



www.sjm06.com

Serbian Journal of Management 8 (1) (2013) 53 - 65

Serbian
Journal
of
Management

IMPROVEMENTS IN ASSESSING THE FORECASTS ACCURACY - A CASE STUDY FOR ROMANIAN MACROECONOMIC FORECASTS

Mihaela (Simionescu) Bratu

Academy of Economic Studies, Faculty of Cybernetics, Statistics and Economic Informatics, B-Dul Anul 1864, No. 61 B, Sector 6, 062372 Bucharest, Romania

(Received 30 January 2013; accepted 1 April 2013)

Abstract

The objective of this study is to introduce new forecasts' accuracy measures for two types of predictions: point forecasts (radical of order n of the mean of squared errors, mean for the difference between each predicted value and the mean of the effective values, ratio of radicals of sum of squared errors (RRSSE- for forecasts comparisons), different versions of U2 Theil's statistic)) and for forecast intervals (number of intervals including the realization, difference between the realization and the lower limit, the upper one, respectively the interval centre). Comparisons are made to present the differences in results determined by the application of the classical measures of predictions accuracy for the inflation and unemployment rate forecasts provided for Romania by Institute for Economic Forecasting (IEF) and National Commission of Prognosis (NCP) on the horizon 2010-2012 and the values of new point forecasts accuracy measures. The hierarchy of predictions provided by the classical indicators and by the new ones are different. A novelty in literature is also brought by the methods of building the forecasts intervals. In addition to the classical interval based on historical error method, some new techniques of building forecasts are used: intervals based on the standard deviation and those constructed using bootstrap technique bias-corrected-accelerated (BCA) bootstrap method.

Keywords: Forecasts, Accuracy, U Theil's statistic, Forecast intervals, Bootstrap technique

1. INTRODUCTION

This research brings as a novelty the calculation of some new accuracy

indicators for assessing the inflation and unemployment rate forecasts of two specialized institutions in Romania: Institute for Economic Forecasting (IEF) and

* Corresponding author: mihaela_mb1@yahoo.com

DOI: 10.5937/sjm8-3235

National Commission of Prognosis (NCP). Our approach consists in reaching objectives like: proposal of some indicators to assess the point forecasts accuracy, the computing of modified versions of U2 Theil's statistic and the proposal of measures of accuracy for interval forecasts, the last contribution bringing a perspective that were not taken in consideration in literature till now. The real reasons for choosing more accuracy measures are the need to assess the predictions' precision from different perspectives and the necessity to have an accuracy indicator for interval forecasts.

The new ways of assessing the predictions' accuracy for comparisons bring different results from those based on classical accuracy indicators. Only the most accurate forecast is indicated by the usual measures of accuracy and also by our proposed indicators. Actually, our objective is to analyse the problem of accuracy from different points of view. Each indicator communicates something regarding the precision of the provided forecasts. It is possible that one indicator suggests one particular forecast as being the best, while other indicator recommends another. More indicators having the lowest value for a certain type of forecasts, more the chances are to be in the case of the most accurate forecast.

2. LITERATURE REVIEW

Many international institutions are specialized in providing their own macroeconomic appreciations. Some researchers were interested in evaluating the accuracy of those predictions (Timmermann for IMF Melander for European Commission, Vogel for OECD), neglecting

the comparison with government's expectations.

Genrea, Kenny, Meylera and Timmermann (2013) made forecasts combinations starting from SPF predictions for ECB and using performance-based weighting, trimmed averages, principal components analysis, Bayesian shrinkage, least squares estimates of optimal weights. Only for the inflation rate there was a strong evidence of improving the forecasts accuracy with respect to the equally weighted average prediction.

Hess and Orbe (2013) studied the association between analyst characteristics and the macroeconomic forecasts accuracy, noticing that the experience and the abilities of the analyst generate a better accuracy.

Clark and McCracken (2013) brought recent and important contribution in this domain: the assessment of point and density forecasts using the Vector Autoregression, direct and iterative forecasts with more steps, the application of accuracy tests on different samples of forecasts.

Bratu (2012) assessed the accuracy of some macroeconomic predictions for Romania made by the Institute of Economic Forecasting and the National Commission of Prognosis, the last institution outperforming the forecasts for: inflation, unemployment, GDP deflator, export rate and exchange rate on the horizon 2004-2011. Novotny and Rakova (2012) assessed the accuracy of macroeconomic forecasts made by Consensus for the Czech Republic, observing an improvement in accuracy from a year to another on the horizon 1994-2009. The authors also proposed a regression for comparing the predictions.

Abreu (2011) was interested in assessing the performance of macroeconomic predictions of IMF, European Commission

and OECD and two private institutions (Consensus Economics and The Economist). The directional accuracy and the ability of predicting an eventual economic crisis were studied.

Dovern and Weisser (2011) used a broad set of individual forecasts to analyze four macroeconomic variables in G7 countries. Analyzing accuracy, bias and forecasts efficiency, resulted large discrepancies between countries and also in the same country for different variables. In general, the forecasts are biased and only a fraction of GDP forecasts are closer to the results registered in reality.

Gorr (2009) showed that the univariate method of prediction is suitable for normal conditions of forecasting while using conventional measures for accuracy, but multivariate models are recommended for predicting exceptional conditions when ROC curve is used to measure accuracy.

Ruth (2008), using the empirical studies, obtained forecasts with a higher degree of accuracy for European macroeconomic variables by combining specific sub-groups predictions in comparison with forecasts based on a single model for the whole Union.

Heilemann and Stekler (2007) explain why macroeconomic forecast accuracy in the last 50 years in G7 has not improved. The first explanation refers to the critic brought to macroeconometrics models and to forecasting models, and the second one is related to the unrealistic expectations of forecast accuracy. Problems related to the forecasts bias, data quality, the forecast process, predicted indicators, the relationship between forecast accuracy and forecast horizon are analyzed.

In order to evaluate the forecast performance, and also to order the predictions, statisticians have developed

several measures of accuracy. Fildes R. and Steckler (2000) analyzed the problem of accuracy using statistics, indicating landmarks in the literature. For comparison between the MSE indicators of the forecasts, Granger and Newbold propose a statistic. Another statistic is presented by Diebold and Mariano in order to compare other quantitative measures of errors. Diebold and Mariano were proposed in 1995 a comparison test of two forecast's accuracy under the null hypothesis that states the lack of difference. The test proposed by them was later improved by Harvey and Ashley, who developed a new statistic based on a bootstrap inference. Later, Christoffersen and Diebold have developed a new way of measuring the accuracy that keeps the cointegration relationship between variables.

Terósvirta, van Dijk and Medeiros (2005) examine the accuracy of forecasts based on linear autoregressive models, autoregressive with smooth transition (STAR) and neural networks (neural network-NN) time series for 47 months of the macroeconomic variables of G7 economies. For each model is used a dynamic specification and it is showed that STAR models generate better forecasts than linear autoregressive ones. Neural networks over long horizon forecast generate better predictions than the models using an approach from private to general.

Armstrong and Fildes (1995) shows that the purpose of measuring forecast error is the provision of information about the shape of errors distribution and proposed a loss function for measuring the forecast error. Armstrong and Fildes show that it is not sufficient to use a single measure of accuracy.

If X is the predicted quantitative variable, the error of forecast is computed as the difference between the registered and the

predicted value: e_x . The frequently used indicators for evaluating the forecasts accuracy, according to Fildes and Steckler (2000), are computed as: Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n e_x^2}, \quad (1)$$

Mean error (ME):

$$ME = \frac{1}{n} \sum_{j=1}^n e_x \quad (2)$$

and Mean absolute error (MAE):

$$MAE = \frac{1}{n} \sum_{j=1}^n |e_x|. \quad (3)$$

These measures of accuracy have some disadvantages. For example, RMSE is affected by outliers. These measures are not independent of the unit of measurement, unless if they are expressed as percentage. If we have two forecasts with the same mean absolute error, RMSE penalizes the one with the biggest errors.

U Theil's statistic, utilized in comparing the forecasts, can be calculated in two variants, specified also by the Australian Treasury. U1 and U2. U1 statistic values are compared and closer to zero is a prediction, higher accuracy has:

$$U_1 = \frac{\sqrt{\sum_{t=1}^n (a_t - p_t)^2}}{\sqrt{\sum_{t=1}^n a_t^2 + \sum_{t=1}^n p_t^2}}. \quad (4)$$

A U2 less than one supposes a superior forecast compared to the naive one that is based on random walk:

$$U_2 = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{p_{t+1} - a_{t+1}}{a_t}\right)^2}{\sum_{t=1}^{n-1} \left(\frac{a_{t+1} - a_t}{a_t}\right)^2}}. \quad (5)$$

The following notations were utilized: a- actual values, p- predicted values, t- reference time, e- error ($e=a-p$), n- number of time periods. We propose the introduction of measures of accuracy as:

- Radical of order n of the mean of squared errors:

$$RnMSE = \sqrt[n]{\frac{1}{n} \sum_{j=1}^n e_x^2} \quad (6)$$

- The mean for the difference between each predicted value and the mean of the effective values on the forecasting horizon:

$$\bar{d} = (\text{mean}(p_t) - \bar{a}) \quad (7)$$

- For comparisons with the naive forecasts a new indicator is computed: ratio of radicals of sum of squared errors:

$$RRSSE = \frac{\sqrt[n]{\sum_{t=1}^n e_t^2}}{\sqrt[n-1]{\sum_{i=2}^n (x_i - x_{i-1})^2}}. \quad (8)$$

In order to compare two forecasts even for different variables, the values of this indicator are compared, a value closer to zero indicating a better accuracy.

3. THE ASSESSMENT OF POINT FORECASTS ACCURACY

The new accuracy measures are computed for the inflation and unemployment rate provided by Institute for Economic Forecasting (IEF) and National Commission of Prognosis (NCP) on the forecasting horizon 2010-2012. The data are published by the two institutions in their papers of forecasting and they are available on the websites (www.ipe.ro and www.cnp.ro). Firstly, some usual accuracy indicators are computed. Amid a severe economic contraction in 2009 in Romania, a relatively

slow rate of reduction in inflation was caused by persistent structural rigidities in the labor market and product market, but also by a variable number of factors acting during the year. In 2010, the volatile prices of food supply have been affected by the influence of external price increases in food goods, because of global supply reduction. Labor market immediately reacted to the crisis in 2009 through higher unemployment, but also slower annual growth in wages. From 2010 the situation is more visible, the unemployment rate continuing its increase. The accuracy indicators are presented in Table 1.

According to U1 indicator IEF unemployment forecasts are the most accurate, while the other indicators (ME, MAE and RMSE) have the lowest values for NCP inflation forecasts. These predictions are also better than the naive ones. The IEF inflation estimations are followed by those of

NCP unemployment rate, IEF inflation and NCP inflation, the results being presented in Table 2.

According to the new accuracy measures, the best forecasts are provided by IEF for the inflation rate, the difference between the mean of registered values and that of the predictions being only of -0.2043. The same value in absolute terms for the mean of the deviation of each predicted value from the average of the effective values supports the persistence of the overestimation of the average. The hierarchy of the predictions regarding the accuracy is evaluated using the RRSSE indicator and it is different as that provided by the evaluation of U1 statistic of Theil: forecasts of IEF for inflation rate, IEF predictions for unemployment rate, NCP forecasts for inflation, respectively for unemployment.

The U2 statistic can be modified in order to make the comparisons with other

Table 1. Classical measures of accuracy for the forecasts made by IEF and NCP for inflation and unemployment rate

Accuracy indicator	IEF- inflation rate	IEF- unemployment rate	NCP- inflation rate	NCP- unemployment rate
ME	0.8700	-0.4000	-0.1558	-0.9270
MAE	0.9250	1.4000	0.5043	1.1770
RMSE	1.1673	1.5732	0.6289	1.3020
U1	0.1194	0.0669	0.1308	0.1023
U2	1.0082	1.6005	0.8714	1.2268

Source: Author's computations

Table 2. The new measures of accuracy for the forecasts made by IEF and NCP for inflation and unemployment rate

Accuracy indicator	IEF- inflation rate	IEF- unemployment rate	NCP- inflation rate	NCP- unemployment rate
RnMSE	0.8079	1.2697	1.1242	1.4470
Bias	-0.2043	-0.9693	0.7967	-0.2333
\bar{d}	0.2043	0.9693	-0.7967	0.2333
RRSSE	0.5867	0.7166	0.8164	0.8167

Source: Author's computations

forecasts: the filtered naïve forecasts; the smoothed naïve forecasts; the values of the new forecasts. The filtered forecasts are gotten using Hodrick-Prescott and the smoothed naïve forecasts are obtained using Holt-Winters technique. The formula for the new U2 is:

$$U_2 = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{p_{t+1} - a_{t+1}}{a_t}\right)^2}{\sum_{t=1}^{n-1} \left(\frac{a^*_{t+1} - a^*_t}{a^*_t}\right)^2}} \tag{9}$$

a*- transformed actual (effective/real) values (filtered/smothered/values of new forecasts based on the proposed model).

The transformed U2 statistic is computed for the new forecasts of reference. The values of U2 are displayed in Table 3, the indicators being denoted U2*, U2** and U3***.

U2* statistic values show that, excepting the NCP inflation rate, all the other predictions are better than the filtered naïve forecasts based on Hodrick-Prescott filter. The indicators forecasted by IEF are more accurate than the smoothed naïve ones in Holt-Winters variant and even than the forecasts resulted from the proposed model. For NCP predictions the situation is exactly the opposite.

4. THE ASSESSMENT OF FORECAST INTERVALS ACCURACY

In this study some methods of building the forecast intervals are proposed.

Firstly, to construct the forecasts intervals we take into account the predictions provided by NCP in the pessimistic and optimistic versions and those of IEF for main, respectively desirable scenario. The samples of forecasts have low volume, so the t-Student distribution is used.

$$\overline{inf} l_{(t)} - t_{\alpha, n-1} \cdot RMSE * t < forecasts < \overline{inf} l_{(t)} + t_{\alpha, n-1} \cdot RMSE * t \tag{10}$$

The average of all the predictions that were proposed by forecasters is computed and the standard deviation coorrected with the number of forecasts will be utilized in constructing new forecast intervals:

$$\overline{inf} l_{(t)} - t_{\alpha, n-1} \cdot \frac{s_t}{\sqrt{n-1}} < forecasts < \overline{inf} l_{(t)} + t_{\alpha, n-1} \cdot \frac{s_t}{\sqrt{n-1}} \tag{11}$$

Other forecasts intervals are constructed using the RMSE of the previous year, where the RMSE is computed differently as the root

Table 3. The U2 transformed statistic for the forecasts made by IEF and NCP for inflation and unemployment rate

Forecasts	U2*	U2**	U2***
IEF inflation rate	0.6773	0.8230	0.6773
IEF unemployment rate	0.6347	0.8829	0.6347
NCP inflation rate	1.0752	1.3064	1.3064
NCP unemployment rate	0.8935	1.2431	1.2431

Source: Author's computations

mean squared of differences between each prediction of a year made by a certain institution and the real value of the indicator.

$$\overline{inf}l_{(t)} - t_{\alpha, n-1} \cdot RMSE_t < forecasts < \overline{inf}l_{(t)} + t_{\alpha, n-1} \cdot RMSE_t \tag{12}$$

A resampling technique is applied to build forecasts intervals and it consists in replicating the sample of predictions a huge number of times. Basically, a proxy population is made starting only from a sample. This is actually an artificial population.

The bias-corrected-accelerated interval (BCA) is a complex bootstrap technique used to construct confidence intervals. In this

case, Davison and Hinkley (1997) showed that estimates for bias and acceleration are provided using the initial sample and the bootstrap samples.

One accuracy measure for forecast intervals could be the number of intervals in which the real value is placed. The results are presented in Table 4.

6 out of 11 values of the unemployment rate are placed in the bootstrap intervals and in the historical RMSE intervals, 7 in the BCA bootstrap intervals and respectively in the intervals based on the standard deviation. The most values (10 out of 11) are located in the intervals based on the previous registered value.

Other measures of accuracy for forecast intervals can be computed. The differences

Table 4. Forecasts intervals for unemployment rate on the forecasting horizon 2001-2012

Year	Forecasts intervals based on:				Real values	Historical RMSE
	Bootstrap technique	BCA bootstrap method	Previous registered value	Forecasts' standard deviation		
2001	8.11-9.9	8.1895-9.9	5.775-9.850	8.547-9.514	8.8	7.416-10.672
2002	7.3-9.2	7.375-9.2	6.315-9.285	7.841-8.808	8.4	6.65-9.99
2003	6.55-8.9	6.725-8.9	4.999-10.541	7.212-8.412	7.4	5.733-9.891
2004	6.8-8.6	6.9-8.25	4.663-9.922	7.354-8.245	6.3	6.256-9.343
2005	6.37-8.4	6.37-8.2775	1.855-12.225	7.313-8.226	5.9	6.188-9.351
2006	5.92-7.8	5.92-7.7625	2.575-10.99	6.563-7.728	4	5.782-8.802
2007	5.54-7.6	5.54-6.38	3.832-10.418	6.563-7.728	4.4	5.389-8.69
2008	5.14-7.4	5.14-7.3475	3.799-9.166	6.26-7.304	5.8	4.973-8.591
2009	4.71-8.4	4.71-8.0475	6.534-8.015	6.297-7.952	7.5	4.259-9.99
2010	4.3-7.4	4.9325-7.4	5.775-9.850	5.778-7.185	6.9	4.048-8.916
2011	6.89-7.5	6.89-7.427	6.315-9.285	7.136-7.412	5.3	6.795-7.753
2012	5.2-6.7	5.2-6.625	4.999-10.541	8.547-9.514	8.8	4.555-7.269

Source: Own computations

between the realization for a specific year and the lower limit of each interval or the upper one or the interval centre could be considered as suitable measures of predictions accuracy. A lower difference implies a better forecast interval. The differences are computed as: $d1 = \text{realization} - \text{lower limit}$, $d2 = \text{realization} - \text{upper limit}$, $d3 = \text{realization} - \text{centre}$. Starting from these deviations we can compute their average or their absolute average on the forecasting horizon, the results being presented in Table 5.

Only 5 of the forecast intervals based on bootstrap method include the real values of inflation rate and 3 of those based on BCA bootstrap technique are suitable. 7 out of 11 forecast intervals based on standard deviation contain the inflation rate values. All the intervals constructed using the

inflation rate of the previous year include the realizations of inflation. Excepting the interval from 2002, all the others built using the previous RMSE include the real values. The new accuracy measures are shown in Appendix 1 for the inflation rate predictions and in Appendix 2 for unemployment rate.

For the forecast intervals of inflation rate based on bootstrap technique, the lowest value is registered by $d2$, but when absolute values of deviation are taken into account $d2$ is the highest.

BCA bootstrap method gave the best results for $d1$ (0.069), the lowest value for all methods. In average, the deviation between the realization and the inferior limit is 0.069% while the one between the realization and the intervals' centres is around 1.23%.

$d3$ registered the lowest value compared

Table 5. Forecasts intervals for inflation rate (2001-2012)

Year	Forecasts intervals based on:				Real values	Historical RMSE
	Bootstrap technique	BCA bootstrap method	Previous registered value	Forecasts' standard deviation		
2001	33.8-37.2	33.8-36.35	-16.7-87.2	33.832-36.667	34.5	32.158-38.341
2002	26-28.3	26-27.85	14.36-39.48	25.995-27.854	22.5	23.752-30.097
2003	17-19	17-18.825	11.32-25.32	17.549-19.1	15.3	9.716-26.933
2004	11.9-14.8	12.3525-14.4	10.3-16.89	12.594-14.61	11.9	7.650-19.554
2005	9-13.74	9-12.6075	5.92-14.54	8.315-12.159	9	6.427-14.047
2006	7-8.6	7-8.2	5.25-10.13	7.02-8.374	6.56	3.191-12.203
2007	5-8.14	5-7.535	3.58-8.34	4.713-7.216	4.84	3.415-8.514
2008	3.6-8.5	3.6-7.275	1.657-10.262	3.715-8.2	7.85	2.668-9.251
2009	4.5-8.25	4.5-7.3125	1.386-10.038	4.257-7.167	5.59	-0.035-11.46
2010	6.2-8.29	6.2-7.8175	4.347-9.197	5.936-7.608	6.09	3.852-9.692
2011	3.8-9.11	3.8-7.7825	1.721-9.184	3.385-7.521	5.8	3.338-7.568
2012	4.9-8.77	5.05-8.0775	3.57-8.814	4.725-7.659	3.6	2.004-10.38

Source: Own computations

to d1 and d2 when the method applied is based on the previous registered value.

If the forecast intervals are based on standard deviation, d1 registers again the best value. It maintains to be the minimum even if the absolute values of the deviations are computed.

According to the values of d1, d2 and d3 and the corresponding values for absolute deviations, the BCA bootstrap technique provided the best intervals for inflation rate.

A negative value but better than d1 and d2 is registered when the historical RMSE is utilized.

For unemployment rate forecast intervals the better value is registered for d3 when BCA bootstrap method is applied. d1 is a good measure of accuracy for this method.

d2 has the lowest value for intervals based on bootstrap method.

So, the new accuracy measures recommend the forecast intervals based on BCA bootstrap technique for inflation (d1) and for unemployment rate (d3).

5. CONCLUSION

This research enlarges the perspective of measuring the forecasts accuracy, by proposing some new measures for point forecasts and also for forecast intervals. The proposed measures draw attention about different results that may be registered when more predictions are compared. U1 Theil's statistic and the new indicator (ratio of

УНАПРЕЂЕЊА У ДОСТИЗАЊУ ТАЧНОСТИ ПРЕДВИЂАЊА- СТУДИЈА СЛУЧАЈА МАКРОЕКОНОМСКИХ ПРЕДВИЂАЊА У РУМУНИЈИ

Mihaela (Simionescu) Bratu

Извод

Циљ овог истраживања је да уведе нови приступ тачности предвиђања за две врсте предвиђања: предвиђања по тачкама (корен реда n -просечне вредности квадратне грешке, аритметичке средине разлике између сваке предвиђене вредности и аритметичке средине ефективне вредности, однос корена суме квадратних грешака (PPCSE - за поређење предвиђања), и различите верзије U_2 "Theil's" статистике)), као и за интервале предвиђања (број интервала укључених у реализацији, разлика између реализоване вредности и доње границе, горње границе, и средине интервала). Како би се представиле разлике у резултатима представљеним у овом раду, поређени су са вредностима одређеним применом класичног мерења тачности предвиђања стопе инфлације и незапослености које врши Институт економског предвиђања Румуније (ИЕФ) као и Национална комисија за прогнозу (НЦП) у периоду 2010 - 2012. Хијерархија предвиђања добијена класичним индикаторима и новим начином, је различита. Преко методе грађења интервала предвиђења искоришћене су и нова достигнућа представљена у литератури. Коришћени су нови интервали засновани на стандардној девијацији као и они конструисани употребом полазних техника убрзања корелације одступања (БЦА).

Кључне речи: Предвиђање, Тачност, "U Theil's" статистика, Интервали предвиђања, Техника убрзања корелације

radicals of sum of squared errors) gave the different results regarding the hierarchy of forecasts. Our indicator reduces the too large weight assigned to large errors. Therefore, the new measure could be used to identify the most accurate forecast. According to RRSSE, the hierarchy of the predictions regarding the accuracy on the horizon 2010-2012 is: forecasts of IEF for inflation rate, IEF predictions for unemployment rate, NCP forecasts for inflation, respectively for unemployment.

The BCA bootstrap techniques gave the best results for the same accuracy measures of the prediction intervals for Romanian inflation and unemployment. The measures of accuracy proposed for forecast intervals are a novelty in this field.

References

- Abreu, I. (2011). International organisations' vs. private analysts' forecasts: an Evaluation, Retrieved from <http://www.bportugal.pt/en-US/BdP%20Publications%20Research/wp201120.pdf>.
- Armstrong, J.S., & Fildes, R. (1995). On the selection of Error Measures for Comparisons Among Forecasting Methods. *Journal of Forecasting*, 14: 67-71.
- Bratu, M. (2012). *Strategies to Improve the Accuracy of Macroeconomic Forecasts in USA*, LAP LAMBERT Academic Publishing.
- Clark, T.E., & McCracken, M.W. (2013). Evaluating the accuracy of forecasts from Vector autoregression, Federal Reserve Bank of St. Louise, Working Paper 2013-010A.
- Davison, A.C., & Hinkley, D.V. (1997). *Bootstrap methods and their applications*, Cambridge University Press.
- Dovern, J., & Weisser J. (2011). Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7. *International Journal of Forecasting*, 27(2): 452-465.
- Fildes, R., & Steckler, H. (2000). *The State of Macroeconomic Forecasting*. Lancaster University EC3/99, George Washington University, Center for Economic Research, Discussion Paper No. 99-04.
- Genrea, V., Kenny, G., Meylera, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, 29(1): 108-121.
- Gorr, W.L. (2009). Forecast accuracy measures for exception reporting using receiver operating characteristic curves. *International Journal of Forecasting*, 25 (1): 48-61.
- Heilemann, U., & Stekler, H. (2007). Introduction to "The future of macroeconomic forecasting". *International Journal of Forecasting*, 23(2): 159-165.
- Hess, D., & Orbe, S. (2013). Idiosyncratic predictability and analyst characteristic. *University of Cologne papers*, 30(1): 3-20.
- Novotny, F., & Rakova, M.D. (2012). Assessment of Consensus forecasts accuracy: The Czech National Bank perspective. *Economic Research Bulletin*, 10(2): 10-13.
- Ruth, K. (2008). Macroeconomic forecasting in the EMU: Does disaggregate modeling improve forecast accuracy? *Journal of Policy Modeling*, 30 (3): 417-429.
- Terósvirta, T., van Dijk, D., & Medeiros, M.C. (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting, macroeconomic

APPENDIX 1

The new accuracy measures for forecast intervals of inflation

Forecasts based on bootstrap technique						
Accuracy measures						
Year	d1	d2	d3	d1	d2	d3
2001	0.7	-2.7	-1	0.7	2.7	1
2002	-3.5	-5.8	-4.65	3.5	5.8	4.65
2003	-1.7	-3.7	-2.7	1.7	3.7	2.7
2004	0	-1.81	-0.905	0	1.81	0.905
2005	0	-0.21	-0.105	0	0.21	0.105
2006	-0.44	-1.63	-1.035	0.44	1.63	1.035
2007	-0.16	-3.36	-1.76	0.16	3.36	1.76
2008	4.25	-0.29	1.98	4.25	0.29	1.98
2009	1.09	-2.66	-0.785	1.09	2.66	0.785
2010	-0.11	-2.2	-1.155	0.11	2.2	1.155
2011	2	-3.31	-0.655	2	3.31	0.655
2012	-1.3	-5.17	-3.235	1.3	5.17	3.235
average	-1.33375	1.270833	2.736667	1.66375	2.736667	1.66375

Forecasts based on BCA bootstrap technique						
Accuracy measures						
Year	d1	d2	d3	d1	d2	d3
2001	0.7	-1.85	-0.575	0.7	1.85	0.575
2002	-3.5	-5.35	-4.425	3.5	5.35	4.425
2003	-1.7	-3.525	-2.6125	1.7	3.525	2.6125
2004	0	-2.5	-1.25	0	2.5	1.25
2005	0	-3.6075	-1.80375	0	3.6075	1.80375
2006	-0.44	-1.64	-1.04	0.44	1.64	1.04
2007	-0.16	-2.695	-1.4275	0.16	2.695	1.4275
2008	4.25	0.575	2.4125	4.25	0.575	2.4125
2009	1.09	-1.7225	-0.31625	1.09	1.7225	0.31625
2010	-0.11	-1.7275	-0.91875	0.11	1.7275	0.91875
2011	2	-1.9825	0.00875	2	1.9825	0.00875
2012	-1.3	-4.4775	-2.88875	1.3	4.4775	2.88875
average	0.069167	-2.54188	-1.23635	1.270833	2.637708	1.639896

time series: A re-examination. International Journal of Forecasting, 21 (4): 755-774.

Forecasts based on previous registered value						
Accuracy measures						
Year	d1	d2	d3	d1	d2	d3
2001	51.2	-52.7	-0.75	51.2	52.7	0.75
2002	8.14	-16.98	-4.42	8.14	16.98	4.42
2003	3.98	-10.02	-3.02	3.98	10.02	3.02
2004	1.6	-4.99	-1.695	1.6	4.99	1.695
2005	3.08	-5.54	-1.23	3.08	5.54	1.23
2006	1.31	-3.57	-1.13	1.31	3.57	1.13
2007	1.26	-3.5	-1.12	1.26	3.5	1.12
2008	6.193	-2.412	1.8905	6.193	2.412	1.8905
2009	4.204	-4.448	-0.122	4.204	4.448	0.122
2010	1.743	-3.107	-0.682	1.743	3.107	0.682
2011	4.079	-3.384	0.3475	4.079	3.384	0.3475
2012	0.03	-5.214	-2.592	0.03	5.214	2.592
average	7.234916667	-9.655416667	-1.21025	7.234916667	9.655416667	1.58325

Forecasts based on forecasts' standard deviation						
Accuracy measures						
Year	d1	d2	d3	d1	d2	d3
2001	0.668	-2.167	-0.7495	0.668	2.167	0.7495
2002	-3.495	-5.354	-4.4245	3.495	5.354	4.4245
2003	-2.249	-3.8	-3.0245	2.249	3.8	3.0245
2004	-0.694	-2.71	-1.702	0.694	2.71	1.702
2005	0.685	-3.159	-1.237	0.685	3.159	1.237
2006	-0.46	-1.814	-1.137	0.46	1.814	1.137
2007	0.127	-2.376	-1.1245	0.127	2.376	1.1245
2008	4.135	-0.35	1.8925	4.135	0.35	1.8925
2009	1.333	-1.577	-0.122	1.333	1.577	0.122
2010	0.154	-1.518	-0.682	0.154	1.518	0.682
2011	2.415	-1.721	0.347	2.415	1.721	0.347
2012	-1.125	-4.059	-2.592	1.125	4.059	2.592
average	0.1245	-2.550416667	-1.212958333	1.461666667	2.550416667	1.586208333

Forecasts based on historical RMSE						
Accuracy measures						
Year	d1	d2	d3	d1	d2	d3
2001	2.342	-3.841	-0.7495	2.342	3.841	0.7495
2002	-1.252	-7.597	-4.4245	1.252	7.597	4.4245
2003	5.584	-11.633	-3.0245	5.584	11.633	3.0245
2004	4.25	-7.654	-1.702	4.25	7.654	1.702
2005	2.573	-5.047	-1.237	2.573	5.047	1.237
2006	3.369	-5.643	-1.137	3.369	5.643	1.137
2007	1.426	-3.674	-1.124	1.426	3.674	1.124
2008	5.182	-1.401	1.8905	5.182	1.401	1.8905
2009	5.625	-5.905	-0.14	5.625	5.905	0.14
2010	2.238	-3.602	-0.682	2.238	3.602	0.682
2011	2.462	-1.768	0.347	2.462	1.768	0.347
2012	1.596	-6.78	-2.592	1.596	6.78	2.592
average	2.949583333	-5.37875	-1.214583333	3.15825	5.37875	1.5875

APPENDIX 2

The new accuracy measures for forecast intervals of unemployment

Forecasts based on						
Accuracy measures						
Average	d1	d2	d3	d1	d2	d3
Bootstrap method	0.5558	-1.5250	-0.4846	1.4925	1.8750	1.3171
BCA bootstrap method	0.4673	-1.3348	-0.4337	1.4207	1.6973	1.2429
Previous forecast value	1.8387	-3.3823	-0.7718	2.0078	3.3823	1.2457
Forecasts' standard deviation	-0.4926	-1.5440	-1.0183	1.0889	1.5440	1.1632
Historical RMSE	0.9547	-2.4797	-0.7625	1.7137	2.7352	1.3896