Management Dynamics in the Knowledge Economy
Vol. 3 (2015) no. 1, pp.133-149; www.managementdynamics.ro
ISSN 2392-8042 (online)

The Evaluation of Global Accuracy of Romanian Inflation Rate Predictions Using Mahalanobis Distance

Mihaela SIMIONESCU
Romanian Academy
Institute for Economic Forecasting
13, Calea 13 Septembrie, Sector 5, Bucharest, Romania
mihaela_mb1@yahoo.com

Abstract. *The purpose of this study is to emphasize the advantages of Mahalanobis distance in assessing the overall accuracy of inflation predictions in Romania when two scenarios are proposed at different times by several experts in forecasting or forecasters using data from a survey (F1, F2, F3 and F4). Mahalanobis distance evaluates accuracy by including at the same time both scenarios and it solves the problem of contradictory results given by different accuracy measures and by separate assessments of different scenarios. The own econometric model was proposed to make inflation rate forecasts for Romania, using as explanatory variables for index of consumer prices the gross domestic product, index of prices in the previous period and the inverse of unemployment rate. According to Mahalanobis distance, in 2012 and 2013 F1 registered the highest forecasts accuracy distance. The average distance shows that F1 predicted the best the inflation rate for the entire period. F2 provided the less accurate prognoses during 2011-2013. According the traditional approach, based on accuracy indicators that were evaluated separately for the two scenarios, F1 forecasts provided the lowest mean absolute error and the lowest root mean square error for both versions of inflation predictions. All the forecasts of the inflation rate are superior as accuracy of naïve predictions. However, according to U1 Theil's coefficient and mean error, F3 outperformed all the other experts and also the forecasts based on own econometric model.*

Keywords: *forecasts, Mahalanobis distance, accuracy, forecast error.*

Introduction

The macroeconomic predictions attract the interest of various categories of public when they refer only to the next period. Few people are interested by the past anticipations in the economic activities, although Öller and Barot (2000) show that it is necessary to evaluate the past performance of predictions in order to have a clear picture of the forecasting process. To achieve this objective statistical measures should be employed.

In this study two distinct approaches are applied in order to assess ex-post the degree of accuracy for inflation rate forecasts in Romania of some forecasters F1, F2, F3 and F4. The data are collected using a survey. Each forecaster provided two scenarios for the inflation rate made at different moments in time during the current year for which the prediction was made. Moreover, the owner econometric model was applied to predict the inflation rate during 2011-2013. Some quarterly forecasts based on the proposed model were aggregated in order to determine the annual inflation rate.

The Mahalanobis distance has the advantage of measuring the overall accuracy, by including both scenarios of each institution. The average distance gives a global accuracy measure for the entire horizon. The traditional approach based on accuracy measures makes the assessment separately for each scenario. Moreover, the accuracy measures provide in many cases contradictory results that are canceled by a global measure like Mahalanobis distance.

After the presentation of the theoretical framework, the evaluation of inflation rate forecast accuracy is made and in the end some conclusions are drawn.

The novelty of this research is given by the fact that we do not consider more variables, but the same variable in two scenarios. Moreover, the evaluation is not done only for one year or more, the average distance being computed for the entire horizon. We also made the comparison with the usual approach based on accuracy measures.

Literature review

There are many ways for measuring forecast accuracy. The prediction error is computed as the difference between the actual value and the forecasted value. Many accuracy measures are based on this error, but in literature some indicators were taken over from other fields and adapted in the context of prediction assessment. Neglecting the error's direction can assess the prediction error. Hyndman and Koehler (2006) identified two types of prediction errors: scale dependent measures and scale independent errors. The first category is used with care when comparisons are made between data based on different scales.

Hyndman and Athanasopoulos (2014) propose several traditional accuracy measures based on the forecast error. These authors build scaled errors to compare the accuracy across data series that have various scales. The errors that are computed using training mean absolute error are scaled and the mean absolute scaled error is proposed.

Swanson, Tayman and Bryan (2011) propose a rescaled variant of mean absolute percent error-MAPE, called MAPE-R. This indicator provides a more significant representation of the mean error when outlying errors are present. MAPE-R has not the advantage of an empirical test.

In many studies, comparisons are made between forecasts provided by different national or/and international institutions. In the following table several examples are provided:

Table 1. Studies in literature regarding the forecast accuracy assessment

Authors	Institutions for which forecasts are evaluated	Prediction accuracy measures
Ash et al. (1998)	OECD (Organization for Economic Co-operation and Development)	Non-parametric procedures
Artis and Marcellino (1998)	European Commission, OECD, IMF (International Monetary Fund)	- Mean error, mean absolute error, root mean, squared error - Theil's statistic - Lagrange Multiplier test - Diebold-Mariano test
Koutsogeorgopoulou (2000) and Vogel (2007)	OECD	- Mean error, mean absolute error, root mean squared error, Theil's statistic - Frequency of positive errors - Errors' auto-correlation - The correlation coefficient between forecasts and realizations - R-squared - Contingency tables for checking directional accuracy
Blix et al. (2001)	250 institutions	- Mean error - Mean of the root mean squared error
Gluck and Schleicher (2005)	OECD and IMF	Smoothed first and second moments of forecast errors
Timmermann (2007)	World Economic Outlook	- Mean error and median error - Standard error - Serial correlation
Bowles et al. (2010)	SPF	- Mean error, mean absolute error, root mean squared error - Theil's statistic
Barnichon and Nekarda (2013)	SPF	- Root mean squared error - Giacomini-White statistic

In several recent papers the Mahalanobis distance is employed to assess the degree of forecast accuracy. An approach based on Mahalanobis distance was employed by Eisenbeis, Waggoner and Zha (2003) and Bauer, Eisenbeis, Waggoner and Zha (2006). This method permits ranker of the forecasts' providers for a single variable than four more years as in Sinclair and Steckler (2013). Recently, Müller-Dröge, Sinclair and Steckler (2014) assessed the predictions of a vector of variables for German economy provided by different institutions. The method was applied only for the predictions for 2013 made by twenty-five different institutions. The authors computed how each competitor performed overall. Bundesbank provided the best macroeconomic forecasts for Germany for 2013. Sinclair, Stekler and Carnow (2014) described a multivariate analysis for the predictions of the Federal Reserve in USA. The authors assessed the FED's predictions by employing univariate methods for ten expenditure types of the real GDP.A vector of forecasts was assessed in order to get a global measure of accuracy for inflation rate, unemployment and growth rate and comparisons are made between SPF (Survey of Professional Forecasters) and FED. There are not high differences in accuracy between the two institutions' forecasts. FED's anticipations are consistent with the global economic conditions.

Only a few studies in literature considered the multivariate properties of the FED's predictions. These studies did not propose a general method for simultaneous assessment of quantitative predictions. The joint directional predictions of inflation rate and GDP were analyzed by Sinclair, Stekler and Kitzinger (2010) who used contingency tables. Moreover, Sinclair, Gamber, Stekler, and Reid (2012) computed the costs of jointly wrong estimation GDP and inflation rate when Taylor rule is applied.

Analyzing the predictions' rationality, Caunedo, DiCecio, Komunjer and Owyang (2013) jointly verified the FED's predictions for GDP growth, inflation rate and unemployment rate.

Eisenbeis, Waggoner and Zha (2003) made an aggregation across variables for only one horizon and one time period for each expert in forecasting. The authors proposed a rank of medium quality of Wall Street Journal Experts in forecasting for every period and horizon for more variables.

In other studies of Clements, Joutz and Stekler (2007) or Davies and Lahiri (1999) there is not a pool across variables, but across horizons for every variable. In the first study, only the FED's forecasts are used, while Davies and Lahiri (1999) employed the predictions of SPF and Blue Chips. Dovern (2014) employed the Mahalanobis distance to evaluate the relative disagreement of every expert and to utilize the determinant of the crosssection sample covariance matrix corresponding to individual vector predictions like a measure of overall multivariate disagreement.

In Romania, there are few studies related to the assessment the forecast accuracy for macroeconomic variables, some papers belonging to Simionescu (2013), who evaluated the accuracy of own predictions based on econometric models. The author has employed several traditional measures for assessing the prediction accuracy like mean error, absolute mean error, root mean square error, U1 and U2 Theil's coefficients. For the same predicted variable, in order to take into account the values of more accuracy measures at the same time, the multi-criteria ranking is proposed by the author. The Mahalanobis distance has not been used till now to assess the forecast accuracy for Romania.

Getting accurate forecasts for inflation is important for policymakers conducting fiscal and monetary policy, for labor negotiating salary contracts, for investors covering the risk of nominal assets; for firms that make investment decisions and setting prices.

There are several studies in literature dedicated to the assessment of inflation rate forecasts. For example, Öller and Barot (2000) obtained that OECD inflation rate forecasts were more accurate than those made for GDP growth during 1971-1988 for the European Union. Root mean squared error and Diebold-Mariano test were employed to assess the predictions accuracy. For USA, Ang, Bekaert and Wei (2007) made and assessed quarterly inflation rate forecasts based on Phillips curve, ARIMA models and term structure measure, covering the horizon 1952-2002.

Recently, Clements (2014) explained that experts that forecast inflation tend to be overconfident on the long-term while for short horizon the problem of small confidence often appears. When the horizon decreases, the ex-ante uncertainty is higher than the ex-post uncertainty. In Romania, the assessment of inflation rate forecasts was made by Simionescu (2013) for own predictions based on quantitative methods and predictions based on Dobrescu macromodel.

Theoretical framework

Our main goal is to select the most accurate forecasts from these alternative ones. We should choose the forecaster with the best forecasts for a year and with the best predictions for an entire period. Therefore, some ways of assessing the degree of accuracy are proposed. Predictions of macroeconomic variables are often based on an overall picture of the economic state. A multivariable framework is developed to jointly assess the predictions of all variables.

For a single variable the forecast errors are measured by taking into account the bias (systematic error) and directional accuracy.

The bias might exist even if there are not high differences between the actual modifications and prediction. In order to check if there is a systematic relationship between actual data and the predictions, two approaches are employed. This relation is tested by using the Mincer-Zarnowitz regression. Another test is used to check if there are systematic errors in accordance to economic state. Mincer-Zarnowitz regression checks the bias in the predictions made for only one variable:

$a_t=\beta_0+\beta_1 p_t+e_t$ (1)

where a_t –actual data

p_t -predictions

e_t- errors

For testing informational efficiency, we check if the constant is zero and the slope is one (null assumption). The rejection of this assumption checks if the predictions are biased or/and inefficient. The Wald test is applied to check this assumption.

It is also important for a prediction to show an exact picture of the direction of movement in the economy. The signs of the forecasted modifications for each variable are compared with the signs of actual modifications.

Sinclair, Stekler and Carnow (2014) proposed a procedure to check the presence of systematic errors in the predictions of each variable. A joint framework is proposed to analyze the forecast error characteristics for all the variables. For predicting each component, a vector autoregression of order 1 for the errors is built. Neither parameter of VAR model is significant if the predictions are not biased estimations of results. This means that the constant terms are zero. The own lags parameters are null and neither forecast errors of the other variables Granger cause the other errors.

For assessing the forecast accuracy, distance measures can be computed. Among these, the most utilized are the Euclidian and the Mahalanobis distance. These distances differ in the case of assumptions of statistical independence of vectors. Starting from two independent vectors, P and A, representing predicted and actual values, for the n variables of each vector. The difference between the two vectors is measured by the Euclidian distance:

$$d(P,A) = \sqrt{(P-A)'(P-A)} \quad (2)$$

A- Mean of actual values in the vector

P- Mean of predicted values in the vector

This distance is computed only for independent vectors and with variance equaled to 1. A generalization of the Euclidian distance allows the scale variability among the variables and a correlation differs from zero between the variables. For measuring the distance between each set of predictions and the registered values we will evaluate the difference between the vectors for each set of data by comparison with a historical variation of the actual values. This is the Mahalanobis value:

$$d^2 = (P-A)'W(P-A) \quad (3)$$

W- inverse of the variance-covariance matrix of the sample

The Mahalanobis distance is zero only when predicted values equal the actual ones. For computing the rank, it is essential the correlation between variables and the historical variance of each variable. The distance is greater if the forecasts are in a direction with lower correlation.

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

The indicators for evaluating the forecast accuracy that will be taken into consideration are the common ones, according to Bratu (2012):

- Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n} e_X^2(T_0 + j, k)} \quad (4)$$

- Mean error (ME)

$$ME = \frac{1}{n}\sum_{j=1}^{n} e_X(T_0 + j, k) \quad (5)$$

- Mean absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{j=1}^{n} \left| e_X(T_0 + j, k) \right| \quad (6)$$

U Theil's statistic can be computed in two ways.

Several distinct notations are employed here:

a- actual results

p- predicted results

t- time index

e- forecast error (e=a-p)

n- horizon length

$$U_1 = \frac{\sqrt{\sum_{t=1}^{n} (a_t - p_t)^2}}{\sqrt{\sum_{t=1}^{n} a_t^2} + \sqrt{\sum_{t=1}^{n} p_t^2}} \quad (7)$$

A value that is close to zero in case of U_1 statistic supposes a higher accuracy.

$$U_2 = \sqrt{\frac{\sum_{t=1}^{n-1} (\frac{p_{t+1} - a_{t+1}}{a_t})^2}{\sum_{t=1}^{n-1} (\frac{a_{t+1} - a_t}{a_t})^2}} \quad (8)$$

If U_2 =1=> there are not differences in terms of accuracy between the two forecasts to compare

If U_2 <1=> the forecast to compare has a higher degree of accuracy than the naive one

If U_2 >1=> the forecast to compare has a lower degree of accuracy than the naive one

Assessing the inflation rate forecasts accuracy

Another aim of this research is to provide the own predictions for the inflation rate. An econometric model that explains the index of consumer prices using other macroeconomic variables represents the forecasting method. The index of consumer prices is used to compute the inflation rate by subtracting 100% from the index of prices expressed as a percentage. We start of the relationship between gross domestic product, index of prices and unemployment rates. Quarterly data were collected in Romania using the database of the National Institute of Statistic and Eurostat, covering the period 2000:Q1-2014:Q3. The data are seasonally adjusted using Tramo/Seats method. In order to ensure the data stationarity, some transformations are made to the data. According to Augmented Dickey-Fuller, the index of prices and real GDP are not stationary. The logarithm was applied to the index of consumer prices and real GDP in order to get stationary data series.

The following model was used to explain the evolution of consumer price index:

$$\log(ICP_t) = \alpha_0 + \alpha_1 \log(ICP_{t-1}) + \alpha_2 \log(GDP_t) + \alpha_3 \frac{1}{U_t} + \varepsilon_t \quad (8)$$

ICP_t- index of consumer prices at time t

ICP_{t-1}- index of consumer prices at time t

ICP_t- real gross domestic product at time t

U_t- unemployment rate at time t

ε_t- error term

$\alpha_0, \alpha_1, \alpha_2, \alpha_3$- parameters

The model is linearized by the introduction of the variable U’ that is the inverse of the unemployment rate:

$$\log(ICP_t) = \alpha_0 + \alpha_1 \log(ICP_{t-1}) + \alpha_2 \log(GDP_t) + \alpha_3 U'_t + \varepsilon_t \quad (9)$$

U'_t- inverse of the unemployment rate at time t

The matrix of correlation between the explanatory variables indicates that there is not a strong correlation between the inverse of unemployment rate, logarithm of index of prices in the previous period and the logarithm of GDP. The following valid regression model was obtained for which the errors are independent and homoskedastic. According to Breusch-Godfrey test, for a level of significance of 5%, there is not enough evidence to reject the assumption of independent errors (the probability associated to LM statistic is 0.159 which is more than 0.05 and the null hypothesis is not rejected). For the same level of significance, White test indicated that the errors are homoskedastic (the probability associated to White statistic is 0.146 which is more than 0.05 and the null hypothesis is not rejected).

Moreover, we do not have reasons to reject the normal distribution of the errors at the 5 % level of significance, according to Jarque-Bera test (the probability associated with JB test is 0.172, which is greater than 0.05) (Appendix 1).

This model is used to make quarterly predictions on the horizon Q1:2011-Q4:2013. Then, these quarterly forecasts are aggregated in order to determine the annual inflation rate. The aggregation consists in computing the geometric mean of the quarterly indices of prices. The predictions based on this econometric model and the forecasts provided by F1, F2, F3 and F4 during 2011-2013 are presented in Table 2. The forecasted data are presented in two versions for each institution, depending on the forecast origin. For our predictions we also chose two variants: one-step-ahead forecasts with an updated model for each 4 quarters.

Table 2. The predicted and actual inflation rate (%) forecasts in Romania (horizon: 2011-2013, source: own computations)

Year	**F1**		**F2**		**F3**		**F4**		**Own econometric model**		**Actual values**
2011	4.7	5	2.72	2.8	5.103	5.2	3.8	4	3.3	3.5	5.8
2012	3.3	3.6	3.22	3.25	3	3.1	2.8	3	2.9	3	3.33
2013	4.6	3.9	2.99	3.1	4.8668	5	2.5	2.8	2.7	3.1	3.98

The sorted absolute values of the errors for each year in the case of the first version of forecasts are shown in Table 3. The hierarchy of forecasters for each year is different, in all cases of the first version of forecasts the fourth rank being assigned to the predictions based on our model.

Table 3. The sorted absolute errors of forecasts (percentage points) for inflation rate in Romania-first version of forecasts (horizon: 2011-2013, source: own computations)

2011	**Absolute error**	**2012**	**Absolute error**	**2013**	**Absolute error**
F3	0.697	F1	0.03	F1	0.62
F1	1.1	F2	0.11	F3	0.8868
F4	2	F3	0.33	F2	0.99
Own model	2.5	Own model	0.43	Own model	1.28
F2	3.08	F4	0.53	F4	1.48

In the case of the first version of the forecasts, in 2011, F3 provided the forecast with the lowest absolute error, while for F2 the prediction error is quite high (3.08 percentage points). In 2012, F1 obtained the lowest absolute error, very closed to 0 (0.03 percentage points). F4 obtained an

Vol. 3 (2015) no. 1, pp.133-149; www.managementdynamics.ro

absolute error close to 0.5 (0.53 percentage points). In 2013, F1 also performed the best, F4 registering a quite high error (almost 1.5 percentage points).

The sorted absolute values of the errors for each year in the case of the second version of forecasts are shown in Table 4. In this second version, we used the one-step-ahead forecasts based on the updated model.

Table 4. The sorted absolute errors of forecasts (percentage points) for inflation rate in Romania- second version of forecasts (horizon: 2011-2013, source: own computations)

2011	Absolute error	2012	Absolute error	2013	Absolute error
F3	0.6	F2	0.08	F1	0.08
F1	0.8	F3	0.23	F2	0.88
F4	1.8	F1	0.27	Own model	0.88
Own model	2.3	Own model	0.33	F3	1.02
F2	3	F4	0.33	F4	1.18

The hierarchy for 2011 for the second version of forecasts kept constant as in the first variant, but with other values of absolute error. In 2012 and 2013, the hierarchies have changed, F2 provides the best prediction in 2012 and F4 the worst one. In 2013 F1 maintained the best anticipation, our model with the updated version proving better forecast.

The sorted Mahalanobis distance is displayed in Table 5, choosing a 3 years weighting matrix. Moreover, the average Mahalanobis distance is computed.

Table 5. The sorted Mahalanobis distance for inflation rate in Romania (horizon: 2011-2013, source: own computations)

2011	Distance	2012	Distance	2013	Distance	2011-2013	Average distance
F3	0.38	F1	0.01	F1	0.04	F1	0.28
F1	0.79	F2	0.1	F2	0.78	F3	0.42
F4	3.24	F3	0.07	F3	0.81	F4	1.66
Own model	5.18	Own model	0.13	Own model	1.01	Own model	2.11
F2	8.32	F4	0.16	F4	1.57	F2	3.04

In 2012 and 2013 F1 registered the lowest Mahalanobis distance. The average distance shows that F1 predicted the best the inflation rate for the entire period. F2 provided the less accurate prognoses during 2011-2013.

Some accuracy measures were assessed for the predictions in the two versions for the entire horizon (2011-2013) in Table 6 and Table 7: U1 and U2 coefficients of Theil, mean absolute error, mean error and root mean square error. There are differences between the Mahalanobis distance approach and this way of assessing the prediction accuracy.

Table 6. The accuracy of inflation rate forecasts (first version) during 2011-2013 (source: own computations)

Forecaster	F2	F4	F3	F1	Own model
Mean error	1.393	1.337	0.047	0.170	1.403
Mean absolute error	1.393	1.337	0.638	0.583	1.403
Root mean square error	3.237	2.544	1.175	1.263	2.841
U1	0.0977	0.1938	0.0936	0.1148	0.1645
U2	0.1510	0.3269	0.1858	0.2233	0.2734

According to mean error value and U1, F3 obtained the lowest value. F1 forecasts provided the lowest mean absolute error and the lowest root mean square error. All the forecasts of the inflation rate are superior as accuracy of naïve predictions. Our model provided the worst predictions.

Table 7. The accuracy of inflation rate forecasts (second version) during 2011-2013 (source: own computations)

Forecaster	F2	F4	F3	F1	Own model
Mean error	1.320	1.103	-0.063	0.203	1.170
Mean absolute error	1.320	1.103	0.617	0.383	1.170
Root mean square error	3.127	2.177	1.206	0.848	2.485
U1	0.1099	0.1883	0.0925	0.1153	0.1596
U2	0.1510	0.3269	0.1858	0.2233	0.2734

For the second version of inflation forecasts, the lowest mean error and U1, are registered by F3 on 2011-2013. According to the mean absolute error and root mean squared error, F1 outperformed the other forecasters. The predictions in the second version continued to outperform the naïve forecasts.

Conclusions

The purpose of this paper is to make an assessment of inflation rate forecasts for Romania by proposing a global measure of accuracy. We started from a problem met in practice: more versions of forecasts are provided by an expert for a variable at different moments in time, but we want to know what the forecaster obtained the best predictions in a certain period, but considering all the scenarios. For Romania, an inflation rate forecast assessment was made by Simionescu (2013), but traditional accuracy measures were used (indicators based on forecast error).

In this study two distinct approaches were applied in order to assess ex-post the degree of accuracy for inflation rate forecasts in Romania, made by 4 forecasters (F1, F2, F3, F4) that provide information in a survey. Each expert provided two scenarios for the inflation rate made at different moments in time during the current for which the prediction was made. Moreover, the own econometric model was applied to predict the inflation rate during 2011-2013. Some quarterly forecasts based on the proposed model were aggregated in order to determine the annual inflation rate.

According to Mahalanobis distance, in 2012 and 2013 F1 registered the highest forecasts accuracy distance. The average distance shows that F1 predicted the best the inflation rate for the entire period. F2 provided the less accurate prognoses during 2011-2013.

According to the traditional approach, based on accuracy indicators that were evaluated separately for the two scenarios, F1 forecasts provided the lowest mean absolute error and the lowest root mean square error for both versions of inflation predictions. All the forecasts of the inflation rate are superior as accuracy of naïve predictions. However, according to U1 Theil's coefficient and mean error, F3 outperformed all the other experts and also the forecasts based on own econometric model.

The Mahalanobis distance has the advantage of measuring the overall accuracy, by including both scenarios of each institution. The average distance gives a global accuracy measure for the entire horizon. The traditional approach based on accuracy measures makes the assessment separately for each scenario. Moreover, the accuracy measures provide in many cases, like this one, contradictory results that are canceled by a global measure like Mahalanobis distance.

Acknowledgements: This paper has been financially supported within the project entitled "*Routes of academic excellence in doctoral and post-doctoral research,* contract number *POSDRU/159/1.5/S/137926,* beneficiary: Romanian Academy, the project being co-financed by European Social Fund through the Sectoral Operational Program for Human Resources Development 2007-2013.

References

Ang, A., Bekaert, G., and Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better?. *Journal of Monetary Economics*, 54(4), 1163-1212.

Artis, M., and M. Marcellino. (2001). Fiscal Forecasting: The track record of the IMF, OECD and EC. *The Econometrics Journal*, 4(1), 30–36.

Ash, J.C.K., Smyth, D.J., and Heravi, S.M. (1998). Are OECD forecasts rational and useful?: A directional analysis. *International Journal of Forecasting,* 14(3), 381–391.

Barnichon, R., and C. J. Nekarda. (2013). The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market. *Finance and Economics Discussion Series (FEDS),* Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington (D.C.), 2013–2019.

Blix, M., Wadefjord, J., Wienecke, U., and M. Ådahl. (2001). How Good is the Forecasting Performance of Major Institutions. *Sveriges Riksbank Economic Review*, 3(2001), 37–67.

Bauer, A., Eisenbeis, R.A., Waggoner, D.F., and Zha, T. (2003). Forecast evaluation with cross-sectional data: The Blue Chip Surveys. *Economic Review-Federal Reserve Bank of Atlanta*, 88(2), 17-32.

Bauer, A., Eisenbeis, R.A., Waggoner, D.F., and Zha, T.A. (2006).Transparency, expectations, and forecasts. *Working Paper*, 5(1), 1-18.

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., and Rautanen, T. (2010). An Evaluation of the Growth and Unemployment Forecasts in the ECB Survey of Professional Forecasters. *Journal of Business Cycle Measurement and Analysis*, 2010(2), 63–90.

Bratu, M. (2012). The reduction of uncertainty in making decisions by evaluating the macroeconomic forecasts performance in Romania. *Economic Research–Ekonomska Istraživanja*, 25(2), 239-262.

Caunedo, J., DiCecio, R., Komunjer, I., & Owyang, M.T. (2013). Federal reserve forecasts: asymmetry and state-dependence. *Federal Reserve Bank of St. Louis Working Paper Series*, 2013-012.

Clements, M.P., Joutz, F., and Stekler, H.O. (2007). An evaluation of the forecasts of the Federal Reserve: a pooled approach. *Journal of Applied Econometrics*, 22(1), 121-136.

Clements, M.P. (2014). Forecast Uncertainty—Ex Ante and Ex Post: US Inflation and Output Growth. *Journal of Business & Economic Statistics*, 32(2), 206-216.

Davies, A., and Lahiri, K. (1999). *Re-examining the rational expectations hypothesis using panel data on multi-period forecasts. Analysis of Panels and Limited Dependent Variable Models*, Cambridge: Cambridge University Press.

Dovern, J. (2014). A Multivariate Analysis of Forecast Disagreement: Confronting Models of Disagreement with SPF Data. *Discussion Papers Series*, 571, 1-32.

Glück, H., and Schleicher, S.P. (2005). Common Biases in OECD and IMF Forecasts: Who Dares to be Different? Paper presented at the workshop *A Real Time Database for the Euro-Area*, 13–14 June 2005, Brussels.

Hyndman, R. J., and Athanasopoulos, G. (2014). *Forecasting: principles and practice*. Melbourne: OTexts.
Hyndman, R. J., and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4), 679-688.
Koutsogeorgopoulou, V. (2000). A Post-Mortem on Economic Outlook Projections. *OECD Economics Department Working Papers*, no. 274, 1-34.
Müller-Dröge, C., Sinclair, T., and Steckler, H.O. (2014). Evaluating Forecasts of a Vector of Variables: a German Forecasting Competition, *Cama Working Paper*, 55(1), 1-38.
Öller, L.E., and Barot, B. (2000). The accuracy of European growth and inflation forecasts. *International Journal of Forecasting*, 16(3), 293-315.
Sinclair, T. M., Stekler, H.O., and Carnow, W. (2015). Evaluating a vector of the Fed's forecasts. *International Journal of Forecasting*, 31(1), 157-164.
Sinclair, T.M., Stekler, H.O., and Kitzinger, L. (2010). Directional forecasts of GDP and inflation: a joint evaluation with an application to Federal Reserve predictions. *Applied Economics*, 42(18), 2289-2297.
Simionescu, M. (2013). *Incertitudinea previziunilor în modelarea macroeconomică [The uncertainty of forecasting in macroeconomic modeling]*. Bucharest: Universiară Publishing House.
Swanson, D.A., Tayman, J., and Bryan, T.M. (2011). MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population forecasts. *Journal of Population Research*, 28(2-3), 225-243.
Timmermann, A. (2007). An Evaluation of the World Economic Outlook Forecasts. *IMF Staff Papers*, 54(1), 1-33.
Vogel, L. (2007). How Do the OECD Growth Projections for the G7 Economies Perform? A Post-Mortem. *OECD Economics Department Working Paper*, no. 573, 1-33.

Appendix 1

The regression models and residuals' tests

MATRIX OF CORRELATION	LOG_CPI (-1)	U	LOG_GDP
LOG_CPI (-1)	1.000000	-0.075553	0.072704
U	-0.075553	1.000000	-0.070044
LOG_GDP	0.072704	-0.070044	1.000000

ADF test for LOG_CPI

Intercept

ADF Test Statistic	-3.697508	1% Critical Value*	-3.5547
		5% Critical Value	-2.9157
		10% Critical Value	-2.5953

Trend and intercept

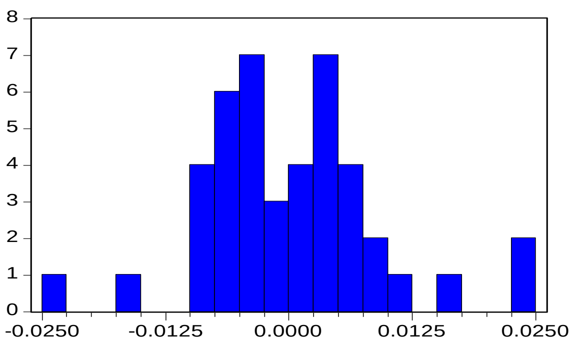
ADF Test Statistic	-3.871370	1% Critical Value*	-4.1348
		5% Critical Value	-3.4935
		10% Critical Value	-3.1753

None

ADF Test Statistic	-3.200552	1% Critical Value*	-2.6055
		5% Critical Value	-1.9467
		10% Critical Value	-1.6190

Dependent Variable: LOG_CPI				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.794547	0.218730	3.632539	0.0008
LOG_GDP	-0.081444	0.022250	-3.660370	0.0007
1/U	0.352571	0.137340	2.567139	0.0142
LOG_CPI(-1)	0.352702	0.148973	2.367566	0.0230
R-squared	0.818282	Mean dependent var		0.026148
Adjusted R-squared	0.804304	S.D. dependent var		0.020985
S.E. of regression	0.009283	Akaike info criterion		-6.432816
Sum squared resid	0.003361	Schwarz criterion		-6.268983
Log likelihood	142.3055	F-statistic		58.53943
Durbin-Watson stat	2.031700	Prob(F-statistic)		0.000000

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	1.836839	Probability	0.183324
Obs*R-squared	1.982689	Probability	0.159107



Series: Residuals	
Sample 2000:2 2010:4	
Observations 43	
Mean	-1.97E-16
Median	-2.50E-06
Maximum	0.024067
Minimum	-0.022623
Std. Dev.	0.008946
Skewness	0.402357
Kurtosis	4.145958
Jarque-Bera	3.513073
Probability	0.172642

White Heteroskedasticity Test:			
F-statistic	1.704131	Probability	0.148293
Obs*R-squared	9.511475	Probability	0.146790

Dependent Variable: LOG_CPI				
Method: Least Squares				
Sample(adjusted): 2000:2 2010:4				
Included observations: 43 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.488100	0.195673	2.494468	0.0168
LOG_GDP	-0.046876	0.018910	-2.478863	0.0175
LOG_CPI(-1)	0.557437	0.134332	4.149696	0.0002
R-squared	0.787575	Mean dependent var		0.026148
Adjusted R-squared	0.776954	S.D. dependent var		0.020985
S.E. of regression	0.009911	Akaike info criterion		-6.323196
Sum squared resid	0.003929	Schwarz criterion		-6.200322
Log likelihood	138.9487	F-statistic		74.15106
Durbin-Watson stat	2.238320	Prob(F-statistic)		0.000000

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	7.542521	Probability	0.009071
Obs*R-squared	6.968432	Probability	0.008296