SOME EMPIRICAL STRATEGIES FOR IMPROVING THE ACCURACY OF UNEMPLOYMENT RATE FORECASTS IN ROMANIA

Assist. Ph.D. Candidate in Statistics **Mihaela BRATU (SIMIONESCU)**
Department of Statistics and Econometrics
Faculty of Cybernetics, Statistics and Economic Informatics
Academy of Economic Studies
E-mail: mihaela_mb1@yahoo.com

Abstract

This study proposed to evaluate some alternative forecasts for the unemployment rate of Romania made by European Commission and two national institutions: National Commission for Prognosis (NCP) and Institute for Economic Forecasting (IEF). The most accurate predictions on the forecasting horizon 2001-2011 were provided by IEF and the less accurate by NCP. These results were obtained using U_1 Theil's statistic and 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 institutions 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, U_1 and U_2 statistics of Theil. The combined forecasts of institutions' 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 NCP expectations. The assessment and improvement of forecasts accuracy have an important contribution in growing the quality of the decisional process.

Keywords: *forecasts, predictions, accuracy, multi-criteria ranking, combined forecasts, Hodrick-Prescott filter, Holt-Winters smoothing exponential technique*

JEL Classification: E_{21}, E_{27}, C_{51}, C_{53}

Introduction

The evaluation of forecasts accuracy is necessary for establishing the decisional process. When more institutions in a country provide forecasts for the same macroeconomic variable, the deciders have to choose the one with the highest accuracy. The term of "accuracy" is put in correlation with the errors that affect the forecasting process, because only by hazard 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 institution according to more accuracy indicators.

On the other hand, the literature reports the necessity of improving the forecasts accuracy. We proposed as a strategy of obtaining 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.

Literature review

The actual objective of the researchers interested in the accuracy of the forecasts is to find out a suitable strategy to improve the accuracy. Therefore, new predictions are built starting from the initial ones. The economic crisis draws attention on the problem of uncertainty minimization.

In order to make comparisons between the MSE indicators of forecasts, Granger and Newbold used a statistic. Diebold and Mariano (1995) compared other quantitative measures of errors. Diebold and Mariano proposed in 1995 a test to check the differences in the accuracy of two forecasts. The test was later improved by Ashley and Harvey, using a bootstrap inference. Subsequently, Diebold and Christoffersen preserved the co-integration relation between variables.

Meese and Rogoff's paper, *Empirical exchange rate models of the seventies*, brought the most important initial contribution on the comparing of accuracy and bias. Recent studies made comparisons for forecasts based on different methods or made comparisons between predictions of the same variable registered in different regions.

Allan (2012) improved the OECD forecasts accuracy by the combined technique for G7 countries (horizon 1984- 2010).

Dovern and Weisser (2011) observed major differences in terms of bias, efficiency and accuracy for G7 countries forecasts and for each country between variables predictions.

Many institutions give their economic forecasts, the researchers being able to make comparisons between alternative forecasts of OECD, IMF, European Commission.

Abreu (2011) compared the performance of forecasts provided by IMF, European Commission and OECD, Consensus Economics and The Economist.

Franses, Kranendonk and Lanser (2011) concluded that the CPB model forecasts for 1997-2008 are in general biased and more accurate than those based on the government model.

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) proposed as strategy of improving the accuracy the use for more models associated to different countries in the European Union instead of one model.

Heilemann and Stekler (2007) provided some reasons for the lack of improvements in G7 predictions: non-useful macro-econometrics models and the unrealistic expectations regarding the accuracy.

Comparisons between unemployment rate forecasts made by different institutions

In this study we used the forecasted values of the annual registered unemployment rate made for Romania by European Commission, National Commission for Prognosis and Institute for Economic Forecasting. The forecasting horizon is 2001-2011. The objective is to assess the accuracy, the biasness and the efficiency of these predictions and determine the best institution with the highest performance.

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 institutions according to the accuracy of their forecasts using multi-criteria ranking.

Two methods 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 institution that provided the best forecasts on the horizon 2001-2011 taking into account, at the same time, all computed measures of accuracy. The multi-criteria ranking can be applied to make a hierarchy of institutions taking into account the performance of forecasts in all its dimensions: accuracy and efficiency.

$\hat{X}_t(k)$ is the forecasted value after k periods, t being the origin. The error at time (t+k) is: $e_t(t+k)$. It is computed as difference between the actual value (a) and the forecasted/ predicted one (p).

The measures of accuracy that were taken into account at the same time for the multi-criteria ranking are:

- Root Mean Squared Error (RMSE)

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

- Mean error (ME)

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

The sign of indicator value provides 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)

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

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

➢ U1 and U2 Theil's statistics

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

If U1 value is close to zero for U_1 (less than 0.5) we have a high degree of accuracy.

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

U1 and U2 Theil's coefficients are used to make comparisons between forecasts. The benchmark when U2 indicator is used is the naïve forecast.

If U_2 =1=> no significant differences as degree of accuracy between the two forecasts

If U_2 <1=> the forecast to compare more accurate than the naive one

If U_2 >1=> the forecast to compare less accurate than the naive one

According to all accuracy indicators for forecasts made on the horizon 2001-2012, excepting the mean error, the Institute for Economic Forecasting that used Dobrescu macromodel, provided the most accurate predictions for the unemployment rate. Only the forecasts of this institution outperformed the naïve predictions based on the random walk. The negative values of the mean error imply too high in average predicted values for all institutions. The less accurate forecasts are made by the National Commission for Prognosis.

Table 1

The accuracy of forecasts made by European Commission, National Commission for Prognosis and Institute for Economic Forecasting for the unemployment rate in Romania (2001-2012)

ACCURACY MEASURE	INSTITUTION		
	European Commission (EC)	National Commission for Prognosis (NCP)	Institute for Economic Forecasting (IEF)
ME	-0.5462	-0.5643	-0.7279
MAE	1.2372	1.6369	1.0916
RMSE	1.4959	1.7638	1.3059
U1	0.1074	0.1249	0.0927
U2	1.1587	1.0978	0.9983

Source: own computations using Excel.

Ranks method has several steps:

1. Ranks assign

The statistical units are the institutions that provided forecasts. The rank for each institution is denoted by: $(r_{i\,ind_j})$, i = 1, 2, 3 and ind_j –accuracy indicator j. We chose 5 indicators: mean error, mean absolute error, root mean squared error, U1 and U2.

2. The sum of ranks and the scores

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

3. Assign final ranks

Table 2

The ranks of institutions according to the accuracy measures (ranks method)

ACCURACY MEASURE	INSTITUTION		
	European Commission	National Commission for Prognosis	Institute for Economic Forecasting
ME	1	2	3
MAE	2	3	1
RMSE	2	3	1
U1	2	3	1
U2	3	2	1
Sum of ranks	10	13	7
Final ranks	2	3	1

Source: own computations using Excel.

The results of the ranks method are the same as those provided by most accuracy measures, especially U1 used in making comparisons between forecasts. Actually, if all the calculated accuracy indicators are taken into account at the same time, the following hierarchy was gotten: Institute for Economic Forecasting, European Commission and National Commission for Prognosis.

The method of relative distance with respect to the maximal performance

1. The distance of each accuracy measure compared with the indicator with the lowest value

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

2. The relative distance for each institution computed as a geometric mean

$$\overline{d_i} = \sqrt[5]{\prod_{j=1}^{5} d_{i\,ind_j}},\ i = 1, 2, 3 \quad (8)$$

3. Assign final ranks According to the values of average relative distances, the final ranks are assigned.

4. The location according to the best institution

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

Table 3

The ranks of institutions according to the accuracy measures (method of relative distance with respect to the best institution)

ACCURACY MEASURE	European Commission	National Commission for Prognosis	Institute for Economic Forecasting
ME	1	1.0338	1.3413
MAE	1.1342	1.550	1
RMSE	1.1465	1.3522	1
U1	1.1597	1.3489	1
U2	1.1623	1.0987	1
Average relative distance	1.1188	1.2628	1.0605
Ranks	2	3	1
Location (%)	105.4970	119.0771	100

Source: own computations using Excel.

The method of relative distance with respect to the best institution gave the same results as the previous methods. The lowest average relative distance was registered by the Institute for Economic Forecasting (1.0592).

The Diebold-Mariano test (DM test) is utilized to check if two forecasts have the same accuracy. The following steps are applied:

✓The difference between the squared errors of forecasts (e^2) to compare and the squared errors of reference forecasts (e^{*2}): $d_{t,t} = \left(e^{2}_{t,t}\right) - \left(e^{*2}_{t,t}\right)$

✓The following model is estimated: $d_{t,t} = a + \varepsilon_t$

✓We test if "a" differs from zero, where the null hypothesis is that a=0 (equal forecasts). A p-value less than 0.05 implies the rejection of the null hypothesis for a probability of 95% in guaranteeing the results.

The following variables are computed: d1, d2 and d3 to make comparisons between EC and NCP forecasts, EC and IEF predictions, respectively NCP and IEF expectations. All the parameters are zero from statistical point of view, so there are not significant differences between the forecasts provided by the three institutions in terms of accuracy. The regression models are estimated in EViews and the results are presented in ***Appendix 1***. So, the accuracy test showed that there are not significant differences between the forecasts provided by the three institutions. If we take into account the results based on accuracy indicators and those of the DM test, we conclude the best predictions are those of IEF, followed by EC and NCP, but the differences between the unemployment rate forecasts are not too big.

By applying qualitative tests for directional accuracy we check if there is a correct prediction of the change. A test of independence between the effective values and the direction of change can be applied in this situation, the null hypothesis showing the independence. A probability less than 0.05 implies the rejection of null hypothesis. All the asymptotic significances are greater than 0.05, according to ***Appendix 2***, fact that makes us to conclude that the directional changes in the outturn are independent from the predictions.

Strategies to improve the accuracy of unemployment rate predictions

Bratu (2012) specify her own strategies of improving the accuracy: (combined forecasts, regressions models, historical accuracy method, use of filters and exponential smoothing techniques).

The most utilized combination approaches are:

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

Bates and Granger (1969) used the predictions f1;t and f2;t, for the same variable Xt, 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: $c_t = m \cdot f_{1t} + (1-m) \cdot f_{2t}$. The error of the combined forecast is: $e_{c,t} = m \cdot e_{1t} + (1-m) \cdot e_{2t}$. The mean of the combined forecast is zero and the variance is:

$\sigma_c^2 = m^2 \cdot \sigma_1^2 + (1-m)^2 \cdot \sigma_{2t}^2 + 2 \cdot m \cdot (1-m) \cdot \sigma_{12}$. The optimal value for m is (m_{opt}):

$$m_{opt} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2 \cdot \sigma_{12}} \quad (10)$$

The initial forecasts are inversely weighted to the relative mean squared forecast error (MSE). $m_{inv} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$ (11)

Equally weighted combined predictions (EW) supposes the same weights to all models.

The U Theil's statistics were computed for the combined forecasts based on the three schemes, the results being shown in the following table (Table 4):

Table 4

The accuracy of combined forecasts for unemployment rate (2001-2011)

Accuracy indicator	EC+NCP forecasts	EC+IEF forecasts	NCP+IEF forecasts
U_1 (optimal scheme)	**0.0846**	**0.0666**	0.1254
U_2 (optimal scheme)	**0.9867**	**0.7130**	1.1063
U_1 (inverse MSE scheme)	**0.0864**	**0.0553**	**0.1105**
U_2 (inverse MSE scheme)	1.0026	**0.5888**	1.0116
U_1 (equally weighted scheme)	**0.0861**	**0.0739**	**0.0888**
U_2 (equally weighted scheme)	**0.9207**	**0.7933**	**0.9134**

Source: author's computations using Excel.

The combined forecasts proved to be a good strategy of improving the accuracy when EC and NCP forecasts, respectively EC and IEF predictions are combined using OPT and INV schemes. Only if equally weighted scheme is utilized we obtained better forecasts for the combined predictions of NCP and IEF. The most accurate forecasts are those resulted from combining EC and IEF expectations. All the combined predictions are better than the naïve ones excepting those of NCP and IEF using OPT scheme.

We test the biasedness of the combined forecasts. Only the combined forecasts based on CE and IEF expectations are biased, all the other predictions being unbiased. So, the combined forecasts are a very good strategy of getting unbiased forecasts.

Each combined forecast based on INV scheme provided different information if we make comparisons of two forecasts from this group. The combined forecasts of CE and IEF and those of NCP and IEF are relative efficient with respect to the combined predictions of CE and NCP. These efficient combined forecasts have a better performance than the original ones of the institutions in what concerns the efficiency.

Another technique of improving the forecasts accuracy used by Bratu (Simionescu) (2013) is the application of filters to the predicted data. The author recommends also the use of exponential smoothing methods like Holts Winters.

Hodrick-Prescott filter and Holt-Winters exponential technique were applied to the original predictions and the accuracy of new forecasts was evaluated. *Holt-Winters Simple exponential smoothing method* is recommended for data series with linear trend and without seasonal variations. The Hodrick-Prescott (HP) filter is very used in macroeconomics to extract the trend of the data series and separate the cyclical component of the time series. The smoothed data obtained are more sensitive to long term changes.

Table 5

The accuracy of filtered and smoothed forecasts for unemployment rate (2001-2011)

Accuracy measure	EC Filtered forecasts	NCP Filtered forecasts	IEF Filtered forecasts	EC smoothed forecasts	NCP smoothed forecasts	IEF smoothed forecasts
U_1	0.1316	0.1049	0.1043	0.1298	0.1291	0.1173
U_2	1.3966	0.9297	1.0721	1.3421	1.1795	1.2626

Source: author's computations using Excel.

Excepting NCP filtered forecasts, all the predictions based on HP filter and HW technique are less accurate than the naïve forecasts. Indeed, the NCP forecasts accuracy is improved, because a smaller value for U_1 was registered for the filtered predictions. The Holt-Winters smoothing technique did not improve the forecasts accuracy. So, the HP filter application is a good strategy of improving only the NCP forecasts. However, the combined predictions remain a better strategy. The filters or the smoothing techniques give god results only if there is not a change in forecasts direction compared to the real values.

Conclusions

In addition to economic analysis, the elaboration of forecasts is an essential aspect that conducts the way of developing the activity al macroeconomic level. But any forecast must be accompanied by macroeconomic explanations of its accuracy. The purpose of this evaluation is related to different aspects: the improvement of the model on which the forecast was based, adjustment of government policies, the planning of results. Basically, performance evaluation in this context refers directly to the degree of trust conferred to the prediction. Although the literature on forecasting methods and techniques used in describing the evolution of an economic phenomenon is particularly rich, surprisingly, few researchers have dealt with the methods used to improve the measurement of forecast uncertainty. The aspect is important, because the macroeconomic predictions must not be easily accepted, taking into account the negative consequences of macroeconomic forecasts failures, consequences that affect the state policies. The decisions of economic policy are based on these forecasts. Hence, there is an evident interest of improving their performance.

In our study, we assessed the unemployment forecasts performance for the predictions provided during 2001-2011 by three institutions: European Commission, National Commission for Prognosis and Institute of Economic Forecasting. The best accuracy is provided by IEF, followed by EC and NCP. This hierarchy resulted from the application of the multi-criteria ranking, but also from the measurement of accuracy indicators, as U_1, used in making comparisons between forecasts.

The combined forecasts using the three classical schemes are a good strategy of improving the accuracy, most of the combined predictions being better than the initial ones. Filtered forecasts based on HP filter or smoothed ones based on Holt-Winters technique succeeded in improving only the NCP forecasts.

The forecasts accuracy should be a priority for the public that uses these predictions in underlying the decisional process. The combined forecasts and in some cases the filtered and smoothed predictions are a very good strategy of getting improvements in accuracy for the unemployment rate predictions.

REFERENCES

Abreu, I., *International Organisations' vs. Private Analysts' Forecasts: An Evaluation*, Banco de Portugal, article available at: http://www.bportugal.pt/enUS/BdP%20Publications%20Research/wp201120.pdf, 2011.

Allan, G., *Evaluating the Usefulness of Forecasts of Relative Growth*, Strathclyde, "Discussion Papers in Economics", No. 12-14, 2012.

Armstrong, J.S. and Fildes, R., *On the Selection of Error Measures for Comparisons among Forecasting Methods*, "Journal of Forecasting", 14, p. 67-71, 1995.

Bates, J. and Granger, C.W.J., *The Combination of Forecasts*, "Operations Research Quarterly", 20(4), 451-468, 1969.

Bratu (Simionescu), Mihaela, *Filters or Holt Winters Technique to Improve the Forecasts for USA Inflation Rate?*, "Acta Universitatis Danubius. Œconomica", Vol. 9, issue no. 1/2013 (forthcoming), 2013.

Bratu, M., *Strategies to Improve the Accuracy of Macroeconomic Forecasts in USA*, LAP LAMBERT Academic Publishing, ISBN-10: 3848403196, ISBN-13: 978-3848403196, 2012.

Diebold, F.X. and Mariano, R., *Comparing Predictive Accuracy*, "Journal of Business and Economic Statistics", 13, p. 253-265, 1995.

Dovern, J. and Weisser, J., *Accuracy, Unbiasedness and Efficiency of Professional Macroeconomic Forecasts: An Empirical Comparison for the G7*, "International Journal of Forecasting", 27 (2), p. 452-465, 2011.

Franses, P.H.; McAleer, M. and Legerstee, R., *Evaluating Macroeconomic Forecasts: A Concise Review of Some Recent Developments,* Working paper/Department of Economics and Finance, University of Canterbury, 2012.

Gorr, W.L., *Forecast Accuracy Measures for Exception Reporting Using Receiver Operating Characteristic Curves*, "International Journal of Forecasting", Volume 25, Issue 1, January-March 2009, p. 48-61, 2009.

Heilemann, U. and Stekler, H., *Introduction to the Future of Macroeconomic Forecasting*, "International Journal of Forecasting", 23(2), p. 159-165, 2007.

Ruth, K., *Macroeconomic Forecasting in the EMU: Does Disaggregate Modeling Improve Forecast Accuracy?*, "Journal of Policy Modelling", Volume 30, Issue 3, May-June 2008, p. 417-429, 2008.

APPENDIX 1

The results of Diebold-Mariano test in EViews

Dependent Variable: D1
Method: Least Squares
Date: 11/22/12 Time: 13:02
Sample: 2001 2011
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.874545	1.187738	-0.736312	0.4785
R-squared	0.000000	Mean dependent var		-0.874545
Adjusted R-squared	0.000000	S.D. dependent var		3.939283
S.E. of regression	3.939283	Akaike info criterion		5.666382
Sum squared resid	155.1795	Schwarz criterion		5.702555
Log likelihood	-30.16510	Durbin-Watson stat		1.518619

Dependent Variable: D2
Method: Least Squares
Date: 11/22/12 Time: 13:02
Sample: 2001 2011
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.530909	0.624816	0.849704	0.4154
R-squared	0.000000	Mean dependent var		0.530909
Adjusted R-squared	0.000000	S.D. dependent var		2.072281
S.E. of regression	2.072281	Akaike info criterion		4.381685
Sum squared resid	42.94349	Schwarz criterion		4.417857
Log likelihood	-23.09927	Durbin-Watson stat		1.521367

Dependent Variable: D3
Method: Least Squares
Date: 11/22/12 Time: 13:03
Sample: 2001 2011
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.405455	0.886219	1.585900	0.1438
R-squared	0.000000	Mean dependent var		1.405455
Adjusted R-squared	0.000000	S.D. dependent var		2.939256
S.E. of regression	2.939256	Akaike info criterion		5.080698
Sum squared resid	86.39227	Schwarz criterion		5.116871
Log likelihood	-26.94384	Durbin-Watson stat		1.686150

APPENDIX 2

The results of tests for directional accuracy

Test Statistics

	ur	Ec
Chi-Square	.818[a]	1.273[b]
Df	9	8
Asymp. Sig.	1.000	.996

Test Statistics

	ur	Ncp
Chi-Square	.818[a]	.000[b]
Df	9	10
Asymp. Sig.	1.000	1.000

Test Statistics

	ur	Ief
Chi-Square	.818[a]	1.273[b]
Df	9	8
Asymp. Sig.	1.000	.996