

THE REDUCTION OF UNCERTAINTY IN MAKING DECISIONS BY EVALUATING THE MACROECONOMIC FORECASTS PERFORMANCE IN ROMANIA

MIHAELA BRATU¹

ARTICLE INFO

JEL classification: E21, E27, C51, C53

Keywords:

- forecasts
- accuracy
- bias
- efficiency
- forecasts intervals

ABSTRACT

The evaluation of macroeconomic forecasts performance does not include only the calculating of some statistical measures, rather controversial in literature, like root mean squares error or absolute mean error. In theory and economic practice, three directions have been traced regarding the evaluation of forecasts performance: the analyse of accuracy, bias and efficiency. Using the forecasted values on medium run of inflation rate and unemployment rate through the period from 2004-2010 in Romania, we get a better degree of accuracy and a lower efficiency for forecasts made by National Commission of Forecasting comparing to those based of Dobrescu model used by Institute of Economic Forecasting. Following the international tendency, the forecasts are, in all cases, biased because of difficulties in precise anticipation of shocks which affect the economy. Forecasts performance is indestructible related by their uncertainty, RMSE, the measure of evaluating the accuracy being used in building forecast intervals based on historical errors. For forecasted values of inflation rate published by National Bank of Romania we propose a new way of building forecast interval in order to take into account the economic shocks.



¹ Faculty of Cybernetics, Statistics and Economic Informatics, Academy of Economic Studies, tel.: 0748112411, mihaela_mbl@yahoo.com, Bucharest

I. INTRODUCTION

In addition to economic analysis, the elaboration of forecasts is an essential aspect that conducts the way of developing the activity at macroeconomic level. But any forecast must be accompanied by macroeconomic explanations of its performance. 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 literature there are 3 directions in evaluating the performance of macroeconomic forecasts: accuracy, bias and efficiency. A large number of articles have considered the problem of comparing the accuracy measures, contributions in the field are related of names like: Leith and Tanner-1990, Makridakis- 1993, Yokum and Armstrong-1995, Tashman-2000, Makridakis and Hibon-2000, Koehler, Martin și Witt -2002, Hyndman -2006 . and Witt -2002, Hyndman-2006.

Meese and Rogoff's paper, „ Empirical exchange rate models of the seventies”, remains the starting point for many researches on the comparing of accuracy and bias. Recently, Dovern J. and J. Weisser (2011) examines in the same article, „ Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7” the three criteria using the empirical data from the G7 economies.

II. FORECASTS ACCURACY IN LITERATURE

Forecast accuracy is a large chapter in the literature related to the evaluation of forecasts uncertainty. There are two methods used in comparing the prediction quality: vertical methods (eg, mean squared error) and horizontal methods (such as distance in time). An exhaustive presentation of the problem taking into account all the achievements in literature is not possible, but will outline some important conclusions.

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) analyze 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.

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.

Mariano R.S. (2000) presents the most significant tests of forecasts accuracy, including the changes of his test- Diebold Mariano (DM). Since the normal distribution is a poor approximation of the distribution of low volume data series, Harvey, Leybourne, and Newbold improve the properties of finite data sets, applying some corrections: the change of DM statistics to eliminate the bias and to make comparison not to normal distribution, but to the t-Student. Clark evaluates the power of some tests of equal forecast accuracy, such as modified versions of DM test or those of Newey and West, which are based on the Bartlett kernel and a fixed length of data series. Meese and Rogoff in their study from 1983, "The empirical exchange rate models of the seventies" compared the RMSE and the bias of exchange rate forecasts, that were based on structural models and they made a conclusion that was later used to improve macroeconomic forecasts performance. They have thus demonstrated that random walk process generates better forecasts than structural models.

In the evaluation of a forecast based on a model, Clements and Hendry (2005) identify six important aspects to be studied: ex-ante and ex post, evaluation, forecast horizon length, the quality of a model to be conditional or not, internal and external standards, testing the stability and the significance of the models, testing parameters stability and assurance of their continuous updating.

In literature, there are several traditional ways of measurement, which can be ranked according to the dependence or independence of measurement scale. A complete classification is made by RJ Hyndman and AB Koehler (2005) in their reference study in the field, "Another Look at Measures of Forecast Accuracy":

- Scale-dependent measures

The most used measures of scale dependent accuracy are:

-> Mean-Square Error (MSE) = average (e_t^2)

-> Root Mean Square Error (RMSE) = \sqrt{MSE}

-> Mean Absolute Error (MAE) = average ($|e_t|$)

-> Median Absolute Error (MdAE) = median ($|e_t|$)

RMSE and MSE are commonly used in statistical modeling, although they are affected by outliers more than other measures.

- Scale-independent errors:

-> *Measures based on percentage errors*

The percentage error is given by: $p_t = \frac{e_t}{X_t} \cdot 100$

The most common measures based on percentage errors are:

* Mean Absolute Percentage Error (MAPE) = average ($|p_t|$)

* Median Absolute Percentage Error (MdAPE) = median ($|p_t|$)

* Root Mean Square Percentage Error (RMSPE) = geometric mean (p_t^2)

* Root Median Square Percentage Error (RMdSPE) = median (p_t^2)

When X_t takes the value 0, the percentage error becomes infinite or it is not defined and the measure distribution is highly skewed, which is a major disadvantage. Makridakis introduced symmetrical measures in order to avoid another disadvantage of MAPE and MdAPE, ie, too large penalizing made to positive errors in comparison with the negative ones.

* Mean Absolute Percentage Error (sMAPE) = average ($\frac{|X_t - F_t|}{X_t + F_t} \cdot 200$)

* Symmetric Median Absolute Percentage Error (sMdAPE) = median ($\frac{|X_t - F_t|}{X_t + F_t} \cdot 200$), where

F_t - forecast of X_t .

-> *Measures based on relative errors*

It is considered that $r_t = \frac{e_t}{e_t^*}$, where e_t^* is the forecast error for the reference model.

* Mean Relative Absolute Error (MRAE) = average ($|r_t|$)

* Median Relative Absolute Error (MdRAE) = median ($|r_t|$)

* Geometric Mean Relative Absolute Error (GMRAE) = geometric mean $(|r_t|)$

A major disadvantage is the too low value for the error of benchmark forecast.

For example, the relative RMSE is calculated: $rel_RMSE = \frac{RMSE}{RMSE_b}$, where $RMSE_b$ is the RMSE of “benchmark model”

Relative measures can be defined for MFA MDAE, MAPE. When the benchmark model is a random walk, it is used rel_RMSE , which is actually Theil’s U statistic. Random walk or naive model is used the most, but it may be replaced with naive2 method, in which the forecasts are based on the latest seasonally adjusted values.

- Free-scale error metrics (resulted from dividing each error at average error)

Hyndman and Koehler introduce in this class of errors “Mean Absolute Scaled Error” (MASE) in order to compare the accuracy of forecasts of more time series.

Other authors, like Fildes R. and Steckler H. (2000) use another criterion to classify the accuracy measures. 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: $e_t(t+k)$. Indicators used to evaluate the forecast accuracy can be classified according to their usage. Thus, the forecast accuracy measurement can be done independently or by comparison with another forecast.

A. Independent measures of accuracy

In this case, it is usually used a loss function, but we can also choose the distance criterion proposed by Granger and Jeon for evaluating forecasts based on economic models. The most used indicators are:

- a) Mean Square Error (MSE)
- b) Root Mean Squared Error (RMSE)
- c) Generalized Forecast Error Second Moment (GFESM)
- d) Mean Absolute Percentage Error (MAPE)
- e) Symmetric Median Absolute Percent Error (SMAPE)
- f) Mean error (ME)

g) Mean absolute error (MAE).

In practice, the most used measures of forecast error 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)

$$MAE = \frac{1}{n} \sum_{j=1}^n | e_x(T_0 + j, k) | \quad (3)$$

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 if they are expressed as percentage. Fair, Jenkins, Diebold and Baillie show that these measures include average errors with different degrees of variability. The purpose of using these indicators is related to the characterization of distribution errors. Clements and Hendry have proposed a generalized version of the RMSE based on errors intercorrelation, when at least two series of macroeconomic data are used. If we have two forecasts with the same mean absolute error, RMSE penalizes the one with the biggest errors.

B. Measures for the evaluation of the relative accuracy of forecasts

Relative accuracy measures are related to the comparison of the forecast with a forecast of reference, found in the literature as the 'benchmark forecast' or 'naive forecast'. However, it remains a subjective step to choose the forecast used for comparison. Problems may occur in this case are related to these aspects: the existence of outliers or inappropriate choice of models used for predictions and the emergence of shocks. A first measure of relative accuracy is Theil's U statistic, which uses as reference forecast the last observed value recorded in the data series. Collopy and Armstrong have proposed instead of U a new similar indicator

(RAE). Thompson improved MSE indicator, suggesting a statistically determined MSE- log mean squared error ratio.

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. U-Theil proposed the calculation of U, that takes into account both changes in the negative and the positive sense of an indicator:

$$U = \sqrt{\frac{\sum (X_{t+k} - \hat{X}_t(k))^2}{\sum X_{t+k}^2}} \quad (4)$$

Hyndman and Koehler proposed scale errors based on the mean absolute error of a naive forecasting method. MAE serves therefore, as denominator. Using this method, it is generated the one-step-ahead forecast. Scale error is defined as:

$$es_t = \frac{e_t}{\frac{1}{n-1} \sum_{i=2}^n |X_i - X_{i-1}|} \quad (5)$$

and mean absolute scale error as: $MASE = \text{mean} | es_t |$.

Naive forecast values are considered to be the current ones recorded during the previous period. MASE is used both to compare forecast methods applied to a given set of data and also to compare the accuracy of several series. If the scale error is less than 1, the compared forecast is better than the reference one (naïve forecast).

One of the business objectives in forecasting was empirical validation. Famous results have been registered by Makridakis and Hibon, who lead research groups around the world to make comparisons between different methods of forecasting. In literature the results are known as "M-competition". Ex-ante forecast errors for 21 methods were compared with predictions based on 1001 economic series. Accuracy criteria used in the M competition were: central tendency error (APE median), MSE, which gives more weight to larger error, MAPE, which is the basic measure. This is the measure recommended in reference books in forecast accuracy domain, written by Hanke and Reitsch or Bowerman, O'Connell and Koehler.

Armstrong and Collopy use MdRAE, MdAPE and GMRAE, the last two measures being also recommended by Fildes, who also uses GRMSE (geometric mean squared relative error). In M3 competition, Makridakis and Hibon recommended MdRAE, sMAPE and sMdAPE.

Clements and Hendry propose a linear transformation of original series and the robustness evaluation of a measure used to quantify the error.

It is made distinguish between errors within the actual data series and errors on forecast horizon, which is enshrined in the literature as in sample errors out sample errors. A part of the forecast errors are generated by the fact that the series is affected by errors and by

the forecast itself. One-step-ahead errors must be, normally, uncorrelated and of zero mean, conditions to be respected also by in sample and out sample errors. Some authors recommend the use of Kullback-Leibler divergence in order to compare the forecasts in terms of accuracy.

Kullback-Leibler distance (divergence) (KL) also known as informational divergence or relative entropy is an essential indicator used in statistics and information theory. It is used in hypothesis testing, adaptive estimation, estimation of maximum or minimum function. KL divergence between two probability densities, A and B, defined on a discrete finite set X is:

$$D(A||B) = \sum_{x \in X} A(x) \cdot \log \frac{A(x)}{B(x)} = E\left(\log \frac{A(x)}{B(x)}\right). \quad (6)$$

In the continuous case, if a and b are two probability densities with the mean m, then:

$$D(a||b) = \int a \log \frac{a}{b} dm \quad (7)$$

Using the KL divergence to measure forecast accuracy, some indicators are proposed:

- KL-N (quadratic loss function is scaled with the variance estimated):
- KL-N1 using the estimated variance only for the last five periods;
- KL-N2 takes into account only the last 10 periods in calculating the estimated variance;
- KL-DE1 based on a double exponential distribution of the error;
- KL-DE2, unlike the KL-DE1 use another estimator for the scale parameter.

Other measures to used to evaluate the accuracy are correlation between forecasted values and actual values (measured by coefficient of determination of changes in the series of values and the ones in forecasted series), the percentage of turning points forecast, calculated for binary variables by rank (score) Kuiper, conditional efficiency (for comparing two different forecasts based on the same regression model).

Recent studies target 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.

Ericsson NR (1992) shows that the parameters stability and mean square error of prediction are two key measures in evaluation of forecast accuracy, but they are not sufficient and it is necessary the introduction of a new statistical test.

Considering the AR (1) process, which is represented as $y_t = \beta y_{t-1} + u_t$, Hoque A., Magnus JR and Pesaran B. show that for small values of β the prediction mean square error is a decreasing function in comparison with the number of forecast periods.

Granger CWJ și Jeon Y. CWJ Granger and Y. Jeon (2003) consider four models for U.S. inflation: an univariate model, a model based on an indicator used to measure inflation, an univariate model based on the two previous models and a bivariate model. Applying the

mean square error criterion, the best prediction made is the one based on an autoregressive model of order 1 (AR (1)). Applying distance-time method, the best model is the one based on an indicator used to measure the inflation.

Ledolter J. (2006) compares the mean square error of ex-post and ex ante forecasts of regression models with transfer function with the mean square error of univariate models that ignore the covariance and show superiority of predictions based on transfer functions.

T. Teräsvirta, van Dijk D., Medeiros MC (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 generated better predictions than the models using an approach from private to general.

U. Heilemann and Stekler H. (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.

Ruth K. (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.

Gorr WL (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.

Dovern J. and J. 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 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.

In Netherlands, experts make predictions starting from the macroeconomic model used by the Netherlands Bureau for Economic Policy Analysis (CPB). For the period 1997-2008 was reconstructed the model of the experts macroeconomic variables evolution and it was compared with the base model. The conclusions of Franses PH, Kranendonk HC and Lanser D. (2011) were that the CPB model forecasts are in general biased and with a higher degree of accuracy.

Many studies in literature refer to the combining of two methods based on the same model (such as eg bayesian mediation model), but French and Insura point out that a combination between model predictions and expert assessments has not been proposed yet.

III. EVALUATION OF MACROECONOMIC FORECASTS ACCURACY IN ROMANIA

In this study we evaluate the accuracy of forecasts made by principal institutions in Romania for inflation rate and unemployment rate: Institute of Economic Forecasting and National Commission of Forecasting

We consider the values of inflation and the unemployment rate projected by National Commission of Forecasting for the period 2004-2010. The indicators mentioned above are calculated for the forecast errors. The values of ME show the underestimation tendency of inflation. Moreover, MAE has the same value, showing the persistence of this underestimation.

The result is confirmed by practice, taking into account the shocks recorded in this period. Thus, in 2004, currency appreciation, the restrictiveness of fiscal policy and greater tendency towards savings have contributed to the disinflation process. However, a number of other factors determined the increase of inflationary pressure: the gross average wage growth, the accumulation of arrears and the increase of consumption. Disinflation process was resumed in 2005 due to weaker dynamics of administered prices and currency appreciation against the euro. A further acceleration of the disinflation process registered in 2006 due to volatile price downturn, the reducing of the basic component and the increased competition in the retail market.

But in 2007, the inflation trajectory changed, the annual inflation increase being attributed shocks like: unexpected increase of volatile prices for agricultural products, increase to food prices, the RON exchange rate correction, all of these in the context of an excessive demand. If the first half of 2008 inflationary pressures were generated by supply shocks (food market tensions, the rise in import prices of agricultural raw materials and unprocessed products) and demand shocks (increase in fuel prices and natural gas). Since August 2008 these factors have started to downsize, but influences on the demand generated by easy fiscal policy persisted, the maintenance of laxity in wage policy, expansion of lending. 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.

TABLE 1 - Values for inflation rate in Romania in 2004-2010

Year	Inflation rate (%)			
	INS	National Commission of Forecasting	Institute of Economic Forecasting (Dobrescu model)	Institute of Economic Forecasting (PEP program)
2004	11,9	9	6,2	12
2005	9	7	13,74	9
2006	6,56	4	6,88	7
2007	4,84	4,5	6,82	5
2008	7,85	3,8	5,88	3,6
2009	5,59	4,5	4,36	
2010	6,09	3,5	4,04	

SOURCE: Author

National Institute of Statistics - www.insse.ro, National Commission of Forecasting- www.cnp.ro, Dobrescu E. (2006), Macromodels of the Romanian Transition Economy, Expert Publishing house, Bucharest

ME indicator shows an underestimation of the annual inflation rate to 0.29 percentage points. The very low RMSE value indicates a low variability of the series of errors.

TABLE 2—The average and standard deviation of the inflation rate, respectively unemployment rate in 2004-2010

Statistical indicators	Inflation Rate (%)	Unemployment rate (%)
Average (2004-2010)	7,38	5,82
Standard deviation	2,62	1,27

SOURCE : own calculations using Excel

TABLE 3— Indicators of forecast accuracy for inflation and unemployment rate in Romania (2004-2010)

Forecast Errors	Inflation rate	Unemployment rate
RMSE*	0,1	0,09
ME*	0,29	-0,05
MAE*	0,29	0,27
MASE*	0,041	0,038
Stat. U Theil	0,33**	0,14***

* Percentage points

** Compared to Dobrescu model

*** Compared to the MA (1) model- dynamic forecast

SOURCE: Author

Unlike the forecast inflation rate, the negative ME value indicates an overestimation of the unemployment rate by 0.05 percentage points. Errors variability is very small, because RMSE has a value of only 0.09 percentage points. Labor market immediately reacted to the crisis in 2009 through higher unemployment, but also slower annual growth in wages.

For inflation, which recorded only positive rates of growth in the analyzed period, naive model have to extrapolate the latest trend. If Moore had proposed the comparison to projections based on an extrapolation method, the development of VAR and ARIMA models, impose their use as benchmark models. A value of U less than one indicates lower forecast errors than those from the naïve model. The same conclusion is reached when it is calculated the scale error proposed by Hyndman and Koehler.

The U statistic was calculated, taking as reference the inflation forecasts based on Dobrescu model and those of unemployment rate based on an MA (1) model.

Unemployment rates from 1991 to 2010 in Romania follow a MA(1) process:

$r_{unemployment} = 7,761 + e_t - 0,805 \cdot e_{t-1}$. We used two forecasting techniques: ex-post (corresponding the dynamic forecast) and ex-ante forecasts used for static forecast.

is the one that forecasts the value in period only based on data up to time then, for all periods that are already projected using data from period makes forecasts based only on registered data.

A dynamic forecasting was made in EViews in 2004-2010. The ex- post technique was applied first, using the first 13 values of the unemployment rate for the model and the rest for prediction. Ex-ante technique of forecasting is based on all values.

Eviews Program displays a set of indicators to evaluate the model reliability:

- RMSE (Root Mean Squared Error), which must have a small value as possible;
- MAE (Mean Absolute Error);
- MAPE (Mean Absolute Percent Error);
- Theil's inequality coefficient (takes values in (0,1), a value close to 0 indicating a good adjustment; in this case has a low value, so, the adjustment is very good);
- Bias Proportion has to be small (in this case is quite large);
- Variance Proportion has to be as small; in this case is rather close to 0;
- Covariance Proportion is desirable to be as large as possible; in this case is very small.

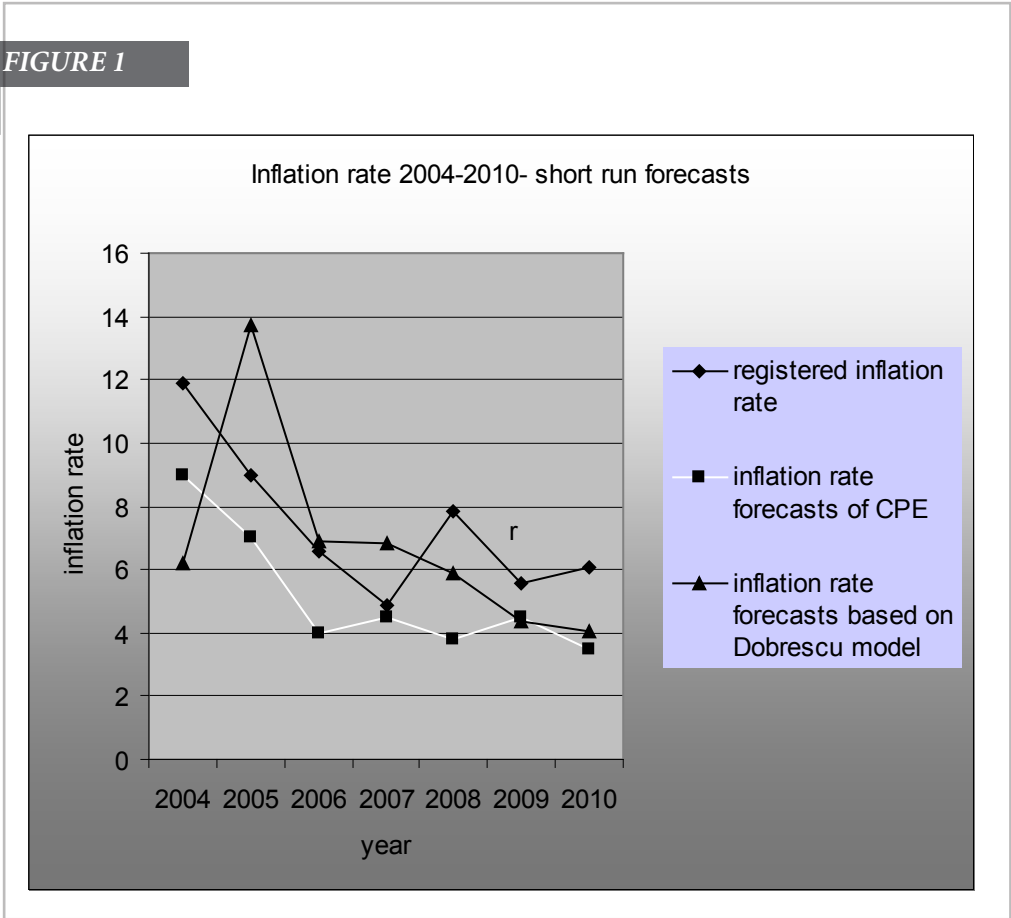
Inflation forecast errors of Dobrescu model are higher than those of the Commission of Forecasting. This conclusion was also reached by comparing RMSE and the Theil's U coefficient modified by changing the benchmark model.

RMSE for forecasted inflation rate of Dobrescu model is 1.03 percentage points compared to 0.29 percentage points, the value calculated for the Commission of Forecasting predictions.

Taking as reference the Dobrescu model forecast inflation rate the U statistics has a value of 0.33, which means that the Commission's forecasts are better. This conclusion is also results from the graphic analysis below.

SHORT RUN FORECASTS FOR THE INFLATION RATE (2004-2010)

FIGURE 1



SOURCE: Table 1

To make a comparison of forecasts characteristics, the loss-function values are analyzed-root mean squared error (RMSE), which calculates the forecast deviation from the actual values recorded. It is estimated that a prediction is much closer to the real evolution as much as RMSE value is lower. Static forecast is superior to the dynamic one for unemployment rate which follows a MA (1), because of the lower value of RMSE.

The reduction of uncertainty in making decisions by evaluating the macroeconomic forecasts performance in Romania

TABLE 4— Unemployment rate in Romania in 2004-2010

Year	Unemployment rate (%)				
	National Commission of Forecasting	Static Forecast MA (1)	Dynamic Forecast MA(1)	Institute of Economic Forecasting (Dobrescu model)	Institute of Economic Forecasting (PEP program)
2004	6,3	7,5599876	7,756	8,02	8
2005	5,9	6,7465086	7,761	7,9	7,9
2006	4	7,0794254	7,761	7,65	7,8
2007	4,4	5,2815697	7,761	7,42	7,6
2008	5,8	7,0511957	7,761	7,19	7,4
2009	7,5	6,7535875	7,761	6,99	
2010	6,9	8,3619821	7,761	6,83	

SOURCE: National Commission of Forecasting - www.cnp.ro, Dobrescu E. (2006), Macromodels of the Romanian Transition Economy, Expert Publishing House, Bucharest

Commission of Forecasting makes predictions using quarterly model of Romanian economy. On long run and short run, index of prices used to calculate inflation is estimated according to the index of M2 and the exchange rate.

RMSE for unemployment rate forecasts of Commission of Forecasting (1,27) is lower than the one of MA(1) model, which has the value 2,24 for a dynamic forecast in EViews.

EViews static forecast is better than the dynamic one taking into consideration all indicators used in reliability evaluation, but the RMSE is higher for predicted values of Commission of Forecasting. The modified U statistic is calculated in order to compare forecasts of the Commission with those based on the MA (1) model and a value of 0.07 resulted when static forecast is chosen as 'benchmark forecast' and 0.14 for a dynamic forecast used as reference forecast. In conclusion, we obtained smaller forecast errors for the unemployment rate provided by the Commission of Forecasting in comparison with the errors resulting from a moving average model.

E. Dobrescu (2003) groups the constraints affecting economic performance in two categories, one of these taking into account demand factors (limited access to foreign markets and the decline in real terms of capacity absorption) and the other one the supply factors (the activity of companies lacking performance). For the period 2003-2007, the academician Dobrescu has developed some scenarios of the evolution of main macroeconomic indicators, the workings being made in PEP program (Pre-Accession Economic Program).

For the period 2004-2010 the unemployment rate projected in the Dobrescu model achieved a RMSE value of 0.46. That means an error greater than the one of the forecast of Commission of forecasting. IPE also used the PEP program (Pre-Accession Economic Programme) to forecast the unemployment rates from 200 to 2008. RMSE has a value of about 0.58, higher than 0.54, the result when Dobrescu model is used.

An optimal macroeconomic forecast is unbiased and efficient, in the macroeconomic

framework taking into account the rational expectations hypothesis. This implies a zero average prediction error.

IV. BIAS AND EFFICIENCY OF FORECASTS

Bias and forecast efficiency have an important role in the literature that examines the problem of forecast performance.

A. Bias forecasts

Corderie (2003) shows that McNees (1978, 1987) and Fair and Schiller (1989) brought among the first contributions in the field of bias and efficiency of the individual forecasts made by consensus. Figlewski and Wachtel noted that early results showed that the projections of private sector are biased and uncorrelated with the rational expectations hypothesis. Batchelor R. (2007) detected the presence of systematic bias in the forecast of real GDP and inflation made by the private sector in the G7 countries during 1990-2005. The measuring and test of bias was based on regression models and nonparametric tests of accuracy of the ranks. Empirical researches have shown a conclusion already presented in the literature, namely, the discrepancy between rational expectations tests and the too pessimistic or too optimistic forecasts.

Bias in this context implies a zero mean forecast error series. In the literature rationality tests are used to check if the forecasts are optimal in relation to a certain criterion, eg, if they are biased or ensure a good informational efficiency. The standard test of forecast bias-test-Mincer-Zarnowitz starts from this model : $A_t = a + b \cdot P_t + e_t$.

A_t -Current values, P_t - predicted values

We test Test the hypothesis of unbiased forecasts: $a = 0$, $b = 1$ using a F test, called Mincer-Zarnowitz test.

Holden and Peel proposed a modified version of the test, which is based on forecast errors by testing whether their mean (m) is zero: $A_t - P_t = m + e_t$.

In the first version of the test for the inflation rate resulted that a and b are not significantly different from 0, so the null hypothesis can not be accepted for the forecast made by the Commission of Forecasting. In the second version of the test was properly calculated a p-value of 0.003 for t test, which is lower than the threshold of 0.05. So, with a probability of 95% hypothesis of unbiasedness is rejected. The result is correct, since all expected values were lower than those actually recorded.

The bias of forecasts can be generated by technological shocks, which might be an important factor of economic growth. Wojciech C. and Svetlana M. (2009) realized an empirical research, breaking down inflation into a linear and a nonlinear component resulting from technological shocks, for which were they demonstrate the positive contribution to economic growth.

Accuracy can be improved if it is known that there is autocorrelation between errors and other data available at the time the forecast is made. The correlation indicates an inefficient

use of information from the past. If X_i are the observed variables that influence the forecast,

$$\text{then: } e_{\gamma}(t-k, k) = \gamma + \sum_i \sum_{j>k} \delta_{i,j} X_i(t-j) + e_t.$$

If γ and $\delta_{i,j}$ are significantly different from zero, the forecasts can be improved if one takes into account the influence of X_i variables. However, Jeong and Maddie have demonstrated that tests of rationality are dependent on assumptions made for regression models. Pain shows that while the data series is nonstationar with unit roots, cointegration tests should be used. In the case of asymmetric loss functions the forecasts are rational, even if the mean error is zero.

Given the shocks in food prices during the chosen period, we check that only a part of the bias was caused precisely by the components of the index of consumer prices of goods used to calculate the inflation. Thus, we calculate a new inflation rate excluding food prices and we use this indicator in the regression model for testing the bias. Using an F test a lower bias than the initial one resulted. So, shocks in the price of food influence inflation forecasts.

B. Forecasts efficiency

Fair and Schiller (1989) propose a test in order to compare the efficiency of two forecasts made by two different institutions for the same variable. To this end, they consider a simple linear regression model: $X_t - X_{t-1} = b_0 + b_1(X_1^p - X_{t-1}) + b_2(X_2^p - X_{t-1})$

X_t - The value recorded for variable X at time t

X_{t-1} - The value recorded at time t-1 for variable X

X_1^p - The predicted value of the first institution

X_2^p - The predicted value of the second institution

If $b_1 > 0$ and $b_2 = 0$, the second institution provides a forecast relatively inefficient, and the first institution forecast contains, in addition to the first one information, an essential piece of information about changes that may occur in the analyzed variable.

If $b_2 > 0$ and $b_1 = 0$, the first institution provides a relatively inefficient forecast.

If both parameters of the regression model are strictly positive, then, each institution brings different information through the forecast.

We compared in terms of efficiency the inflation rate forecasts of Commission of Forecasting and Institute of Economic Forecasting for 2004-2010. After the estimation of model parameters, it resulted that $b_1 = 0.06$, value that does not differ statistically to 0, and for b_2 a value of 1.538 $b_2 > 0$. Therefore, the Commission for Forecasting prediction is inefficient compared with that of the Institute of Economic Forecasting.

V. BUILDING FORECASTS INTERVALS

The problem of building forecast intervals and the determination of distributions was approached quite late in the literature, notable works in this area being written by Cogley (2005), Adolfson (2007), Clark (2009) and Jore (2009), Giordani and Villiani (2009). The results showed an important conclusion: in order to build a forecast interval with a certain probability, the model has to include variances deviation in time.

Kjellberg D. and M. Villani (2010) numbered the advantages and disadvantages of both types of forecasts, the ones based on models and those built by the experts. Forecast methods based on models describe the complex relationships using endogenous variables by its transparence making easy the identification of mistakes that generated wrong predictions. The disadvantages are related to the difficulty of adapting the model to recent changes in the economy, as well as the too simple form of the models. Chatfield shows that forecast intervals are often too narrow not taking into account the uncertainty related to model specification, problem that is encountered also in the experts' assessment. Unlike the forecasts based exclusive on models, expert assessments modify immediately to any change of information related to the predicted phenomenon. Disadvantages in experts assessments are related just to the low degree of transparency, the difficulty of using many explanatory variables outside an explicit model.

A. Building prediction intervals based on historical forecast errors

The build of intervals taking into account the forecasts accuracy is an effective way to highlight the uncertainty that accompanies any forecast made. In following, we used historical forecast errors to determine the forecast interval for inflation. We also used the projected inflation rates at the end of the year published by the National Bank of Romania for each quarter from 2007 to2010. Forecast errors are calculated as the difference between expected value and the registered value. Forecast errors for each quarter are calculated by RMSE.

Forecast intervals are built considering that the forecast error series is normally distributed of zero mean and standard deviation equal to the RMSE corresponding to historical forecast errors. For a probability of $(1-\alpha)$, forecast interval is calculated:

$$(X_t(k) - z_{\alpha/2} \cdot RMSE(k), X_t(k) + z_{\alpha/2} \cdot RMSE(k)), k = 1, \dots, K$$

$X_t(k)$ - punctual forecast for variable X_{t+k} at time t

$z_{\alpha/2}$ - the $\alpha / 2$ quintile of standardized normal distribution.

The table below displays the RMSE and lower and upper limits of the forecast interval for inflation predicted by the central bank with a quarter before ("one-step-ahead").

TABLE 5— The limits of the inflation rate forecast intervals in Romania from 2007 Q1 to 2010 Q4 (based on historical forecast errors)

Quarter	RMSE	Lower limit	Upper limit
2007 T1	0,67	3,18	5,82
2007 T2	0,51	3,31	5,29
2007 T3	0,19	4,42	5,18
2007 T4	1,99	0,79	8,61
2008 T1	1,65	3,06	9,54
2008 T2	2,36	1,57	10,83
2008 T3	2,72	0,07	10,73
2008 T4	2,51	-0,62	9,22
2009 T1	0,77	4,49	7,51
2009 T2	0,59	4,35	6,65
2009 T3	0,11	4,88	5,32
2009 T4	0,06	4,38	4,62
2010 T1	0,43	3,35	5,05
2010 T2	0,02	4,34	4,41
2010 T3	0,27	7,24	8,30
2010 T4	0,31	7,56	8,78

SOURCE: calculations made using data from reports of inflation of National Bank of Romania between 2006-2010 - www.bnr.ro

The forecast intervals based on RMSE are independent of the state of the economy. Therefore, Blix and Sellin proposed the change of the method, so that the interval takes into account of changes in the economy, multiplying RMSE by a factor of uncertainty subjective chosen by the expert in forecasting. Another approach uses, for the series of observations, a model in which time varies. The series of quarterly inflation rates follows an autoregressive AR process in which the series has a residual variance of stochastic type. It is assumed the hypothesis that errors are identically distributed and follows a standardized normal distribution. Then, the regression model can be written:

$$ri = m + \sum_{k=1}^K \varphi_k (ri_{t-k} - m) + \alpha_t \cdot e_t, \text{ where } \alpha_t \text{ is the standard deviation of errors}$$

$\ln \alpha_t^2 = \ln \alpha_{t-1}^2 + \varepsilon_t$, where ε_t follows a normal distribution and $\ln \alpha_t^2$ is a random walk
 We introduce a new statistical measure called the relative volatility or relative variance (variance of T moment in relation with the geometric mean of variances corresponding to the

interval used to calculate RMSE), calculated by the formula: $\beta_r = \frac{\hat{\alpha}_r}{n^{-1} \prod_{t=t_1}^{t_2} \hat{\alpha}_t^{\frac{1}{n}}}$ t_1 and t_2 are the

initial moment and the final one of the period for which RMSE is calculated, the time of the interval bounded of the two moments is: $n = t_1 + t_2 - 1$, and $\hat{\alpha}_T$ is a bayesian estimation.

The new intervals of variation of forecast values will be calculated as follows:

$$(X_t(k) - z_{\alpha/2} \cdot \alpha_t \cdot RMSE(k), X_t(k) + z_{\alpha/2} \cdot \alpha_t \cdot RMSE(k)), k = 1, \dots, K$$

The relative volatility is 0,279.

Relative volatility of Q4 of 2010 was 1.279, which means that a 62.1% decrease in the value of RMSE is necessary to take into account the changes in the economy.

B. The proposal of a new way to build forecast interval for Romania

Between 2007-2010 inflation rates calculated at the end of the quarter may be represented by an AR process of order 1 (AR (1)). To determine the interval of variances of BNR predictions taking into account the state of the economy in each of the periods for which data were recorded, the coefficient which multiplies RMSE is calculated in different way than that recommended in the literature. Inflation is modeled in 2007-2010 as:

$$r_inf_t = 6,917 + 0,714 \cdot r_inf_{t-1} + e_t$$

For an AR process ($X_t = \phi_1 \cdot X_{t-1} + e_t$), the variance is: $\text{var}(X_t) = \frac{\sigma_e^2}{1 + \phi_1^2}$, where

σ_e^2 – AR error variance.

$$\text{The variance of inflation is: } \text{var}(r_inf) = \frac{\sigma_e^2}{1 + 0,714^2} = \frac{1,232}{1,509} = 0,816$$

We introduce as a measure of economic state the indicator δ -relative variance of the phenomenon at a specific time in relation with the variance on the entire time horizon, which

$$\text{for T moment is calculated as: } \delta_T = \frac{[e_T - E(e_t)]^2}{\text{var}(r_inf)} = 0,339$$

TABLE 6—The limits of the inflation rate forecast intervals in Romania from 2007 Q1 to 2010 Q4 (based on own method)

Quarter	e_t	$[e_t - E(e_t)]^2$	δ_t	RMSE	Lower limit	Upper limit
2007 T2	-0,921	0,849	1,040	0,507	3,267	5,333
2007 T3	0,307	0,094	0,116	0,193	4,756	4,844
2007 T4	1,149	1,320	1,618	1,993	-1,621	11,021
2008 T1	1,195	1,428	1,750	1,653	0,629	11,971
2008 T2	0,905	0,819	1,003	2,363	1,552	10,848
2008 T3	0,029	0,001	0,001	2,720	5,394	5,406
2008 T4	-0,967	0,934	1,145	2,510	-1,333	9,933
2009 T1	-0,071	0,005	0,006	0,770	5,991	6,009
2009 T2	-0,722	0,521	0,639	0,587	4,765	6,235
2009 T3	-1,336	1,785	2,188	0,113	4,614	5,586
2009 T4	-0,980	0,961	1,177	0,063	4,354	4,646
2010 T1	-0,603	0,364	0,445	0,434	3,817	4,575
2010 T2	-0,923	0,852	1,044	0,017	4,342	4,412
2010 T3	2,410	5,808	7,118	0,270	3,999	11,535
2010 T4	0,526	0,277	0,340	0,311	7,960	8,374

SOURCE: calculations made using data from reports of inflation of National Bank of Romania between 2006-2010 - www.bnr.ro

In this case, we obtained a relatively large variance, which means that it is necessary a decrease of RMSE value with 66.1% if one takes into the state of the economy in the last quarter of 2010.

VI. CONCLUSIONS

Forecast performance evaluation is an important indicator of the extent to which projections made accomplished their purpose to be closer as much as possible of the registered values. Forecasts accuracy in Romania for inflation and unemployment rate was evaluated for each an institution specialized in the elaboration of forecasts, and comparisons were made showing the superiority of forecasts made by National Commission of Forecasting. However, in terms of efficiency, Institute of Economic Forecasting inflation provides better predictions. Regardless of the institution, all forecasts are biased, an important part of this bias being caused by shocks in the food prices in 2004-2010.

Based on data of inflation forecasts provided quarterly by the Central Bank, forecast intervals were built using the method of historical forecast errors. For Romania, when inflation rates follows an AR (1), we have improved the technique of building forecast intervals taking into account the state of the economy in each period for which data were recorded.

In conclusion, macroeconomic forecasts evaluation is necessary to inform the public about the way in which state institutions predicted the economic phenomenon. Further,

according a certain degree of reliability by studying the results, in the future the public attention will focus on a particular institution in accord with the criterion followed, accuracy or efficiency, whereas the forecasts are generally biased at national and internationally level because of the difficulty to anticipate the structural shocks.

REFERENCES

- Armstrong, Scott J., and Collopy, Fred.** 2000. Another Error Measure for Selection of the Best Forecasting Method: The Unbiased Absolute Percentage Error. *International Journal of Forecasting* 8 (May): 69-80.
- Armstrong, Scott J., and Fildes, Robert.** 1991. On the selection of Error Measures for Comparisons Among Forecasting Methods. *Journal of Forecasting* 14 (August): 67-71.
- Athanasopoulos, George, and Vahid, Farshid.** 2005. A Complete VARMA Modelling Methodology Based on Scalar Components. Monash University: Department of Econometrics and Business Statistics.
- Bowles, Carlos, Friz, Roberta, Genre, Veronique, Kenny, Geoff, Meyler, Aidan, and Rautanen, Tuomas.** 2007. The ECB Survey of Professional Forecasters (SPF): A review after eight years' experience. *Occasional Paper Series*, 59(4): 3-19.
- Clements, Michael P., and Hendry, David F.** 2005. Evaluating a Model by Forecast Performance. *Oxford Bulletin of Economics and Statistics* 16 (s1): 931-956.
- Corder, Kewin J.** 2003. Managing Uncertainty. The Bias and Efficiency of Federal Macroeconomic Forecasts. Department of Political Science, Western Michigan University, the 7th National Public Management Conference. Georgetown University. Washington DC, Octombre.
- Diebold, Francis X., and Mariano, Robert.** 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13: 253-65.
- Dobrescu, Emilian.** 2006. *Macromodels of the Romanian Transition Economy*. Expert Publishing House, Bucharest.
- Dobrescu, Emilian.** 2003. Macroeconomic estimators for the Romanian "Pre-accesion Economic Program"(the 2003 version). *Romanian Journal of Economic Forecasting* 4(4): 3-10.
- Dovern, Jonas, and Weisser, Johannes.** 2011. Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7. *International Journal of Forecasting* 27 (2): 452-465.
- Fildes, Robert, and Steckler, Herman.** 2000. The State of Macroeconomic Forecasting. Lancaster University EC3/99. George Washington University: Discussion Paper No. 99-04.
- Gorr, Wilpen, L.** 2009. Forecast accuracy measures for exception reporting using receiver operating characteristic curves. *International Journal of Forecasting* 25 (1): 48-61.
- Granger, Clive W. J., and Jeon, Yongin.** 2003. [Comparing forecasts of inflation using time distance](#). *International Journal of Forecasting* 19 (3): 339-349.
- Harvey, David I., and Newbold, Paul.** 2003. The non-normality of some macroeconomic forecast errors, *International Journal of Forecasting* 19 (4): 635-653.
- Hyndman Rob J., and Koehler, Anne B.** 2005. Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting* 22(4): 679-88.
- Kjellberg, David, and Villani, Mattias.** (2010), The Riksbank's communication of macroeconomic uncertainty. *Economic Review* 24(1): 5-41
- Lanser, Debby, and Kranendonk, Heck.** 2008. Investigating uncertainty in macroeconomic forecasts by stochastic simulation. CPB Discussion Paper, 112.
- Makridakis, Spyros.** 1984. *Forecasting: Methods and Applications*, Wiley & Sons, New York, p.

122.

Makridakis, Spyros, Wheelwright, Steven C. and Hyndman Rob J. 1998. Forecasting: Methods and Applications. third edition. John Wiley & Sons, New York.

Mariano, Robert S. 2000. Testing forecast accuracy. University of Pennsylvania, article available at <http://projects.chass.utoronto.ca/link/200010/papers/testforecast.pdf>.

Ruth, Karsten. 2008. Macroeconomic forecasting in the EMU: Does disaggregate modeling improve forecast accuracy ?. Journal of Policy Modeling 30 (3): 417-29.

Teräsvirta, Timo, van Dijk, Dick, and Medeiros, Marcelo C. 2005. Linear models, smooth transition autoregressions, and neural networks for forecasting, macroeconomic time series: A re-examination. International Journal of Forecasting 21 (4): 755-774.

Wojciech, Charemza, and Svetlana, Makarova. 2009. Nonlinear Inflationary Persistence and Growth: Theory and Empirical Comparative Analysis. Romanian Journal of Economic Forecasting 10(2): 5-22.

National Commission of Forecasting. 2010. Data base. [online] Available at <http://www.cnp.ro/ro/prognoze> [accessed on May 2011].

National Institute of Statistics. Data base. [online] Available at <http://www.insse.ro/cms/rw/pages/anuarstatistic2009.ro.do> [accessed on May 2011].

SMANJENJE NEODLUČNOSTI PRI DONOŠENJU ODLUKA VREDNOVANJEM UČINKA MAKROEKONOMSKIH PREDVIĐANJA U RUMUNJSKOJ

Sažetak

Vrednovanje performansi makroekonomskih predviđanja ne uključuje samo izračun nekoliko statističkih mjera, prilično kontroverznih u literaturi, kao što su korijen srednje kvadratne greške ili srednje apsolutne greške. U teoriji i ekonomskoj praksi, možemo pratiti tri smjera u vezi s vrednovanjem performansi predviđanja: analizu točnosti, pristranosti i efikasnosti. Koristeći prognozirane srednje vrijednosti stope inflacije i stope nezaposlenosti u periodu od 2004. do 2010. u Rumunjskoj, dobivamo bolji stupanj točnosti i manju efikasnost za prognoze koje daje Nacionalna komisija za prognoziranje u usporedbi s onima baziranim na Dobrescu modelu kojeg koristi Institut za ekonomska prognoziranja. Slijedeći međunarodni trend, predviđanja su u svakom slučaju pristrana radi poteškoća u preciznom predviđanju šokova koji utječu na ekonomiju. Performansa predviđanja je neraskidivo vezana za njihovu nesigurnost, RMSE, mjeru vrednovanja točnosti koju koristimo u stvaranju prognostičkih intervala zasnovanih na povijesnim pogreškama. Za predložene vrijednosti stope inflacije koje objavljuje Rumunjska Narodna Banka, predlažemo nov način kreiranja prognostičkog intervala kako bi se uzeli u obzir ekonomski šokovi.

Ključne riječi: makroekonomsko predviđanje, točnost, pristranost, učinkovitost

