USE OF MACROECONOMETRIC MODELS IN GDP ANALYSIS

Prof. Constantin ANGHELACHE PhD *(actincon@yahoo.com)*
Bucharest University of Economic Studies / „Artifex" University of Bucharest
Prof. Radu Titus MARINESCU PhD *(radu_titus_marinescu@yahoo.com)*
„Artifex" University of Bucharest
Doina AVRAM Ph.D Student *(doina.avram@gmail.com)*
Bucharest University of Economic Studies
Ec. Ștefan Gabriel DUMBRAVĂ *(stefan.dumbrava@gmail.com)*

Abstract

In presenting and analyzing this issue, the authors considered a number of issues that are relevant in macroeconomic analyzes. Researchers, analysts, express, without reservation, the fact that a sustainable growth of the national economy is achieved principally by increasing investment. Clearly, this is because new investments mean new companies or the development of existing companies, the emergence of new jobs, the recruitment of new employees by absorbing the unoccupied population or by absorbing unemployed people, increasing the tax base wages that lead to increased revenues to the consolidated state budget, improved management, improved living conditions and much more. This is a positivist, realistic study, but does not take into account the relationships that need to exist or which texist at one time. Therefore, the authors have proposed to highlight some macroeconomic models that can be used in the gross domestic product analysis as well as in the analysis of the structural evolution of the gross domestic product. A theoretical presentation is carried out in order to emphasize the existence and the perspective of the use of these econometric models by making an overview of the aspects that the use of econometric models raises. In this respect, the simple or multiple linear regression model is emphasized and the hypotheses that a regression model raises, but it is finally concluded that the regression parameters are those that have statistical significance and give consistency to the hypothesis that rises on their account.

Keywords: *macroeconomic model, GDP, regression model, dynamic analysis, statistical indicators*

JEL Classification: C50, C61, E01

Introduction

In the study, the authors focused on some aspects that are significant from the point of view of the structural analysis of the gross domestic product. Theoretical concrete theoretical references are made to the usable macroeconomic models, there are presented aspects related to the linear and simple linear regression model, it explains the hypotheses to be fulfilled and in general terms it is emphasized that regression parameters with statistical significance after testing are consistent elements can lead to a judicious analysis and interpretation of the economic situation underlying this type of analysis. References and how to explain the theoretical aspects of using macroeconomic models of structural analysis of gross domestic product are considered to be sufficient and fully enlightening for those who want to use these models subject to the researchers' perceptible eye by the authors.

Literature review

Anghelache and Anghel (2017) conducted a study of GDP evolution in Romania based on deflated data. Anghelache, Partachi, Sacala and Ursache (2016) analyzed the correlation between GDP and foreign direct investment. Anghelache et al (2007) presented the fundamental aspects of the macroeconomic analysis. Anghelache and Anghel (2016) presented the statistical instrument used in economic analyzes. Anghelache (2013) highlighted the econometric models used in economic analyzes. Bardsen, Nymagen and Jansen (2005) studied the use of econometric models in macroeconomic modeling. Dobrodolac (2011) addressed aspects of the importance of forecasting in the decision-making process. Guner, Ventura and Yi (2008) studied the macroeconomic effects of adopted policies. Koulakiotis, Lyroudi and Papasyriopoulos (2012) analyzed the correlation between inflation and gross domestic product in some European countries. Lucas (2003) addressed a number of issues related to macroeconomic priorities. Lucas and Moll (2014) studied aspects of economic growth. Woodford (2010) analyzed the role of financial intermediation in macroeconomic analyzes.

Methodology, data, results and discussions

To analyze the impact of foreign direct investment on economic growth in Romania, we used a series of complex macroeconomic models.

In our research, we used models specific to the prognosis of the evolution of an economic system, the simple and multiple linear regression model.

Using these models it is possible to determine the value of some coefficients according to the values recorded by other macroeconomic indicators considered as variables depending on the evolution of the time factor.

In this economic analysis, it is necessary to establish the relationship between two or more variables, using two techniques specific to statistics, namely:

- correlation - shows how strong the link between the variables analyzed is;

- regression - Explains and allows the prediction of the value of one of the factors on the account of the other person / others.

Thus, we obtain a factor (variable) which is the determinant factor, including the other influences in the residual variable.

The linear linear regression model can be transcribed as a mathematical function in the form of:

$$Y = f(x) + \varepsilon \qquad (1)$$

Within this model the interdependence between y and x can be written by the function:

$$y_i = b + a \cdot x_i + \varepsilon_i \qquad (2)$$

in which:

y_i – the resulting variable;

x_i – the factorial variable;

ε_i – residue variable.

In choosing the linear linear regression model, the corresponding econometric assumptions will be taken into account. Regression Model Hypotheses:

I1: the data series are not affected by measurement errors;

I2: the residual variable has a zero average;

I3: the dispersion of the residual variable is invariant over time, ie it has homoscedasticity property;

I4: residues are not autocorrelated;

I5: the factorial variable (explanatory) is not correlated with the residual variable;

I6: model errors are normally distributed after a scatter of zero mean and dispersion σ^2.

• Some aspects of the linear regression function

In statistical modeling, regression analysis is a statistical process for estimating relationships between variables that includes several techniques for modeling and analyzing variables when emphasizing the correlation between two variables. The regression analysis expresses how the „variable criterion" changes if the variable (x) shows a variation, while the other independent variables remain unchanged. Most frequently, the regression analysis estimates the dependent expectation of the dependent variable when the value of the independent variables is given - that is, the mean value of the dependent variable when the independent variables are fixed.

The notions and models presented in this final chapter have taken into account the theoretical and practical researches and theoretical and practical aspects of some important authors from the literature. In this way, econometric modeling is developed by Bardsen and others (2005), but also by other authors. Granger and Haldrup (in 1997) develop the theoretical elements regarding cointegration systems, with extensive results in the published works in their continuation.

The earliest form of regression was the smallest square method, which was published by Legendre in 1805 and by Gauss in 1809.

Both Legendre and Gauss have applied this method to determine, from astronomical observations, the orbits of the bodies around the Sun (especially comets, but also later, the newly discovered smaller planets). Gauss published a further development of the least square theory in 1821, including a version of the Gauss-Markov theorem.

The term „regression" was invented by Francis Galton in the nineteenth century to describe a biological phenomenon. The phenomenon studied analyzes the fact that the height of the descendants of the high population tends to regress down to a normal average (a phenomenon also known as regression to mean). For Galton, this regression study only had this biological significance, but his work was subsequently extended by Udny Yule and Karl Pearson in a broader statistical context. In Yule and Pearson's work, the shared distribution of response and explanatory variables is assumed to be Gaussian. This assumption was weakened by R.A. Fisher in his works of 1922 and 1925. Fisher assumed that the conditional distribution of the response variable was of the Gaussian type, but the joint distribution should not fulfill the same condition. In this sense, the Fisher hypothesis is closer to Gauss's 1821 formulation.

Regression analysis is also of interest in characterizing variation of the dependent variable around the regression function that can be described by a probability distribution.

Regression analysis is widely used for prediction and prognosis and also for understanding which of the independent variables are related to the dependent variable, and to explore the forms of these relationships. In small circumstances, regression analysis can be used to infer causal relationships between dependent and independent variables. However, this can lead to false illusions or relationships, so caution is recommended (for example, the correlation does not imply causality).

Many techniques have been developed to perform regression analysis; the most common methods (linear regression and least squares method) are parametric methods, because the regression function is defined in terms of a finite number of unknown parameters that are estimated by data. Nonparametric regressions refer to techniques that allow the regression function to stretch into a certain set of functions, which can be infinite-dimensional.

The performance of regression analysis methods in practice depends on the form of the data generation process and the regression method used. Since the true form of the data generation process is not generally known, regression analysis often depends to a certain extent on assumptions about this process. These assumptions can sometimes be tested if there is enough available data.

Predictive regression models are often useful even when hypotheses are moderately violated, in the sense that they can not always be optimally met. However, in many applications, especially those that refer to causality based on observational data, regression methods can produce misleading results.

In a narrower sense, the regression analysis can specifically refer to the estimation of the continuous response variables, unlike the discreet response variables used in the classification (the classification is the problem of identifying which set of categories (sub-populations) belongs to a new observation, based on a set of data training containing observations, whose belonging to a certain category is known).

Regression methods continue to be an area of active research, with new robust regression methods being developed in recent decades, regressions involving correlated responses such as chronological and growth curves, regressions in which the predictor (the independent variable) or the variables response data are curves, images, graphics, or other complex data objects, regression methods that involve estimating different missing data types, non-parametric regressions, Bayesian regression methods, regressions in which predictor variables are error-prone, multi-variable regressions predictor than observations and causal inference with regression.

• **Linear regression models involve the following variables:**

-unknown parameters, denoted by β, with scalar or vector value

-the independent variables, X.

- dependent variable, Y.

- all other factors, which have a random or constant action on the endogenous variable, expressed by the residual variable Ɛ.

Different terminologies are used in various application domains instead of dependent and independent variables.

A regression model involves reporting the dependent variable Y to a function of X (the independent variable) and β (the free term).

$Y \approx f(X, \beta)$ (3)

Approximation is typically formalized as $E(Y \mid X) = f(X, \beta)$. In order to perform a regression analysis, the form of the function f must be specified. Sometimes the form of this function is based on the knowledge about the relation between Y and X; if no such information is available, we choose a flexible and convenient form for the f function.

Let us assume now that the vector of unknown parameters β is of length k. To be able to perform a regression analysis, the user must provide information about the dependent variable Y:

- if the N data points of the form (Y, X), where N <k are observed, most classical approaches of the regression analysis can not be achieved: since the system of equations defining the regression model is undetermined, there is enough data to determine b.

- if exactly N = k data points are observed, and the function f is linear, the equations of type $Y = f(X, \beta)$ can be solved concretely. This is reduced to solving a set of unknown N equations (elements b), which has a unique solution as long as X are linearly independent.

-the most common situation is where N> k data points are observed. In this case, there is enough information to make it possible to estimate a single value for β, and the regression model when applied can be viewed as an overdetermined β system.

In the latter case, regression analysis provides tools for:

-find a solution for unknown beta parameters to minimize the distance between the measured and predicted values of the Y dependent variable (also known as the least squares method).

- In some statistical assumptions, the regression analysis uses the surplus of information to provide statistical information on unknown parameters and to predict values of the dependent variable.

When the number of measurements N is greater than the number of unknown parameters, k, and measurement errors are normally distributed, then the excess of information contained in (N - k) measurements is used to make statistical predictions about unknown parameters. This excess of information is expressed as degrees of freedom of regression.

• **Hypotheses of the linear regression model**

- any regression model is based on a series of assumptions. In this regard, we recall:
- the sample is representative of the population for which statistical inference is used.

- the error is a random variable with a mean of zero, conditioned by the explanatory variables.
- Independent variables are measured without error.
- the independent variables (predictors) are linearly independent, meaning that any predictor can not be expressed as a linear combination of the other variables.
- the errors are uncorrelated, which means that the variance-covariance matrix of the errors is diagonal and each element other than zero represents the error variation.
- error variance is constant in observations (homoscedasticity).

These are sufficient conditions for estimating parameters using the smallest square method; Particularly, these assumptions imply that the parameter estimates will be impartial, consistent and effective in the class of linear impartial estimators. It is important to note that real data rarely satisfy hypotheses, but also in these situations the method is used even if assumptions are not true. Reports of statistical analyzes usually include analyzes of sample data tests and methodology for appropriate model selection and utility.

In linear regression, the model specification is that the dependent variable yi is a linear combination of parameters (but must not be linear in the independent variables). For example, in linear linear regression for modeling n data points there is an independent variable: xi and two parameters, **β0 şi β1:**

Equation of regression line: $y_i = a + bx_i + \varepsilon i$, i = 1, ..., n. (4)

In multiple linear regression, we consider at least two factorial variables, ie functions of independent variables.

Adding a X_i^2 term to the previous regression leads to the parable:

$y_i = a_0 + b_1x_1 + b_2x_2 + \varepsilon i$, i = 1, ..., n. (5)

This is also a linear regression; although the expression on the right is quadratic in xi independent variable Xi, is linear in parameters a0, b1 and b2. In both cases, it is a term of error and the index i is attributed to each particular observation. Turning our attention to the linear regression model, taking into account a random sample of the population, we estimate the parameters of the population and obtain the linear regression model:

$\hat{Y}_i = \beta_0 + \beta_1 X_i$ (6)

The residual value, ei = Yi - Ŷi is the difference between the value of the dependent variable predicted by model Ŷi, and the real value of the dependent variable, Yi. An estimation method is the smallest square method. This method obtains parameter estimates that minimize the sum of squares of residues, SSE:

$$SSE = \sum_{i=1}^{n} e_i^2 \qquad (7)$$

Minimizing this equation results in a set of normal equations, a set of simultaneous linear equations based on parameters that are resolved to determine the parameter estimators β0 şi β1.

$$\beta_1 = \frac{\sum(xi - \bar{x})(yi - \bar{y})}{\sum(xi - \bar{x})^2} \tag{8}$$

If the term of the population error has a constant variation, the estimation of this variance is given by:

$$\sigma^2\varepsilon = \frac{SSE}{n-2} \tag{9}$$

the mean square error (MSE) of regression.

The denominator is the sample size of the sample with the number of model parameters estimated at the same data, (n-p) for regressions p or the value (n-p-1) if the term intercept is used; in this case p = 1, so the denominator is (n-2).

• **Regression parameters and their statistical significance**

Once the regression model has been built, it is important to confirm the model's suitability and the statistical significance of the estimated parameters. Commonly used checks include hypothesis testing and residual values. Statistical significance can be verified by a test F as a general test, followed by t-tests for individual parameters.

Interpretations of these diagnostic tests are based in particular on the model assumptions. Although the examination of residues can be used to invalidate a model, the results of a test or test F are sometimes more difficult to interpret if the model assumptions are violated.

For example, if the error term does not have a normal distribution, in the case of an analysis based on a small number of parameters, they will not track normal distributions and complicate inference. However, if the relatively large samples are used, the central limit theorem may be used so that hypothesis testing can continue using asymptotic approximations. The response variable may be non-continuous („limited" to cover a subset of linear regression data). For binary variables, if the least square regression analysis is applied, the model is called a linear probability model.

The regression models predict a value of the Y variable given the known values of the X variables. The prediction within the value range in the data set used for the model used is known informally as the interpolation. Forecast outside of this data range is known as extrapolation, which strongly relies on the regression model assumptions.

The more extrapolation takes place outside the existing data, the higher the chances for the model to fail because of the differences between assumptions and sample data or real values.

It is generally recommended that when an extrapolation is made, the estimated value of the dependent variable shall be accompanied by a prediction interval representing the uncertainty. Such time intervals tend to expand rapidly as the values of the independent variable are shifted out of the range covered by the observed data.

However, this does not completely cover the modeling errors that may arise: in particular the hypothesis of a particular form for the relationship between Y and X. A correctly performed regression analysis will include an assessment of how the proposed form is offset as a model of the observed data, but this is only possible within the range of available variables of the independent variables. This means that any extrapolation is particularly relevant if it is based on assumptions made about the structural form of the regression relationship. Best practices recommend that choosing a linear model in variables or linear parameters should not be chosen conventionally, but consider and exploit all the knowledge available to implement a regression model. The implications of this step of choosing a proper functional form for regression may be great when making the extrapolation decision. At a minimum, it can be assured that any extrapolation resulting from a suitable model is „realistic".

Conclusion

The author's study reveals important theoretical conclusions that a pertinent macroeconomic analysis based on the macroeconomic indicators of aggregate results and the factorial evolution of global indicators is possible by using macroeconomic or more simplified models of statistical-econometrics used at macroeconomic level. These models, by taking into account sufficiently long series of indicators that are deflated and brought to comparable price levels, give concrete figures on which to interpret the data. Also, based on these macroeconomic models, predictions can be made using the regression parameters derived from a previous data survey, extrapolated over time during the period under review. A final conclusion is that the study is broader, making only some concretizations and this leads to the fact that it can be extended and used in any other perspective where the variables analyzed by these macroeconomic models should first be studied, logically interpreted to identify the perspective the possibility of a correlation between them.

References

1. Anghelache, C. and Anghel, M.G. (2017). Analysis of the evolution of the Gross Domestic Product of Romania using deflated data, *International Conference „Economic Experiences”*, The Romanian Foundation for Business Intelligence, Transilvania University of Brașov, The Romanian Academy, Brașov, 145-153
2. Anghelache, C., Partachi, I., Sacală, C. and Ursache, A. (2016). Using econometric models in the correlation between the evolution of the Gross Domestic Product and Foreign Direct Investments. *Romanian Statistical Review, Supplement*, 10, 124-129
3. Anghelache, C. (2013). *Elemente de econometrie teoretică,* Editura Artifex, București
4. Anghelache, C. şi colaboratorii (2007). *Analiza macroeconomică: Sinteze şi studii de caz*, Editura Economică, București
5. Anghelache, C. and Anghel, M.G. (2016). *Bazele statisticii economice. Concepte teoretice şi studii de caz*, Editura Economică, București
6. Bardsen, G., Nymagen, R. and Jansen, E. (2005). *The Econometrics of Macroeconomic Modelling,* Oxford University Press
7. Dobrodolac, T. (2011). Forecasting by econometric models as support to management. *Perspectives of Innovations, Economics & Business, 7* (1), 72-76
8. Guner, N., Ventura, G. and Yi, X. (2008). Macroeconomic Implications of Size-Dependent Policies. *Review of Economic Dynamics, 11* (4), 721–744
9. Koulakiotis, A., Lyroudi, K. and Papasyriopoulos, N. (2012). Inflation, GDP and Causality for European Countries. *International Advances in Economic Research, 18* (1), 53-62
10. Lucas, R.E. Jr. (2003). Macroeconomic Priorities. *American Economic Review, 93* (1), 1-14
11. Lucas, R.E. and Moll, B. (2014). Knowledge Growth and the Allocation of Time, *Journal of Political Economy, 122* (1), 1–51
12. Woodford, M. (2010). Financial Intermediation and Macroeconomic Analysis. *Journal of Economic Perspectives*, *24* (4), 21-44