The Regression Model for the Statistical Analysis of Albanian Economy

Bederiana Shyti[*1], Dhurata Valera[*2]
[#] *Department of Mathematics, FNS, University of Elbasan "Aleksandër Xhuvani", Albania*

Abstract—*In practice, as well as in economic theory, fulfilling strategic, managerial and operational decisions is the ultimate goal of any company. Economic performance should be assessed on the basis of the effectiveness and efficiency of resource use. The purpose of this paper is to study economic phenomena through a statistical standpoint. Our motive is to emphasize the validity of regression analysis method on economic performance. Multiple regression analysis (MRA) will be applied to demonstrate the economic trend of the transition economy of Albania. All paper will be a case study for a limited liability company in the production industry. There are several processes from raw material to final product. We have used multiple linear regression model (MLRM) and three types of statistical technique for statistical analysis (SA): regression analysis (RA), correlation analysis (CA),analysis of variance (Two-way ANOVA) to represent a mathematical-statistical model for analyzing the company's economic performance depending on the indicators used. The main purpose of SA used in this paper is to show the type of relationship between variables presented by a multiple linear regression (MLR) and to explain it as a mathe matical-statistical model (MSM - in our case with one dependent variable and two independent variables). The Statistical Package for Social Sciences, version 17.5 (SPSS) will be used to analyze and evaluate statistical outcomes.*

Keywords—*Multiple linear regression, economic performance, statistical analysis, statistical correlation method.*

I. INTRODUCTION

In the continuous progress of an enterprise strategic, managerial and operational management is very important. In order to evaluate the extant at which the enterprise goals are achieved it is imperative to define an integrated system of indicators that can tell us at any time if our business is going in right direction or not.

Our main reason for compelling the financial statements is to provide a better decision taking, diagnosis of the current situation of a business and future perspective.

Plain cost and profit are not the main aim of the enterprise, we want to study how qualitative the earning is, what is the sale revenue, which one of the indicators best reflects the economic progress of the enterprise.

For this purpose, the *multiple regression analysis* (MRA) will be used to demonstrate the economic trend of transitory economy of Albania. To have a better understanding, we will apply the multiple regression model (MRM) in a case study. Our case will be a company with foreign capital in the construction and trade industry [1], [5], [10], [12], [13], [15], [19].

II. RESEARCH METHOD

Multiple regression analysis (MRA) is a statistical method that correlates the behavior or variation of a number of factors (independent variables), in order to ascertain their individual and combined impact upon a single factor (called the dependent variable) [13].

In this paper, we have used multiple linear regression model (MLRM) and three types of *statistical technique* for statistical analysis (SA):

- regression analysis (RA);
- correlation analysis (CA);
- analysis of variance (Two-way ANOVA).

A two-way ANOVA refers to an ANOVA using two independent variables and can be used to examine the interaction between the two independent variables. 2-way ANOVAs are also called factorial ANOVAs [11], [13], [14], [15].

The objective of SA is to control existing of the relationship between the dependent and the independent variables and if so, intends to use the existing information for dependent and independent variables, in order to improve the correctness of prediction of the values of dependent variable.

So, by using the multiple regression model (MRM), the statistical correlation method (SCM) and the statistical analysis (SA), our aim is to achieve:

- the valuation of the connection between the variables - two independent and one dependent variable;
- to prove the correlation of the relationship between these two variables.

Analysis of the multiple linear regression is used to valuate a variable (dependent) in dependence to the values of independent variables.

The mathematical-statistical model (MSM)

The multiple regression equation (MRE) as a mathematical model, describes the average relationship between these variables, and this relationship is used to predict or control the dependent variable.

Let's make the mathematical explanation of this problem of multiple linear regression.

Suppose, that our variable $\boldsymbol{y}$ depends from $\boldsymbol{n}$ variables "non-statistically independent" $x_1, x_2, \dots, x_n$, that are called independent variables $\left(\boldsymbol{y} = f(x)\right)$ [1]-[5], [9], [12], [15].

The problem of the multiple regression is considered to be a mathematical *linear* model where is showed the function of these variables in the following form $\boldsymbol{y}$:

$$(\text{MLRF})\ \boldsymbol{y} = f(x_1, x_2, \dots, x_n) + \varepsilon, \qquad (1)$$

where $\varepsilon - random\ variable\ error\ or\ residue$ has the normal distribution $N(0, \sigma^2)$ and expresses the deviations of expected observed values of y, calculated through the regression equation. The dependent variable $\boldsymbol{y}$ is also a casual unit normally distributed.

In the mathematical-statistical model of multiple linear regression (MLRM):

- all the independent variables take part *in the same time* in the statistical model;
- the value of R and R^2 determine *the strength of the correlations* between the independent variables and the dependent variable.

The Statistical Instrument Test (SIT):

- *Fisher test* is used to show if this connection determined by the election, may or no generalize for whole the population;
- *t-test* is used to evaluate the individual link between every independent variable and the dependent variable.

The Statistical Package Analyses (SPA) for developing our MLRM:

- Statistical Package for the Social Sciences, version 17.5 (SPSS) will be used as instruments to analyze and evaluate the statistical results [2].

Determining and testing the correlation ratio:

One of major problems that appears during the creation of multiple regression model is *the definition and test the correlation* ratio between variables. For this purpose, in multiple regression is defined a unit that is called the coefficient of multiple correlation (R) that defines the amount of linear correlation between the dependent variable $\boldsymbol{y}$ and the *n*-independent variables $(x_1, x_2, \dots, x_n)$.

This coefficient takes defined values. If the distribution values are near the medium, the linear dependence is perfect so, $\boldsymbol{y}$ can be expressed exactly as a linear combination of the independent variables $x_1, x_2, \dots, x_n$ and in this case the statistical model of linear regression is too strong.

This coefficient is depended from the degree of measurement of random variables that are part of the model. The bigger is the number of variables, the bigger will be the value of R^2, because the expected error of prediction will be smaller.

Beside the values of $R - squared$, in consideration should be taken also the values of $adjusted - \mathrm{R}^2$ that define the part of variance of $\boldsymbol{y}$ that is "explained" from relationship of $x_1, x_2, \dots, x_n$.

Even, if we are using reliable data, a large adjusted value of R^2 does not necessarily means that there is a strong causal relationship, it follows that, on the basis of regression, it is not possible to determine causation [1], [5], [12].

III.STATISTICAL ANALYSIS USING MULTIPLE LINEAR REGRESSION - CASE STUDY

Method of multiple linear regression analysis (MLRAM) is used to show the connection that exists among variables and to make concrete this relation.

For this purpose, we are considering a *case study*, based on some data provided by *opendata.al* for an economic organization in Albania with foreign capital in the construction and trade industry in 10-fiscal years. All the data are using in following table and gives the incomes in dependence of the number of employees and the price of the product [6], [7], [8], [13], [15].

In this regard, will be taken into consideration the following financial ratios derived from the organization studied during years 2005-2015: the incomes, the number of employees and the product price, as are describing in the table below [6].

IV.TABLE I

DATA OF AN ALBANIAN ECONOMIC ORGANIZATION

Year	Incomes (/000000 ALL)	No. of employees	Product Price
2015	56	201	15,3
2014	79	255	17,5
2013	92	233	21,7
2012	94	235	22,5
2011	82	215	19,9
2010	89	232	21,4
2009	75	234	18,6
2008	81	214	18,6
2007	78	269	18,7
2006	71	272	15,7

Thus, it was established as:

- dependent variable: the total incomes of the financial indicator;
- independent variables: the number of employees and the product price.

The mathematical-statistical Model **(MSM):**

SM - *the statistical model* of our case study is a multiple linear regression model (MLRM) between a dependent variable (in our case incomes) and independent variable (in our case the number of employees and the price of the product).

T*he mathematical explanation of our model* is used to explain the variation of the dependent variable (the total revenue) by its covariance with independent variables (the number of employees and the product price). Our model is mathematical-statistical one.

So, it is defined by the multiple linear regression equation (MLRE) bellow:

$$(\text{MLRE})\ \ y = a_0 + a_1 x_1 + a_2 x_2 + \varepsilon. \qquad (2)$$

Now, we are giving the *mathematical explanation* of selecting variables and parameters in MLRE for our problem:

- y = the explained variable – the total income of the financial indicator;
- x_1 = the explanatory variable 1 – the number of employees;
- x_2 = the explanatory variable 2 – the product price;
- a_0 = constant (free term of equation);
- a_1, a_2 = model parameters (coefficients of independents variables)
- ε = specification error (error term of equations, unknown);
- n = total number of observations – 10 fiscal years.

We mark that the model parameters give the change values of y that corresponds with one unit change values of x. Also, ε - the value of error is unknown and shows the difference between the true and the specified model [1], [5], [12], [17], [19].

For *the statistical analyses* (SA) and developing our MLRM, we are using the specific Statistical Package for the Social Sciences, version 17.5 (SPSS instruments) to analyze and evaluate the statistical results [2]. So, the tables below are taken from the results registered by using our specific SPSS and give the links between the variables in our case study.

TABLE 2.
MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,960[a]	,922	,900	3,52156

a. Predictors: (Constant), Product_Price, No_of_employees

The table 2 indicates that the correlation report $R - squared = 92.2\%$ tending towards 1 and the variance of the incomes can be explained by our regression model.

So, in other words, we have a very good multiple linear regression model that explains the economic phenomenon.

TABLE 3.
TWO-WAY ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1025,290	2	512,645	41,338	,000[a]
	Residual	86,810	7	12,401		
	Total	1112,100	9			

a. Predictors: (Constant), Product_Price, No_of_employees

b. Dependent Variable: Incomes_000000_ALL

From the Table 3 of two-way ANOVA, specially by the $F - Fisher\ theoretical\ value$ and the $\propto - significance\ level$, we can see that our linear regression model is significant because $F_{theorical} = F_{\propto}(k-1, n-k)$ with probability $\propto = 5\%$ and $(two, seven) - degree$ of freedom from FINV function:

$$F_{theoretical} = F_{0,05}(2,7)$$

$$F_{0,05}(2,7) = 19,371.$$

Further, our $F - Fisher\ value$ from two-way ANOVA table is:

$$F^* = 109,121 \text{ and } F^* > F_{theoretical}.$$

So, we can say that our MLRM is well built [1, 5, 12].

TABLE 4.
COEFFICIENTS[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-29,528	15,911		-1,856	,106
	No_of_employees	,111	,051	,233	3,193	,064
	Product_Price	4,375	,485	,958	9,016	,000

a. Dependent Variable: Incomes_000000_ALL

From the table 4 of coefficients, the multiple linear regression model (MLRM) describes the *relationship between these indicators* of the economy subject [18], [20].

In our case, the $t_{theoretical}$ with the probability 5% and nine degree of freedom, is $t_{theoretical} = 2,262$, that is less then each calculated values in interval $[3,193; 9,016]$.

So, the number of employees and the product price explain very well our $y - incomes$. The "sig." values show us that the variables are making a statistically significant contribution to the power of the model [1], [5], [12], [21]-[24].

The figures bellow are representing the graphics of the direct correlations between each the independent variables and the dependent variable.

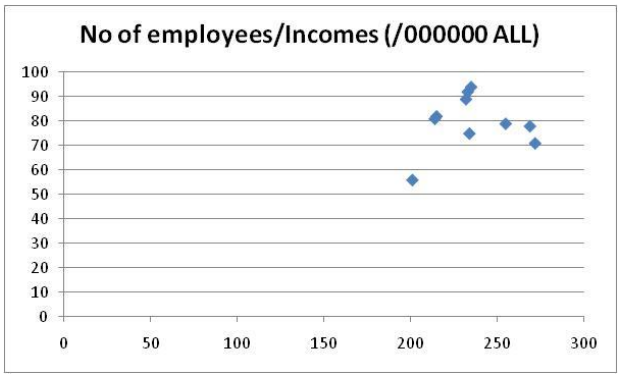

Fig.1 The relation between the number of employees and the incomes

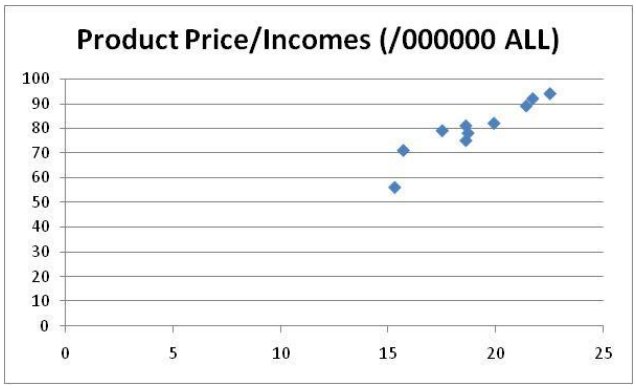

Fig. 2 The relation between the product price and the incomes

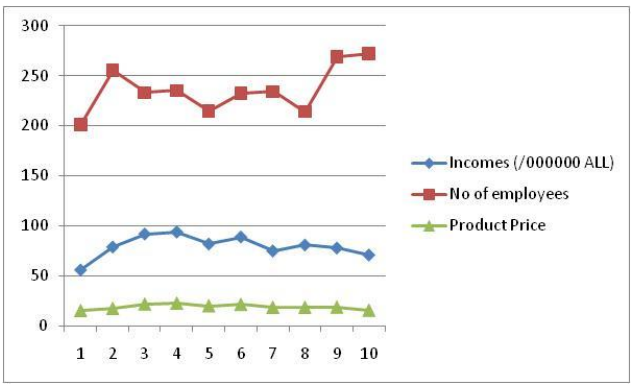

Fig. 3 The relation between the number of employees, the product price and the incomes

The mathematic-statistical model (MSM) *of the multiple linear regression* of our case study in this research, may be given by the following MLRE:

$$(\text{MLRE}) \quad y = -29{,}528 + 0{,}111\, x_1 + 4{,}375\, x_2. \qquad (3)$$

As it can be observed by our MLRE, *the number of employees* (x_1) *and the product price* (x_2) *are the factors that significantly influence in the economy trends.*

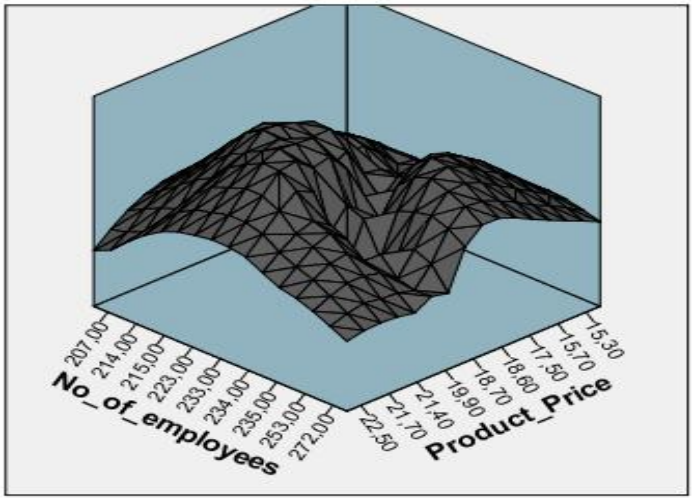

Fig. 4 The evolution of the incomes for the Albanian economic organization in 10-fiscal years (2005-2015)

Finely, the more detailed study of MLRE, SCM, ST results specifically that from:

- MLRE – the *negative value* of the constant a_0 $(a_0 = -29{,}528)$ represents *the fixed costs* of a business.
- SCM – the *correlation* between the number of employees and the price of the product as independent variables is *positive* with the dependent variable. They are both significant from a statistical point of view for the model MLR. In other words, the significance values lower then 0.05 $(\propto = 5\%)$ for each calculated coefficients (valued from ST), suggest that there is a significant correlation between the analyzed variables.
- MLRE – the value of x_1, x_2 *influence in the economy trends.*
- ST– the value of $\mathrm{R-squared}$ and $\mathrm{adjusted-\ R-squared}$ define the correctness of the created *econometric model*.

V. CONCLUSIONS

Albanian companies operate in a dynamic environment that are changing in cooperation and compliance with the influence of other factors or not. Because of this dynamism, it is mandatory to have a good strategy in disposal, especially during financial crisis.

The complexity of making decisions during statistical analysis of albanian economy increases the need for the statistical correlation method (SCM) to explain the relationship between the variables.

Results (outputs of the tasks) and the behaviors (work process entries) are important in the process of assessing and measuring of performance. In this economic situation, the product incomes are the results of the conjugation of many influencing variables, but not all the defined ratios have the same importance, the action of some are compensating the others.

Our case study model (MLRM) is a representative one. The number of employees and the price of the product have a big impact on the level of sales revenue.

REFERENCES

[1] F. Fechete, A. Nedelcu, "Analysis of the economic performance of a organization using multiple regression", International *Conference of scientific paper, AFASES May 2014, Brasov*, p. 22-24.
[2] J. Lani, (2010) The Multiple Linear Regression Analysis. The SPSS website. [Online]. Available: http://www.statisticssolutions.com.
[3] C. Anghelache, M. G.R. Pagliacci, E. Bugudui, D. C. Christian, B. Dragomir, "Multiple Regression Used in Macro-economic Analysis", Revista Română de Statistică – Supliment Trim II/2013.
[4] C. Anghelache, A. Manole,, M.G. Anghel, "Analysis of final consumption and gross investment influence on GDP – multiple linear regression model*", Theoretical and Applied Economics Volume* XXII (2015), No. 3(604), Autumn, pp. 137-142.
[5] C. Anghelache, A. Manole,, M.G. Anghel, M.Popovici, "Multiple Regressions Used in Analysis of Private Consumption and Public Final Consumption Evolution", *International Journal of Academic Research in Accounting, Finance and Management Sciences* Vol. 5, No.4, October 2015, pp. 69–73 E-ISSN: 2225-8329, P-ISSN: 2308-0337. ©2015 HRMARS www.hrmars.com (IJARA, AFMS).
[6] http://open.data.al
[7] http://www.monitor.al

[8] http://www.financa.gov.al
[9] http://web.stanford.edu/~hastie/Papers/ESLII.pdf
[10] http://www.lexjansen.com/nesug/nesug90/NESUG90009.pdf
[11] http://www.statisticssolutions.com/manova-analysis-anova/
[12] L. Buşe, M. Ganea, D. Cîrciumaru, "Using linear regression in the analysis of financial-economic performances", Craiova, Romania, Available: http://feaa.ucv.ro/AUCSSE/0038v2-002.pdf.
[13] R. A. Bell, O. C. Anderson, M. V. Sanders, "Real Estate Damages", *2nd Edition. Chicago, Illinois: Appraisal Institute*, 2008: 31-37.
[14] Spurr, William A. & Bonini, Charles P., Statistical Analysis for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc., 1973: 495.
[15] Isakson, Hans R., "Using Multiple Regression Analysis in Real Estate Appraisal." The Appraisal Journal (October, 2001): 424-430
[16] Spurr, William A. & Bonini, Charles P., Statistical Analysis for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc., 1973: 503.
[17] Isakson, Hans R., "Using Multiple Regression Analysis in Real Estate Appraisal." The Appraisal Journal (October, 2001): 424-430
[18] C. Anghelache, M. G. Anghel,., L. Prodan, C. Sacală, C., M. Popovici, (2014). "Multiple Linear Regression Model Used in Economic Analyses. Romanian Statistical" *Review Supplement*, Volume (Year): 62 (2014), Issue (Month): 10 (Octomber), 120-127.
[19] C. Anghelache, Economic status on the recovery path. Bucharest: *Economica Publishing House*, Romania 2014.
[20] C. Anghelache, (coord.) et al. (2014). "Statistical-econometric Models used to study the Macroeconomic Correlations", *Romanian Statistical Review* (Supplement), December 2014.
[21] C. Anghelache, (2008) "Treaty on economic and theoretical statistics", Bucharest: *Economica Publishing House*.
[22] R. Bachman, (2011) "Public Consumption over the Business Cycle", *Society for Economic Dynamics* in 2011 Meeting Papers with number 701.
[23] C. Gourieroux, J. Jasiak, (2001) "Financial econometrics: problems, models and methods", *Princeton University Press*, Princeton.
[24] M. G.R. Pagliacci, G. V. Anghelache, I. M. Pocan, R. T. Marinescu, A. Manole "Multiple Regression – Method of Financial Performance Evaluation", ART ECO – Review of Economic Studies and Research, Editura Artifex, Vol. 2/No.4/2011, pp. 3-9.