The Accuracy Analysis of Inflation Rate Forecasts in Euro Area

MIHAELA SIMIONESCU
Institute for Economic Forecasting
Romanian Academy
Calea 13 Septembrie, No. 13, District 5, Bucharest
ROMANIA
mihaela_mb1@com

Abstract: - The main objective of this study is to provide a comparative analysis of the accuracy associated to the inflation forecasts for euro area made by International Monetary Fund (IMF) and Organisation for Economic Co-operation and Development (OECD). On the horizon from 2000 to 2013, IMF provided significantly more accurate inflation rate forecasts compared to OECD, according to Diebold-Mariano test and U1 Theil's statistic value. Moreover, the predictions provided by the two institutions are better than the naïve ones. All the predictions do not provide valuable information for future decisional process.

Key-Words:-forecasts, accuracy, directional accuracy, U Theil's statistic, Diebold-Mariano test, inflation rate.

JEL Classification: C52, C53, E27, E37

1. Introduction

The general public is very interested in macroeconomic predictions made for a future time, but the degree of interest decreases when the horizon became a period in the past. However, we should know the past performance of the forecasts in order to anticipate the quality of the next predictions.

The main aim of this study is to assess the accuracy of inflation rate forecasts for euro area by making a comparative analysis of predictions provided by two international providers: International Monetary Fund (IMF) and Organisation for Economic Co-operation and Development (OECD). Therefore, the U Theil's statistics are computed and the Diebold Mariano test is applied. Moreover, the directional accuracy was assessed, using only the predictions' signs, and the final values. The errors' magnitude has been neglected. In order to make the predictions robust to the presence of outliers, the high and the low errors received the same importance.

The paper is structured as it follows. After this brief introduction, a short literature review is provided. After the description of the methodology, the assessment and comparison of forecasts' accuracy are made for inflation rate predictions in euro area. The last section concludes.

2. The accuracy of forecasts provided by international institutions

Granger (1996) considered that the point predictions should be followed by forecast intervals based on the past performance of the point predictions. Many international institutions provided own predictions of the macroeconomic variables, among these being International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), World Bank (WB), Survey of Professional Forecasters (SPF). Important studies regarding IMF and OECD predictions' accuracy are made by Artis (1996), Ash et al. (1998), McNees (1992), Mills and Pepper (1999), Abreu (2011), Allan (2012), Heilemann and Stekler (2013). Pons (2000) compared in terms of accuracy the OECD's predictions for 13 European countries and the national predictions of each country.

Heilemann and Stekler (2007) concluded that there is a low accuracy of G7's forecasts, the causes beingthe improper forecasting methods and non-realistic assumptions regarding the accuracy of the forecasts.

Abreu (2011) evaluated various macroeconomic forecasts made by different institutions like: European Commission, Consensus Economics, OECD, IMF, and The Economist.

Allan (2012) proposed the combination technique as method for improving the OECD forecasts made for GDP in the G7 countries. The accuracy assessment of these predictions supposed the application of qualitative and quantitative techniques.

González Cabanillas and Terzi (2012) studied the forecasts accuracy of the predictions provided by European Commission before and during the recent economic crisis. They compared these forecasts with those provided by Consensus Economics, IMF and OECD. The Commission's forecasts errors have increased because of the low accuracy from 2009 for variables as GDP, inflation rate, government budget balance, and investment.

Heilemann and Stekler (2013) analyzed the forecasts' accuracy for inflation and real GDP growth rate in case of the Germany predictions made by OECD and 3 professional forecasters from Germany. In the last 10 years, the accuracy forecasts for Germany's inflation and GDP did not improved too much.

Frenkel, Rülke and Zimmermann (2013) described the strategic behavior of the private forecasters that placed their expectations away from OECD's and IMF's ones, this duration of this event being 3 months.

Liu and Smith (2014) concluded that that Greenbook inflation forecasts are more accurate than those of the private forecasts, the authors making comparisons between the predictions provided by Survey of Professional Forecasters, Greenbook and other private forecasters.

Freedman (2014) analyzed the IMF forecasts' accuracy, concluding that there is a qualitative statistical analysis, but the researches were not too documented in some fields like: reference period, comparisons with previous studies, the review of changes, management response.

3. Methodology

For making comparisons between forecasts in terms of accuracy the U theil's statistic is used in two variants: U1 and U2.

U1Theil's coefficient is utilized to compare two predictions made for the same variable or for different variables.

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

a- actual values
p- predicted values
t- time index
e- error (e=a-p)
n- horizon lenght

The forecast for which U1 is closer to zero it is more accurate.

U2 Theil's coefficient allows the comparison with the naïve forecast based on random walk. If U2 is less than zero the forecast is more accurate than the naïve one.

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

The comparison between forecasts can also be made using accuracy tests, the Diebold-Mariano test (DM test) being the most used approach in literature. If the two competing predictions are denoted as $y^1_{t+h/t}$and $y^2_{t+h/t}$, the forecasts' errors are computed as:

$$\varepsilon^1_{t+h/t} = y_{t+h} - y^1_{t+h/t}$$
$$\varepsilon^2_{t+h/t} = y_{t+h} - y^2_{t+h/t} \quad (3)$$

The following loss function is computed for measuring the predictions' accuracy:

$$L(y_{t+h}, y^i_{t+h/t}) = L(\varepsilon^i_{t+h/t}),\ \text{i=1,2} \quad (4)$$

The null hypothesis of DM test states that there are not significant differences between the two forecasts regarding the degree of accuracy.

The DM test uses the loss differential: $d_t = L(\varepsilon^1_{t+h/t}) - L(\varepsilon^2_{t+h/t})$

The null hypothesis is equivalent to $E(d_t) = 0$. The statistic of the test is:

$$S = \frac{\bar{d}}{\sqrt{\frac{\widehat{LRV}_{\bar{d}}}{T}}} \quad (5)$$

where

$$\bar{d} = \frac{1}{T_0}\sum_{t=t_0}^{T} d_t$$

$$LRV_{\bar{d}} = cov(d_t, d_{t-j})$$

$\widehat{LRV}$- consistent estimate of the long-run variance of $\sqrt{T}d$

At 5% level of significance the null hypothesis of accuracy equality is rejected for a value grater than 1,96 of the absolute value of S.

In order to check if the predictions are 'valuable' the comparison is made with the naïve forecast that supposes that the value in the actual period will remain the same in the next period. Schnader and Stekler (1990) and Stekler (1994) used the contingency table approach in order to check the probabilistically independence between the sign of the predicted, respectively actual change. The null hypothesis of this directional accuracy test assumes the independence between the actual and the predicted value. The forecasts are valuable if the independence hypothesis is rejected. The real and the forecast values of the variable changes are presented in a 2×2 contingency table. Different tests are use in this case: Fisher's exact test, chi-square test, and the test proposed by Pesaran and Timmermann (1992).

Table 1: Contingency Table for macroeconomic forecasts

Actual (A) / Forecasted (F)	negative change	positive change	Subtotal
negative change	n_{00}	n_{01}	$n_{0\square}$
positive change	n_{10}	n_{11}	$n_{1\square}$
Subtotal	$n_{\square0}$	$n_{\square1}$	N

Source: author's construction

Note: there is a total number of N observations, subscript i for n_{ij} shows the forecasted outcome, subscript j for n_{ij} shows the actual result, $i(j)=0$ implies negative change, and $i(j)=1$ implies positive change.

The most used test is based on the contingency tables (chi-square test). The statistic of this test is:

$$\hat{\chi}^2 = \sum_{i=0}^{1}\sum_{j=0}^{1}\frac{(n_{ij} - n_{i\square}n_{\square j}/N)^2}{n_{i\square}n_{\square j}/N} \quad (6)$$

Wickens (1989) concluded that this test can become too conservative because the independence assumption can be wrongly accepted. Therefore, it is recommended the use of Yates' (1934) continuity correction based on the following statistic:

$$\chi^2_{\text{Yates}} = \frac{N(|\, n_{00}n_{11} - n_{01}n_{10}\,| - N/2)^2}{n_{0\square}n_{1\square}n_{\square0}n_{\square1}} \quad (7)$$

Another problem of the chi-square test is the continuous distribution hypothesis for the chi-square, but the computation uses discrete categories. The discrete frequencies approximation can generate an inaccurate approximation of the test statistic in case of very low expected frequencies. For an accurate test requires no more than 20% of the cells should have frequencies less than 5 and all cells should have frequencies greater than 1 .

In order to solve the problem of low expected frequencies, the Fisher's exact test for contingency tables is employed. This test is based on a hyper-geometric repartition for directly computing the independence probability. This probability for a 2×2 contingency table is computed as:

$$p=\frac{\binom{n_{0\square}}{n_{00}}\binom{n_{1\square}}{n_{10}}}{\binom{N}{n_{\square 0}}}=\frac{n_{0\square}!n_{1\square}!n_{\square 0}!n_{\square 1}!}{n_{00}!n_{01}!n_{10}!n_{11}!N!} \tag{8}$$

Pesaran and Timmermann (1992) proposed a non-parametric test on the correct forecast of the directional accuracy. It supposes the estimation of the probability of independence between results and predictions. This statistic of this test follows a chi-square distribution with one degree of freedom. The general standardized test statistic for assessing the predictive performance has the following form:

$$S_n^2=\frac{(\hat{p}-\hat{p}^*)^2}{\mathrm{Var}(\hat{p})-\mathrm{Var}(\hat{p}^*)}\ \square\ \chi^2(1) \tag{9}$$

$\hat{p}=(n_{00}+n_{11})/N$: Sample's estimate of the probability of a correctly signed prediction

$\mathrm{Var}(\hat{p})=[\hat{p}^*(1-\hat{p}^*)]/N$

$\hat{p}_f=n_{1\square}/N$: probability of positive change in predicted outcomes

$\hat{p}_a=n_{\square 1}/N$: probability of positive change in actual results

$\hat{p}^*=\hat{p}_f\hat{p}_a+(1-\hat{p}_f)(1-\hat{p}_a)$: estimator under the null hypothesis $\mathrm{Var}(\hat{p}^*)=[(2\hat{p}_f-1)^2\hat{p}_a(1-\hat{p}_a)+(2\hat{p}_a-1)^2\hat{p}_f(1-\hat{p}_f)+4\hat{p}_a\hat{p}_f(1-\hat{p}_a)(1-\hat{p}_f)/N]/N$.

Pesaran and Timmermann (1994) provied also the generalization of their test when actual values and predictions are grouped in more than two classes. The test is useful when a joint assessment of two predictions is made, no requirement being necessary regarding the forecasts' independence.

4. The evaluation of forecasts' accuracy for inflation rate in euro area

The annual inflation rate forecasts provided by International Monetary Fund (IMF) and Organisation for Economic Co-operation and Development (OECD) for euro area (changing composition) are compared in terms of accuracy. The forecasts' horizon is: 2000-2013.

The World Economic Outlook (WEO) database presents the IMF projections regarding the evolution of different variables at the global level, in some groups of countries and in a lot of individual countries. OECD also provides The Economic Outlook Annex Tables with the main projections on macroeconomic variables in individual countries and in certain regions.

Some statistics are computed for the predictions of inflation rate in euro area: U1 and U2 statistics. The value of DM test is 2.03, which is greater than 1.96. This implies that the null hypothesis is rejected and the differencies between forecasts' accuracy are statistically significant. From Table 2 we can conclude that the forecasts provided by IMF are more accurate than those of the OECD. Moreover, the predictions of the two international institutions are superior to naïve forecasts, the U2 statistic values being less than 1.

Table 2: Accuracy Measures

Statistic	IMF	OECD
U1 Theil's statistic	0,153371	0,186492
U2 Theil's statistic	0,481601	0,663481

Source: author's computations

The directional accuracy approach is based on the acceleration (deceleration) of growth forecast. The directional predictions usually consider no change in government policies, nominal exchange rates, and dollar-denominated oil prices.

The data are organized in a contingency table.

n_{00} - negative change in registered values and negative change in predictions

n_{01} - negative change in registered values and positive change in predictions

n_{10} - positive change in registered values and negative change in predictions

n_{11} - positive change in registered values and positive change in predictions

It was computed the number of correct (n_{00} and n_{11}) and incorrect (n_{01} and n_{10}) direction forecasts that were predicted by the two institutions. According to contingency tables made for all the institutions, the cells frequencies are very low, this method being unsuitable for this particular case.

Table 3: Contingency tables

Forecasts' provider	n_{00}	n_{01}	n_{10}	n_{11}
IMF	6	1	2	4
OECD	4	4	4	2

Source: author's computations

If the sum of inputs in the two cells of the leading diagonal ($n_{00} + n_{11}$) is high, four statistics could be computed, the null assumption of the tests stating that the prediction change is independent from probabilistic point of view of the actual change .

Table 4: Tests for directional accuracy

Forecasts' provider	$\hat{\chi}^2_{\text{Yates}}$	p	S_n^2	Chi-square
IMF	0,029167	0,005828	0,00346	0,343812
OECD	0,027778	0,34965	0,00573	0,633629

Source: author's computations

The four statistics were computed to study the directional accuracy of inflation prediction. The results show that for all forecasts the null hypothesis was not rejected, which means that the forecasts are not valuable in the directional predictions.

5. Conclusion

The results of the comparisons between forecasts' accuracy show that the IMF forecasts of the inflation rate in the euro area are better than those provided by OECD. The DM test and also the U coefficients confirm this assumption. According to directional accuracy tests, the predictions provided by the two institutions are not valuable.

A future direction of analysis would be the assessment of forecasts' bias and efficiency, two other dimensions of forecasts' performance. Therefore, some tests should be applied to check the presence of bias and determine the "most" efficient prediction.

References:

[1] Abreu I. (2011)"International organizations' vs. private analysts' forecasts: an Evaluation@,*Banco de Portugal*, article available at: http://www.bportugal.pt/en-US/BdP%20Publications%20Research/wp201120.pdf

[2] Allan, G. (2012)"Evaluating the usefulness of forecasts of relative growth", *Discussion Papers in Economics,* 12-14: 34-58.

[3] Artis, M. M. J. (1996). "How Accurate are the IMF's Short-Term Forecasts? Another Examination of the World Economic Outlook (EPub)", International Monetary Fund.

[4] Ash, J. C. K., Smyth, D. J., & Heravi, S. M. (1998). "Are OECD forecasts rational and useful?: a directional analysis", *International Journal of Forecasting,14*(3): 381-391.

[5] Dovern, J., &Weisser, J. (2011). "Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7" *International Journal of Forecasting, 27*(2): 452-465.

[6] Freedman, C. (2014). "An Evaluation of Commissioned Studies Assessing the Accuracy of IMF Forecasts.", IEO Background Paper BP/14/02 (Washington: Independent Evaluation Office of the International Monetary Fund).

[7] Frenkel, M, Rülke, J.C., and Zimmermann, L. (2013). "Do private sector forecasters chase after IMF or OECD forecasts?",*Journal of Macroeconomics,* 37: 217-229

[8] Granger, C. W. (1996) "Can we improve the perceived quality of economic forecasts?". *Journal of Applied Econometrics*, *11*(5): 455-473.

[9] González Cabanillas, L. and Terzi, A. (2012)"The accuracy of the European Commission's forecasts re-examined", *Economic Papers* 476, http://ec.europa.eu/economy_finance/publications/economic_paper/2012/pdf/ecp476_en.pdf

[10] Heilemann, U. and Stekler, H. O. (2013). "Has The Accuracy of Macroeconomic Forecasts for Germany Improved?".*German Economic Review,* 14: 235–253.

[11] McNees, S. K. (1992). "How large are economic forecast errors". *New England Economic Review*, *1*: 25-33.

[12] Mills, T. C., & Pepper, G. T. (1999). "Assessing the forecasters: an analysis of the forecasting records of the Treasury", the London Business School and the National Institute. *International Journal of Forecasting*, *15*(3): 247-257.

[13] Pesaran, M. H. and Timmermann A. G., (1994). "A Generalization of the Edited the Non-Parametric Henriksson-Merton Test of Market Timing". *Economics Letters*, 44, 1-7.

[14] Pesaran, M. H. and Timmermann A. (1992). "A Simple Nonparametric Test of Predictive Performance".*Journal of Business and Economic Statistics*, 10:.461-465.

[15] Pons, J. (2000). "The accuracy of IMF and OECD forecasts for G7 countries".*Journal of Forecasting*, *19*: 53-63.

[16] Schnader, M. H. and Stekler H. O., (1990)"Evaluating Predictions Of Change". *Journal of Business*, 63, 1, pp.99-107.

[17] Stekler, H.O., (1994)"Are Economic Forecasts Valuable?"*Journal of Forecasting*, 13: 495-505.

[18] Wickens, T. D., (1989). "Multiway Contingency Tables Analysis for the Social Sciences". Lawrence Erlbaum Associates, Hillsdale, NJ.

[19] Yates, F. (1934). "Contingency Table Involving Small Numbers and the χ^2 Test"*Supplement to the Journal of the Royal Statistical Society*, 1: 217-235.