DYNAMIC ECONOMETRIC MODELS
Vol. 9 – Nicolaus Copernicus University – Toruń – 2009

Mariola Piłatowska
Nicolaus Copernicus University in Toruń

Combined Forecasts Using the Akaike Weights

A b s t r a c t. The focus in the paper is on the information criteria approach and especially the Akaike information criterion which is used to obtain the Akaike weights. This approach enables to receive not one best model, but several plausible models for which the ranking can be built using the Akaike weights. This set of candidate models is the basis of calculating individual forecasts, and then for combining forecasts using the Akaike weights. The procedure of obtaining the combined forecasts using the AIC weights is proposed. The performance of combining forecasts with the AIC weights and equal weights with regard to individual forecasts obtained from models selected by the AIC criterion and the *a posteriori* selection method is compared in simulation experiment. The conditions when the Akaike weights are worth to use in combining forecasts were indicated. The use of the information criteria approach to obtain combined forecasts as an alternative to formal hypothesis testing was recommended.

K e y w o r d s: combining forecasts, weighting schemes, information criteria.

1. Introduction

The development of time series analysis and computing power of computers made that many different forecasts can be obtained when forecasting the same economic variable with different methods. Many selection criteria based on the performance of ex post forecasts are used to choose the best forecast (Armstrong, 2001). Combining forecasts can be treated as an alternative approach to the selection of the best individual forecast. Since the seminal paper of Bates and Granger (1969) has been known that combining forecasts can produce a forecast superior to any element in the combined set[1]. Hence, instead of seeking the best forecasting model the combined forecasts based on competing models are received.

Moreover, the reason for combining forecasts (or model averaging) is that the data generating model (true model) is unknown. Therefore each model

[1] The paper of Bates and Granger (1969) caused the development of research on combining forecasts (for overview see Timmermann, 2006).

should be treated as an approximation of unknown data generating model. These models may be incomplete (or incorrectly specified) in different ways; forecast based on them might be biased. Even if forecasts are unbiased, there will be covariances between forecasts which should be taken into account. Then, combining forecasts produced by misspecified models may improve the forecast in comparison to any individual forecast obtained from the underlying models. As a consequence, the problem of selecting the individual forecasts over the set of available forecasts and the weighting schemes is occurred. Especially, the selection of weighting scheme is important.

Most frequently the following weighting schemes can be distinguished: equal weights (Stock, Watson, 2004, 2006; Marcellino, 2004), Akaike weights (Atkinson, 1980; Swanson, Zeng, 2001; Kapetanios et al., 2008), optimized and constrained weights (Jagammathan, Ma, 2003), Bayesian weights (Min, Zellner, 1993; Diebold, Pauly, 1980; Wright, 2003).

In the paper the focus is on the information criteria approach, especially the Akaike information criterion which is used to produce the Akaike weights. This approach enables to obtain not only one, but several plausible models for which the ranking can be built using the Akaike weights. The individual forecasts, calculated from the considered models, are aggregated with the Akaike weights. The paper propagates the application of the Akaike weights, previously unknown in Polish literature, and evidence ratios in selecting a model over the underlying set of models and in producing the combined forecasts.

The purpose of the paper is to propose the procedure of combining forecasts using the Akaike weights, and also to compare the combined forecasts (with the Akaike weights and equal weights) with individual forecasts obtained from the best model selected according to: (1) the Akaike information criterion and (2) traditional hypothesis testing. The analysis will be conducted in the simulation study in which autoregressive models and causal models are taken as approximating models provided that the data generating model is unknown.

The structure of the paper is as follows. In section 2 the construction of the Akaike weights will be presented. In section 3 the procedure of combining forecasts using the Akaike weights will be depicted. Next, the results of simulation experiment will be showed and at the very end – some conclusions.

2. The Akaike Weights

The Akaike information criterion is applied to select the best model from among the candidate models considered. The Akaike's (1973) seminal paper proposed the use of the Kullback-Leibler information or distance as a fundamental basis for model selection. The Kullback-Leibler (K-L) information between models f (true model or probability distribution) and g (approximating model in terms of a probability distribution) is defined for continuous functions as the integral:

$$I(f,g) = \int f(x) \ln \frac{f(x)}{g(x \mid \theta)} dx, \tag{1}$$

where $I(f,g) \geq 0$, $I(f,g) = 0$ only if $f = g$, $I(f,g) \neq I(g,f)$.

$I(f,g)$ denotes the information lost when the model *g* is used to approximate the model *f*. The purpose is to seek an approximating model that loses as little information as possible. This is equivalent to minimizing $I(f,g)$ over *g*. Akaike (1973) found a rigorous way to estimate K-L information based on the empirical log-likelihood function at its maximum point. This result took the form of an information criterion:

$$\text{AIC} = -2 \ln L(\hat{\theta}) + 2K, \tag{2}$$

where $L(\hat{\theta})$ is the maximum likelihood for the candidate model i, which is corrected by K the number of estimated parameters.

Akaike has showed that choosing the model with the lowest expected information loss (i.e. the model which minimizes the expected Kullback-Leibler discrepancy) is asymptotically equivalent to choosing the model M_i $(i = 1, 2, ..., R)$ that has the lowest AIC value.

To obtain the Akaike weights a simple transformation of the raw AIC values should be performed. For each model the difference in AIC with respect to the AIC of the best candidate model is computed:

$$\Delta_i = \text{AIC}_i - \text{AIC}_{\min}. \tag{3}$$

These Δ_i are easy to interpret and allow a quick comparison and ranking of candidate models. The best model over the candidate models has $\Delta_i \equiv \Delta_{\min} = 0$. The larger Δ_i is, the less plausible is that the fitted model M_i is the K-L best model, given the data. For nested model some rough rules of thumb are available in selecting the model (Burnham, Anderson, 2002), i.e. models with $\Delta_i < 2$ have substantial support, models with $4 < \Delta_i < 7$ – considerable less support. Models with $\Delta_i > 10$ have either essentially no support and might be omitted from further consideration, because they fail to explain some substantial explainable variation in the data. In empirical data analysis the models with $\Delta_i < 4$ are accepted.

From the differences Δ_i we can obtain the relative plausibility of model M_i over the set of candidate models by estimating the relative likelihood $L(M_i \mid x)$ of model M_i given the data x (Akaike, 1983):

$$L(M_i \mid x) \propto \exp(-0.5\Delta_i), \tag{4}$$

where $\propto$ stands for „is proportional to”.

Finally, the relative model likelihoods are normalized (divided by the sum of the likelihoods of all models) and the Akaike weights w_i are obtained:

$$w_i = \frac{\exp(-0.5\Delta_i)}{\sum_{r=1}^{R}\exp(-0.5\Delta_r)}, \qquad \sum_{i=1}^{R} w_i = 1. \tag{5}$$

Weight w_i can be interpreted as the probability that M_i is the best model (in the AIC sense, i.e. the model minimizing the K-L information) given the data and the set of candidate models.

Additionally weights w_i can be useful in evaluating the relative strength of evidence for the best model (with biggest weight) over the other in the set of R models. Thus, the evidence ratios or the ratio of Akaike weights w_i / w_j (in particular the ratio w_1 / w_j, where w_1 is the weight for the best model, and w_j – weights for models in the set) are calculated. It is worth pointing out that this approach does not assume that any of the candidate models is necessarily true, but rather the ranking of models in the sense of K-L information is considered[2].

The AIC weights (5) can be generalized into the form (Burhnham, Anderson, 2002, 2004):

$$w_i = \frac{\exp(-0.5\Delta_i)q_i}{\sum_{r=1}^{R}\exp(-0.5\Delta_r)q_r}, \tag{6}$$

where q_i is a prior probability of our prior information (or lack thereof) about which of the R models is the K-L best model for the data. By the AIC weights we mean the expression (6) with the equal prior probabilities, i.e. $q_i = 1/R$. The inclusion of prior probabilities in (6) makes that the AIC weights can be treated as an approximation of the Bayesian posterior model probability (Burnham, Anderson, 2002, 2004). However, it is not a true Bayesian approach. The full Bayesian approach to model selection requires both the prior q_i on the model and a prior probability distribution on the parameter θ in model M_i for each model. Then the derivation of posterior results requires integration (usually achievable only by Markov chain Monte Carlo methods). In that context the AIC weight seem to be useful because they are much easier to compute and additionally the researcher is not required to determine prior densities for the parameters.

[2] It is the main difference in the comparison with the Bayesian model averaging which assumes that the true generation model is in the set of candidate models and measures the degree of belief that a certain model is the true data-generating model.

3. The Procedure of Obtaining The Combined Forecasts Using The Akaike Weights

When calculating combined forecasts using the Akaike weights some conditions should be satisfied. Namely, all models in the set of candidate models should be fitted to exactly the same set of data because the inference based on information criteria is conditional on the data in hand. Moreover, all models in the set should represent the same response variable. A common type of mistake is to compare models of y_t with models of transformed variable, e.g. $\ln y_t$ or Δy_t.

The steps in the procedure of obtaining forecasts aggregated with the Akaike weights are following.

Step 1. Establishing the initial set of R models describing a given variable and their specification. The guidelines on specifying causal models should be derived from an economic theory explaining the phenomenon in interest. In the case of large number of variables it is not recommended to run all possible regressions because the set of candidate models should be plausible with respect to the economic theory, and not be automatically selected. The true generation model does not have to be included in the set of models.

Step 2. Fitting the models $(i = 1, 2, ..., R)$ to the data, calculating the AIC values and differences Δ_i. Models should satisfy statistical and economic requirements.

Step 3. Creating the reduced set of models $(i = 1, 2, ..., R^*)$ for which $\Delta_i < 4$, i.e. models plausible in the sense of K-L information.

Step 4. Calculating the Akaike weights (eq. (4)) and combined forecasts according to formula:

$$\hat{y}_{t,t+h} = \sum_{i=1}^{R^*} w_{i,h} f_{i,t+h}, \quad \sum_{i=1}^{R^*} w_{i,h} = 1, \tag{7}$$

where $\hat{y}_{t,t+h}$ – combined forecast, $w_{i,h}$ – the weight assigned to the forecast $f_{i,t,t+h}$ based on the *i*th individual model.

When combining forecasts the problem is to estimate the weights $w_{i,h}$, so as to minimize a penalty function depending on the forecast errors. Very often, the penalty function is simply the mean square forecast error (*MSFE*).

4. Simulation Experiment Results

The purpose of simulation experiment is to compare combining forecasts (using the Akaike weights and equal weights) with individual forecasts obtained from the best model selected according to: (1) the Akaike information criterion and (2) traditional hypothesis testing. In experiment the autoregressive models and causal models are taken as approximating models provided that the data generating model is unknown.

Simulation scenario is following.

The data-generating model of Y_t has the form:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon), \;\; \sigma_\varepsilon = 1, 2, 3,$$

with parameters: $\beta_0 = 10, \beta_1 = 1.5, \beta_2 = 1.2, \beta_3 = 2$, for samples: $n = 100, 50$ (number of replications $m = 1000$). Processes $X_{1,t}, X_{2,t}$ and $X_{3,t}$ have following structure:

$$X_{1,t} = 12 + 0.8 X_{1,t-1} + \zeta_t + 0.6\zeta_{t-1}, \;\; \zeta_t \sim N(0,1),$$

$$X_{2,t} = 14 + 0.7 X_{2,t-1} + \eta_t + 0.8\eta_{t-1}, \;\; \eta_t \sim N(0,1),$$

$$X_{3,t} = 8 + 1.2 X_{3,t-1} - 0.4 X_{3,t-2} + \xi_t, \;\; \xi_t \sim N(0,1).$$

As approximating models are taken: the autoregressive models:

$$Y_t = \gamma_0 + \gamma_1 Y_{t-1} + \ldots + \gamma_q Y_{t-q}, \quad q = 1, 2, 3, 4,$$

and causal models[3]:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \alpha_4 Y_{t-4} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \tag{M1}$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \tag{M2}$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t \tag{M3}$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma_1 X_{1,t} + \gamma_2 X_{1,t-1} + v_t. \tag{M4}$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \gamma_1 X_{1,t} + v_t, \tag{M5}$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma_1 X_{1,t} + v_t. \tag{M6}$$

It is assumed that the true generation model is unknown, therefore specifying the causal models the variables $X_{2,t}$ and $X_{3,t}$ were omitted.

The analysis was carried out separately for autoregressive models and causal models. In each replication only the models with $\Delta_i < 4$ were taken. For those models the Akaike weights and equal weight $(1/R^*)$ were received, and having calculated the individual[4] forecasts, the *h*-period ahead combined forecasts were obtained. To compare forecasts the mean square forecast error (*MSFE*) was calculated for combined forecasts and individual forecasts obtained from the

[3] Having conducted initial simulation, these models are accepted as plausible.

[4] Individual forecasts were dynamic, and as values of explanatory variables X_{1t} in forecast period the generated values are taken.

best model (selected by: the AIC and *a posteriori* selection method applied to the causal model[5] M1). The results present Table 1 (for $n = 100$) and Table 2 (for $n = 50$).

Results presented in Table 1 and 2 show that the differences between *MSFE* obtained for combined forecasts (using AIC weights and equal weights) and individual forecasts (from the best model selected by the minimum of AIC and by *a posteriori* selection method) are small. However, certain regularities indicating the usefulness of combined forecasts can be observed.

In the case of small size of disturbance (σ_ε = 1) the combined forecasts (with the Akaike weights, w_{AIC}) obtained from causal models give smaller mean square forecast error (*MSFE*) than forecasts combined with equal weights ($w_{EQ.}$) at the whole forecast horizon (Table 1). This slight dominance of combined forecast with the *AIC* weights is hold for σ_ε = 2, 3 at longer horizons ($h > 5$), and for shorter horizons – the forecasts combined with equal weights ($w_{EQ.}$) have lower *MSFE*.

In general the combined forecasts (with AIC weights and equal weights) outperform the individual forecasts obtained from the best model (selecte*d* by *a posteriori* method, $m_{SEL.}$), because the *MSFE* for combined forecasts are visibly lower than the *MSFE* for individual forecasts; this occurs for all sizes of disturbance σ_ε (Table 1). Such performance indicates the dominance of combined forecasts. However this dominance is not complete, because the lowest *MSFE* are obtained for individual forecasts calculated from the best model selected by the minimum of AIC (for σ_ε=1). For bigger size of disturbance, i.e. σ_ε =2 and 3, the lower *MSFE*s at the horizon $h \le 5$ give the forecasts combined using equal weights, and at the horizon $h \ge 6$ – forecasts combined using the AIC weights. These lower *MSFE*s for forecasts from models selected by the minimum of AIC (for σ_ε=1) refer to the cases when the set of candidate models is small (in the considered experiment it were models M3 and M4), and additionally one model in the set has the dominating AIC weight ($w_{AIC} > 0.7$). For the bigger size of disturbance, i.e. σ_ε = 2 and 3, the set of competing models, being used in combining forecasts, consisted most frequently of models M3, M4, M5, M6, and none had the dominating AIC weight. Then, the combined forecasts (using the AIC weights or equal weights) outperformed the individual forecasts from model selected by the minimum of AIC, i.e. they gave the lower *MSFE*s.

[5] Variables elimination in *a posteriori* selection method was realized at the 5% significance level.

Table 1. Mean square forecast errors (*MSFE*) for sample n = 100

	Causal models				Autoregressive models		
h	w_{AIC}	$w_{EQ.}$	min_{AIC}	$m_{SEL.}$	w_{AIC}	$w_{EQ.}$	min_{AIC}
				$\sigma_\varepsilon = 1$			
1	2.950	2.973	2.935	3.050	3.102	3.117	3.096
2	2.290	2.301	2.287	2.357	2.530	2.547	2.523
3	2.471	2.500	2.456	2.604	2.568	2.595	2.548
4	3.280	3.320	3.254	3.436	2.569	2.602	2.537
5	4.064	4.097	4.043	4.192	3.706	3.748	3.653
6	3.949	3.985	3.926	4.092	3.413	3.452	3.367
7	3.807	3.842	3.785	3.944	3.192	3.225	3.156
8	3.589	3.622	3.568	3.717	3.038	3.073	3.001
9	3.503	3.528	3.488	3.607	2.931	2.960	2.906
10	3.903	3.915	3.900	3.961	3.266	3.280	3.265
				$\sigma_\varepsilon = 2$			
1	1.958	1.944	2.008	2.035	2.220	2.217	2.229
2	2.200	2.175	2.280	2.207	3.111	3.105	3.121
3	3.236	3.221	3.284	3.266	2.890	2.885	2.899
4	4.795	4.794	4.823	4.901	3.200	3.198	3.204
5	6.334	6.342	6.350	6.497	4.696	4.698	4.697
6	6.333	6.341	6.347	6.484	4.836	4.838	4.835
7	5.971	5.978	5.984	6.107	4.596	4.598	4.595
8	5.675	5.681	5.689	5.792	4.359	4.361	4.358
9	5.397	5.403	5.410	5.508	4.172	4.173	4.171
10	5.224	5.229	5.237	5.324	4.087	4.088	4.086
				$\sigma_\varepsilon = 3$			
1	3.172	3.177	3.168	3.222	4.430	4.415	4.459
2	3.456	3.455	3.459	3.477	5.061	5.048	5.087
3	5.283	5.251	5.321	5.211	6.934	6.921	6.961
4	7.320	7.320	7.331	7.361	9.907	9.897	9.933
5	8.008	8.014	8.010	8.066	11.199	11.189	11.225
6	7.621	7.624	7.625	7.670	10.966	10.957	10.992
7	7.171	7.172	7.178	7.215	10.380	10.371	10.405
8	6.794	6.796	6.802	6.836	9.846	9.837	9.870
9	6.514	6.518	6.520	6.563	9.339	9.331	9.362
10	6.260	6.262	6.267	6.305	8.993	8.985	9.015

Note: In columns w_{AIC}, $w_{eq.}$ are the *MSFE*s for forecasts combined using the Akaike weights and equal weights, and in columns min_{AIC}, m_{SEL} – the *MSFE*s for individual forecasts obtained from model selected by the minimum of AIC and model received after applying *a posteriori* selection method to model M1 at the 5% significance level.

Table 2. Mean square forecast errors (*MSFE*) for sample n = 50

	Causal models				Autoregressive models		
h	w_{AIC}	$w_{EQ.}$	min_{AIC}	$m_{SEL.}$	w_{AIC}	$w_{EQ.}$	min_{AIC}
				$\sigma_\varepsilon = 1$			
1	1.019	1.018	1.018	1.023	1.069	1.069	1.076
2	1.502	1.587	1.409	1.888	2.712	2.705	2.721
3	1.564	1.576	1.570	2.002	2.804	2.796	2.815
4	2.440	2.393	2.493	2.494	2.489	2.482	2.498
5	3.599	3.486	3.722	3.251	2.288	2.283	2.296
6	3.685	3.609	3.761	3.274	2.136	2.132	2.144
7	3.455	3.391	3.518	3.144	2.219	2.212	2.229
8	3.279	3.212	3.352	3.093	2.443	2.435	2.456
9	3.164	3.110	3.224	2.988	2.342	2.334	2.355
10	3.669	3.642	3.699	3.412	2.569	2.565	2.576
				$\sigma_\varepsilon = 2$			
1	3.522	3.508	3.539	4.230	5.419	5.399	5.414
2	3.854	3.803	3.918	4.644	5.739	5.711	5.744
3	4.199	4.100	4.318	4.900	5.669	5.641	5.677
4	4.832	4.671	5.018	5.261	5.483	5.458	5.489
5	5.379	5.222	5.555	5.668	5.679	5.660	5.685
6	5.493	5.324	5.683	5.616	5.356	5.340	5.362
7	6.133	5.955	6.329	5.949	5.273	5.262	5.280
8	6.171	6.016	6.346	5.890	5.139	5.130	5.146
9	6.311	6.195	6.456	6.048	5.362	5.357	5.369
10	6.286	6.189	6.411	6.034	5.380	5.376	5.387
				$\sigma_\varepsilon = 3$			
1	5.088	5.096	5.085	5.335	7.026	6.997	7.064
2	6.378	6.371	6.389	6.498	8.233	8.203	8.276
3	6.968	6.952	6.987	6.993	8.696	8.683	8.720
4	6.578	6.562	6.597	6.599	8.389	8.386	8.399
5	6.193	6.175	6.213	6.205	7.944	7.945	7.948
6	5.803	5.788	5.823	5.821	7.391	7.393	7.396
7	5.599	5.579	5.627	5.593	6.938	6.939	6.942
8	5.504	5.484	5.532	5.492	6.593	6.594	6.597
9	5.325	5.308	5.351	5.334	6.456	6.458	6.459
10	5.181	5.166	5.206	5.228	6.601	6.603	6.603

Note: See Table 1.

In the case of autoregressive models the forecasts from models selected by the AIC gave the lowest *MSFE* for small size of disturbance σ_ε=1 for the same reasons as in the case of causal models, i.e. the set of candidate models contained the small number of models (here AR(2), AR(3) and AR(4)) and one model had a dominating AIC weight ($0.5 < w_{AIC} < 0.6$). Hence, in the sense of the *MSFE*, the individual forecasts outperformed the combined forecasts. For the bigger size of disturbance, i.e. $\sigma_\varepsilon = 2$ (and bigger uncertainty) and at the shorter horizon ($h \le 5$) the dominance of forecasts combined using equal weights is observed, and at the longer horizons ($h \ge 6$) the dominance of forecasts combined using the AIC weights occurs. For the disturbance $\sigma_\varepsilon = 3$ the forecasts combined using equal weights slightly outperform the forecast combined using the AIC weights and individual forecast ($\min_{AIC}$) – see Table 1. Generally, the combined forecasts gave the lower *MSFE* than the individual forecasts. This refers to the cases when the set of models consisted of many autoregressive models of different order and none of them had the dominating AIC weight. Then, the lower *MSFE*s are received for combined forecasts.

The results tabulated in Table 2, for sample $n = 50$ indicate that for the causal models the *MSFE*s are lower for forecasts combined using equal weights than those using the *AIC* weights for all size of disturbance (except $\sigma_\varepsilon = 1$ and $h = 1, 2$). The dominance of combined forecasts (w_{AIC}, $w_{EQ.}$) or individual forecasts ($\min_{AIC}$, $m_{SEL.}$) depends on the forecast horizon and size of disturbance σ_ε. For horizon $h \le 4$ and disturbance $\sigma_\varepsilon = 1$ (also $\sigma_\varepsilon = 2$ and $h \le 7$; $\sigma_\varepsilon = 3$) the *MSFE*s for combined forecasts are lower than for forecasts from the best model selected by *a posteriori* method ($m_{SEL.}$), but for longer horizon $h \ge 5$ (for $\sigma_\varepsilon = 1$) and $h \ge 7$ (for $\sigma_\varepsilon = 2$) the *MSFE*s are lower for forecasts from $m_{SEL.}$. Forecasts from models selected by the minimum of AIC have the higher *MSFE*s than combined forecasts and in general also higher than forecasts from $m_{SEL.}$.

In the case of autoregressive models the *MSFE*s for combined forecasts are always lower than for individual forecasts ($\min_{AIC}$) – see Table 2. Simultaneously the combined forecasts using equal weights ($w_{EQ.}$) outperform those using the AIC weights (w_{AIC}).

5. Summary

From comparison of forecasts combined using the AIC weights and equal weights results that in the case when the set of candidate models contains the model with dominating AIC weight ($w_{AIC} > 0.7$) the combination of forecasts using the AIC weights is not effective. Then the *MSFE*s are higher than those for forecasts combined using equal weights. However, in such case the AIC weights can be useful in building ranking of models, and additionally in calculating the evidence ratios informing about the relative strength of evidence for the best model (with biggest weight) in the sense of AIC over the other models in the set of candidate models. The benefits from applying the AIC weights occur when the number of candidate models in the set is big and none has the dominating weight w_{AIC}.

The results of experiment indicate that the combined forecasts outperform the individual forecasts in the case of autoregressive models. For causal models this dominance of combined forecasts is hold at the shorter horizon for disturbance $\sigma_\varepsilon = 1, 2$ and at the whole horizon for bigger size of disturbance $\sigma_\varepsilon = 3$.

Summing up, the information criteria approach, particularly the use of AIC weights to build the ranking of models and to calculate the combined forecasts, can be treated as alternative to the traditional hypothesis testing approach directed to select the best model and calculate individual forecasts.

References

Akaike, H. (1973), Information Theory as an Extension of the Maximum Likelihood Principle, [in:] Petrov, B. N., Csaki, F., *Second International Symposium on Information Theory*, Akademia Kiado, Budapest.

Akaike, H. (1978), On the Likelihood of a Time Series Model, *The Statistician*, 27, 217–235.

Armstrong, J. S. (2001), *Principles of Forecasting*, Springer.

Atkinson, A. C. (1980), A Note on the Generalized Information Criteria for Choice of a Model, *Biometrika*, 67 (2), 413–418.

Bates, J. M., Granger, C. W. J. (1969), The Combinations of Forecasts, *Operations Research Quarterly*, 20, 415–468.

Burnham, K. P., Anderson, D. R. (2002), *Model Selection and Multimodel Inference*, Springer.

Burnham, K. P., Anderson, D. R. (2004), Multimodel Inference. Understanding AIC and BIC in Model Selection, *Sociological Methods and Research*, vol. 33 (2), 261–304.

Jagannathan, R. Ma, T. (2003), Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps, *The Journal of Finance*, 58 (4), 1651–1684.

Kapetanios, G., Labhard, V., Price, S. (2008), Forecasting using Bayesian and Information-theoretic Model Averaging: an Application to U.K. Inflation, *Journal of Business and Economics Statistics*, 26 (1), 33–41.

Kitchen, J., Monaco, R. (2003), Real-Time Forecasting in Practice, *Business Economics*, 38 (4), 10–19.

Marcellino, M. (2004), Forecast Pooling for Short Time Series of Macroeconomic Variables, *Oxford Bulletin of Economic and Statistics*, 66, 91–112.

Min, C. K., Zellner, A. (1993), Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates, *Journal of Econometrics*, 53 (1–2), 89–118.

Stock, J. H., Watson, M. (2004), Combination Forecasts of Output Growth in a Seven-Country Data Set, *Journal of Forecasting*, 8, 230–251.

Stock, J. H., Watson, M. (2006), Forecasting with Many Predictors, [in:] Elliott, G., Granger, C. W. J., Timmermann, A. (ed.), *Handbook of Economic Forecasting*, Elsevier.

Swanson, N. R., Zeng, T. (2001), Choosing Among Competing Econometric Forecasts: Regression-based Forecast Combination using Model Selection, *Journal of Forecasting*, 20, 425–440.

Timmermann, A. (2006), Forecast Combinations, [in:] Elliott G., Granger C. W. J., Timmermann A. (ed.), *Handbook of Economic Forecasting*, ch. 4, Elsevier.

Prognozy kombinowane z wykorzystaniem wag Akaike'a

Z a r y s t r e ś c i. W artykule uwaga jest skupiona na podejściu wykorzystującym kryteria informacyjne, a w szczególności kryterium Akaike'a, które jest wykorzystywane do wyznaczenia wag Akaike'a. Podejście to umożliwia otrzymanie nie jednego, a kilku wiarygodnych modeli, dla których można stworzyć ranking stosując wagi Akaike'a. Modele te stanowią podstawę obliczenia prognoz indywidualnych, a te z kolei służą do wyznaczenia ostatecznej prognozy kombinowanej, przy formułowaniu której wykorzystuje się wagi Akaike'a.

Słowa kluczowe: prognozy kombinowane, systemy wag, kryteria informacyjne.