DYNAMIC ECONOMETRIC MODELS
Vol. 11 – Nicolaus Copernicus University – Toruń – 2011

Mariola Piłatowska
Nicolaus Copernicus University in Toruń

Information and Prediction Criteria in Selecting the Forecasting Model

A b s t r a c t. The purpose of the paper it to compare the performance of both information and prediction criteria in selecting the forecasting model on empirical data for Poland when the data generating model is unknown. The attention will especially focus on the evolution of information criteria (AIC, BIC) and accumulated prediction error (APE) for increasing sample sizes and rolling windows of different size, and also the impact of initial sample and rolling window sizes on the selection of forecasting model. The best forecasting model will be chosen from the set including three models: autoregressive model, AR (with or without a deterministic trend), ARIMA model and random walk (RW) model.

K e y w o r d s: information and prediction criteria, accumulated prediction error, model selection.

Introduction

The model selection literature has recently emphasized the necessity of considering the choice of model depending on the purpose of econometric modeling. In modeling approach the two aims are mentioned the most frequently, namely searching 'true' model and selecting the best forecasting model (optimizing prediction).

That first aim of modeling is hard to realize because the economic reality is seen as a complex and dynamically evolving structure whose mechanism is hidden and almost impossible to uncover. Therefore the model is an approximation (or simplification) of reality which represents the relevance of a particular phenomenon. It is advocated to assume that each model is not true by definition (Taub, 1993; deLeew, 1998) or that "all models are wrong, but some are useful" (Box, 1976). Having in mind that none of models cannot reflect all of reality, the debate concerning true models should be completed because it seems to be unproductive. Hence, the second aim of modeling, i.e. selecting the best forecasting model acquires relevance from the practical point of view.

In predictive approach the goal of selecting the true model is abandoned and the attention focuses on seeking a model with as small predictive errors as possible. It should be emphasized that in forecasting situation the misspecified model are allowed because such a model may yield excellent forecasts. On the other hand good prediction is treated as a test of any subsidiary aim, i.e. if the purpose of an analysis is to estimate parameters, then the best estimated model should give the best prediction; if the purpose of an analysis is hypothesis testing, then any rejected model should give worse forecasts than any accepted model (Clarke, 2001; De Luna, Skouras, 2003; Kunst, 2003).

To select the forecasting model the different model selection methods can be used, for instance information and prediction criteria. However, question may arise whether the performance of both criteria is the same with regard to the choice of model. It has been shown on simulated data[1] (Kunst, 2003) that information criteria should be rather used if one is interested in finding 'true' model. If the purpose of analysis is to choose a forecasting model, the prediction criteria are preferred because they select the model yielding the smallest prediction error, although sometimes it may be an incorrect choice (not true model). However, in economic reality the true model is unknown, therefore it is worth checking the performance of information and predictiion criteria in practical context (empirical data).

The purpose of the paper it to compare the performance of both information and prediction criteria in selecting the forecasting model on empirical data when the data generating model is unknown. The attention will especially focus on the evolution of information criteria (AIC, BIC) and accumulated prediction error (APE) for increasing sample sizes and rolling windows of different size, and also the impact of sample and rolling window sizes on the selection of forecasting model. The best forecasting model will be chosen from the set including three models: autoregressive model, AR (with or without a deterministic trend), ARIMA model and random walk (RW) model on the basis of empirical data for Poland. The choice of model will be carried out using information (AIC, BIC) and prediction (APE and MSE, MAPE, U) criteria. The decision of selecting a model by information and prediction criteria is checked in out-of-sample forecasting by comparing accuracy measures for given forecast models.

[1] The true data were generated from ARMA(1, 1) models with 100+100+10 observations (first 100 observations were discarded; 1000 replication were conducted). The set of models included: ARMA(1, 1), AR(1) and MA(1) models. To select a model the AIC information criterion and the mean squared error (MSE) based on prediction error from 10 one-step-ahead forecasts were applied (see Kunst, 2003).

1. Information and Prediction Criteria

Generally, information criterion takes the form:

$$IC = -2\ln L(\hat{\theta}) + q,$$

where $L(\hat{\theta})$ is the likelihood function, and q is a penalty term that is a function of the number of parameters K and the number of observations n; this penalty guards against overfitting, i.e. using too many parameters. For Akaike's information criterion (AIC) the penalty is equal to $q = 2K$, for Schwartz (Bayes) information criterion − $q = K\ln(n)$, for Hannan-Quinn information criterion − $q = 2K\ln(\ln n)$.

For small samples the AIC criterion is biased and may suggest a model with a high number of parameters compared with the number of observations $(n/K \le 40)$. Thus a bias-corrected version, AIC_c, is increasingly used. The latter is given by adding the quantity $2K(K+1)/(n-K-1)$ to the ordinary AIC. The *BIC* (like the AIC_c criterion) penalizes the addition of extra parameters more severely than the *AIC*, and should be preferred to the ordinary *AIC* in time-series analysis especially when the number of parameters is high compared with the number of observations.

Applying information criteria in model selection the model with the minimum of a given information criterion is chosen.

To traditional prediction criteria, used both in the accuracy evaluation and selection of forecasting model, belong:

Mean Absolute Error	$MAE = \dfrac{\sum_{t=1}^{T} \lvert e_t \rvert}{T},$
Mean Square Error	$MSE = \dfrac{\sum_{t=1}^{T} e_t^2}{T},$
Root Mean Square Error	$RMSE = \sqrt{MSE},$
Mean Absolute Percentage Error	$MAPE = \dfrac{\sum_{t=1}^{T} \lvert e_t / y_t \rvert}{T} 100\%,$
Theil's Inequality Coefficient	$U = \dfrac{RMSE(\text{'new' model})}{RMSE(\text{'benchmark' model})},$

where e_t denotes prediction error, $e_t = y_t - \hat{y}_t$, y_t − realization of y in period t, $\hat{y}_t$ − forecast of y for period t.

Applying usual accuracy measures the model with the smallest value of given measure is selected what corresponds to the smallest prediction error.

The Theil's inequality coefficient U indicates whether a given model is worse (U > 1) or better (U < 1) than the random walk model ($\hat{y}_{t+1} = y_t$) considered as a benchmark model.

It is worth highlighting that the choice of accuracy measure can affect the ranking of forecasting methods, and also models (Armstrong, 2001; Armstrong, Fildes, 1995). For instance, the MSE depends on the scale in which the variable is measured. This means that the MSE is appropriate only for assessing the results for a single time series, and should be avoided to assess accuracy across (many) different series. In that case the scale-independent measures are required[2], e.g. MAPE or Theil's inequality coefficient (U). Besides, for the reason that the loss function may be asymmetric (e.g. underforecasting is worse than overforecasting), it may be important to forecast the direction of movement or to predict large movements (Chatfield, 2000). However, it is not possible to select a measure of forecast accuracy that is scale-independent and yet satisfies the demands of the appropriate loss function. In conclusion, there is no measure suitable for all types of data and all contexts. There are many empirical evidence that a method which is 'best' under one criterion need not be 'best' under alternative criteria (e.g. Swanson and White, 1997).

The choice of forecasting model may also be carried out by the accumulated prediction error, APE, (Rissanen, 1986). According to the APE the most useful model is the model with the smallest out-of-sample one-step-ahead prediction error.

The APE method proceeds by calculating sequential one-step-ahead forecasts based on gradually increasing sample. For model M_j the APE is calculated as follows (Wagenmaker, Grünwald, Steyvers, 2006):

1. Determine the smallest number s of observations that makes the model identifiable. Set $i = s + 1$, so that $i - 1 = s$.
2. Based on the first $i - 1$ observations, calculate a prediction $\hat{p}_i$ for the next observation i.
3. Calculate the prediction error for observation i, e.g. squared difference between the predicted value $\hat{p}_i$ and the observed value x_i.
4. Increase i by 1 and repeat steps 2 and 3 until $i = n$.
5. Sum all of the one-step-ahead prediction errors as calculated in step 3. The result is the APE.

For model M_j the accumulated prediction error is given by:

$$APE(M_j) = \sum_{i=s+1}^{n} d[x_i, (\hat{p}_i | x_{i-1})],$$

[2] Mentzer and Kahn (1995) found in a survey of 207 forecasting executives in US that MAPE was the most commonly used measure (52%) while only 10% used MSE.

where d indicates the specific loss function that quantifies the discrepancy between observed and predicted values.

In the case of point predictions one typically uses the squared error $(x_i - \hat{p}_i)^2$, but another choice would be to compute the absolute value loss $|x_i - \hat{p}_i|$, or more generally, an α-loss function, $|x_i - \hat{p}_i|^\alpha$, $\alpha \in [1, 2]$ (Rissanen, 2003).

2. Selection of Forecasting Model – Empirical Examples

To compare the performance of information and prediction criteria in selecting the best forecasting model the monthly data on consumer price index CPI (corresponding period of previous year =100) and industry production IP (in billion PLN zl) in Poland were used[3] (in the period 2002:01−2010:12, 108 observations, data are seasonally adjusted). The reason for such selection of time series is the desire to check the performance of selection criteria with regard to time series of different properties, i.e. in the above case CPI is expected to be rather an integrated process, and IP − rather a stationary process around deterministic trend[4].

The set of candidate models for CPI includes: autoregressive model, AR(12), ARIMA(12,1,0) model and random walk model (RW) as a benchmark model; for industry production IP this set was as follows: linear trend model with autoregression of twelfth order (further denoted as AR for the convenience of presentation), ARIMA(12,1,0) and random walk model as a benchmark model[5].

Two versions of estimation procedures are considered:

version I: the models are iteratively estimated beginning with the initial sample size $n(s)$ which is being increased by one until n = 108 (until 2010:12); three sizes of initial sample $n(s)$ are taken: 40, 60, 80;

version II: the models are iteratively estimated for rolling window size of 40, 60 and 80 observations.

The question is to what extent the size of initial sample and rolling window have the impact on the choice of forecast model when the information criteria

[3] Data have been taken from the Statistical Bulletin of the Central Statistical Office in Poland.

[4] Many empirical studies conclude that the consumer price index (CPI) as a financial series is rather integrated and then the ARIMA model is more appropriate than the AR model (with or without deterministic trend). Whereas, the industry production (IP) is treated rather as a stationary process around deterministic trend and then the AR model (with or without deterministic trend) is often taken as a more correct model. However, it may occur in forecasting that an inappropriate model will yield better forecasts. To take into account this possibility, both models are used in empirical study.

[5] The order of autoregression was fixed at 12 as the potentially highest order reflecting monthly frequency of data.

(AIC, BIC) and prediction criteria (APE_SE, APE_AE) are used as selection criteria. Notations APE_SE and APE_AE denote the accumulated prediction error (APE) that uses the squared error and absolute error respectively as a loss function.

To make a choice of best forecasting model the information criteria (AIC and BIC) and prediction criteria (APE_SE and APE_AE) were calculated at each iteration. The results are presented in Figures 1−8 as difference in a given criterion for pairs of models, i.e. AIC(M_i)−AIC(M_j), BIC(M_i)−BIC(M_j), APE_SE(M_i)−APE_SE(M_j), APE_AE(M_i)−APE_AE(M_j), and also in tables presenting the choices of forecasting model for all pairs of models. These differences in selection criteria are interpreted as follows: the positive differences favor the second model over the first one in a pair of models (this means either a lower value of information criterion or smaller prediction error for the second model), and the negative differences indicate that the first model outperforms the second one. The sign of differences in criteria may change in time which indicates that one model has become outdated. However, from the forecasting point of view the most important is the sign of differences in criteria at the end of studied period, therefore the choice of the best forecasting model has been made basing on the performance of differences in criteria at the end of sample (at least three observations with the same sign of difference in a given criterion).

Figure 1 (row 1) demonstrates that independently of initial sample size the performance of differences in AIC criterion for pairs of models is similar, i.e. the AIC criterion favors the AR model over the ARIMA and RW models, and the ARIMA model is better in sense of AIC criterion than the RW model. However, the results for different rolling window sizes are different (Fig. 1, row 2). While the dominance of the AR model over ARIMA model is maintained, the AIC criterion prefers the RW model over AR and ARIMA model (differences in AIC for pairs of models, AIC(ARIMA)−AIC(RW) and AIC(AR)−AIC(RW), are positive) what is opposite to results obtained for increasing by one sample size (Fig. 1, row 1). The different results for window size of 80 observations in comparison with those obtained for window size of 40 and 60 observations could suggest the influence of window size on a choice of model by the AIC criterion, but in that case it is rather the problem of too large window size with regard to the number of observations.

The results of model selection for consumer price index (CPI) by the BIC criterion seem to be more stable and insensitive as well to the initial sample size (Fig. 2, row 1) as rolling window size (Fig. 2, row 2) in comparison with the selection by the AIC criterion. The BIC criterion favors the RW model over the AR and ARIMA models. As previously the AR is preferred over the ARIMA model.

Summing up, the choices by information criterion differ, i.e. the AIC criterion prefers the AR model for CPI in the case of increasing sample size (version

I), and model RW − in the case of rolling window size, version II (except the window size of 80 observations), while the BIC criterion prefers the RW model in both version.

Figures 3 and 4 demonstrate the differences in prediction criteria APE_SE and APE_AE respectively. The APE_SE criterion favors the RW model over the AR and ARIMA model (Fig. 3, row 1 and 2) except the initial sample size and rolling window size of 80 observations when the ARIMA model is preferred over the RW and AR models. The lack of support for the RW model in sample of 80 observations shows rather the influence of initial sample size and rolling window size on selecting the model. Hence, the size of initial sample or rolling window should not be to large with regard to total number of observations.

Observing the differences in APE_AE the influence of the size of initial sample and rolling window is much more distinct (Fig. 4, row 1 and 2). For initial sample size and window size of 40 observations the APE_AE criterion prefers the RW model over the ARIMA and AR models. However, for sample (or window) of 60 observations this criterion favors the ARIMA model over the AR and RW models, and for sample (or window) of 80 observations – the AR model over the ARIMA and RW models.

Table 1. Results of model selection for CPI in Poland using information (AIC, BIC) and prediction criteria (APE_SE, APE_AE)

Pairs of models	Selection criteria	Version I			Version II		
		n=40	n=60	n=80	n=40	n=60	n=80
ARIMA vs. AR	AIC	AR	AR	AR	AR	AR	AR
	BIC	AR	AR	AR	AR	AR	AR
	APE_SE	ARIMA	ARIMA	ARIMA	AR	ARIMA	ARIMA
	APE_AE	AR	ARIMA	AR	AR	ARIMA	AR
ARIMA vs. RW	AIC	ARIMA	ARIMA	ARIMA	RW	RW	ARIMA
	BIC	RW	RW	RW	RW	RW	RW
	APE_SE	RW	RW	ARIMA	RW	RW	ARIMA
	APE_AE	RW	ARIMA	ARIMA	RW	ARIMA	ARIMA
AR vs. RW	AIC	AR	AR	AR	RW	RW	AR
	BIC	RW	RW	RW	RW	RW	RW
	APE_SE	RW	RW	AR	RW	RW	AR
	APE_AE	RW	AR	AR	RW	AR	AR

Table 2. Accuracy measures for one-step-ahead forecasts of CPI from different models in the period 2011:01−20011:06 − version I

Accuracy measures	Models		
	ARIMA	AR	RW
MSE	0.34	0.35	0.26
RMSE	0.58	0.59	0.51
U	1.31	1.36	1.00
MAPE (%)	0.48	0.46	0.41

Table 3. Accuracy measures for one-step-ahead forecasts of CPI from different models in the period 2011:01−2011:06 − version II

Accuracy measures	n=40			n=60			n=80		
	ARIMA	AR	RW	ARIMA	AR	RW	ARIMA	AR	RW
MSE	0.314	0.292	0.265	0.305	0.365	0.276	0.297	0.342	0.279
RMSE	0.560	0.540	0.515	0.552	0.604	0.526	0.545	0.585	0.528
U	1.183	1.099	1.000	1.104	1.322	1.000	1.063	1.224	1.000
MAPE	0.443%	0.434%	0.423%	0.453%	0.470%	0.421%	0.444%	0.450%	0.432%

Generally, the RW model should be chosen as the best model for CPI because it is preferred by all criteria in the case of rolling windows (except window of 80 observations for APE_SE and APE_AE and window of 60 observations for APE_AE which seem rather too large with regard to number of observations) and also in the case of increasing sample size except the AIC criterion which favored the AR model and prediction criteria for initial sample size of 80 observations (for both criteria APE) and of 60 observations (for APE_AE) − see Figure 1−4 and Table 1.

If the RW model is really the best one, its predictive performance should be also confirmed in out-of-sample evaluation. Out-of-sample forecast evaluation (i.e. in the period 2011:01−2011:06) has been realized by the measures of accuracy (MSE, RMSE, U, MAPE) − see Table 2 and 3.

The results in Table 2 and 3 indicate that one-step-ahead forecasts of CPI made from the RW model have the smallest prediction errors independently of the versions (iteratively increasing sample size by one − version I, or rolling window of given size − version II). The out-of-sample performance of RW model is in rough agreement with choices received by APE criteria but also BIC criterion, thus providing evidence of correct model selection by these criteria.

The results of selecting the best model for industry production (IP) in Poland are presented in Figures 5−8. Figure 5 shows that the AIC criterion favors the AR model over the ARIMA and RW models independently of initial sample size and rolling windows. Hence, in that case the selection of a model by the AIC criterion seems to be insensitive to the size of initial sample and rolling windows. When the BIC criterion has been used the results were similar but only in version I (initial sample size increased by one) − see Figure 6, row 1 and for window size of 80 observations (version II, see Figure 6, row 2). For window sizes of 40 and 60 observations the BIC criterion prefers the RW model over the ARIMA and AR model (Figure 6, row 2).

The choice of model of industry production (IP) in Poland by prediction criteria is different than by information criteria. Namely, the APE_SE criterion favors the ARIMA model over the AR and RW model except the initial sample of 80 observations (version I) and window size of 80 observations (version II) − Figure 7. This dominance of model ARIMA is maintained when the APE_AE criterion is used to select the best model (except the initial sample size of 60 observations, version I) − Figure 8.

Summing up, according to information criteria (AIC, BIC) the AR model should be chosen as the best model for industry production, IP, (except the choice of BIC criterion for rolling window of 40 and 60 observations when the RW model is preferred) − see Figure 5−6 and Table 4. Whereas the choice of prediction criteria (APE_SE, APE_AE) is the ARIMA model (except initial sample of 60 observations) − see Figure 7-8 and Table 4.

Table 4. Results of model selection for IP in Poland using information (AIC, BIC) and prediction criteria (APE_SE, APE_AE)

Pairs of models	Selection criteria	Version I			Version II		
		n=40	n=60	n=80	n=40	n=60	n=80
ARIMA vs. AR	AIC	AR	AR	AR	AR	AR	AR
	BIC	AR	AR	AR	AR	AR	AR
	APE_SE	ARIMA	ARIMA	AR	ARIMA	ARIMA	AR
	APE_AE	ARIMA	AR	ARIMA	ARIMA	ARIMA	AR
ARIMA vs. RW	AIC	ARIMA	ARIMA	ARIMA	ARIMA	ARIMA	ARIMA
	BIC	ARIMA	ARIMA	ARIMA	RW	RW	ARIMA
	APE_SE	ARIMA	ARIMA	ARIMA	ARIMA	ARIMA	ARIMA
	APE_AE	ARIMA	ARIMA	ARIMA	ARIMA	ARIMA	ARIMA
AR vs. RW	AIC	AR	AR	AR	AR	AR	AR
	BIC	AR	AR	AR	RW	RW	AR
	APE_SE	AR	AR	AR	RW	AR	AR
	APE_AE	AR	AR	AR	AR	AR	AR

Out-of-sample evaluation of IP forecasts (i.e. in the period 2011:01−2011:06) has been realized by the measures of accuracy (MSE, RMSE, U, MAPE) − see Table 5 and 6.

The results in Table 5 and 6 show that both the ARIMA model and AR model have similar predictive value because the accuracy measures of one-step-ahead forecasts of IP do not differ much, so these model may complete with each other. In case of version I (iteratively increasing sample size by one) the RMSE and U measure prefer the ARIMA model as a model with the smallest prediction error but the MAPE indicates the AR model (Table 5). The opposite result is obtained in version II for rolling window of 40 observations (Table 6).

Table 5. Accuracy measures for one-step-ahead forecasts of IP from different models in the period 2011:01−20011:06 − version I

Accuracy measures	Models		
	ARIMA	AR	RW
RMSE	2673.30	2679.50	3390.70
U	0.622	0.624	1.00
MAPE (%)	2.12	2.03	3.52

Table 6. Accuracy measures for one-step-ahead forecasts of IP from different models in the period 2011:01−2011:06 − version II

Accuracy measures	n=40			n=60			n=80		
	ARIMA	AR	RW	ARIMA	AR	RW	ARIMA	AR	RW
RMSE	2916.6	2908.4	3411.8	2716.42	2863.65	3386.73	2787.6	2697.1	3395.6
U	0.731	0.727	1.000	0.643	0.715	1.000	0.674	0.631	1.00
MAPE (%)	2.390	2.460	3.510	2.12%	2.210	3.500	2.220	1.990	3.520

For the window of 60 and 80 observations the ARIMA model and AR model give the smallest prediction errors. It is worth noting that the ARIMA and AR models substantially outperform the RW model. On the whole, the predictive performance of ARIMA and AR models is in agreement with choices obtained by APE criteria and AIC criterion.

Conclusions

The results of choosing the forecasting model on empirical data for Poland showed that the size of initial sample (version I) and rolling windows (version II) do have an impact on the choice of forecasting model (within a given version), especially it concerns the APE. The size of initial sample and rolling windows should be relatively small with regard to sample size because too large initial sample (or window) enables to follow the evolution of APE for sufficient number of periods. There are no relevant differences between the selection of forecasting model for the initial sample and rolling windows (within the comparative number of observations). Therefore it seems sufficient to calculate the APE only for some small initial sample size increasing iteratively by one observation. Both prediction and information criteria are useful in selecting the forecasting model, however the choice of models by prediction criteria is supposed to be much more in agreement with their best out-of-sample performance.

References

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

Armstrong, J. S., Fildes, R. (1995), *On the Selection of Error Measures for Comparisons Among Forecasting Methods*, *Journal of Forecasting*, 14, 67−71.

Box, G. E. P. (1976), Science and Statistics, *Journal of the American Association*, 71, 791−799.

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

Chatfield, C. (2000), *Time-series Forecasting*, Chapman & Hall/CRC, Boca Raton-London-New York-Washington.

Clarke, B. (2001), Combining Model Selection Procedures for Online Prediction, *Sankhya: The Indian Journal of Statistics*, 63, Series A, 229−249.

De Leeuw, J. (1988), Model Selection in Multinomial Experiments, in: T. K. Dijkstra (ed.), *On Model Uncertainty and its Statistical Implication*, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, New York.

Dawid, A. P. (1984), Statistical Theory: the Prequential Approach, *Journal of Royal Statistical Society Series B*, 147, 278−292.

De Luna, X., Skouras, K. (2003), Choosing a Model Selection Strategy, *Scandinavian Journal of Statistics*, 30, 113−128.

Kunst, R. M. (2003), Testing for Relative Predictive Accuracy: A Critical Viewpoint, *Economics Series*, 130, 1−45, Department of Economics and Finance, Vienna.

Mentzer, J. T., Kahn, K. B. (1995), Forecasting Technique Familiarity, Satisfaction, Usage, and Application, *Journal of Forecasting*, 14, 465−476.

Rissanen, J. (1986), Order estimation by Accumulated Prediction Errors, *Journal of Applied Probability*, 23A, 55-61. Skouras, K., Dawid, A. P. (1998), On Efficient Point Prediction Systems, *Journal of Royal Statistical Society B*, 60, 765−780.

Swanson, N. R., White, H. (1997), Forecasting Economic Time Series Using Flexible Versus Fixed Specification and linear Versus Nonlinear Econometric Models, *Journal of Forecasting*, 13, 439−461.

Taub, F. B. (1993), Book Review: Estimating Ecological Risks, *Ecology*, 74, 1290−1291.

Wagenmaker, E-J., Grünwald, P., Steyvers, M. (2006), Accumulative Prediction Error and the Selection of Time Series Models, *Journal of Mathematical Psychology*, 50, 149−166.

Kryteria informacyjne i predykcyjne w wyborze modelu prognostycznego

Z a r y s t r e ś c i. Celem artykułu jest porównanie zachowania się kryteriów informacyjnych i predykcyjnych w wyborze modelu prognostycznego na podstawie danych empirycznych dla Polski, przy założeniu nieznajomości modelu generującego dane. Uwaga będzie poświęcona śledzeniu zmian kryteriów informacyjnych (AIC, BIC) oraz skumulowanego błędu prognoz (APE) dla próby powiększanej iteracyjnie o jedną obserwację i ruchowych okien (o różnej wielkości), a także ocenie wpływu wielkości próby (startowej) i ruchomego okna na wybór modelu prognostycznego. Wybór najlepszego modelu prognostycznego jest dokonywany spośród następującego zestawu modeli: model autoregresyjny (AR, z trendem i bez trendu deterministycznego), model ARIMA, model błądzenia przypadkowego (RW).

S ł o w a k l u c z o w e: kryteria informacyjne, kryteria predykcyjne, skumulowany błąd predykcji, wybór modelu.

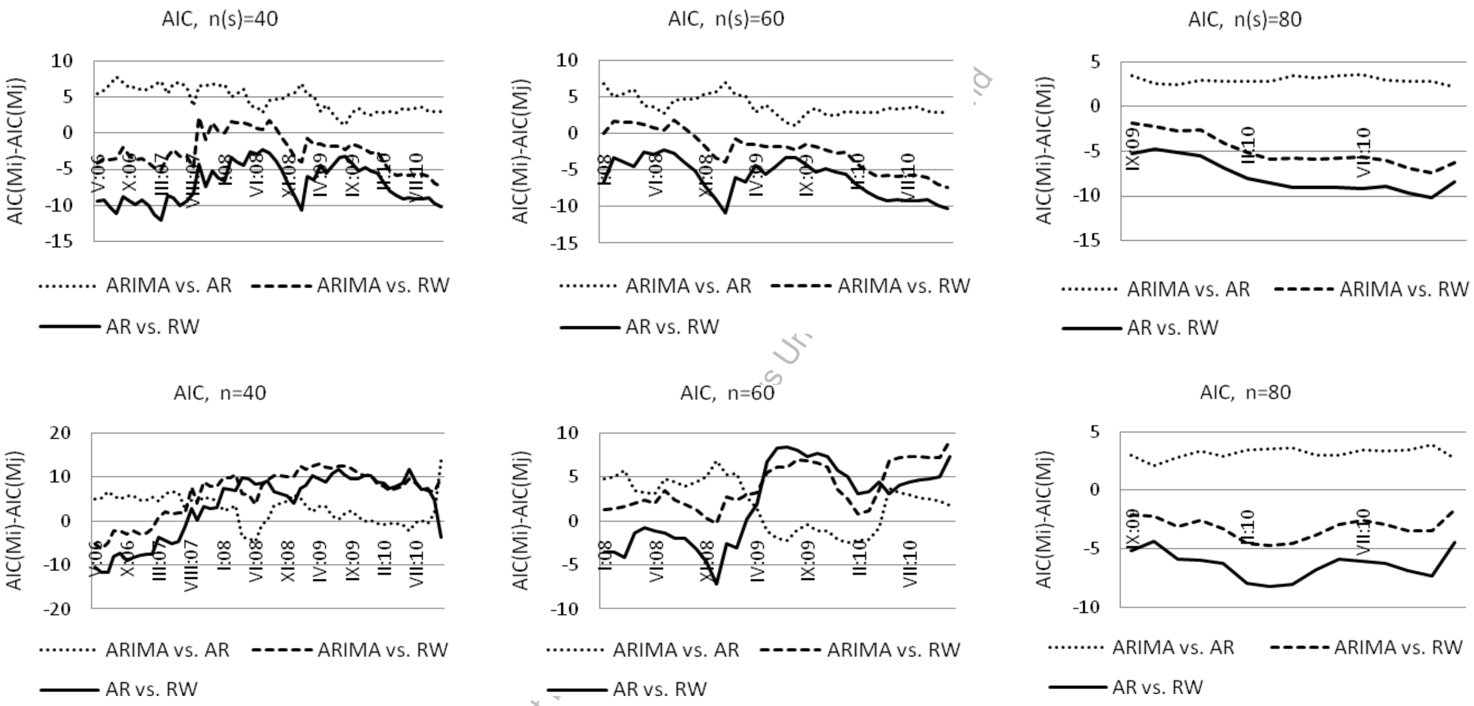

Figure 1. Differences in AIC information criterion (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for CPI in Poland

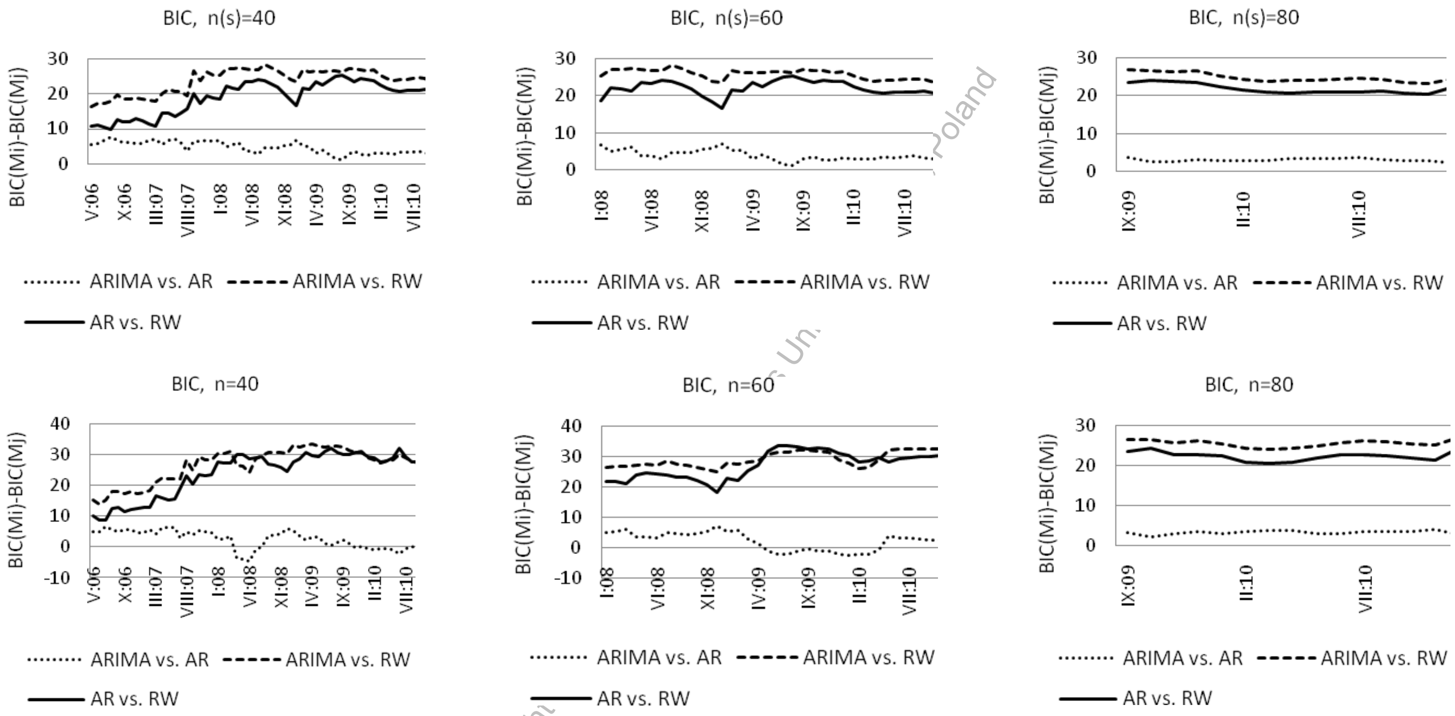

Figure 2. Differences in BIC information criterion (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for CPI in Poland

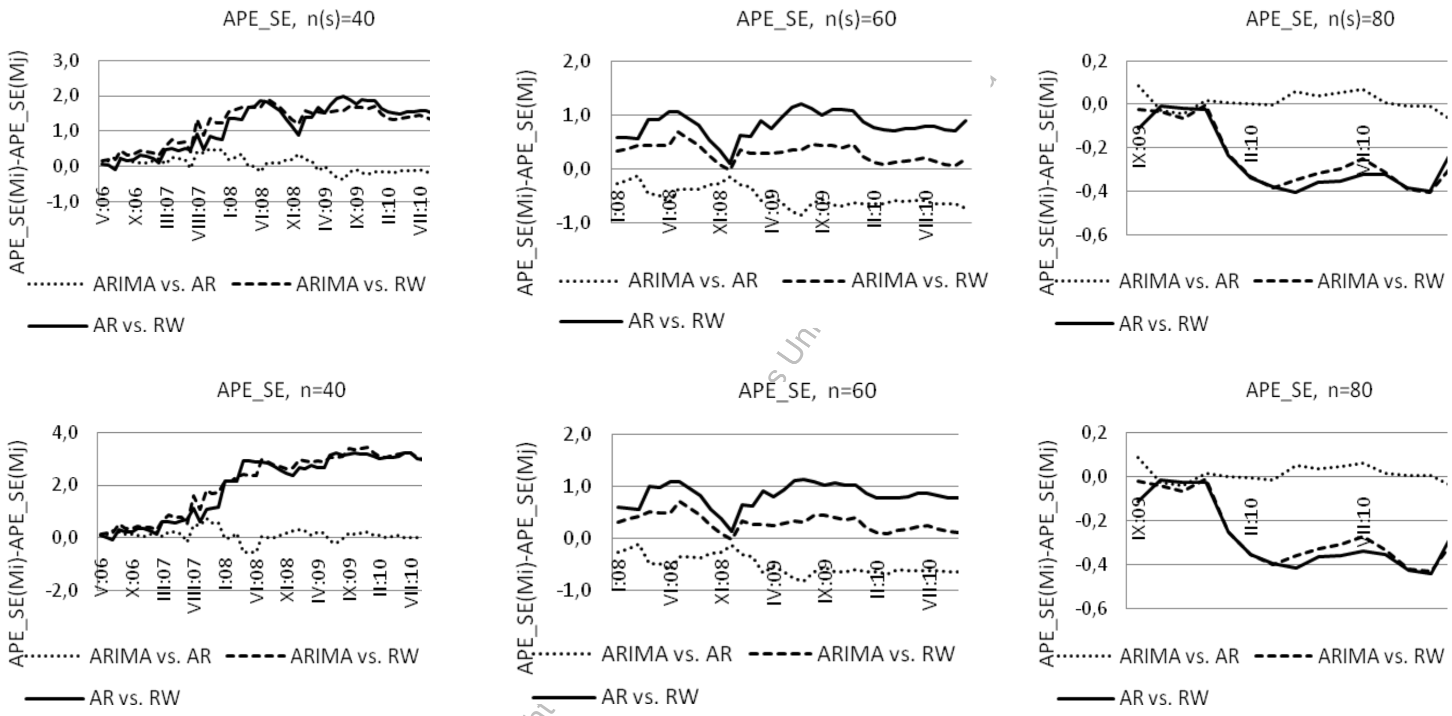

Figure 3. Differences in prediction criterion APE_SE (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for CPI in Poland

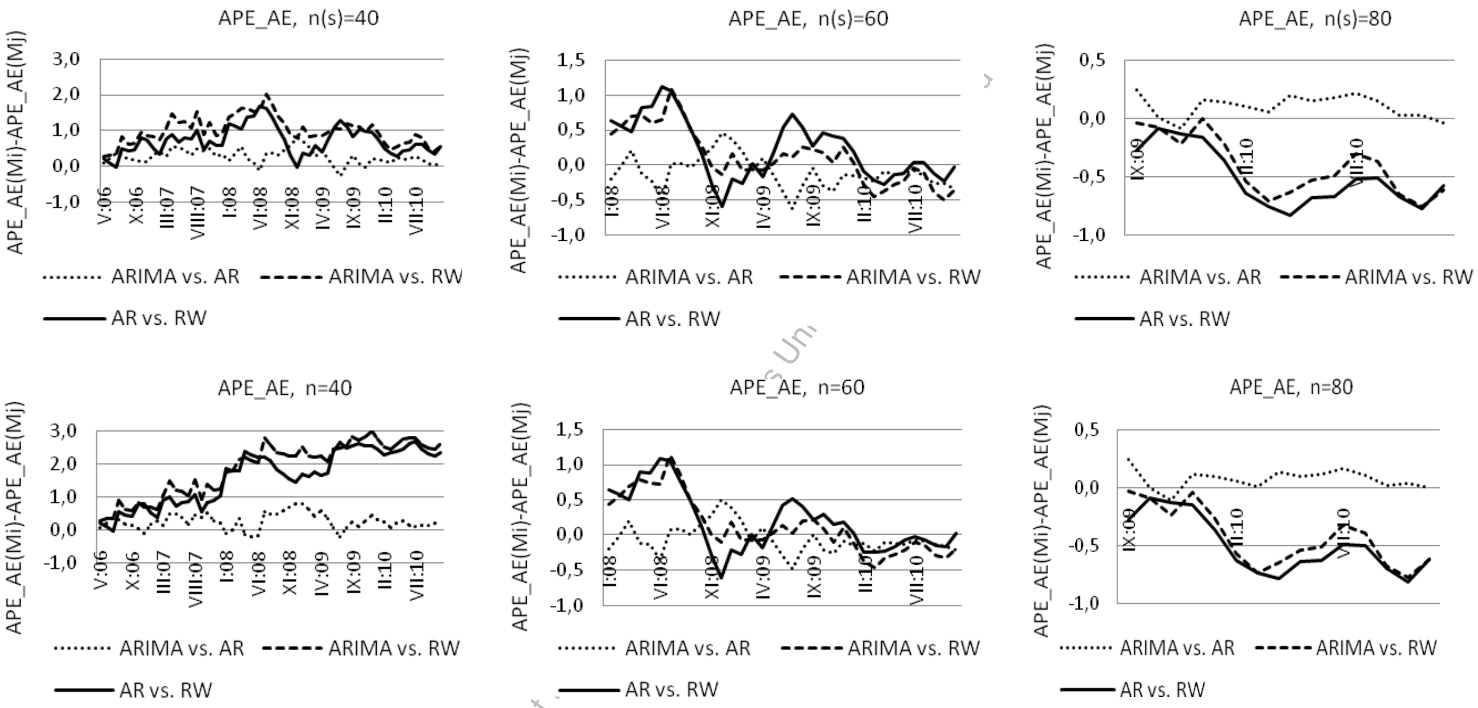

Figure 4. Differences in prediction criterion APE_AE (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for CPI in Poland

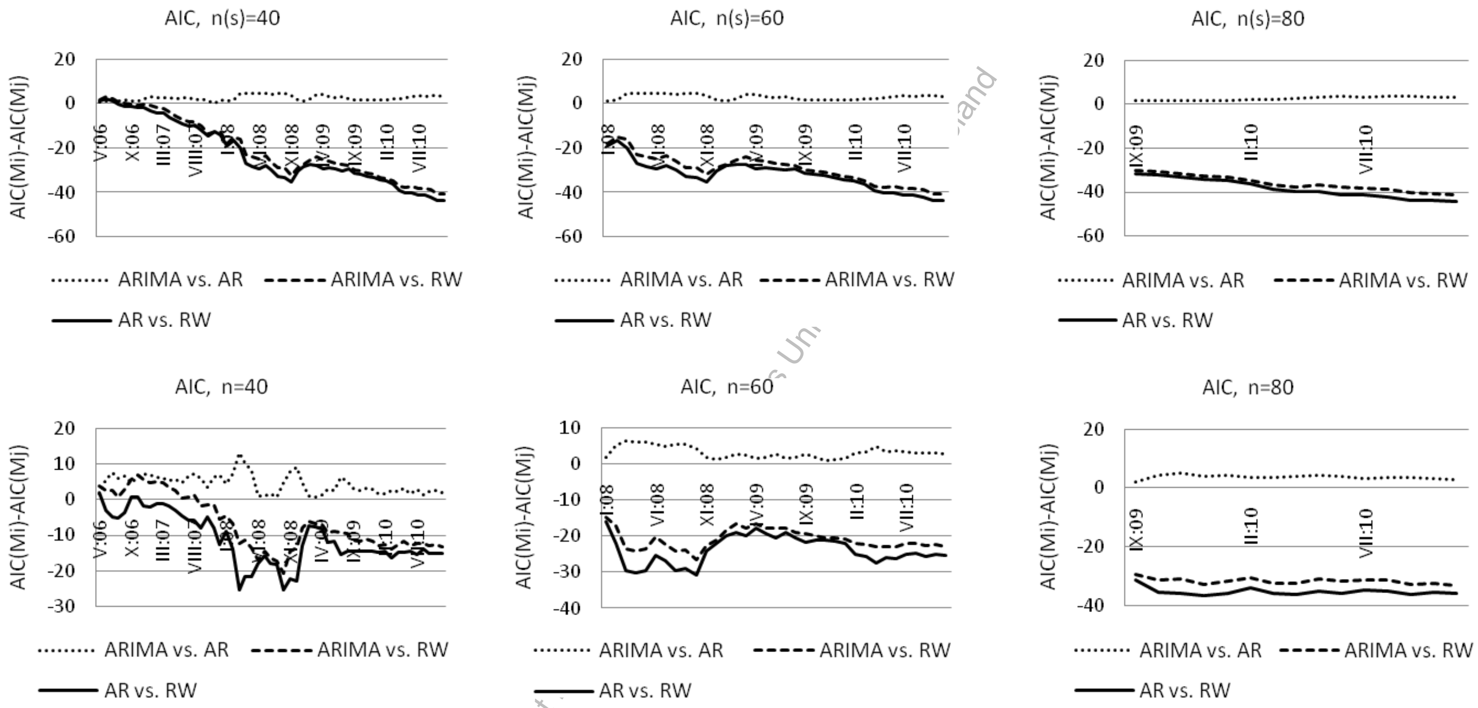

Figure 5. Differences in AIC criterion (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for IP in Poland

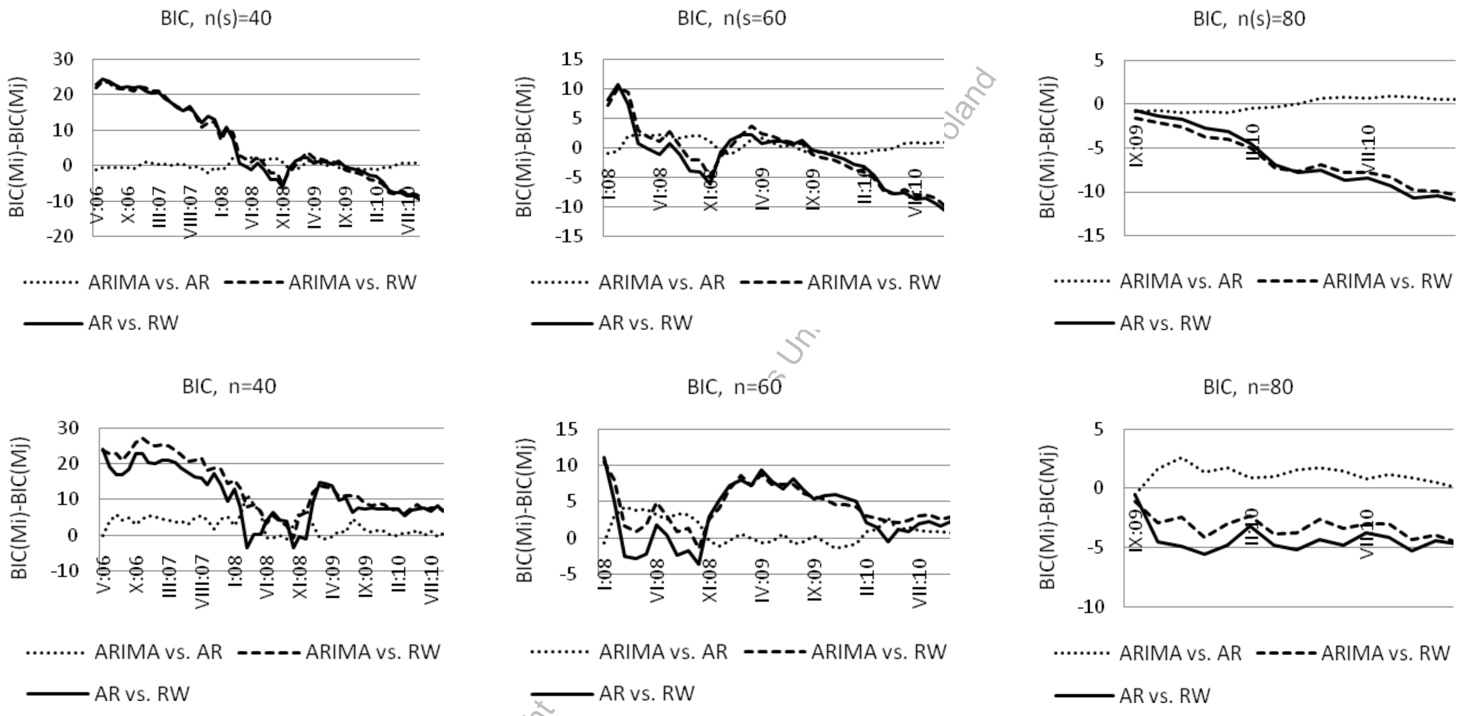

Figure 6. Differences in BIC criterion (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for IP in Poland

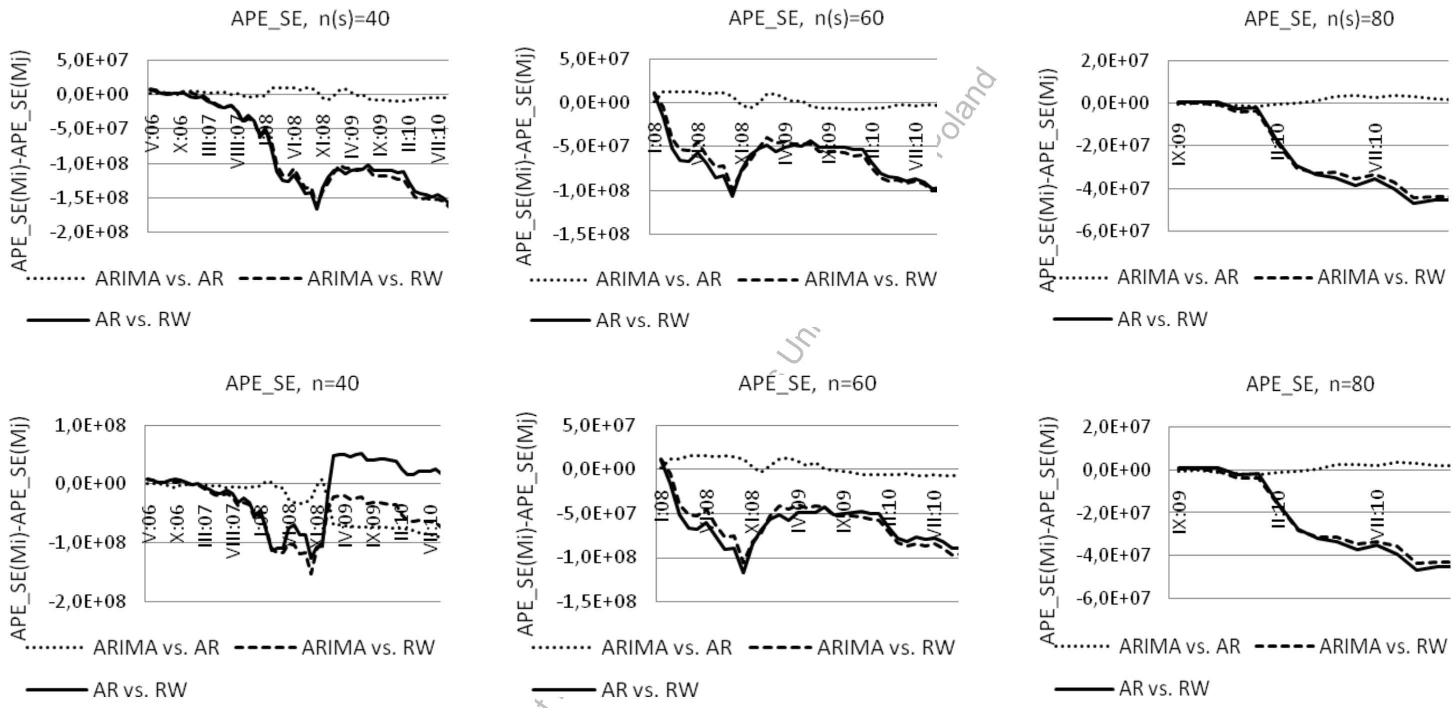

Figure 7. Differences in prediction criterion APE_SE (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for IP in Poland

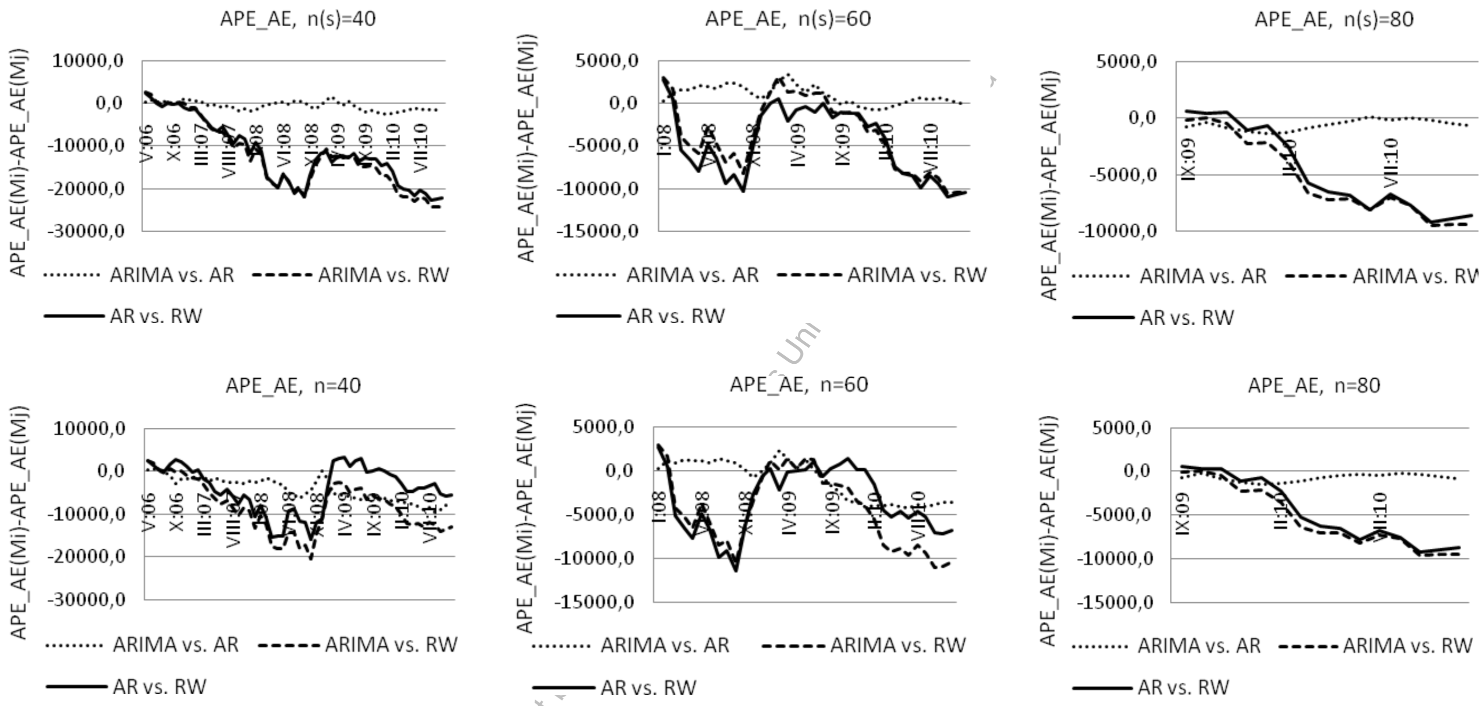

Figure 8. Differences in prediction criterion APE_AE (version I - row 1, version II - row 2) for pairs of models (ARIMA vs. AR, ARIMA vs. RW, AR vs. RW) depending on starting sample size and size of rolling window for IP in Poland