Theoretical and Applied Economics
Volume XXI (2014), No. 10(599), pp. 59-70

Econometric model used in the capital market analysis

Mădălina Gabriela ANGHEL
„ARTIFEX" University of Bucharest, Romania
madalinagabriela_anghel@yahoo.com

Abstract. *In the frame of this article, by applying the linear regression model, we proceed to the analysis of the existing dependence between the value recorded by the Bucharest Exchange Trading index (meaning the overall evolution of the capital market in Romania) and that of the stock exchange capitalization. The application of this econometric model provides to the capita investors a series of necessary information meant to fix their behaviour during the forthcoming periods.*

Keywords: regression model, capitalization, BET index, statistical tests, the least squares method.

JEL Classification: G10, G14, C58.
REL Classification: 11B.

1. Introduction

The financial markets and, mainly, the stock exchange markets, are complex structures where a very large number of economic entities are acting wishing, first and foremost, to get a profit as high as possible. In this respect, the modern financial markets are distinguishing themselves by a very conspicuous dynamics of the investment activity and – particularly – as far as the portfolio investments are concerned. In this context, it is required that statistical-mathematical models are formulates, as support for modelling the main processes occurring on these markets as well as to allow a easier management of all the activities involved by the investment activity. A significant aspect when analysing the performances of the capital market is given by the establishment of the existing correlations between the various indicators and/or indices this one.

2. Literature review

The first econometric models which could be applied on the financial markets, generally speaking and, particularly, on the capital market, have been conceived during the second half of the XX century. This is the period when starting from the existing theoretical models, the first models with strict application in the financial and capital investment domains have been conceived and implemented. This kind of econometric models can be found out in the works of certain well-known specialists in the field, such as: A. Goldberg *(Econometric theory* – 1964), H. Theil (*Principles of econometrics* – 1971), L. Klein (*A textbook of econometrics* – 1974) etc.

By the years 1990 and the beginning of the years 2000, the econometric models meant to the financial and banking analysis reached a new stage of development which allowed them to be matched with the new requirements of the capital market. In this respect, to note the efforts paid by researchers such T. Milles (*The econometric modelling of financial time series* – 1993), E. Berndt *(The practice of econometrics: classic and contemporary* – 1991), W. Green (*Econometric analysis* – 2000), F. Diebild *(Elements of forecasting* – 2002) or C. Dougherty *(Introduction to econometrics* – 2007), He, Changli; Silvennoinen, Annatiina; Terasvirta, Timo *(Parameterizing Unconditional Skewness in Models for Financial Time Series* –2008).

The econometric model of regression and the actual possibilities to use it for economic analysis kept on being permanently analysed during the previous century , a series of reference works for this domain to be mentioned such as those signed by Franklin Graybill (*An introduction to linear statistical models* – 1961), Peter Sprent *(Models in regression and related topics* – 1969), David Belsley, Edwin Kuh şi Roy Welsh *(Regression diagnostics* – 1977), Damodar Guijarati *(Basic econometrics* – 2005) or Gregory Chow *(Econometrics* – 1983).

3. The simple regression model

From a theoretical point of view, the simple regression model is defined through a mathematical relation built up on the basis of the economic theory which assumes that the economic phenomenon, as an effect, is the outcome of the cumulated action of two categories of factors (Dougherty, 2008):

- a main, determinant factor;
- all the other factors which can be considered as not essential, of casual or constant action, invariable of the economic effect phenomenon.

Consequently, by way of abstracting, these ones are not taken into account for the aimed analysis, their effect being concentrated in the residual variable.

From a mathematical point of view, the simple regression model implies a direct relation between two variables, of which one is considered as a factorial one (causal or explicative) while the second one, is a resulting one. Such an econometric model can be transcribed, in a mathematical for, as follows (Andrei, Bourbonnais, 2008):

$Y = f(X) + \varepsilon$.

Out of the category of the uni-factorial regression models the largest utilization in the economic analyses goes to the linear model. In the case of this model, the relation between the resulting (Y) and the factorial variable (X) can be synthetized through a function of the form (Voineagu, Ţiţan, 2007):

$y_i = b + a \cdot x_i + \varepsilon_i$

where:

y_i – the resulting characteristic (explained);

x_i – the factorial characteristic (explicative, causal);

ε_i – the residual variable.

In the situation that the linear dependence is found again as a result of making transformations on the two variables, we shall say that the regression model is a linear one as against its parameters.

For instance, the model $Y = b + a\ln X$ is non-linear as against the factorial variable but it is a linear model as against the two parameters. On the contrary, the regression model $Y = b + \ln aX$ is linear as against the factorial variable, but it is not linear as against the two parameters a şi b, but as against b and $\ln a$.

From a theoretical point of view, it must be mentioned that there are various uni-factorial non-linear models, which are linearized through transformations applied to the variables of the regression model. This category of non-linear models transformed into linear models can include the following (Guijarati, 2005)

- $y_i = a \cdot x_i^b$ turns into a linear model by logarithmic procedure applied to the two terms of the equality above:

$\log y_i = \log a + b \cdot \log x_i$

On the basis of this transformation, we get as result a linear model as against the variables $\log y_i$ and $\log x_i$.

- *The exponential model* or *the log model* defined by the relation:

$y_i = a + b^x$

it is linearized by logarithmic procedure, resulting the linear model:

$\log y_i = \log a + x_i \cdot \log b$

A series of non-linear models cannot be written in the form of linear models by applying certain elementary transformations. In other situations, there are different techniques of estimations which are applied in order to estimate the parameters. As they cannot be linearized through elementary transformations, the estimation of the parameters is made through numerical methods (Anghelache, Anghelache, 2009).

In order to get correctly estimates results out of applying the linear regression model it is necessary to define a set of standard hypotheses for the model out of the general population. These hypotheses are formulated as follows (Anghelache, 2013):

- I_1: the data series are not affected by measurement errors;
- I_2: the residual variable has a zero mean;
- I_3: the dispersion of the residual variable is invariant in time, meaning that it gets the homoscedasticity property;
- I_4: the residuum are not self-correlated;
- I_5: the factorial (explicative) variable is not correlated with the residual variable;
- I_6: the model's errors are normally distributed according to a co-distribution of a zero mean and a dispersion σ^2 ($\varepsilon_t \rightarrow N(0, \sigma_t^2)$).

In order to test these hypotheses the statistical tests are applied.

The estimation of the linear regression model is made on the basis of the data series for the two characteristics. From the mathematical point of view, these series are represented by the column vectors of the variables x_i and y_i:

When defining the linear regression, the above mentioned hypotheses are taken into consideration. Considering the relation $y_i = a * b^{x_i}$ it is to notice that the estimated value of the resulting variable, of the estimators of the model parameters and their properties depend on the characteristics of the independent variable and the properties of the residual variable (Anghelache et al., 2010).

In order to estimate the parameters of the linear model of regression, we can use the least squares method. In this case, the values of the resulting characteristic are estimated by using the relation:

$$\hat{y}_i = \hat{b} + \hat{a}x_i$$

where â and $\hat{b}$ are the estimators of the parameters of the regression line.

The real values of the resulting characteristic are equal to the estimation we get with the help of the regression model, corrected with the residual error, namely:

$$y_i = \hat{y}_i + e_i$$

The estimation of the parameters is based on the condition that the sum of the squares of the differences between the real value and the estimated one through the regression model is a minimum one (Anghelache, 2013)*:*

$$F(\hat{\alpha},\hat{\beta}) = \min\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 = \min\sum_{t=1}^{n}(y_t - \hat{\alpha} - \hat{\beta}\cdot x_t)^2$$

$$\min_{\hat{a},\hat{b}} \phi\left(\hat{a},\hat{b}\right) = \min_{\hat{a},\hat{b}} \sum_{i}^{n} e_i^2 = \min_{\hat{a},\hat{b}} \sum_{i}^{n}\left(y\hat{i} - \hat{b} - \hat{a}x_i\right)^2 .$$

The optimum conditions lead to the following system of equations:

$$\begin{cases} \dfrac{\partial\left(\hat{a},\hat{b}\right)}{\partial\left(\hat{b}\right)} = -\sum_{i} 2\left(y\hat{i} - \hat{b} - \hat{a}x_i\right) = 0 \\ \dfrac{\partial\left(\hat{a},\hat{b}\right)}{\partial\left(\hat{a}\right)} = -2\sum_{i}\left(y\hat{i} - \hat{b} - \hat{a}x_i\right)\cdot x_i = 0 \end{cases}$$

In order to establish the two estimators the linear system of equations is solved:

$$\begin{cases} n\hat{b} + \hat{a}\left(\sum_{i} x_i\right) = \sum_{i} y_i \\ \left(\sum_{i} x_i\right)\cdot\hat{b} + \hat{a}\left(\sum_{i} x_i\right) = \sum_{i} y_i \end{cases} .$$

The computing relations of the two, $\hat{a}$ and $\hat{b}$, are coming out of the settlement of the linear system of equations:

$$\hat{a} = \frac{\begin{vmatrix} n & \sum y_t \\ \sum x_t & \sum (x_t \cdot y_t) \end{vmatrix}}{\begin{vmatrix} n & \sum x_t \\ \sum x_t & \sum x_t^2 \end{vmatrix}}$$

The coefficient of the regression line slope is obtained out of the relation (Anghelache, 2013):

$$\hat{a} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}.$$

$$w_i = \frac{x_i - x}{\sum_{i=1}^{n} (xi - \bar{x})^2}.$$

The formula for calculating the estimator of the loose term of the regression line is set up by solving the equations system or by considering the fact that the regression line is passing through the centre of the points cloud, namely:

$$\hat{b} = \bar{y_t} - \hat{\beta} \cdot \bar{x_t}$$

$$\hat{b} + \hat{a}\bar{x} = \bar{y}$$

The method of the least squares holds certain drawbacks as to its utilization in estimating the parameters of the linear regression model, as follows:

- It does not provide acceptable outcomes if the formulated hypotheses are not satisfied:
- If noting by $\hat{a}^n, \hat{b}^n$ the estimators set up on the basis of the series $(x_i, y_i), i = \overline{1, n}$ and by $\hat{a}^{n+1}, \hat{b}^{n+1}$ those evaluated through the values series (x_i, y_i), $i = \overline{1, n+1}$, it is resulting that between the two pairs of estimators there is not a simple relation of recurrence:
- The estimators are affected in case that the data series is showing major changes, in the form of breaks of level.

In order to facilitate the utilization of the regression models for economic and financial analyses, specific statistical tests have been created which afterwards have been included in specialized informatics packages. For instance, the Eviews programme which allows getting econometric models of quality with a minimum effort from the side of the user of such a software solution (Andrei, 2008).

4. The application of the regression model on the capital market in Romania

In order to take advantage of the submitted theoretical aspects, we used the regression econometric model meant to evaluate the relation between the value recorded by the BET index and the value of the stock exchange capitalization.

The analysis of the Bucharest Stock Exchange can be achieved through the support of a complex system of indices. Among them, a significant role as to generate a true image on the performances of the major stock exchange institution from Romania is played by the BET index *(Bucharest Exchange Trading)* (Anghel, 2008). The BET index has been launched on September 19th, 1997 and represents the reference index of the capital market in our country. BET is a prices index weighted by the free-float capitalization (free-float of a company included in the structure of the BET index represents the number of shares issued in circulation which are available for transactions to the public), which is reflecting the overall tendency of the equities issued by the first ten companies, classified depending on their liquidity (Anghelache, 2009). Thus, the prices recorded during every session of transactions are reported at the corresponding prices recorded during the transactions session of reference, by applying the following calculation relation in this respect (Bucharest Stock Exchange – BET index manual):

$$BET_T = BET_{T-1} \cdot \frac{\sum_{i=1}^{N} p_{i,T} \cdot q_{i,T} \cdot Ff_i \cdot R_i \cdot c_{i,T}}{\sum_{i=1}^{N} p_{i,T-1} \cdot q_{i,T} \cdot Ff_i \cdot R_i \cdot c_{i,T-1}}$$

where:

BET_T – value of BET index at the current moment, T;

BET_{T-1} – value of BET index at the previous moment, T-1;

$p_{i,T}$ – price corresponding to the equities of the company „i” by the current moment T;

$p_{i,T-1}$ – closing price corresponding to the equities of the company „i” by the moment T-1;

$q_{i,T}$ – number of equities by the current moment T;

Ff_i – free float factor corresponding to the company „i” from the index being calculated with two digits, bearing one of the following values: {0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 1};

R_i – representation factor at maximum 20% of the weight of the components of the index cost, corresponding to the equities of the company „i”; it is calculates with three digits and belongs to (0,1];

$c_{i,T}$ – correction factor of the price corresponding to the equities of the company; „i" at the moment T, in the days of operational revision; it is calculated with six digits;
$c_{i,T-1}$ – correction factor of the price corresponding to the equities of the company; „i" at the moment T-1, in the days of operational revision; it is calculated with six digits;
N – number of companies included in the index.

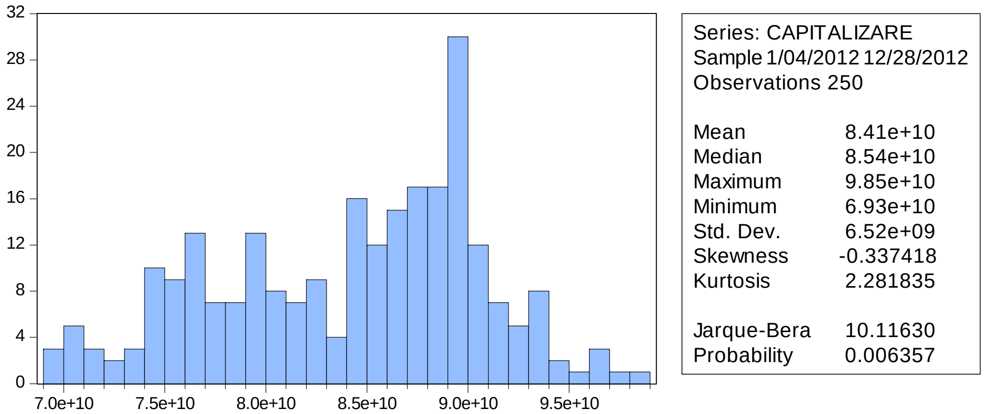
One of the factors of significant influence on the value of the BET index is given by the stock exchange capitalization. This one represents an indicator of one stock exchange potential reflecting the market value of the listed companies, which can be established as a sum of the products between the number of shares issued by each of the listed companies and the corresponding market price (Anghel, 2013).

In order to analyse the existing relation between the values recorded by the Bucharest Exchange Trading index and the stock exchange capitalization value, both statistical-mathematical models and, mainly, regression econometric models can be used. In this respect, we submitted to the analysis a series of data (daily frequency) concerning the BET index value and the stock exchange capitalization for the year 2012. (366 calendar days/ 250 actual transacting sessions – the data being used are daily surveys over a time horizon between January 1st 2012 and December 31st 2012, respectively 250 surveys, excepting week-end days and legal holidays).

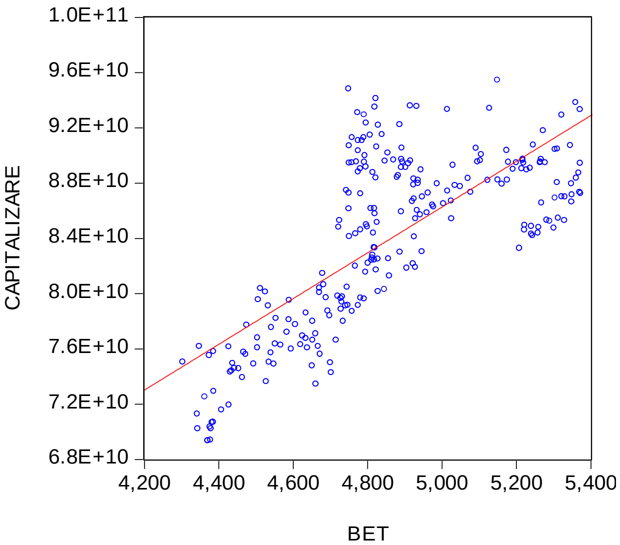
As a first stage of the analysis of the considered series of data, we used the informatics package Eviews in order to generate a series of statistical tasts typical for each one of the two indicators taken into account: the BET index value and the stock exchange capitalization value.

The statistical tests applied to the series of data referring to the BET index evolution during the analysed period evidenced the fact that average value of this index counts for 4,857.236 lei. Meantime, it is to notice that the distribution of this series of data do not correspond exactly to the normal distribution, a situation which, however, should be not considered as very grave as it is typical to most of the data series in the financial domain.

In the case of the analysis of the series of data referring to the evolution of the stock exchange capitalization, it is to state out that this indicator is recording an average value of 84,146.510 million lei, the distribution of the recorded values being also different from the normal distribution.



In order to set up the type of regression model to be utilized for establishing the dependence between the BET index value and the value of the stock exchange capitalization, the two series of data have been represented as a points graph on which the regression line has been also plotted.



Out of the analysis of the above graph, we state out that there is connection of linear type between the two variables. This dependence can be synthetically showed up by means of an econometric model typical for the simple linear regression, namely:

BET = $\alpha + \beta$ * CAPITALIZATIO + ε

where:

BET – BET index value (the dependent variable);

CAP – value of Bucharest Stock Exchange capitalization (the independent variable);

α, β – the parameters of the linear regression model;
ε – the residual value of the regression model.

Based on the previous mentioned elements, with the help of the informatics package, the parameters of the previously submitted regression model have been estimated, through the method of the least squares.

Dependent Variable: BET
Method: Least Squares
Date: 11/20/13 Time: 18:51
Sample (adjusted): 1/04/2012 12/18/2012
Included observations: 245 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2141.023	162.4726	13.17775	0.0000
CAPITALIZARE	3.24E-08	1.93E-09	16.75345	0.0000

R-squared	0.535975	Mean dependent var	4855.378
Adjusted R-squared	0.534065	S.D. dependent var	278.6151
S.E. of regression	190.1811	Akaike info criterion	13.34196
Sum squared resid	8789032.	Schwarz criterion	13.37054
Log likelihood	-1632.390	Hannan-Quinn criter.	13.35347
F-statistic	280.6781	Durbin-Watson stat	1.982817
Prob(F-statistic)	0.000000		

Estimation Command:
=========================
LS BET C CAPITALIZARE

Estimation Equation:
=========================
BET = C(1) + C(2)*CAPITALIZARE

Substituted Coefficients:
=========================
BET = 2141.02316511 + 3.23613403506e-08*CAPITALIZARE

Based on the previously submitted data, the values of the two parameters of the regression model, α and β, are set up as follows:

$\alpha = 2141{,}123$

$\beta = 3{,}24 \cdot 10^{-8}$

The regression model defining the relation between the BET index and the value of the stock exchange capitalization is transcribed as follows:

$$\text{BET} = 2141{,}123 + 0{,}0000000324 \cdot \text{CAPITALIZARE}$$

The value being recorded by Prob (0.0000) evidences the fact that the variable is significant from statistical point of view. The connection between the BET index value and that of the stock exchange capitalization is a direct one which can be explained as follows: by an increase of 1% of the capitalization value, the BET index value is recording an increase 0.00000324% only. Both R-squared (R^2) and the Adjusted R-squared (R^2 adjusted) show the extent to which the dependent variable is explained by the independent variable. The recorded values are comprised between 0 and 1. The difference between the two coefficients consists of the fact that Adjusted R-squared is presenting, from statistical point of view, a more significant result since it is sanctioning the inclusion of independent variable of a low relevance for explaining the dependent variable. In the situation here submitted, the value recorded by R-squared shows that in a proportion of 53.59% of the BET index value is explained by the value of the capitalization, the difference up to 100% representing the influence of other factors not included into the present mode (a rather low influence). A comparison made between the model set up for the values recorded by the two indicators during the year 2012 and those for the years 2010, respectively 2007 (which made the object of previous analyses) evidenced the fact that the significance of the stock exchange capitalization for setting up the BET index value diminished significantly. This situation can be explained by going deeply into the effects of the economic and financial crisis which led to the occurrence of rapid and difficult to anticipate alterations of the factors which are influencing the activity of the main companies quoted on the capital market from our country and, implicitly, the evolution of the major specific indicators of the Bucharest Stock Exchange. The connection between the BET index value and that of the stock exchange capitalization can be explained as follows: for the increase of 1% of the capitalization value, the BET index value is recording an increase of 0.00000324% only.

The validity of the analysed regression model can be studied by means of the tests implemented in the informatics package Eviews. Thus, based on the values of the test R-squared and of the test F-statistic (which value of 280.6781 is exceeding the reference level of the tables) we can state out that the model describing the relation between the stock exchange capitalization and the value of the main stock exchange index of the Romanian capital market is a correct one.

5. Conclusions

The simple linear regression model submitted above is confirming, on the basis of the real data recorded on the Romanian capital market, the allegations from the specialized literature according to which this type of econometric model can be successfully utilized in order to define and systemize in a mathematical form the existing relations between the various financial indicators specific to the stock exchange system. The regression models allow the identification of certain functional dependences between the various components of the capital market, which provides the real possibility to forecast the phenomena submitted to the analysis over a large established time horizon. The simple linear regression model can be utilized with good results in other analyses concerning the evolution of the financial market.

References

Andrei, T., Bourbonnais, R. (2008). *Econometrics*, Economica Publishing House, Bucharest

Andrei, T., Stancu, S., Iacob, A.I. (2008). *Introduction in Econometrics using Eviews,* Economica Publishing House, Bucharest

Anghel, M.G. (2008). "Using the Regression Model in the Analysis of the Capital Market", *Romanian Statistical Review*, Supplement „Romania in the Process of European Integration", issue 12

Anghel, M.G. (2013). *Models for the Management and Analysis of Portfolios*, Economica Publishing House, Bucharest

Anghelache C. (2013). *Elements of Theoretical Econometrics*, Artifex Publishing House, Bucharest

Anghelache C., Anghelache G.V. (2009). *Utilization of the econometric models in the analysis of the assets rates*, Metalurgia International Vol. XIV, Special Issue 9, F.M.R. Scientific

Anghelache, C., Marin, D. and Barbu, C.M. (2010). *Estimation of the regression functions transformation*, Metalurgia International, Vol. XV, issue. 1

Anghelache, G.V. (2009). *Capital Market in an European Context*, Economica Publishing House, Bucharest

Dougherty, C. (2008). *Introduction to econometrics. Fourth edition*, Oxford University Press

Guijarati, D. (2005). *Basic Econometrics*, The McGraw-Hill Companies

Voineagu, V., Ţiţan, E. (2007), *Theory and Practice in Econometrics*, Meteor Press Publishing House, Bucharest

Wooldrige, J. (2006). *Introductory econometrics. A modern approach,* 2nd edition, MIT Press

www.bvb.ro (Bucharest Stock Exchange)