A STRATEGY TO IMPROVE THE INFLATION RATE FORECASTS IN ROMANIA

Mihaela SIMIONESCU
Institute for Economic Forecasting of the Romanian Academy
mihaela.simionescu@ipe.ro

Abstract
The main goal of this research is to improve the degree of accuracy for inflation rate forecasts in Romania. The inflation was forecasted using a vectorial-autoregressive model. According to Granger test for causality, the relationship between the two variables is reciprocal. The inflation rate volatility is due mainly to the evolution of this indicator, the influence decreasing insignificantly in time, not descending under 96%. More than 87% of the variation in unemployment rate is explained by the own volatility for all lags. For the first lag the inflation is explained only by its evolution, the contribution of the unemployment rate to inflation variation being null. The inflation rate dynamic simulations (deterministic and stochastic) on the horizon 2011-2013 were more accurate than the predictions based on Dobrescu model. The combined forecasts proved to be a good strategy of improving the VAR forecasts and those based on Dobrescu model only if the dynamic and deterministic simulations were combined with Dobrescu's anticipations on the horizon 2011-2013.

Keywords: forecasts accuracy, combined forecasts, Granger causality, VAR model, inflation rate

JEL Classification*:* C51, C52

1. Introduction

The main objective of this research is to construct a VAR (vectorial-autoregressive model) model for inflation rate and unemployment rate in order to make ex-post forecasts of these indicators. The VAR approach allows us to evaluate the variance decomposition of each indicator. In this way, we can determine if the variation in the variable's evolution is mainly due to the other variable or to its own evolution. The model was applied for the Romanian economy on the period from 1994 to 2013, predictions being made for 2011-2013.

Inflation rate is an important macroeconomic indicator used by many institutes in decision-making process. The central bank is directly interested by the most accurate inflation rate forecasts in its targeting. Predictions for inflation rate in Romania are made by national institutes like Centre of Macroeconomic Forecasting of Academician Emilian Dobrescu and National Commission for

Prognosis, but also by international organizations like European Commission, OECD or International Monetary Fund.

2. Literature

The forecasts performance is an important objective for many specialists in forecasting. Our objective is to evaluate the performance in order to apply a suitable strategy for growing the degree of predictions performance. In economic crisis the performance decreases, the necessity of assessing the performance growing. The forecasts performance is a very large domain of research, an exhaustive presentation of it being impossible. But, some of the recent results will be described.

Bratu (2013) proved that the filters and Holt Winters technique could be used as strategies to get more accurate predictions for inflation rate in USA, when the initial expectation are provided by SPF. The Holt-Winters method gave better results. According to Bratu (Simionescu) (2012), the combined forecasts are a suitable way of improving the unemployment forecasts in Romania.

Deschamps and Bianchi (2012) concluded that there are large differences between macroeconomic forecasts for China regarding the accuracy measures for consumption and investment, GDP and inflation. The slow adjustment to structural shocks generated biased predictions, the information being utilized an inefficient way.

Allan (2012), who used quantitative and qualitative techniques to assess the forecasts accuracy, proved that combined forecasts are a good strategy to improve the OECD predictions for GDP in G7 countries. Dovern and Weisser (2011) showed that G7 countries' forecasts are in general biased because of shocks, for accuracy and efficiency the results being different from a country to another.

Most international institutions provide their own macroeconomic forecasts. It is interesting that many researchers compare the predictions of those institutions (Melander for European Commission, Vogel for OECD, Timmermann for IMF) with registered values and those of other international organizations, but it is omitted the comparison with official predictions of government.

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

In Netherlands, experts made 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, Kranendonk and Lanser (2011) were that the CPB model forecasts are in general biased and with a higher degree of accuracy.

Edge, Kiley and Laforte (2009) evaluated the performance of forecasts made by Federal Reserve staff and of those based by a time-series model and a DSGE model. Gorr (2009) recommended the use of classical accuracy measures when a normal evolution of the economy is expected, while the ROC curve is more suitable for crisis times.

Lam, Fung and Yu (2008) compared the predictions performance for the exchange rate, showing that combined forecasts are better than the predictions based on a single model.

Meese and Rogoff (1983) in their study, " 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.

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.

Heilemann and Stekler (2007) gave some reasons for the lack of accuracy of G7 predictions in the last 50 years. There is a continuous critique brought to macro-econometrics models and to forecasting techniques, but also the accuracy expectations are not realistic. Other aspects for the forecasts failure are related to: forecasts bias, data quality, the forecasting procedure, type of predicted indicators, the relationship between forecast accuracy and forecast horizon.

The accuracy of forecasts based on VAR models can be measured using the trace of the mean-squared forecasts error matrix, according to or generalized forecasts error second moment, according to Clements and Hendry (2003).

Robinson (1998) got a better accuracy for predictions based on VAR model for some macroeconomic variables with respect to other models like transfer functions.

Lack (2006) found out that combined forecasts based on VAR models are a good strategy of improving the predictions accuracy.

Bratu (2012) utilized some strategies to improve the forecasts accuracy (combined predictions, regressions models, historical errors method, application of filters and exponential smoothing techniques). The limitation of these strategies is related to the fact that they are empirical, being dependent by the type of forecasts. A strategy could improve a forecast or not. The researcher has to check if his assumption is valid on a particular set of data. On the other hand, we consider that a particular valid strategy on a past horizon might give better results for future short run predictions.

3. Methodological framework

Firstly, we consider that the vector y_t has "m" variables. Each of these variables has "p" lags. The rest of the variables (the deterministic variables and the constant) are placed in a vector denoted by y_{*t} that has m* elements. The VAR model has the following form:

$$y_t = A(L)y_{t-1} + Cy*_t + e_t \,, \quad e_t \rightarrow \left(0, \sum_e \square\right)$$

Number of regressors: k= mq+m*

Number of coefficients: c=mk

The VAR model is written in two equivalent forms (X- a Tk matrix, Y and E- Tm matrices, Im- identity matrix, α- mk vector, y and e- mT vectors):

Y=XA+E (1)

$$y = (Im*X)\alpha + e, \quad e \rightarrow \left(0, \sum_e \square * I_T\right) \quad (2)$$

For the selection of optimal lag a likelihood ratio is applied with the assumptions:

H0: VAR(p0)

H1: VAR(p1)

Different informational criteria are chosen for the optimal lag selection, the most known one being Akaike and Schwartz- Buniakovsky criterion. It is chosen the lag that minimizes the information criterion value for p=1, ..., P.

$$AIC = \ln|\omega| + \frac{2(n^2 p + n)}{T}$$

$$SIC = \ln|\omega| + \frac{(n^2 p + n)\ln(T)}{T}$$

Let us consider two random variables X and Y.

According to Granger (1969), X is cause for Y considering that the information given by X improves the prediction of Y.

Let us consider the lag length p.

$$X_t = c_1 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + e_t$$

OLS is used to test the assumptions:

H0: $\beta_1 = \beta_2 = \cdots = \beta_p = 0$

H1: $\beta_i \neq 0$

The restricted sum of squares is: $RSS_1 = \sum_t \hat{e}_t^2$

The unrestricted sum of squares is: $RSS_0 = \sum_{t=p} \hat{e}_t^2$

$$F = \frac{RSS_2 - RSS_1}{\frac{RSS_1}{T - 2p - 1}}$$

F statistic follows a chi-square distribution with p degrees of freedom. The variance decomposition shows the contribution of the orthogonalized innovation j to MSE- mean square error for the s-step-ahead prediction.

The last relationship shows the contribution of the first innovation to MSE. The residuals decomposition for a standard VAR in a triangular way is called Choleski decomposition. Favero (2001) explained the differences between Choleski identification and Sims-Bernake one.

The optimal approach for combined forecasts is the most used in literature.

Bates and Granger (1969) used the case of two predictions denoted by f1;t and f2;t, for the indicator Xt. In the case of unbiased predictions, the error is: $e_{i,t} = X_{i,t} - f_{i,t}$. The normal repartition of parameters 0 and sigma(i) square. The covariance is computed as $\sigma_{12} = \rho\sigma_1\sigma_2$. The combined forecast is actually a weighted average: $c_t = mf_{1t} + (1 - m)f_{2t}$. The error of the combined forecast is:. $e_t = me_{1t} + (1 - m)e_{2t}$.

The mean of the combined forecast is zero and the variance is:

$$\sigma_c^2 = m^2\sigma_1^2 + (1 - m)^2\sigma_2^2 + 2m(1 - m)\sigma_{12}$$

By minimizing the error variance, the optimal value for m is determined:

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

U Theil's statistic, used in making comparisons between predictions, can be used in two variants, presented also by the Australian Treasury.

The next notations are used:

a-actual/registered value of the analyzed variable

p- value for the predicted variable

t- time

e- error (difference between actual value and the forecasted one)

n- number of periods

U1 takes value between 0 and 1, a closer value to zero indicating a better accuracy for that prediction. If there are alternative forecasts for the same variable, the one with the lowest value of U1 is the most accurate.

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

4. The construction of a VAR model used in forecasting

The data are represented by the unemployment rate provided by Eurostat and the inflation rate given by National Institute of Statistics. The data series covers the period from 1994 to 2013. The data are not stationary, this property being achieved by first differencing the unemployment rate and by applying the logarithm and the first differentiation of the inflation rate. The results of ADF test are presented in the Appendix 1.

The Granger causality test is applied for data series in order to establish if a variable is cause for the other one. In Granger acceptance, a variable X is cause for Y if better predictions result when the information provided by X is taken into account.

The results of Granger causality test show that unemployment rate is the cause of inflation rate and the inflation rate is the cause of unemployment.

Table 1: VAR Granger causality tests

Dependent variable: LOG_INFLATIE

Excluded	Chi-sq	df	Prob.
D_SOMAJ	1.238387	1	0.2658
All	1.238387	1	0.2658

Dependent variable: D_SOMAJ

Excluded	Chi-sq	df	Prob.
LOG_INFLATIE	0.092818	1	0.7606
All	0.092818	1	0.7606

Source: author's computations

Almost all the lag length criteria, excepting logL, at 5% level indicate that a VAR(1) model is the best model.

Table 2: Lag length criteria

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-44.70026	NA	0.834157	5.494148	5.592174	5.503892
1	-28.43990	26.78177*	0.198586*	4.051753*	4.345828*	4.080985*
2	-27.01908	2.005869	0.276715	4.355185	4.845311	4.403905

Source: author's computations

All the tests necessary to be applied for checking the validity of the estimated VAR(1) model are displayed in the following tables. The form of the VAR model is the following:

LOG_INFLATON = 0.923470137451*LOG_INFLATON(-1) - 0.165340287516*D_UNEMPLOYMENT(-1) + 0.0878349862224

D_UNEMPLOYMENT = 0.0576505625253*LOG_INFLATON(-1) - 0.242190928323*D_UNEMPLOYMENT(-1) - 0.194352037925

VAR Residual Portmanteau Tests are used to test the errors' autocorrelation for both identified model. The assumptions of the test are formulated as:

H0: the errors are not auto-correlated

H1: the errors are auto-correlated

For the lag 1 up to 12, the probabilities (Prob.) of the tests are greater than 0.05, fact that implies that there is not enough evidence to reject the null hypothesis (H0). So, we do not have enough reasons to say that the errors are auto-correlated. So, after the application of Residual Portmanteau Test, the conclusion is that there are not autocorrelations between errors for VAR(1) model.

Table 3: Residual Portmanteau test for errors auto-correlation

1	1.492683	NA*	1.580488	NA*	NA*
2	2.056683	0.7253	2.214987	0.6963	4
3	6.787521	0.5597	7.891994	0.4441	8
4	8.590948	0.7374	10.21069	0.5975	12
5	11.74823	0.7611	14.58230	0.5554	16
6	15.40329	0.7529	20.06490	0.4539	20
7	16.40587	0.8729	21.70549	0.5968	24
8	17.48713	0.9383	23.65174	0.6998	28
9	19.13175	0.9647	26.94100	0.7206	32
10	19.80744	0.9869	28.46130	0.8104	36
11	23.49561	0.9825	37.94516	0.5631	40
12	24.93374	0.9909	42.25956	0.5464	44

Source: author's computations

The homoscedasticity is checked using a VAR Residual LM test for the VAR(1) model. If the value of LM statistic is greater than the critical value, the errors series is heteroscedastic. LM test shows that there is a constant variance of the errors, because of the values greater than 0.05 for the probability. The Residual Heteroskedasticity test is applied in two variants: with cross terms and without cross terms.

Table 4: VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

Joint test:

Chi-sq	df	Prob.
20.38520	12	0.0601

Individual components:

Dependent	R-squared	F(4,13)	Prob.	Chi-sq(4)
res1*res1	0.443622	2.591353	0.0859	7.985198
res2*res2	0.282112	1.277169	0.3286	5.078018
res2*res1	0.545093	3.894313	0.0272	9.811668

VAR Residual Heteroskedasticity Tests: Includes Cross Terms

Joint test:

Chi-sq	df	Prob.
24.07323	15	0.0639

Individual components:

Dependent	R-squared	F(5,12)	Prob.	Chi-sq(5)
res1*res1	0.582911	3.354171	0.0399	10.49240
res2*res2	0.291800	0.988876	0.4639	5.252409
res2*res1	0.580233	3.317456	0.0412	10.44419

Source: author's computations

The normality tests are applied under the Cholesky (Lutkepohl) orthogonalization. If the Jarque-Bera statistic is lower than the critical value there is not enough evidence to reject the normal distribution of the errors.

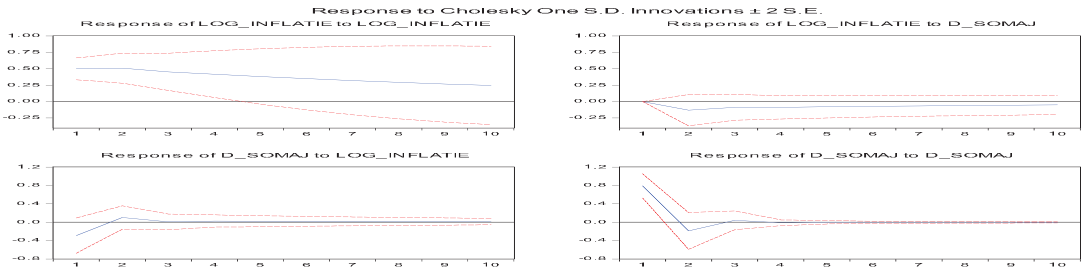
Table 5: VAR Residual Normality Tests Orthogonalization: Cholesky (Lutkepohl)

Component	Jarque-Bera	df	Prob.
1	7.296560	2	0.0260
2	1.099606	2	0.5771
Joint	8.396167	4	0.0781

Source: author's computations

The Residual normality test provided probabilities greater than 0.05, fact that implies that the errors series has a normal distribution when Cholesky (Lutkepohl) orthogonalization is applied. The impulse-response analysis and the decomposition of error variance are made.

Figure 1: The responses of each variable to own shocks or the other variable shocks



Source: author's graph

The inflation rate volatility is due mainly to the evolution of this indicator, the influence decreasing insignificantly in time, not descending under 96%. More than 87% of the variation in unemployment rate is explained by the own volatility for all lags. For the first lag the inflation is explained only by its evolution, the contribution of the unemployment rate to inflation variation being null.

Table 6: Variance decomposition of the variables

Variance Decomposition of LOG_INFLATIE: Period	S.E.	LOG_INFLATIE	D_SOMAJ
1	0.498138	100.0000	0.000000
2	0.723505	96.75490	3.245096
3	0.858191	96.62304	3.376960
4	0.958401	96.44153	3.558466
5	1.034525	96.35929	3.640709
6	1.094298	96.30221	3.697788
7	1.141946	96.26380	3.736196
8	1.180382	96.23599	3.764013
9	1.211640	96.21532	3.784677
10	1.237215	96.19956	3.800439

Variance Decomposition of D_SOMAJ: Period	S.E.	LOG_INFLATIE	D_SOMAJ
1	0.840551	12.05223	87.94777
2	0.867671	12.62276	87.37724
3	0.868550	12.60084	87.39916
4	0.869026	12.66877	87.33123
5	0.869215	12.70637	87.29363
6	0.869404	12.74191	87.25809
7	0.869554	12.77079	87.22921
8	0.869682	12.79514	87.20486
9	0.869788	12.81549	87.18451
10	0.869878	12.83254	87.16746

Source: author's computations

The VAR model is used to make inflation rate forecasts on the horizon 2011-2013. For the VAR predictions four types of scenarios are considered: S1 scenario (Dynamic-Deterministic Simulation), S2 scenario (Dynamic-Stochastic Simulation), S3 scenario (Static-Deterministic Simulation) and S4 scenario (Static-Stochastic Simulation).

Table 7: Predictions of inflation rate (%) based on VAR(1) models

Year	VAR(1) model (S1)	VAR(1) model (S2)	VAR(1) model (S3)	VAR(1) model (S4)
2011	5.4199	5.4553	7.7203	7.7509
2012	5.3629	5.1797	9.8158	9.3453
2013	5.1932	5.1005	12.1870	11.4340

Source: own computations

The second scenario generated the lowest forecast errors for the inflation rate compared to the other three scenarios. If we make the comparison with real data, this scenario generated the most accurate predictions and it could be used to make forecasts for 2014 and 2015.

Table 8: Combined predictions of inflation (%) based on VAR(1) models and Dobrescu model anticipation

Year	S1+Dobrescu model	S2+Dobrescu model	S3+Dobrescu model	S4+Dobrescu model
2011	5.1746	5.2260	4.7626	4.7486
2012	4.1514	3.3715	11.0927	10.5960
2013	5.1837	5.4339	3.1784	3.1106

Source: own computations

All the forecasts were assessed using U1 Theil's coefficient. Not all the combined predictions have improved from accuracy point of view.

Table 9: The values of U1 Theil's statistic for the mentioned forecasts on 2011-2013

Forecasts	U1 value
S1	0.0977
S2	0.0973
S3	0.3885
S4	0.3734
Dobrescu model	0.1004
S1+Dobrescu model	**0.0955**
S1+Dobrescu model	0.5151
S1+Dobrescu model	0.6162
S1+Dobrescu model	0.6072

Source: own computations

The combined forecasts based on the first scenario of simulation and Dobrescu model are more accurate than all the proposed forecasts. The U1 value in this case is the lowest one. The other combined predictions are less accurate than the initial predictions. The first and the second scenario based on VAR model provided a higher degree of accuracy compared to Dobrescu model.

5. Conclusions

According to this analysis based on VAR model, we can conclude that for the inflation in Romania during 1994-2013, the relationship between inflation rate and unemployment rate is reciprocal.

The inflation rate volatility is due mainly to the evolution of this indicator, the influence decreasing insignificantly in time, not descending under 96%. More than 87% of the variation in unemployment rate is explained by the own volatility for all lags. For the first lag the inflation is explained only by its evolution, the contribution of the unemployment rate to inflation variation being null. The inflation rate dynamic simulations (deterministic and stochastic) on the horizon

2011-2013 were more accurate than the predictions based on Dobrescu model. The combined forecasts proved to be a good strategy of improving the VAR forecasts and those based on Dobrescu model only if the dynamic and deterministic simulations were combined with Dobrescu's anticipations on the horizon 2011-2013.

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, number 12-14.
[3] Bates, J. and Granger C. W. J. (1969), *The Combination of Forecasts*, Operations Research Quarterly, 20(4), 451-468.
[4] Bratu M. (2012), *"Strategies to Improve the Accuracy of Macroeconomic Forecasts in USA",*LAP LAMBERT Academic Publishing, ISBN-10: 3848403196, ISBN-13: 978-3848403196.
[5] Bratu (Simionescu), M.,(2012), *The reduction of uncertainty in making decisions by evaluating the macroeconomic forecasts performance in Romania*, Economic research- scientific journal, Vol. 25, no. 2, June 2012, Original Scientific Paper, p. 239-262.
[6] Deschamps, B. and Bianchi, P. (2012), *An evaluation of Chinese macroeconomic forecasts*, Journal of Chinese Economics and Business Studies, Volume 10, Number 3, 1 August 2012 , pp. 229-246(18).
[7] Dovern, J. and Weisser J. (2011), *Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7*, International Journal of Forecasting, 27 (2), p. 452-465.
[8] Edge R.M., Kiley M.T. and Laforte J.-P. (2009), A comparison of forecast performance between Federal Reserve Staff Forecasts simple reduced-form models and a DSGE model, *Finance and Economics Discussion Series.*
[9] Favero, C.A. (2001), *"Applied macroeconometrics"*, Oxford University Press, Oxford.
[10] Franses P.H. , Paap R., Vroomen B. (2004), Forecasting unemployment using an autoregression with censored latent effects parameters, International Journal of Forecasting, Volume 20, Issue 2, April–June 2004, Pages 255–271
[11] Granger (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica, 37.*
[12] Heilemann, U. and Stekler, H. (2007), *Introduction to "The future of macroeconomic forecasting"*, International Journal of Forecasting, 23(2), p. 159-165.
[13] Lack C. (2006), Forecasting Swiss inflation using VAR models, *Swiss National Bank Economic Studies*, No. 2/2006.
Meese, R., and RogoffK.."Empirical Exchange Rate Models of the Seventies." Journal of International Economics, 1983, 14 pp. 3-24.

APPENDIX 1

Null Hypothesis	Augmented Dickey-Fuller test statistic	t-Statistic
D(LOG_INFLATIE) has a unit root	-5.982483	-3.857386 -3.040391 -2.660551
D(LOG_INFLATIE) has a unit root	-5.787331	-4.571559 -3.690814 -3.286909
D(LOG_INFLATIE) has a unit root	-5.753883	-2.699769 -1.961409 -1.606610
D(SOMAJ) has a unit root	-5.108500	-3.857386 -3.040391 -2.660551
D(SOMAJ) has a unit root	-5.054070	-4.571559 -3.690814 -3.286909
D(SOMAJ) has a unit root	-5.259380	-2.699769 -1.961409 -1.606610

The t-statistic is computed at 1%, 5% and 10% level of significance