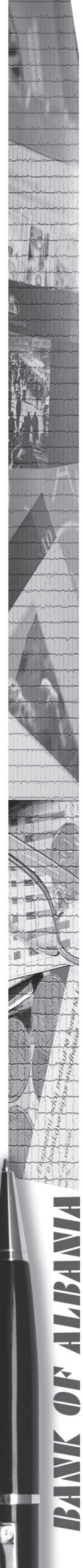
WORKING PAPER



03 (34) 2012

EMPIRICAL INVESTIGATION OF FORECAST UNCERTAINTY WITH MONTE CARLO SIMULATION

Altin Tanku
Elona Dushku
Kliti Ceca*

BANK OF ALBANIA

*Altin Tanku, Elona Dushku, Kliti Ceca, Research Department, Bank of Albania

The views expressed in this paper are those of the authors and do not necessarily reflect the views and policy of the Bank of Albania. We have benefited from invaluable comments by Hilda Shijaku and Vasilika Kota (at the Bank of Albania), as well as from participants at the "5th Workshop on Economic Research in Southeastern Europe", organized by the Bank of Albania in November 2011.

CONTENTS

1. Introduction 5

2. An Analytical Approach to Investigate Forecast Uncertainty 9

3. Deterministic and Stochastic Simulation of the Model 13

4. The Monte Carlo simulation technique 17

5. Experiment with MEAM and the main results 21

6. Conclusions 29

References 31

Annex 33

"The only difference between forecasts of economists and fortune-tellers is that economists are better in explaining why they were wrong"
(F. J. H. Don, 2001).

1. INTRODUCTION

In 2006, the Bank of Albania decided to abandon its Monetary Targeting regime in favour of Inflation Targeting one. Against this setting, the Bank of Albania focused its research and analytic efforts on the development of empirically-based models for forecasting inflation and, in addition, one macroeconomic model, MEAM (macro econometric model for Albania). This model aims to analyze different scenarios and shocks in the economy, thus enabling non-naive forecast of main macroeconomic variables, based on current and expected developments.

This relatively short experience makes modelling a new activity to the Bank of Albania. Yet it has not prevented the decision-making process to be largely based on the results of several models that are used for forecasting or shock analysis, with successful results. The models provide important quantitative input for policy makers and at the same time yield a framework for the discussion of expected developments in the Albanian economy. Despite the detailed analysis of the broader economic developments, the models (in particular the MEAM) represent many-dimensional structures and frequently generate results welcomed by scepticism and stir debate.

In most cases, the scepticism originates from the fact that policy makers are reported only point or deterministic estimates (forecasts) of the expected future developments as reported by the model. Therefore, the uncertainties surrounding the economic transmission mechanism in general, and the empirical estimates in particular, cause the generated point results not to always provide the best possible forecast to match the expectations of the policy makers. In addition, the uncertainty of results grows even higher since they are presented in the form of a point forecast, but not with the uncertainty that surrounds the forecasts for the above reasons.

Both the model builder and the policy makers are aware of the fact that the reported point forecasts or scenarios are uncertain. The model is as much real and rigid generalization of reality as possible and, being so, it holds some level of uncertainty. Point estimates hide this uncertainty as long as they are not presented in the form of probability distribution. Consequently, without an estimation of this uncertainty, it is difficult to understand whether an alternative forecast is or is not different from the deterministic one, associated with probability distribution. One of the most frequent questions asked by the policy makers in the presentations is: How much do you believe in the results? To reframe this question in the econometric framework it would be: What is the probability of uncertainties in the estimation of the forecast, or better, what is the distribution of the stochastic estimation?

The purpose of this paper is to partly answer this particular question and reduce this information gap between the policy makers and the modellers by measuring and providing a measure of the uncertainty that surrounds the results of forecasts and shock analysis produced by the MEAM. The contribution of the material is:

- To apply stochastic simulation and Monte Carlo technique to analyze the sensitivity of the macro model based on random disturbances.
- To identify and quantify the sources of the uncertainty in the model and provide a clear understanding of the sources and size of the forecast error.
- To evaluate if the forecast results are systematically underestimated or overestimated based on the shape of forecast distribution.

Several other institutions have used similar techniques to measure model uncertainties, albeit different sources of uncertainties, using Monte Carlo method. The Bank of Canada [(Amano et al.

(2002)], the Bank of England [(Garratt et al. (2003)], are just a few examples. However, stochastic evaluation of uncertainty has a long and successful history and is a well-established method developed and applied by many influential authors, Canova (1995) and Fair (2003).

Uncertainty has many sources and it may be displayed in four different forms. Based on the structure of the model, the nature and objective of the exercise, we can identify several sources of uncertainty in forecasts: The first source concerns the data uncertainty that come from statistical information. The second source involves the uncertainty associated with the forecast of the exogenous data series since their future realisation is uncertain at the time of model simulation. The third source of uncertainty includes the uncertainty in estimating the model parameters. The last source concerns the uncertainty associated with the error term that may derive from random events or misspecification of the model. Consequently, besides the above four sources, Clements and Hendry (1998), and Ericsson (2001) include another source of uncertainty that stems from the wrong selection of model, variables and equations included in the final model.

However, it is important to emphasize that the analysis provided in this material does not evaluate all sources of uncertainty but rather focuses on the fourth source of uncertainty, in particular the analysis of the error term, that is, other random events unexplained by the model.

Furthermore, as the interest of monetary policy is price stability and some economic variables like the performance of domestic output, interest rates, exchange rate etc., this material focuses on the investigation and analysis of forecast error. Hence, we consider the effects of monetary policy decision-making or future events, taking into account the uncertainty in the variables that are of particular interest to monetary policy.

Given that different projections and analyses obtained from these models are an integral part of the implementation of economic policies, the evaluation of the performance of these models is part of a very important process. Its evaluation through a stochastic simulation approach provides fairly complete information even on the probability distribution of the error term, therefore, also on the distribution forecast obtained from the model, with the assumption that the functional form is well-specified, the assumptions of the exogenous variables are accurate and will materialize in the future, and the estimated parameters are the real ones. Furthermore, based on the type and distribution form of the stochastic forecasts, we can evaluate whether the predictions made by a particular model are overestimated or not.

This paper aims to assess explicitly the performance of MEAM model by Monte Carlo technique of stochastic stimulations, focusing only on the uncertainty that stems from the error term in the model, and the identification of those variables that lead to greater uncertainty in the projection of real GDP.

The outline of the paper is as follows. Section 2 describes model uncertainties and discusses the analysis of expected forecast errors based on analytical formulas. Section 3 defines deterministic and stochastic simulations. In Section 4, we present the Monte Carlo simulation technique. In Section 5, we provide a brief description of MEAM model and summarise the empirical results of stochastic simulations with additive equation disturbances. Finally, Section 6 presents concluding remarks and discussion.

2. AN ANALYTICAL APPROACH TO INVESTIGATE FORECAST UNCERTAINTY

Uncertainty is an inherent attribute of any forecast, which is a result of a number of factors, including our knowledge or preference for the theoretical model most representative of our economic reality, the dynamics of past and expected developments in exogenous variables, and the long experience or not of empirical assessments. All these elements are based on different economic experiences, specified in time and space, in which the understanding process and individual analysis of any person or group of persons is formed. In this context, it is very unlikely that all the policy makers base on the same model and economic theory, have the same expectations of different economic shocks and base the estimation on the same exogenous variables and instruments, etc. Therefore, based on the experience and the beliefs of policy makers, which very likely are different from the ones of the model builder, the outcome of the model is neither sufficiently nor necessary going to describe the economy in exactly the same way as predicted by the policy maker. Every single difference in perception for one or more of the above discussed elements in the data generating process (DGP) is a potential source in the expected results of policy makers and modellers. They are potential sources of uncertainty in the results of all models.

Every model is an acceptable simplification of economic reality. It incorporates all the theoretical basis and useful information on economic data and relationships at the time when the model is built. Using these resources, the modeller tries to specify the model so that it provides a better approximation to the data generating process (DGP) or rather the local data generating process (LDGP) as defined by Hendry (2011). Despite the countless efforts and attempts of the modeller, the model is always left with a vector of error, which for every time period t measures the divergence of the data generated by the model from the real observed values of the variable. In general, any model that tries to imitate the reality can be represented in the following general form:

$$f(y_t, x_t; \theta) = u_t \quad (1)$$

Where y_t is the vector of endogenous variables that we would like to estimate or forecast, x_t is the vector of exogenous variables, θ is the vector of structural parameters of the model and u_t is the vector of shocks that hit the economy or random errors. In all those cases when the system has a single solution of a closed form, we can predict the value of the endogenous variables using function g as follows:

$$y_t = g(u_t, x_t; \theta) \quad (2)$$

With the following mean and variance:

$$Ey_t = \beta(x_t; \theta) \quad (3)$$

$$E[y_t - \beta(x_t; \theta)][y_t - \beta(x_t; \theta)]' = \Omega(x_t; \theta) \quad (4)$$

The problem with the economic modelling is that the functional form g is not known to the modeller or the policymaker; therefore we cannot get a closed form proxy of g. Yet, taking for a given x_t, a given vector of structural parameters θ, and under the assumption that the disturbances for the period under review are zero $u_t = 0$ for all t, and with an appropriate functional form, the modeller can estimate and report predicted values of our endogenous variables according to the following form:

$$\hat{y}^c_* = \hat{\beta}(x_*; \hat{\theta}) \quad (5)$$

The hat stands to show that the set of structural parameters is estimated such that it replicates the LDGP whether they are estimated or calibrated, yielding a vector of endogenous variables for each t in the period of estimation. The reported $\hat{y}^{c}{}_{*}$ is in fact the deterministic estimate of our vector of endogenous variables for all t under review, based on the assumption that $\hat{\theta}$ is a consistent estimator of the real parameters of the model, exogenous variables, x_* and the functional form g.

However, we know that our endogenous variables are stochastic, which is observed in the process of model estimation. Regardless of how successful the modeller is in the above assumptions, the model will never be able to replicate the true DGP, resulting in an estimation error e_t, for each t, as the model generated data will diverge from the observed ones:

$$e_t = y_t - \hat{y}^{c}{}_{*} \tag{6}$$

The modeller should consider all the possibilities to transform the vector e_t of the sample period for $t = 1, 2, 3... t$, which have resulted from the model estimation process in the sample period, into random vectors of normal independent components with distribution $N(0, \delta)$, with known dispersion and covariance matrix Γ.

In fact, these diversions from the true DGP do incorporate not only the random shocks in our variables but also every possible error made in the estimation (calibration) of structural parameters, model coefficients, in the assumption regarding the exogenous variables or the functional form that are included in the model. Therefore, every modelling attempt that involves assumptions, measurement, model selection estimates and expert judgement incorporates uncertainty that comes with each and every one of these steps. This uncertainty flows into the model as a source of uncertainty in the results and turns into the source of observed errors, beyond and in addition to random shocks. This uncertainty incorporates useful information for the modeller and the policymaker, where ignoring it might be potentially dangerous.

In fact, this was also one of the first and major criticisms made to economic models in early 1980s. The main disadvantages consider that models represent a considerable simplification of reality, and suffer from non-exact specification of the main variables and economic ties that exist in the model.

From this prospective, it becomes important to know not only the deterministic estimate of forecasts but also the uncertainty that accompanies (surrounds) these estimates. Hence, the probability distribution of the error is very informative since errors incorporate unexplained shocks of the variables that are part of the true DGP but are not included in the model. Observing and investigating the probability distribution of the error around the point estimate of future developments does not only provide a better understanding of the risk, and uncertainty of the forecast among the variables of the model (providing a measure of forecast error), but also provides a reference distribution to compare the quality of the deterministic forecast (providing information on the form of the distribution and nonlinear nature of the model).

3. DETERMINISTIC AND STOCHASTIC SIMULATION OF THE MODEL

Macroeconomic forecasts are crucial to help and guide the decision making process in terms of further analysis. Since these decisions are made in an environment characterized by uncertainty, their recognition and evaluation is very important. Don (2001) defines two Evaluation Criteria's forecasts respectively for the Statistical and non Statistical ones.

The main basis for the evaluation of forecast, based on the statistical criterion, is that the forecast errors have mean zero and a minimal standard deviation. A similar criterion initially requires some knowledge of the error distribution form, which is subject to a number of untested assumptions taking place in a given economy.

Unlike the statistical criterion, the non-statistical criterion requires the model to have logical and economic coherence in the forecast. Moreover, it should be coherent over time. Logical coherence implies that the forecast model is based on accepted economic identities. While economic coherence goes beyond logical coherence, as the forecast should not only be consistent with economic theory, but it should also replicate the historical performance data observed in reality.

Consistency or the sustainability of forecasting process is another non-statistical criterion, used to evaluate the error performance. This criterion implies that there should not be very major changes between forecasts, in the cases when the new information added does not make great difference from what we had placed originally. This criterion lies between economic coherence on one hand and the successful prediction on the other, clearly explaining the deviation of the forecast to maintain the consistency in decision-making of policymakers [Britton et al. (1998)].

MEAM has been used in the process of policy making providing forecasts and shock analysis for more than three years. This period has shown that MEAM has passed the tests for non-statistical criteria and that it is a reliable tool for analysis and forecast in the policy process. On the other hand, the statistical criterion of results or forecasts obtained from the model has been partial. While the evaluation of the statistical criterion of individual equations was made during the model design and estimation, based mainly on standard root mean square errors (RMSE), these being undocumented attempts before. However, to this moment, the statistical criterion has not been estimated yet and this is the first attempt testing the performance of MEAM as a whole, from the simultaneous interaction of all variables and equations in the model.

One way to do this is to investigate the performance of uncertainty through stochastic simulation rather than deterministic simulation and the results of stochastic simulation can provide information on the distribution of model forecast variables. An explicit description of stochastic simulation is given as below:

The procedure of stochastic simulation requires first to generate a random vector $\tilde{u}_s$ of the serially independent random disturbances; second, to insert these random disturbances in the model in the form of random shocks to generate out of sample forecast s periods ahead, where s is defined as in the previous paragraph and is different from $t = 1, 2, 3, \ldots T$, assuming precise knowledge of exogenous variables x in the entire forecast period s and structural parameters using and solving the model or the system that results from equation 2. If we were able to replicate the process a sufficiently large number of times, using a finite but large number of random disturbances $\tilde{u}_s$ under the assumption of serial independence and similar distribution, based also on the variance covariance matrix of the model, we would be able to generate a sufficiently large number of vectors $\hat{y}_s$. Therefore, if we were able to provide a known parametric distribution for the stochastic disturbances, under the assumption of serial independence and similar distribution, and known variance covariance error matrix of the model, it would be possible to estimate a stochastic estimation of $\hat{y}_s$ in the following general form:

$$\hat{y}_*^c = \int_u \beta\left(u, x_*; \hat{\theta}\right) pdf_*(u)\, du \qquad (7)$$

Where the $PDF\,(\cdot)$ is the distribution density function of $u \equiv \widetilde{u}$, with $\widetilde{u}$ representing independent draw from the pre-specified distribution of e_t the vector of forecast errors. The difference between the observed values of our endogenous variables for the entire period of forecast s with the simulated ones that result from the stochastic forecast given by $y_s - \hat{y}_s$ represents the population of the stochastic results.

Summarizing, this evaluation goes through four steps.

First, a random vector $\widetilde{u}$ of pseudo numbers is generated by Monte Carlo procedure from a previously known distribution and variance covariance matrix.

Second, the $\widetilde{u}$ vectors are introduced into MEAM and the model is solved in the forecast mode s periods ahead, with $s = (T+1, T+2, ..., T+S)$, using the known values of the exogenous variables.

Third, stages 1 and 2 are repeated a sufficient number of times (1000 in our case), shocking the system with a different vector of disturbances $\widetilde{u}$ each time and generating the differences $y_s - \hat{y}_s$ for all endogenous variables of interest.

Fourth, the generated differences for each endogenous variable are used to calculate the moments of their distribution and analyse the results relative to the deterministic forecast along the following lines:

- Linearity versus non-linearity in forecast. To test whether a model is linear or not, we calculate the difference between the stochastic and deterministic mean, known as the bias-coefficient. A high value of this coefficient indicates that the model can be non-linear. In the case of linear models, deterministic and stochastic simulations of the model provide unbiased forecasts, while in the case of non-linear models, a deterministic simulation gives an over or underestimated model solution in terms of forecast's mean [Brown and Mariano (1989 a, 1989, b)].

- Model robustness (sensitivity) to random disturbances. The analysis of the standard deviation of stochastic forecast can help us to analyse the expected error and give us information about the stochastic forecast distribution. Thus, if the analysis reveals that the deviations of stochastic forecast are relatively different from the deterministic mean, then this indicates that the model is unstable and sensitive to external shocks.

- Finally, the last issue we want to address is the shape of the stochastic forecast distribution, which provides essential information on the characteristics of forecasting error. Hence, in a non-linear model, Bianche, Calzolari and Corsi, (1979, 1981) underline that the normal distribution of errors can produce a skewed forecast distribution, which is associated with a different mean and median. Such mean of distribution investigation gives an idea on where would be the stochastic and deterministic forecast error in terms of forecast mode, revealing the possibility of overshooting or undershooting in the deterministic forecasts.

4. THE MONTE CARLO SIMULATION TECHNIQUE

Several analytical, numerical and empirical methods are proposed in the econometric literature for the estimation of the contribution of errors to the forecast, for several or all sources of uncertainty [Bianchi and Calzolari (2010)]. One of the ways becoming more applicable is called "Monte Carlo Methods of simulation of random variables". In the following, we will provide a general description of these methods, focusing on a concrete example that will be used in the next section.

To our best knowledge and understanding, this is not the first time that Monte Carlo simulation technique is used to generate uncertainty in the working papers of the Bank of Albania. Shijaku and Ceca (2009) use Monte Carlo simulation to measure uncertainty, however, they do not provide a discussion of the technique. Due to this and the fact that this is the first time Monte Carlo techniques are used to estimate the efficiency of forecast of the Bank of Albania macroeconomic model MEAM, we are going to provide a general description of the Monte Carlo technique.

The use of the Monte Carlo methods is based on the simulation of the random variable with uniform distribution in the interval]0; 1[. Base algorithms of generating numbers from this distribution are mentioned in the literature (Lemieux, Ch, 2009) and also, some of them are incorporated into the different statistical software. Statistical procedures and tests are also created to show the statistical significance of the sets of numbers created using those algorithms (Lemieux, Ch, 2009).

Afterwards, the use of those methods consists in the modelling of a random variable through the uniform distribution. In that way "switching algorithms" are created. Theoretically, it is proved that any probability distribution may be functionally expressed by the uniform distribution. On the other hand, the practical solution of this expression consists in setting up concrete "switching algorithms" as mentioned above.

While as far as mentioned is applied for parametric distributions, other types of methods are available for nonparametric distributions (sometimes also known as "special methods of modelling of random variables"). The last to be mentioned in that framework is that modelling of random vectors is a normal generalization of the scalar case. The following provides a concrete example of modelling of a random variable and random vector with normal distributions that are applied in our case.

In many cases, the probability distribution of the uncertainty is accepted as normal (based on the Central Limit Theorem). Let us suppose that we are given a random vector Z, with multi-dimensional normal distribution $Z_n \sim N(\mu, \mathbf{\Omega})$, with mean vector μ and covariance matrix $\mathbf{\Omega}$, symmetric, positively definite and non-singular. Applying Cholesky decomposition matrix, Ω can be given by lower triangle squared matrix H, n-dimensional, in the form: $\mathbf{\Omega}=HH^T$. For the details on the calculations of the matrix $H = \{h_{i,j}\}$, when $= \{\omega_{i,j}\}$ is given, the following equalities can be used:

$$\begin{cases} h_{i,j}=0, \, for \; 1 \leq j \leq i \leq n \\ \sum_{k=1}^{j} h_{i,k} h_{j,k} = \omega_{i,j,} \; for \; 1 \leq j \leq i \leq n \end{cases} \tag{8}$$

The matrix H can be used to express random vector Z in the following form:

$$Z = H \bullet U + \mu, \tag{9}$$

Where $U=(U_1, \dots U_n)^T$ is a random vector n – dimensional, with normal distribution $U \sim Nn\,(0, I)$, with probability density:

$$f_u(u) = \frac{1}{\sqrt{(2\pi)^n}} exp\left(-\frac{1}{2} u^T \; I^{-1} u\right) \tag{10}$$

Based on the fact that the covariance matrix of the random vector U is n-dimensional identical matrix I, it results that all the components of vector U are independent and with similar distribution, given by the probability density function: $f_u(u) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{u_i^2}{2}\right), \, u_i \in R$.

In the meantime, it is easy to prove the opposite, meaning that if the random vector U has the probability distribution $U \sim N_n\,(0, I)$, then its linear transformation has the normal distribution. The calculations of the parameters can be done based on the above transformations. In that manner, the transformation (7) gives a normal distributed vector $Z \sim N(\mu, \mathbf{\Omega})$, where $\mathbf{\Omega} = HH^T$.

We mention here that the distribution density of n-dimensional normal random vector Z is:

$$f_z(z) = \frac{1}{\sqrt{(2\pi)^n |\Omega|}} exp\left(-\frac{1}{2}(Z-\mu)^T \Omega^{-1} (Z-\mu)\right), z \in R^n. \qquad (10\text{-}1)$$

Expression (8) and the above transformation are the main concept of the “switching algorithm” for the multi-dimensional normal distribution.

For the simulation of the random numbers with distribution $N (a, \sigma^2)$, the following Box – Müller transformation (10-2) can be used. It generates pseudo random numbers $N(a, \sigma^2)$. For the simulation of the random vector with normal distribution, the transformations (10-3) can be used, which set those generated pseudo normal distributed numbers into a vector with n-dimensions.

Transformation Box – Müller for normal distribution
$N (a, \sigma^2)$: (10-2)

- Simulate two random numbers from random independent variables U_1 and U_2, with probability distribution $U (0, 1)$.
- Simulate two random numbers from the probability distribution $N(a, \sigma^2)$, based on the following transformations:

$$\begin{cases} X = \sigma \cdot (-2lnU_1)^{(1/2)} cos(2\pi U_2) + a \\ Y = \sigma \cdot (-2lnU_1)^{(1/2)} sin(2\pi U_2) + a \end{cases}$$

Transformation for the multivariate n-dimensional normal distribution $Z \sim Nn\ (\mu, \mathbf{\Omega})$: (10-3)

- Simulate n independent values of the random variable $N(0, 1)$, based on the transformations (10-2). In this way, the vector $U \sim N_n(0, I)$ is created.
- Calculate matrix H, as mentioned above.
- Generate vector Z, based on the equality: $Z = H \bullet U + \mu$.

Practically, in the case of our application the generator of $N(0, 1)$ pseudo numbers is used from Eviews package. Then, the above mentioned transformations are used to receive pseudo numbers with $N(a, \sigma^2)$ distribution and values from vector with normal distribution.

5. EXPERIMENT WITH MEAM AND THE MAIN RESULTS

Given the role of MEAM in economic activity analysis and forecasting, it is important for the policy making process to have a measure of the accuracy of these forecasts. The following experiment is designed and implemented to provide answers for the above issues. Given the fact that during the last couple of years the MEAM[1] model has been used to project the country's economic activity, by using the experiment given below, we have tried to explain the above mentioned questions related to the quality of forecasting.

MEAM is a quarterly model with a full structure of the economy that includes the real sector, fiscal sector, external sector, foreign and domestic price equations, as well as some equations on the labour market. Monetary policy is modelled through a simple Taylor rule, which reacts to the deviation of inflation and output gap. The model is estimated through the co-integration approach, which differentiates between short and long-term developments. The long-run relationships are determined by the co-integration relations, while short-run dynamics include current and lagged values of the variables included in the co-integration relation, and also other exogenous variables. Single equations are estimated individually, with the long-term relations being determined by the Johansen method. Generally, equation parameters are estimated, but we have also used some calibration of the long-run relationships. In total, MEAM has 45 equations, where 11 are estimated equations and the remaining part are mainly identities, such as the public sector (state sector), aggregate supply, disposable income, public debt and national account identities etc.

The main purpose of our experiment is to assess the model performance of MEAM with particular focus on GDP and its components. We also investigate for linearity vs. non-linearity in forecast, how the model is sensitive to external shocks and we

[1] For a detailed description of this model, refer to the material prepared by Dushku, Kota and Binaj, 2006 "Macroeconomic Model of Albania" as well as Kota and Dushku, 2007," Macroeconomic Model of Albania, A Follow-Up".

also examine the shape of distribution of the stochastic forecast. In this model, the GDP is determined from the demand side, so the development of economic activity is a function of total consumption, total investment (private and public investment), and of net exports.

By using Monte Carlo technique, we generated about 1000 random variables, particularly for GDP, real total consumption, real private investment, real exports and real imports. In total, we have a matrix of random numbers with dimension of 5x1000. Given the insufficient number of projections (only 8 quarters of forecasts, mainly for the GDP), in order to analyze the forecast error characteristics, we have assumed that these error terms are normally distributed with zero mean and standard deviation equal to the standard deviation of their historical residuals. This assumption is based on the knowledge in the predetermined variables. Worth underlining is that the assessment of economic relations and model solution cover the period 1996-2006, while the model performance has been assessed outside this period. This implies that the rest of the potential sources of uncertainty are treated as fixed; in other words, we are assuming real or fixed assessment coefficient, as well as an accurate structure of the model specification. Therefore, we have a forecast conditional on the performance of exogenous variables included in the model, parameters of equations and model specification accuracy. This means that the expected forecast error is a result of shocks or “noise” in the error term of each equation [Bianchi and Calzolari (1982)].

The experimented results through the MEAM model are presented in Table 1, which contains information related to the percentage bias of results, measured as the difference between the mean stochastic forecast and the mean deterministic forecast, and expressed as a percentage of the latter. The large difference between the mean stochastic forecast and mean deterministic forecast indicates a high degree of non-linearity of the model, thus providing evidence that the stochastic forecast is better than the deterministic forecast.

Another result we have presented is the coefficient of variation of the stochastic forecast, measured as a standard deviation of the forecast as a percentage of the mean stochastic forecast. Through this statistic, we can calculate the deviation of our forecast series compared to the mean forecast. The bigger the dispersion of the stochastic forecast around its mean, the more vulnerable our model is towards external shocks. The period of the model simulation includes altogether 3 years (or 12 quarters). The results related to percentage bias and the variation coefficient are presented for the GDP, consumption, private and total investments, exports, imports, all in real term, as well as CPI and inflation. Each result is shown for the first, middle and last period, in order to see how these statistics change over longer forecast periods.

Table 1 Summary of results for stochastic simulations with additive random disturbance added to the private consumption equation (1000 simulations)

Variable name	Percentage bias [1]			Coefficient of variation[2], %		
	I	II	III	I	II	III
Real GDP	0.0043	-0.0745	-0.1419	2.41	4.27	4.63
Real Total Consumption	0.0055	-0.0918	-0.1833	2.87	5.35	6.18
Real Total Investment	-0.0001	-0.0235	-0.0854	0.26	1.94	3.05
Real Private Investment	-0.0001	0.0145	-0.0992	0.30	2.33	3.60
Real Exports	-0.00001	-0.00002	0.00002	0.23	0.28	0.27
Real Imports	0.0005	-0.0288	-0.0946	0.79	2.37	3.41
Inflation rate	0.0048	0.0145	-0.0139	2.17	3.77	4.80

[1] Percentage bias = (mean stochastic forecast/deterministic forecast -1).
[2] Coefficient of variation = (stochastic forecast standard deviation/mean stochastic forecast).
Source: Authors' own calculations

Some of the obtained results show that the forecast (or projections) derived from the MEAM model in general is linear, based on the low values of the coefficient of percentage bias, which indicates that the difference between the mean deterministic forecast and the mean stochastic forecast is very small. On the other hand, it means that the deterministic forecast obtained from MEAM model is an accurate predictor of the mean. If we analyze error accumulation over time, we will notice that this error is larger at the end of the period than in the beginning, but still insignificant. There is greater error accumulation mainly in consumption, GDP and private investment, ranging between [0.10-0.18].

Another statistic measured for 3 sample years is the coefficient of variation, for each of the endogenous variables. The obtained results show that during the first period of the forecast, for 5 out of 7 variables taken into consideration, random disturbances affect them by less than 1%. Meanwhile, consumption and GDP are excluded, where this coefficient reaches values 3 and 2% respectively. Considering the fact that consumption plays an important role in GDP, it is expected that a part of these uncertainties coming from this variable be transmitted to it as well. This result shows that more attention should be paid on consumption equation and its determining variables. In addition, the longer the forecast period, the higher is the uncertainty accompanying these forecasts.

In the case of the inflation variable, we should not misinterpret the large coefficient of variation for the four periods taken into consideration, varying from 2.17 to 4.8, as inflation is expressed as a percentage. It implies that a variation coefficient varying from 2 to 5% is interpreted as an error in inflation, respectively with 0.03 and 0.8 percentage point error from the respective level of the inflation rate.

Another statistic we have analyzed is the form and type of our forecast distribution, by examining skewness and kurtosis of the forecast data, where the former is a measure of symmetry, and the latter is a measure of the peakedness or flatness of the data in terms of the normal distribution.

The chart below makes a summary of the form of distribution of 1000 simulations for each of the variables mentioned above, detailed for the first, second and third year of forecast, in order to view how this forecast varies over time. Since we examine the monetary policy response and its impact on the main macroeconomic indicators over a two-year period, our forecast distribution analysis will focus only on the first two years of forecast.

Chart 1a Forecast distribution of real GDP in different time

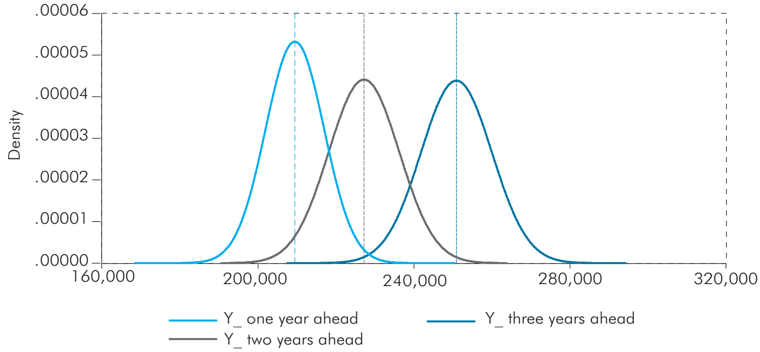


Chart 1b Forecast distribution of real Consumption in different time

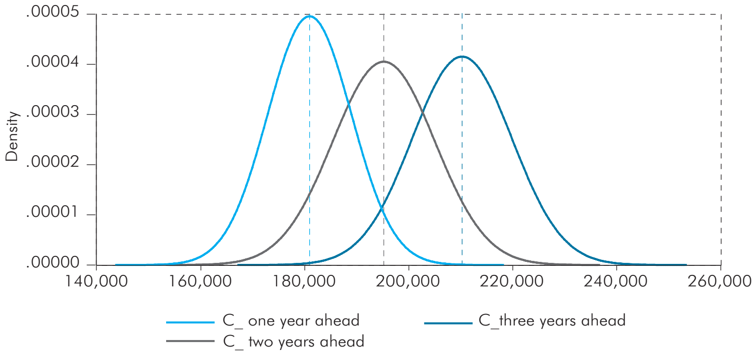


Chart 1c Forecast distribution of real total investment in different time

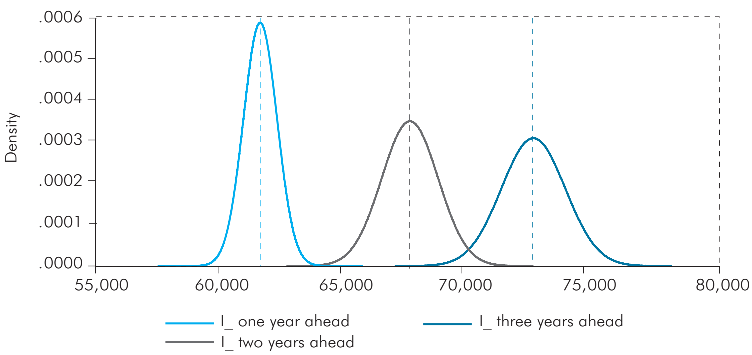


Chart 1d Forecast distribution of real private investment in different period

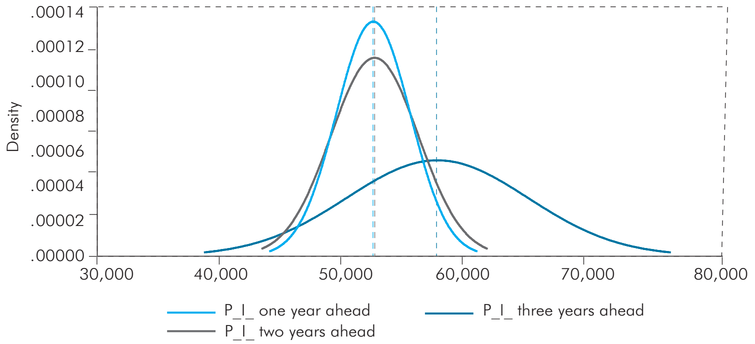


Chart 1e Forecast distribution of real imports in different period

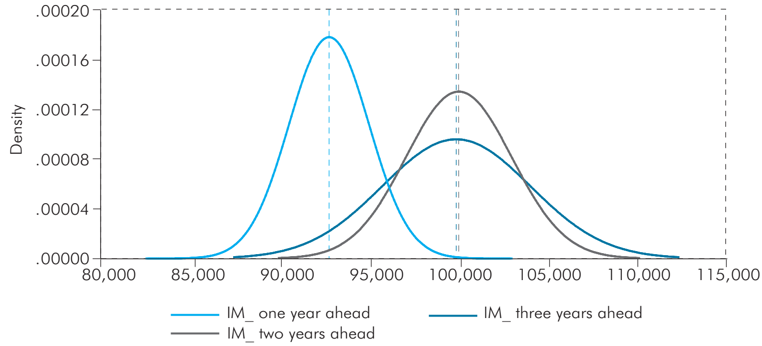


Chart 1f Forecast distribution of real exports in different period

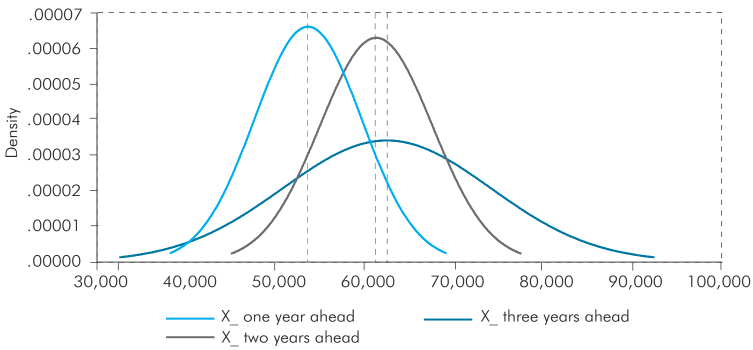
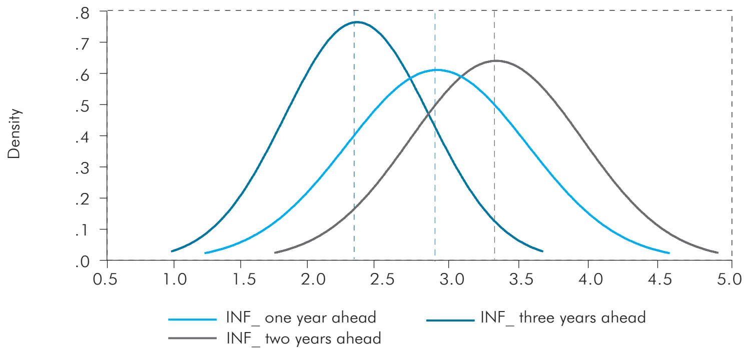


Chart 1g Forecast distribution of inflation in different period



Source: Authors' own calculations

To each of the above variables, we have approximated the normal distribution of 1000 forecast values for respectively the first, second and third year of forecasts. Based on the above figures, we note that the longer the forecast period, the higher is the uncertainty accompanying these forecasts. The same result is also confirmed by the coefficient of variation. We have also tested whether these distributions have normal distribution or not. Table 3[2] makes a detailed summary of Lilliefors test of normality[3] (1967). The test results show that all the forecast series related to the main indicators have normal distribution, thus enjoying all the characteristics of this form of distribution.

While Table 2 contains some general characteristics on these variables, their forecast value in terms of mean, median, maximum, minimum, standard deviation, skewness and kurtosis for the entire simulation period.

[2] See Annex 1.

[3] The basic hypothesis of this test is that the series has a normal distribution, while the alternative hypothesis is that the series does not have a normal distribution.

Table 2 Descriptive Statistics.*

Variable name	Mean	Median	Max.	Min.	Std. Deviation	Skewness	Kurtosis
Real GDP	205632	205558	254970	171190	10693.84	0.137034	3.24406
Real Total Consumption	168400	168719	212456	128681	10417.21	-0.06826	3.22732
Real Total Investment	61174	61125	72487	54543	2338.247	0.323942	2.91077
Real Private Investment	48865	49695	65279	35150	4639.117	-0.28722	2.44040
Real Exports	55929	54282	81504	44037	9778.727	0.942931	3.37017
Real Imports	85790	86220	103307	70286	5478.868	0.024247	2.18572
Inflation	2.55	2.48	4.53	0.62	0.808551	0.202729	2.11792

* Data for real GDP, real consumption, private investments and real total investments, real imports and exports are expressed in millions of ALL. CPI is an index, while inflation is expressed in % and in quarterly data.
Source: Authors' own calculations

Based on the results regarding the mean and the median, we note that they are different for all the variables, hence confirming the fact that the forecasts are asymmetrical. This result is also confirmed by the skewness values for each variable.

Regarding the above findings, it is clear that the series of real GDP, total real investments, exports and inflation are positively skewed, while the other part of the variables is negatively skewed. This means that, until a certain point, data suffer from asymmetry and that the mean and median are notably different. Hence, when we consider the forecasts obtained for real GDP, real total investment, exports, imports and inflation, we have to know that they are overestimated, while the forecasts of consumption and private investment show that they are underestimated.

The overall conclusion is that for most of the variables taken into consideration, the differences between stochastic forecast and deteministic forecast are not significant and the policy implication of such results is that the deterministic forecasts that are generated by the MEAM are efficient and accurate and useful for the purpose of monetary policy. The distribution of forecasts for all variables is assessed as normal, but the different mean and median values confirm that the forecasts are asymmetrical. In the light of these findings it is important that stochastic rather than deterministic forecasts are used in the policy decision-making process.

6. CONCLUSIONS

Macroeconomic forecasts are very important for the decision-making process and, as such, they should be approached as best as possible to the losses distribution function or to the profits maximization of decision-makers. Despite the way how the mistakes are evaluated in a model, those are not very preferred from the decision-makers; this because unexpected results obtained from the forecast of exogenous shocks that come from financial markets are inevitable, making the decisions made to be situated in a very insecure environment. By this, it derives that we should live with the mistakes of our forecasts, but the most important is to understand the source of uncertainty in all the decision-making process [Don (2001)].

In this material, we have investigated the uncertainty of forecast that derives from the mistake terms of the MEAM model through stochastic simulations of the Monte Carlo technique, and we have tried the so-called “weak links”. The estimated results showed that the percentage bias coefficient (measured as a difference between stochastic and deterministic mean of forecast) obtained the low values of it, showing that the deterministic forecast is a good forecaster of the mean of overall variables and that the model is linear.

While the coefficient of variation showed that the biggest piece of uncertainty in the forecasting of real GDP value stemmed from uncertainties that come from consumption and private investments equation. This result is in line with the expectations, taking into consideration the fact that these two components take a considerable part in the total weight of GDP. Also, the investigation of determinant factors of each of the variables above remains a duty for future research.

Another result that we analysed had to do with the symmetry of data, analyzing the form of distribution for stochastic forecasts. Results showed that all forecasts follow a normal distribution pattern, but the biggest part of them is skewed positively from the right, hence the forecast made is overestimated. In addition, the data distribution is asymmetrical given the different mean and median values, so the forecasts are asymmetrical. Based on the results obtained from stochastic simulations through the MEAM model, one of the most important conclusions is that the deterministic forecasts of this model are good forecasters of the mean of the variables and that the model is not subject to external shocks.

Hence, the deterministic estimates obtained from this model for at least the first two years are good forecasters of the behaviour of the main macro indicators. However, in order to enhance the reliability and efficiency of the decisions made on MEAM projections, but also on other models or estimates, it is suggested that these forecasts take the form of probability distribution.

This material analyzed only one of the main sources of uncertainty, as a future research will measure the model performance that stems from other uncertainties mainly the uncertainty arising from the estimated parameters, as well as those from the behaviour of exogenous variables.

REFERENCES

Adelman, I & Adelman, F.L: "The Dynamic Properties of the Klein-Goldberger Model", Econometrica, Vol. 27.4, October, 1959.

Amano, R, McOhail, K, Piora, H & Rennison, A: "Evaluating the Quarterly Projection Model: A Preliminary Investigation", Bank of Canada, Working Paper 2002-20.

Bianchi, C & Calzolari, G: "Evaluating Forecast Uncertainty due to Errors in Estimated Coefficients; Empirical Comparison of Alternative Methods", MPRA Paper, No.22559, May 2010.

Bianchi, C, Calzolari, G & Weihs, C: "Parametric and Nonparametric Monte Carlo Estimates of Standard Errors of Forecasts in Econometric Models", MPRA Paper, No. 24, February 2011.

Brown, Bryan W & Mariano. R.S: "Residual-Based Procedures for Prediction and Estimation in a Nonlinear Simultaneous System", Econometrica, Vol.52, No.2, March 1984.

Canova, F: "Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies", International Economic Review, Vol.36, No.2, 1995.

Don, F.J.H: "Forecasting in Macroeconomics: A Practitioner's View", De Economist 149, No.2, 2001.

Dushku, E, Kota, and Binaj, G: "Modeli Makroekonometrik i Shqipërisë", Discussion Paper, Bank of Albania, 2006.

Ericsson, N.R: "Forecast Uncertainty in Economic Modelling", Board of Governors of the Federal Reserve System, No.697, February 2001.

Fair, R.C: "Bootstrapping Macro Econometrics Models", Studies in Nonlinear Dynamics & Econometrics, Vol.7, Issue 4, 2003.

Hendry, D. F., 2011: "Empirical Economic Model Discovery and Theory of Evaluation", University of Oxford, No. 529, ISSN 1471-0498.

Garratt, A, Lee, K, Pesaran, M.H & Shin, Y: "Forecast Uncertainties in Macro Econometric Modelling: An Application to the UK Economy", 2003.

Gentle, J.E.: "Random Number Generation and Monte Carlo Methods, 2nd ed., 2003, Springer.

Kota, V & Dushku, E: "Modeli Makroekonometrik i Shqipërisë, Material Vijues", Discussion Paper, Bank of Albania, 2009.

Lancer, D, & Kranendonk, H: "Investigating Uncertainty in Macroeconomic Forecasts by Stochastic Simulation", CPC paper, 2008.

Lemieux, Ch: "Monte Carlo and Quasi-Monte Carlo Sampling", 2009, Springer.

Rubinstein, R.Y. & Kroese, D.P.: "Simulation and the Monte Carlo Method, 2nd ed., 2008, Wiley.

Shijaku, H & Ceca, K: "A Credit Risk Model for Albania", Bank of Greece, 2009.

ANNEX:

Critical Value of Lilliefors Test for $\alpha=0.05$ is estimated as:

$L_K = \frac{0.895}{f_N} = 0.028288$, where $f_N = 0.83+1000/\sqrt{1000}$ and $N=1000$.

Table 3 Normality tests for forecast distribution

Name of variable	Test values for the first year	Test values for the second year
GDP	0.020738***	0.016955***
C_real	0.016705***	0.014608***
I	0.017148***	0.020920***
P_I	0.017148***	0.020920***
IM	0.020434***	0.020767***
X	0.021088***	0.015335***
INF	0.020552***	0.019982***

Note: *** refers to statistical significance at $\alpha=0.05$.

CIP Katalogimi në botim BK Tiranë

Altin Tanku, Elona Dushku, Kliti Ceca
Empirical investigation of forecast uncertainty
with Monte Carlo simulation- /
/Tanku Altin, Dushku Elona, Ceca Kliti - Tiranë:
Bank of Albania, 2012

-36 f; 15.3 x 23 cm.

Bibliogr.
ISBN: 978-99956-42-56-4.

You may find this paper in the following address:

www.bankofalbania.org

If you want to receive a hard copy of this paper, write to us at:

Bank of Albania
Sheshi "Avni Rustemi", Nr. 24, Tiranë, Shqipëri
Tel.: + 355 4 2419301/2/3; + 355 4 2419409/10/11
Fax: + 355 4 2419408

or send an e-mail to:

public@bankofalbania.org

Printed in 500 copies