THE IMPROVEMENT OF UNEMPLOYMENT RATE PREDICTIONS ACCURACY

DOI: 10.18267/j.pep.519

Mihaela Simionescu*

Abstract:
This research is related to the assessment of alternative unemployment rate predictions for the Romanian economy, the forecasts being provided by three anonymous forecasters: F1, F2 and F3. F3 provided the most accurate forecasts for the horizon 2001–2014, while F2 predictions are the less accurate according to U1 Theil's statistic and according to a new method that has not been used before in literature in this context. The multi-criteria ranking was applied to make a hierarchy of the forecasters regarding the accuracy and five important accuracy measures were taken into account at the same time: mean errors, mean squared error, root mean squared error, U1 and U2 statistics of Theil. The combined forecasts of forecasters' predictions are the best strategy to improve the forecasts accuracy. The filtered and smoothed original predictions based on Hodrick-Prescott filter, respectively Holt-Winters technique, are a good strategy of improving the accuracy only for F2 expectations. The assessment and improvement of forecasts accuracy have an important contribution in growing the quality of decision-making process.

Keywords: forecasts, accuracy, multi-criteria ranking, combined forecasts, Hodrick-Prescott filter, Holt-Winters smoothing exponential technique
JEL Classification: E21, E27, C51, C53

1. Introduction

The forecasts accuracy evaluation is essential for a better establishing of the decision--making process. In the case when more forecasters build predictions for certain macroeconomic variable, the policy-makers have to select the prediction or predictions with the highest degree of accuracy. The term of "accuracy" is put in correlation with the errors that affect the forecasting process, because it would be a risky business to assume that the predicted value of an indicator is exactly equal with its real value.

The original contribution of this research is related to the proposal of a new method of assessing the forecasts accuracy, taking into account more accuracy measures at the same time. The multi-criteria ranking let us make a classification of the forecasters according to more accuracy indicators.

On the other hand, the literature reports the necessity of improving the forecasts accuracy. We proposed as strategy of getting better predictions than the original ones the combined forecasts and the filtered and smoothed predictions and we made comparisons with the original predictions to measure the degree of improvement.

* Mihaela Simionescu, Romanian Academy, Bucharest, Romania (mihaela_mb1@yahoo.com).

2. Literature Review

The forecasts accuracy assessment is one of the most important concerns of forecasters. One purpose of this assessment is related to the need of improving the predictions. The current economic and financial crisis emphasized the struggles of uncertainty reduction. The forecasts accuracy is a very large domain of research, an exhaustive presentation of it being impossible. But, some of the recent results will be described. Bilan (2012) showed the necessity of accurate forecasts for labour market.

To assess the forecast accuracy, as well as their ordering, statisticians have developed several measures of accuracy. For making comparisons between the mean square errors of the predictions a statistic was proposed. Another one is presented by Diebold and Mariano (1995) - Diebold-Mariano test known as DM test - for comparison of other quantitative measures of errors. This test makes comparisons regarding the accuracy of two forecasts under the null hypothesis that assumes no differences in accuracy. The DM test was later improved by developing another statistic based on a bootstrap inference. Subsequently, a new way of measuring the accuracy was indicated under the assumption of preserving the co-integration relationship between variables.

Chen and Yang (2004) made a retrospective presentation of accuracy measures and they used the Kullback-Leibler divergence. For cross-series comparison of forecasts, individually tailored measures could improve the performance of differentiating between good and poor predictions. Simionescu (2013) proposed some new methods for assessing the uncertainty in forecasts based on opinion survey. Moreover, Simionescu (2014) made an evaluation of forecasts performance for macroeconomic variables in Romania, when the predictions are based on Dobrescu model for Romanian economy.

Todd and McCracken (2013) presented recent results regarding the assessment of density and point predictions based on Vector Autoregressions. Their researches used Monte Carlo simulations for tests regarding the accuracy equality making VAR and AR comparisons. Gürkaynak, Kisacikoglu and Rossi (2013) concluded that macroeconomic forecasts based on DSGE models are more accurate than random walk forecasts or Bayesian VAR forecasts.

Meese and Rogoff's paper "Empirical Exchange Rate Models of the Seventies" remains the starting point for many researches regarding the comparing of accuracy and bias. Recent studies target the accuracy analysis using as comparison criterion different models used in making predictions or the analysis of forecasted values for the same macroeconomic indicators registered in several countries.

Allan (2012) obtained a good accuracy for the OECD forecasts combined with outturn values of GDP growth for G7 countries between 1984 and 2010. The same author mentioned two groups of accuracy techniques used in assessing the predictions: quantitative forecasts accuracy statistics and qualitative accuracy methods.

Dovern and Weisser (2011) used a broad set of individual forecasts to analyse four macroeconomic variables in G7 countries. Analysing accuracy, bias and forecasts efficiency resulted in large discrepancies between countries and also in the same country for different variables.

Most international institutions provide their own macroeconomic forecasts. It is interesting that many researchers compare the predictions of those institutions (Melander for the European Commission, Vogel for the OECD, Timmermann for the IMF) with registered values and those of other international organizations, but it is omitted the comparison with official predictions of government. Kucur and Shinji (2006) checked the accuracy of the IMF macroeconomic predictions in the years 1994–2003. Timmermann (2007) applied the statistical tests to assess the performance of predictions provided by the World Economic Outlook. The predictions were inefficient and unbiased, but the performance of these predictions is similar to that of the consensus forecasts. Stekler and Zhang (2012) assessed the performance of predictions provided by the International Monetary Fund, the Organization for Economic Cooperation and Development, and the private forecasters for real GDP growth and inflation between 1999 and 2010. The long run forecasts were biased and less accurate than the naïve predictions.

Abreu (2011) evaluated the performance of macroeconomic forecasts made by the IMF, the European Commission and the OECD and two private institutions (the Consensus Economics and the Economist). The author analysed the directional accuracy and the ability of predicting an eventual economic crisis.

In the Netherlands, experts made predictions starting from the macroeconomic model used by the Netherlands Bureau for Economic Policy Analysis (CPB). For the period 1997–2008 there was reconstructed the model of the experts macroeconomic variables evolution and it was compared with the base model. The conclusions of Franses, McAleer and Legerstee (2012) were that the CPB model forecasts are in general biased and with a higher degree of accuracy.

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 criticism brought to macro-econometrics 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, the predicted indicators, the relationship between forecast accuracy and the forecast horizon are analysed.

The inflation rate predictions provided by Greenbook are more accurate than those made by private forecasts, Liu and Smith (2014) comparing the predictions provided by the SPF (Survey of Professional Forecasters), Greenbook and some other private experts in forecasting.

Müller-Dröge, Sinclair and Steckler (2014) evaluated the accuracy of the forecasts for a vector of variables for Germany in 2013, where the forecasts were given by 25 institutions. For assessing the degree of accuracy the Mahalanobis distance was employed and the Bundesbank gave the most accurate macroeconomic forecasts for German economy for 2013.

3. Comparisons between Unemployment Rate Forecasts Made by Different Forecasters

In this study we used the forecasted values of the annual registered unemployment rate made for Romania by F2, F2, and F3. The forecasting horizon is 2001–2014. The aim is to evaluate the accuracy of these experts' forecasts and to figure out the forecaster that had the best performance in the mentioned period.

Armstrong and Fildes (1995) showed that it is not sufficient to use a single measure of accuracy. Therefore, more accuracy indicators were computed for the three types of forecasts on the specified horizon.

To make comparisons between forecasts we propose to determine the hierarchy of forecasters according to the accuracy of their forecasts using multi-criteria ranking.

Two techniques of multi-criteria ranking (ranks method and the method of relative distance with respect to the maximal performance) are used in order to select the forecaster that provided the best forecasts in the years 2001–2014 taking into account at the same time all computed measures of accuracy. The multi-criteria ranking can also be applied to make a hierarchy of forecasters taking into account the performance of forecasts in all its dimensions: accuracy, unbiasedness and efficiency.

If we consider $\hat{X}_t(k)$ the predicted value after k periods from the origin time t, then the error at future time (t+k) is: $e_t(t+k)$. This is the difference between the registered value and the predicted one.

The indicators for evaluating the forecasts accuracy that will be taken into consideration when the multi-criteria ranking is used are:

- Root Mean Squared Error (RMSE)
  Formula for mean error

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n} e_X^2 (T_0 + j, k)} \tag{1}$$

- Mean error (ME)
  Formula for mean absolute error

$$ME = \frac{1}{n}\sum_{j=1}^{n} e_X (T_0 + j, k) \tag{2}$$

The sign of indicator value provides an important information: if it has a positive value, then the current value of the variable was underestimated, which means expected average values too small. A negative value of the indicator shows expected values too high on average.

- Mean absolute error (MAE)
  Formula for root mean squared error

$$MAE = \frac{1}{n}\sum_{j=1}^{n} \left| e_X (T_0 + j, k) \right| \tag{3}$$

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

A common practice is to compare the forecast errors with those based on a random-walk. "Naïve model" method assumes that the variable value in the next period is equal to the one recorded at actual moment. Theil proposed the calculation of U statistic that takes into account both changes in the negative and the positive sense of an indicator:

U Theil's statistic can be computed in two variants, specified also by the Australian Treasury.

The following notations are used:

a - the registered results
p - the predicted results
t - reference time
e - the error ($e = a$-p)
n - number of time periods

Formula for U_1

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

A value close to zero for U_1 implies a higher accuracy.

Formula for U_2

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

If U_2 =1=> no differences in accuracy between the predictions
If U_2 <1=> the compared forecast better than naïve prediction
If U_2 >1=> the naive prediction better than compared forecast

Table 1 | The Assessment of Unemployment Predictions Accuracy during 2001–2014 for the Three Forecasters

ACCURACY MEASURE	FORECASTS MADE BY:		
	F1	F3	F2
Mean Error (ME)	-0.5602	-0.7302	-0.5833
Mean Absolute error (MAE)	1.2745	1.102	1.6592
Root Mean Square Error (RMSE)	1.5203	1.4028	1.9037
U_1	0.1124	0.1129	0.1430
U_2	1.2281	0.9983	1.1025

Source: own computations

The lowest value for mean error was registered by F1, but for the rest of accuracy measures the F3 obtained the lowest values, performing better than F1 and F2 in the period 2001-2014. All types of predictions are overestimated. F2's forecasts for unemployment rate are less accurate. Only the F3 predictions are superior to naïve forecasts.

The application of ranks method implies the following stages:

1. The assignment of ranks to accuracy measure value (rank 1 for the lowest value of the indicators);

   The rank for the forecaster i is: $(r_{i_{ind_j}})$, $i = 1,2,3$ and ind_j the j-th accuracy measure/indicator. 5 accuracy indicators are employed: mean absolute error, mean error, root mean squared error, U1 and U2 Theil's statistics.

2. The calculation of sum of ranks and the score's computation;

   Formula for the sum of ranks

$$S_i = \sum_{j=1}^{5}(r_{i_{ind_j}}),\ i = 1,2,3 \tag{6}$$

3. The establishment of the final rank (rank 1 for the lowest score).

Table 2 | The Forecasters' Final Ranks Using Ranks Method

ACCURACY MEASURE	FORECASTS MADE BY:		
	F1	F3	F2
Mean error (ME)	1	3	2
Mean Absolute error (MAE)	2	1	3
Root Mean Square Error (RMSE)	2	1	3
U1 Theil's Statistic	2	1	3
U2 Theil's Statistic	3	1	2
Ranks' sum	10	7	13
Final ranks	2	1	3

Source: own computations

The ranks method gave the same results as all the accuracy indicators, excepting mean error. If we take into consideration all the accuracy indicators, the best forecaster was the F3, being followed by F1 and finally, the F2.

The method of relative distance compared to the best performance consists in assigning for each accuracy indicator the forecaster's distance compared to that forecaster that performed better. The relative distance is calculated as:

Formula for the relative distance

$$d_{i_{ind_j}} = \frac{ind_i^j}{\left\{\min abs(ind_i^j)\right\}_{i=1,\dots,4}},\ i = 1,2,3 \text{ and } j = 1, 2, .., 5 \tag{7}$$

The denominator in the relative distance formula is the lowest value of the accuracy measure for all the forecasters.

The geometric average of the distances is computed, meaning the mean of relative distances of forecaster *i*.

Formula for the average relative distance

$$\bar{d}_i = \sqrt[5]{\prod_{j=1}^{5} d_{i_{indj}}} \ , \ i = 1,2,3 \tag{8}$$

The final ranks are given using the values of the average relative distances, rank 1 being given to the forecaster that has the lowest average relative distance. The forecaster location compared to the best one is determined as the ratio between average relative distance and the lowest mean of relative distances.

Formula for the position of each statistical unit in the hierarchy

$$loc_i^{\%} = \frac{\bar{d}_i}{\min(d_i)_{i=\overline{1,4}}} \cdot 100 \tag{9}$$

Table 3 | The Forecasters' Final Ranks Using the Method of Relative Distance Compared to the Most Accurate Forecaster

ACCURACY MEASURE	FORECASTS MADE BY:		
	F1	F3	F2
Mean error (ME)	1	1.4023	1.0503
Mean Absolute error (MAE)	1.1554	1	1.5403
Root Mean Square Error (RMSE)	1.1602	1	1.3779
U_1 Theil's Statistic	1.1658	1	1.3672
U_2 Theil's Statistic	1.1709	1	1.1103
Average relative distance	1.1284	1.0699	1.2761
Ranks	2	1	3
Location (%)	105.466	100	119.265

Source: own computations

The method of relative distance compared to the forecaster with the most accurate forecasts leads us to the same conclusion as the previous techniques. F3 registered the lowest value for the average relative distance (1.0699).

The Diebold-Mariano test - DM test - is employed in order to verify if two predictions have the same degree of accuracy. Let us consider the actual values of the unemployment rate $\{u_t\}, t = 1, 2, \ldots, T$ and two predictions for it $\{\hat{u}_{t1}\}, t = 1, 2, \ldots, T$ and $\{\hat{u}_{t2}\}, t = 1, 2, \ldots, T$. The prediction errors are computed as: $e_{it} = \hat{u}_{it} - u_t$, i = 1, 2. The loss function is: $g(u_t, \hat{u}_{it}) = g(\hat{u}_{it} - u_t) = g(e_{it})$.

In most cases this function is a square-error loss or an absolute error loss function.

Two predictions being given, the loss differential is: $d_t = g(e_{1t}) - g(e_{2t})$.

The two predictions have the same degree of accuracy if the expected value of loss differential is 0. For Diebold-Mariano (1995) test, the null assumption of equal accuracy checks if the expected value of differential loss is zero: $E(d_t) = 0$ The covariance stationary being given, the distribution of differential average follows a normal distribution. The DM statistic, according to Diebold and Mariano (1995), under null hypothesis is:

Formula for DM statistic

$$S_1 = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \to N(0,1)$$

$$\bar{d} = \frac{\sum_{t=1}^{T} d_t}{T}$$

$$\hat{V}(\bar{d}) = \frac{\hat{\gamma}_0 + 2\sum_{k=1}^{T-1}\hat{\gamma}_k}{n}$$

$$\hat{\gamma}_k = \frac{\sum_{t=k+1}^{T}(d_t - \bar{d})(d_{t-k} - \bar{d})}{T} \qquad (10)$$

The results of DM test indicated that there are not significant differences between the predictions of the three mentioned forecasters. If we combine the results of DM test and the values of accuracy measures, we can conclude that the best forecasts are provided by F3, being followed by F1 and finally by F2. However, the differences in accuracy between the three types of unemployment forecasts are not statistically significant.

Table 4 | The Forecasts Accuracy Comparisons Based on Diebold-Mariano Test

Forecasters whose predictions are compared	DM statistic value	Expert with the more accurate forecasts
F1-F2	S(1) = 0.571 p-value = 0.56	F2
F1-F3	S(1) = 0.56 p-value = 0.58	F3
F2-F3	S(1) = 0.321 p-value = 0.72	No differences in accuracy between forecasts

Source: Own computations

We also applied some qualitative tests for checking directional accuracy. We actually want to verify if there is correct change anticipation. In this case, a test of independence between actual values and the modification direction is applied. The null hypothesis assumes the independence. If the p-value is less than 0.05, then the independence hypothesis is rejected. All the asymptotic significances are more than 0.05, according to ***Appendix 1***. Therefore, we can conclude that directional changes in the outturn are independent from the forecasts.

4. Strategies to Improve the Accuracy of Unemployment Rate Forecasts

The importance of combining forecasts to get more accurate predictions is well-established in literature, many researches demonstrating the utility of this approach in the last 50 years, especially the works of Granger and his colleagues. Bratu (2012) utilized some strategies to improve the forecasts accuracy, among them being the combined predictions, regressions models, historical errors method, application of filters and exponential smoothing techniques.

The methods proposed by Bates and Granger (1969) supposed a linear combination of two forecasts. The combined forecasts can be, in some cases, a strategy of getting more accurate predictions. The most utilized combination approaches are:

- optimal combination (OPT);
- equal-weights-scheme (EW);
- inverse MSE weighting scheme (INV).

A forecast combination that starts from K alternative predictors ($\hat{X}_t^{(1)}, \hat{X}_t^{(2)}, \ldots, \hat{X}_t^{(K)}$) of the variable X_t and it is based on the information that is available till (t-1) period is written as: $\hat{X}_t^C = \sum_{k=1}^{K} \omega_k \cdot \hat{X}_t^{(k)}$.

A recent adaptive method (Aggregated Forecast Through Exponential Re-weighting-AFTER) was employed by Yang (2004), where the weights are computed recursively:

Formula for weights in AFTER algorithm

$$\omega_{t,k} = \frac{\hat{v}_{t-1,k}^{-0.5} \cdot \exp\left(-\frac{e_{t-1,k}^2}{2\hat{v}_{t-1,k}}\right)\omega_{t-1,k}}{\sum_{i=1}^{K} \hat{v}_{t-1,i}^{-0.5} \cdot \exp\left(-\frac{e_{t-1,i}^2}{2\hat{v}_{t-1,i}}\right)\omega_{t-1,i}} \tag{11}$$

where $e_{t-1,k} = X_{t-1} - \hat{X}_{t-1}^{(k)}$, $\omega_{1,k} = \frac{1}{K}$, $0 \le \omega_{t,k} \le 1$, $\sum_{k=1}^{K} \omega_{t,k} = 1$ and $\hat{v}_{t-1,k} = \frac{1}{t-1} \sum_{\tau=1}^{t-1} e_{\tau,k}^2$ which is the estimator of forecast variance.

Bates and Granger (1969) started from two forecasts f 1;t and f 2;t, for the same variable X_t, derived h periods ago. If the forecasts are unbiased, the error is calculated as: $e_{i,t} = X_{i,t} - f_{i,t}$. The errors follow a normal distribution of parameters 0 and σ_i^2. If ρ is the correlation between the errors, then their covariance is $\sigma_{12} = \rho \cdot \sigma_1 \cdot \sigma_2$. The linear combination of the two predictions is a weighted average $c_t = m \cdot f_{1t} + (1 - m). f_{2t}$. The error of the combined forecast is: $e_{c,t} = m \cdot e_{1t} + (1 - m) . e_{2t}$. The mean of the combined forecast is zero and the variance is:

$\sigma_c^2 = m^2 \cdot \sigma_1^2 + (1 - m)^2 \cdot \sigma_{2t}^2 + 2 \cdot m \cdot (1 - m) \cdot \sigma_{12}$. By minimizing the error variance, the optimal value for m is determined (m_{opt}):

Formula for the optimal value of m

$$m_{opt} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\cdot\sigma_{12}} \tag{12}$$

The individual forecasts are inversely weighted to their relative mean squared forecast error (MSE) resulting in INV. In this case, the inverse weight (m_{inv}) is:

Formula for the inverse weight

$$\mathrm{m}_{inv} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \tag{13}$$

Equally weighted combined predictions (EW) are gotten when the same weights are given to all models. Smith and Wallis (2009) concluded that the best solution is the use of an equally weighted average for determining the combining weights.

The U Theil's statistics in both variants were provided for the combined forecasts based on four schemes, the results being in Table 5:

Table 5 | The Unemployment Rate Combined Predictions Assessment (2001–2014)

Accuracy indicator	Combined forecasts of F1 and F2	Combined forecasts of F1 and F3	Combined forecasts of F2 and F3
U1 Theil's Statistic (optimal scheme)	0.0863	0.07823	0.1458
U2 Theil's Statistic (optimal scheme)	0.9891	0.7349	1.2037
U1 Theil's Statistic (inverse MSE scheme)	0.0882	0.05873	0.1237
U2 Theil's Statistic (inverse MSE scheme)	1.102	0.6037	1.1978
U1 Theil's Statistic (equally weighted scheme)	0.0871	0.0822	**0.09302**
U2 Theil's Statistic (equally weighted scheme)	0.9334	0.8582	0.9672
U1 Theil's Statistic (AFTER algorithm)	**0.0855**	**0.0492**	0.09836
U2 Theil's Statistic (AFTER algorithm)	0.9034	0.8339	0.9834

Source: author's computations

All the combined forecasts, excepting the combined predictions of F2 and F3 under optimal and inverse schemes, proved being a good strategy of improving the accuracy of initial experts' predictions. The highest gain in accuracy was brought by combined forecasts of F1 and F3 using AFTER algorithm. The best combined forecasts of F1 and F2, respectively F1 and F3, appeared for AFTER scheme. For F2 and F3 combined predictions, the equally weighted scheme provided the best forecasts. All the combined predictions are better than the naïve anticipations excepting those of F2 and F3 using optimal and inverse scheme and those of F1 and F2 forecasts in inverse scheme.

Other ways of getting more accurate forecasts used by Bratu (Simionescu) (2013) are the application of filters to the forecasted values of the variables and the use of exponential smoothing methods.

Hodrick-Prescott (HP) filter and Holt-Winters (HW) exponential technique were used to the initial forecasts and the accuracy of these forecasts was assessed. *Holt-Winters method* is employed for data series that has linear trend but no seasonal variations. The Hodrick–Prescott (HP) filter is employed in macroeconomics to get the data trend and to separate the cyclical component of a chronological series.

Table 6 | The Accuracy of Transformed Unemployment Forecasts (2001-2014)

Accuracy measure	F1 filtered forecasts	F2 filtered forecasts	F3 filtered forecasts	F1 smoothed forecasts	F2 smoothed forecasts	F3 smoothed forecasts
U1 Theil's Statistic	0.1423	**0.1078**	**0.1098**	0.1302	**0.1296**	0.1188
U2 Theil's Statistic	1.4403	**0.9304**	1.0873	1.3852	**0.9523**	1.3294

Source: author's computations

The F2 smoothed and filtered predictions and F3 filtered forecasts are more accurate than the original anticipations of these forecasters. Excepting F2's filtered and smoothed forecasts, all the predictions based on HP filter and Holt-Winters technique are less accurate than the naïve forecasts. So, the HP filter application is a good strategy of improving only the F2 and F3 forecasts. However, the combined predictions improved the accuracy more than HW model and HP filter. In general, the filters and smoothing techniques give better results if there are not changes in forecasts direction with respect to the actual values.

5. Conclusions

Beside the economic analysis, the forecasts construction is an important element for conducting the way of developing the macroeconomic activities. However, the predictions should be followed by an evaluation of accuracy. There are advantages of this assessment: the model improvement, the government policies improvement, the planning activity. The accuracy evaluation is related to the prediction degree of trust. There is a rich literature related to the prediction methods, but only few studies are dedicated to the accuracy assessment and ways to improve it. This is a crucial aspect, because the economic predictions should be cautiously accepted, mostly because of the negative effects of forecasts failures. The economic policy decisions are built on these forecasts. In the context of economic crisis, the need of more accurate predictions is more important.

In this research, the assessment of unemployment forecasts is made for the forecasts provided during 2001–2014 by three anonymous experts in forecasting. The best accuracy in forecasts was obtained by F3, being followed by F1 and F2. This ranking is given by the multi-criteria ranking, but also by the assessment of accuracy indicators, as U1 Theil's statistic, used in comparing the forecasts.

The combined forecasts based on several schemes (optimal scheme, inverse scheme, equally weighted scheme, AFTER algorithm) proved a good strategy of increasing the accuracy, almost all combined predictions being better than the original anticipations of forecasters. Filtered and smoothed forecasts based on HP filter, namely Holt-Winters technique increased the F2 forecasts accuracy.

The forecasts evaluation should be a priority for the government that makes its decisions using these predictions. The combined predictions and in certain situations the filtered and smoothed anticipations were a good way for improving the accuracy of unemployment predictions in Romania during 2001–2014. For the next few years, these methods can also improve the forecasts accuracy.

References

Abreu, I. (2011), "International Organizations' vs. Private Analysts' Forecasts: An Evaluation." *Bank of Portugal Review*, Vol. 4, No. 2, pp. 4–13.

Allan, G. (2012), "Evaluating the Usefulness of Forecasts of Relative Growth." *Strathclyde, Discussion Papers in Economics*, Vol. 1, No. 12–14, pp. 6–12.

Armstrong, J. S., Fildes, R. (1995), "On the Selection of Error Measures for Comparisons among Forecasting Methods." *Journal of Forecasting*, Vol. 14, No. 1, pp. 67–71.

Bates, J., Granger, C. W. J. (1969), "The Combination of Forecasts." *Operations Research Quarterly*, Vol. 20, No. 4, pp. 451–468.

Bilan, Y. (2012), "Specificity of Border Labour Migration." *Transformations in Business & Economics*, Vol. 11, No. 2, pp. 82–97.

Bratu, M. (2012), *Strategies to Improve the Accuracy of Macroeconomic Forecasts in USA*. Munich: Lap Lambert Academic Publishing.

Bratu Simionescu, M. (2013), "Filters or Holt Winters Technique to Improve the Forecasts for USA Inflation Rate?" *Acta Universitatis Danubius. Economica,* Vol. 9, No. 1, pp. 56–68.

Chen Z, Yang Y. (2004), "Assessing Forecast Accuracy Measures." *Annals of Statistics*, Vol. 28, No. 1, pp. 75–87.

Diebold, F. X., Mariano, R. (1995), "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics*, Vol. 13, No. 1, pp. 253–265.

Dovern, J., Weisser J. (2011), "Accuracy, Unbiasedness and Efficiency of Professional Macroeconomic Forecasts: An Empirical Comparison for the G7." *International Journal of Forecasting,* Vol. 27, No. 2, pp. 452–465.

Franses, P. H., McAleer, M., Legerstee, R. (2012), "Evaluating Macroeconomic Forecasts: A Concise Review of Some Recent Developments." Working Paper/ Department of Economics and Finance, University of Cantenbury, Vol. 2, No.1, pp. 13–26.

Gorr, W. L. (2009), "Forecast Accuracy Measures for Exception Reporting Using Receiver Operating Characteristic Curves." *International Journal of Forecasting*, Vol. 25, No. 1, pp. 48–61.

Gürkaynak, R. S., Kisacikoglu, B., Rossi, B. (2013), "Do DSGE Models Forecast More Accurately out-of-Sample than VAR Models?" CEPR Discussion Papers 9576, C.E.P.R. Discussion Papers.

Heilemann, U., Stekler, H. (2007), "Introduction to the Future of Macroeconomic Forecasting" *International Journal of Forecasting*, Vol. 23, No. 2, pp. 159–165.

Liu, D., Smith, J. K. (2014), "Inflation Forecasts and Core Inflation Measures: Where Is the Information on Future Inflation?" *The Quarterly Review of Economics and Finance*, Vol. 54, No. 1, pp. 133–137.

Kucur, H., Shinji T. (2006), "Testing the Accuracy of IMF Macroeconomic Forecasts, 1994–2003." IEO Background Paper Number BP/06/01. Available at http://www.ieo-imf.org/ieo/pages/CompletedEvaluation114.aspx

Müller-Dröge, C., Sinclair, T., Steckler, H. O. (2014), "Evaluating Forecasts of a Vector of Variables: A German Forecasting Competition." Cama Working Paper, Vol. 1, No. 55, pp. 1–23.

Ruth, K. (2008), "Macroeconomic Forecasting in the EMU: Does Disaggregate Modeling Improve Forecast Accuracy?" *Journal of Policy Modeling*, Vol. 30, No. 3, pp. 417–429.

Simionescu, M. (2013), "Methods for Assessing the Uncertainty in Forecasts Based on Opinion Survey." *International Journal of Academic Research*, Vol. 5, No. 6, pp. 56–67.

Simionescu, M. (2014), "The Performance of Predictions Based on the Dobrescu Macromodel for the Romanian Economy." *Journal for Economic Forecasting*, Vol. 3, pp. 179–195.

Smith, J., Wallis, K. F. (2009), "A Simple Explanation of the Forecast Combination Puzzle." *Oxford Bulletin of Economics and Statistics*, Vol. 71, No. 3, pp. 331–355.

Stekler H. O., Zhang H. (2013), "An Evaluation of Chinese Economic Forecasts." *Journal of Chinese Economic and Business Studies*, Vol. 11, No. 4, pp. 251–259.

Timmermann, A. (2007), "An Evaluation of the WEO Forecasts." IMF Staff Papers, Vol. 54, No. 1, pp. 2–34.

Todd E. C., McCracken, M. W. (2013), "Evaluating the Accuracy of Forecasts from Vector Autoregressions." Working Paper 2013–010A.

Yang, Y. (2004), "Combining Forecasting Procedures: Some Theoretical Results." *Econometric Theory*, Vol. 20, No. 1, pp. 176–222.

APPENDIX 1

The Directional Accuracy Tests

	Unemployment rate	F1	F3	F2
Chi-square	0.763	1.303	1.303	0.521
Degrees of freedom	12	11	11	11
Asymp. Sig.	1.000	0.994	0.994	1.000