.29 | 0.35 |
| P10 | 0.06 | 6.22 | 6.17 | 0.0051 | −0.08 | 0.15 |
| P11 | 0.07 | 6.47 | 6.50 | 0.0054 | 0.00 | 0.07 |
| P12 | 0.06 | 6.07 | 6.01 | 0.0051 | 0.00 | 0.06 |

Source: own study.

**Table 3.** Values of Diebold-Mariano test statistics for expired forecasts

| | P1 | P2 | P3 | P4 |
|---|---|---|---|---|
| P1 | x | | | |
| P2 | –2.52* | x | | |
| P3 | –2.91* | –0.53 | x | |
| P4 | –3.19* | –1.04 | –0.51 | x |

* means rejection of null hypothesis on the significance level of 5%.

Source: own study.

**Table 4.** Weights for individual ex post forecasts

| Combining variant | Variant of individual forecast | | |
|---|---|---|---|
| | P1 | P2 | P3 |
| P5 | 0.33 | 0.33 | 0.33 |
| P6 | 0.03 | 0.93 | 0.04 |
| P7 | 0.08 | 0.87 | 0.05 |
| P8 | 0.02 | 0.73 | 0.25 |
| P9 | 0.09 | 0.89 | 0.02 |
| P10 | 0.03 | 0.93 | 0.04 |
| P11 | 0.04 | 0.84 | 0.12 |
| P12 | 0.09 | 0.89 | 0.02 |

Source: own study.

**Table 5.** Weights for individual expired forecasts

| Combining variant | Variant of individual forecast | | | |
|---|---|---|---|---|
| | P1 | P2 | P3 | P4 |
| P5 | 0.25 | 0.25 | 0.25 | 0.25 |
| P6 | 0.00 | 0.01 | 0.31 | 0.68 |
| P7 | 0.03 | 0.00 | 0.17 | 0.79 |
| P8 | 0.13 | 0.28 | 0.36 | 0.24 |
| P9 | 0.00 | 0.00 | 0.13 | 0.87 |
| P10 | 0.00 | 0.01 | 0.31 | 0.68 |
| P11 | 0.00 | 0.18 | 0.24 | 0.58 |
| P12 | 0.00 | 0.00 | 0.27 | 0.73 |

Source: own study.

## References

Agnew C., *Bayesian consensus forecasts of macroeconomic variables*, Journal of Forecasting 1985, Vol. 4.

Armstrong S., *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer Academic Publishers, Dordrecht 2001.

Armstrong S., Collopy F., *Causal forces: Structuring knowledge for time-series extrapolation*, Journal of Forecasting 1993, Vol. 12.

Armstrong S., Lusk E., *Commentary on the Makridakis time series competition* (*M-competition*), Journal of Forecasting 1983, Vol. 2.

Bates J., Granger C., *The combination of forecasts*, Operations Research Quarterly 1969, Vol. 20.

Bunn D.W., *A Bayesian approach to the linear combination of forecasts*, Operational Research Quarterly 1975, Vol. 26.

Chen Z., Yang Y., *Assessing forecast accuracy measures*, Preprint Series 2004, No 2004-10, Department of Statistics, Iowa State University, 2004.

Cieślak M. (ed.), *Prognozowanie gospodarcze. Metody i zastosowania*, Wydawnictwo Naukowe PWN, Warszawa 2000.

Clemen R.T., *Combining forecasts. A review and annotated bibliography*, International Journal of Forecasting 1989, Vol. 5.

Clements M., Hendry D., *A Companion to Economic Forecasting*, Vol. 2, Blackwell, 2002 [published online].

Collopy F., Armstrong J., *Rule-based forecasting: Development and validation of an expert systems approach to combining time series extrapolations*, Management Science 1992, Vol. 38.

Costantini M., Pappalardo C., *Combination of forecast methods using encompassing tests: An algorithm-based procedure*, Reihe Ökonomie. Economics Series 2008, Vol. 228.

De Menezes L.M., Bunn D.W., Taylor J.W., *Review of guidelines for the use of combined forecasts*, European Journal of Operational Research 2000, Vol. 120.

Diebold F., Mariano R., *Comparing predictive accuracy*, Journal of Business and Economic Statistics 1995, Vol. 13.

Diebold F., Pauly P., *Structural change and the combination of forecasts*, Journal of Forecasting 1987, Vol. 6.

Dittmann P., *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie*, Wolters Kluwer, Warszawa 2008.

Donaldson R.G., Kamstra M., *Forecast combining with neural networks*, Journal of Forecasting 1996, Vol. 15, No 1.

Flores B.E., White E.M., *A framework for the combination of forecasts*, Journal of the Academy of Marketing Sciences 1988, Vol. 6. No 3-4.

Granger C., Newbold P., *Forecasting Economic Time Series*, Academic Press, New York 1986.

Greszta M., Maciejewski W., *Kombinowanie prognoz gospodarki Polskiej*, Gospodarka Narodowa 2005, nr 5-6.

Hansen P., *A test for superior predictive ability*, Journal of Business and Economic Statistics 2005, Vol. 23.

Harvey D.I., Leybourne S.J., Newbold P., *Tests for forecast encompassing*, Journal of Business and Economic Statistics 1998, Vol. 16.

Harvey D., Newbold P., *Test for multiple forecast encompassing*, Journal of Applied Econometrics 2000, Vol. 15.

Lemke C., Gabrys B., *Forecasting and forecast combination in airline revenue management applications*, [in:] N.T. Nguyen, G. Kolaczek, B. Gabrys (eds.), *Knowledge Processing and Reasoning for Information Society*, EXIT, Warszawa 2008.

Li F., Tkacz G., *Combining forecasts with nonparametric kernel regressions*, Studies in Nonlinear Dynamics and Econometrics 2004, Vol. 4.

Makridakis S., Andersen A., Carbone R., Fildes R., Hibon M., Lewandowski M., Newton J., Parzen E., Winkler R., *The accuracy of extrapolation (time series) methods: Results of a forecasting competition*, Journal of Forecasting 1982, Vol. 1.

Makridakis S., Chatfield C., Hibon M., Lawrence M., Mills T., Ord K., Simmons L., *The M2-competition: A real-time judgmentally based forecasting study*, International Journal of Forecasting 1993, Vol. 9.

Makridakis S., Hibon M., *The M3-Competition: results, conclusions and implications*, International Journal of Forecasting 2000, Vol. 16.

McNees S.K., *The uses and abuses of consensus' forecasts*, Journal of Forecasting 1992, Vol. 11.

Milo W. (ed.), *Prognozowanie i symulacje*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2002.

Shi S., Da Xu L., Liu B., *Improving the accuracy of nonlinear combined forecasting using neural networks*, Expert Systems with Applications 1999, Vol. 16, No 1.

Taylor J., Bunn D., *Combining forecast quantiles using quantile regression: Investigating the derived weights, estimator bias and imposing constraints*, Journal of Applied Statistics 1998, Vol. 2.

Taylor J., Bunn D., *Investigating improvements in the accuracy of prediction intervals for combinations of forecasts: A simulation study*, International Journal of Forecasting 1999, Vol. 3.

Zeliaś A., Pawełek B., Wanat S., *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, PWN, Warszawa 2003.

## WIELOKRYTERIALNE PROGNOZY KOMBINOWANE

**Streszczenie:** stosowanie kombinacji prognoz wynika z przekonania, że pojedynczy model nie może opisać w całości rzeczywistego, często skomplikowanego, procesu ekonomicznego. Referat jest próbą spojrzenia na kombinowanie prognoz jako wielokryterialny proces decyzyjny. Wielokryterialność wynika z konieczności wyboru: (a) indywidualnych prognoz cząstkowych (ilości, jakości, rodzaju), (b) wartości wag z nimi związanych (wyznaczenia numerycznego lub przyjęcia arbitralnego), (c) postaci matematycznej kombinacji czy wreszcie (d) miary oceniającej jakość prognozy kombinowanej. W referacie dokonujemy przeglądu metod kombinowania prognoz oraz przedstawiamy wyniki badania własnego polegającego na opracowaniu prognozy kombinowanej inflacji w Polsce. Analiza tych wyników wskazuje na celowość kombinowania, zwłaszcza jeśli wagi dla prognoz indywidualnych będą wyznaczane w drodze optymalizacji wielokryterialnej. Badanie pokazuje również, że kombinowanie prognoz jest użyteczne, także w sytuacji, gdy występuje dominująca prognoza indywidualna.