THE CORRELATION BETWEEN GDP/ CAPITA AND EMPLOYMENT RATE OF PEOPLE- ECONOMETRIC MODEL ANALYSIS

PhD Candidate Ligia PRODAN
Academy of Economic Studies, Bucharest

Abstract

It is presented the evolution of Employment rate of work age population (aged 15-64) by gender and area, as well as statistical correlation between Gross Domestic Product/ capita and Employment rate of population in Romania (during 2002-2011)

Key words: e*mployment rate, gross domestic product, per capita, coefficient of determination, correlation ratio, correlation coefficient.*

***

There is a direct connection and linear type: y = a + bx + ε between Gross Domestic Product / capita and Employment rate of population – as also results from the graphs inside the article. By the calculations performed using linear regression model function we get parameters a = - 633,663.4 and b = 11,150.08. The adjustable single regression is $\hat{y} = -633.663{,}4 + 11.150{,}08 \bullet x$.

Determination report confirms that the employment rate is a significant factor ($R^2 > 50\%$) for GDP growth/capita, it's influence on the GDP variation is 60,95%. Correlation coefficient, $r_{y/x}$= 0,7806 indicates a strong link between the two variables. To verify the significance of the linear correlation the t test (Student) coefficient is applied witch verify the hypothesis of the significance of correlation, so $r_{y/x}$ is statistically significant and analysis model is correctly specified.

Regression and correlation method emerged as a result of intensive research in biology that were then extended to socio- economic phenomena. Method depends on the specific phenomena being studied and the amount of data used. To determinate the trend of a statistical link we use appropiate analytical function estimating equation that expresses the form of the relationship between the characteristic factor and the result. This function is called regresion function and its graphic representation is line/ curve regression. Regression function indicates how the characteristic result of „Y" changes in conditions where only characteristics of values „X" changes, other factors being considered with constant action in all cases subject to observation. Gra-

fical representation of distribution series is required for reliable regression function choosing. So, we can appreciate visually the trend of variables correlations. So, the goal of regression is to identify the mathematical relationship that exist between two variables. To determine the intensity of the relationship between two variables (how good is the data set using regression function) the correlation between them is established. The correlation indicates the intensity of the relationship between variables **by measuring the scattering data recorded** around the regression line. The closer a correlation between the variables analyzed, the more we can predict with greater confidence different sizes based on the regression equation.

In 2011, the employed population had 9,138,000 people (down from the previous year by 102,000 persons).

General employment rate as a ratio between population and total population.

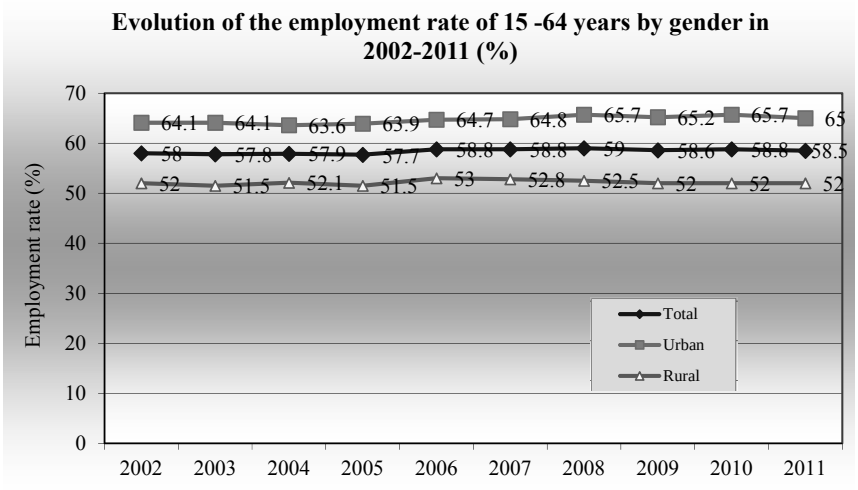
$$RO_{general} = \frac{PO}{PT} \times 100 = \frac{9{,}138{,}000 \text{ people}}{21{,}354{,}000 \text{ people}} \times 100 = 42{,}79\%$$

where: RO = employment rate
PT = total population
PO = employment population.

Evolution of the employment rate of working age population (15- 64 years) by sex (%) in 2002-2011

Year	Employment rate (%)		
	Total	Male	Female
2002	58,0	64,1	52,0
2003	57,8	64,1	51,5
2004	57,9	63,6	52,1
2005	57,7	63,9	51,5
2006	58,8	64,7	53,0
2007	58,8	64,8	52,8
2008	59,0	65,7	52,5
2009	58,6	65,2	52,0
2010	58,8	65,7	52,0
2011	58,5	65,0	52,0

Source: [1]

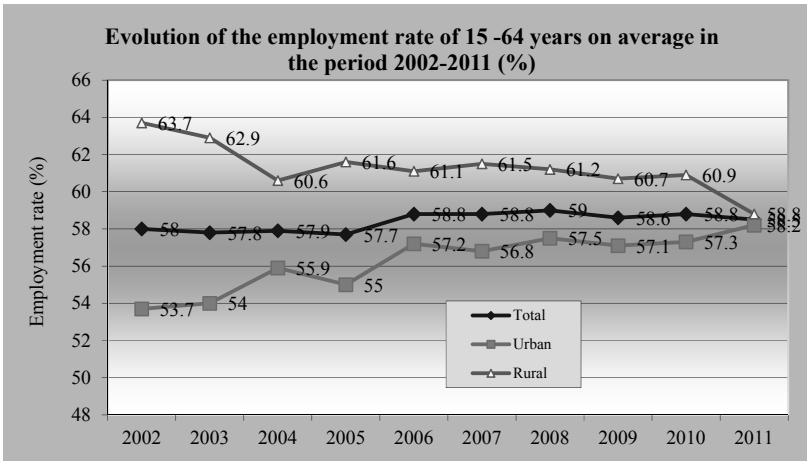


The employment rate of working age population (15- 64 years) registered in 2011 levels of 58.5%, with higher values for the employed male 65,0% compared to 52,0% for the females and those in rural areas 58.8% to 58.2% for those in urban areas.

Evolution of the employment rate of working age population (15-64 years) by area (%)

Year	Employment rate (%)		
	Total	Urban	Rural
2002	58,0	53,7	63,7
2003	57,8	54,0	62,9
2004	57,9	55,9	60,6
2005	57,7	55,0	61,6
2006	58,8	57,2	61,1
2007	58,8	56,8	61,5
2008	59,0	57,5	61,2
2009	58,6	57,1	60,7
2010	58,8	57,3	60,9
2011	58,5	58,2	58,8

Source: [1]



The intensity degree of the correlation between phenomena can be determined by dividing the dispersion formed on the base of the registered factor ($\sigma^2_{y/x}$) on total dispersion factor recorded (σ^2_y).

$$R^2_{y/x} = \frac{\sigma^2_{y/x}}{\sigma^2_y};\ R^2_{y/x} = \text{coefficient of determination.}$$

In other words, the coefficient of determination estimates the proportion in which the variable factor X influences the resultative variation Y. The degree of indetermination or residual influence of residual dispersion is calculated by dividing ($\sigma^2_{y/r}$) to the total variance (σ^2_y).

$$K^2_{y/r} = \frac{\sigma^2_{y/r}}{\sigma^2_y};\ K^2_{y/r} = \text{coefficient of indetermination.}$$

Clearly, the sum of the two coefficients is 1.

$R^2_{y/x} + K^2_{y/r} = 1$

From the relation, coefficient of determination $R^2_{y/x}$ can be expressed:

$R^2_{y/x} = 1 - K^2_{y/r}$

If you extract the square root of the coefficient of determination the correlation ratio $R_{y/x}$ is obtained. Correlation ratio takes values in the interval $\overline{0,1}$. As the correlation ratio takes values closer to 1, the relationship between the two variables is more intense; the trend towards zero of the correlation ratio indicates the independence of the variables.

The correlation ratio will mark correlation parameter „b" of the regression equation:

- ✓ if b>0, $R_{y/x}$ will be positive;
- ✓ if b<0, $R_{y/x}$ will be negative;
- ✓ if b=0, $R_{y/x} = 0$.

For the linear correlation, correlation ratio is equal to the correlation coefficient.

$$r_{y/x} = R_{y/x}$$

Simple linear correlation coeffiecient takes value in the interval $\overline{-1,+1}$.

If $r_{y/x}$ takes values in the interval $\overline{-1,0}$, the relationship between the two variables is opposite and is even stronger as close to -1. If $r_{y/x}$ takes values in the interval $0,+1$, the relationship between the two variables is direct and is even stronger as one approaches +1. If $r_{y/x}$ approaches zero, the phenomena are independent.

Employment rate and GDP/capita in 2002-2011

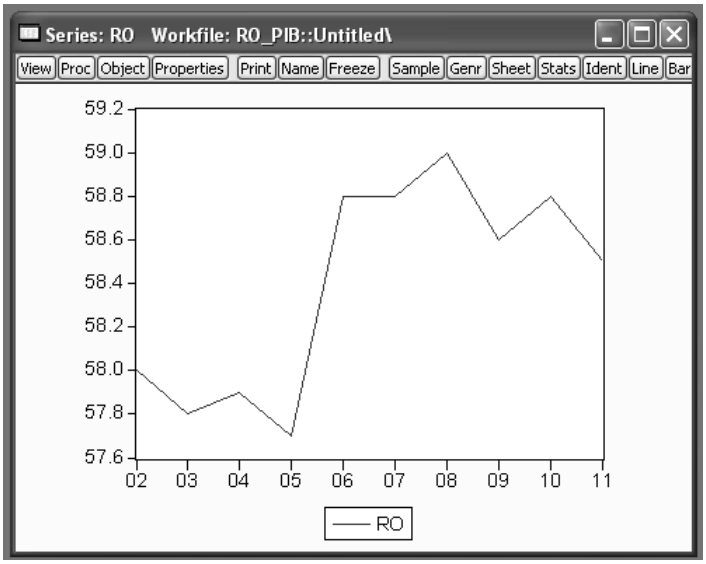
Year	Employment rate (%)	GDP/capita (current prices)
2002	58,0	6974,9
2003	57,8	9084
2004	57,9	11413,5
2005	57,7	13362,8
2006	58,8	15967,6
2007	58,8	19315,4
2008	59,0	23934,6
2009	58,6	23341,4
2010	58,8	24435,9
2011	58,5	26070,0

Source: [1]

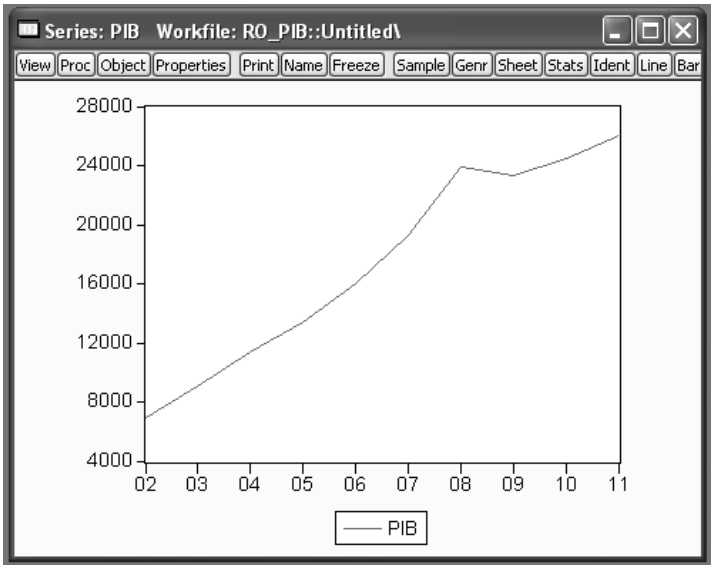
The analysis considered data sets and the estimation the parameters of regression model was performed using specialized software package Eviews 5.1. Thus, in a first stage of analysis peculiarities of the two data sets previously considered were studied.

As can be seen from the above table, the evolution of the two macroeconomic indicators is very similar, with sharp increases for the period 1990 - 2008 and a decrease of approximately 4 to 5% in 2009.

Evolution of employment rate

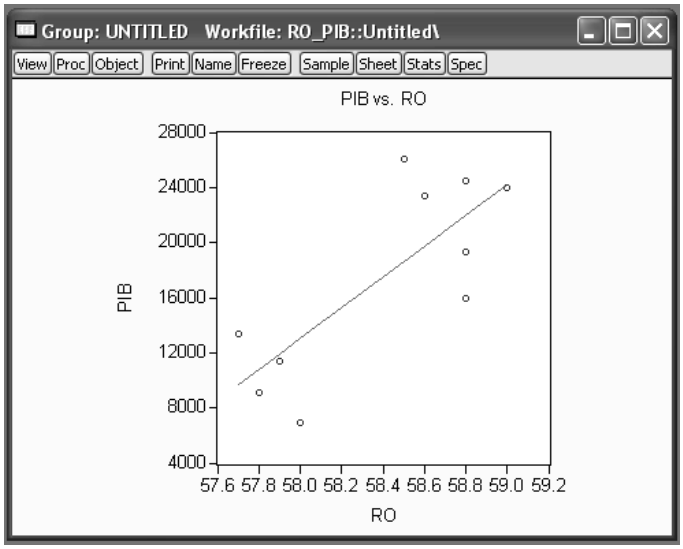


GDP evolution



To confirm the above hypothesis, we performed the graphical representation of data series (using at this specific commands implemented within the software package Eviews 5.1), this tool is particularly useful for identifying a typology that defines the correlation function of the two indicators:

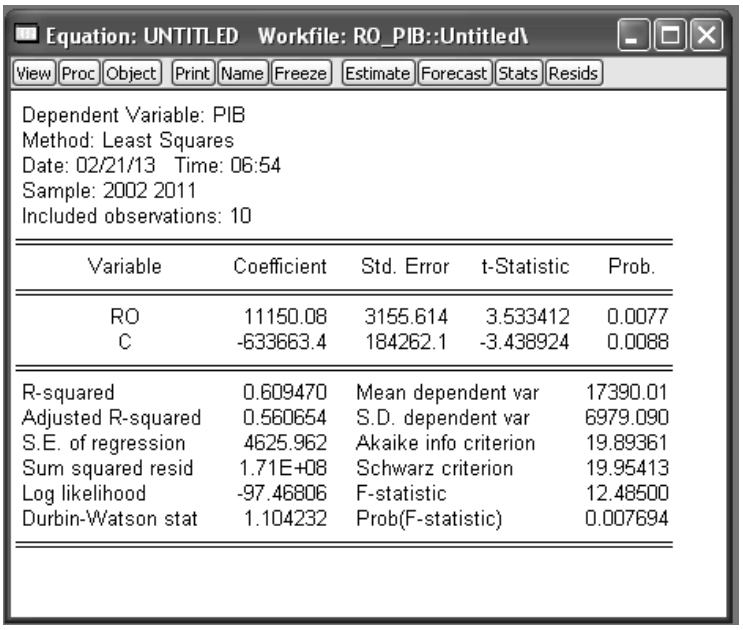
The correlation between employment rate and GDP per capita



The graphic above evidences a direct linear connection between the two indicators, which allows us to affirm that it is possible to use simple linear regression model to study the dependence between the GDP per capita and the employment rate of the population.

Estimating the parameters of regression model using the variable as the employment rate of people employed and the GDP per capita value as variable dependence was performed automatically (Figure 6), using specialized software package Eviews 5.1. Its framework is implemented least-squares method (Least Squares) as a method for estimating the model parameters, requiring only define two variables (PIB_L and E) and the constant term (C).

Parameter estimation results of the regression model that analyzes the dependency between GDP per capita and population employment rate



Equation: UNTITLED Workfile: RO_PIB::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: PIB
Method: Least Squares
Date: 02/21/13 Time: 06:54
Sample: 2002 2011
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RO	11150.08	3155.614	3.533412	0.0077
C	-633663.4	184262.1	-3.438924	0.0088

R-squared	0.609470	Mean dependent var	17390.01
Adjusted R-squared	0.560654	S.D. dependent var	6979.090
S.E. of regression	4625.962	Akaike info criterion	19.89361
Sum squared resid	1.71E+08	Schwarz criterion	19.95413
Log likelihood	-97.46806	F-statistic	12.48500
Durbin-Watson stat	1.104232	Prob(F-statistic)	0.007694

***

Analyzing the results obtained it is possible to formulate practical conclusions concerning the dependence between the value of GDP per capita and employment rate of the population:

• The probability that this is a correct model is relatively high - about 61%, this conclusion can be made based on the values determined using Eviews program for testing R - squared (0.6094) and Adjusted R - squared (0.5606);

• The validity of this model is confirmed by regression test values F - statistic (12.48500 - higher value than table level is considered to be the benchmark in analysis validity of econometric models) and the degree of risk almost zero (reflected by test value Prob F - statistic). We consider the regression model describing the relationship between the GDP per capita and the employment rate of the population as fair, which faithfully reflect the real evolution of the two macroeconomic indicators.

• Based on estimated values previously considered regression model may be given as follows:

PIB = - 633,663.4 + 11,150.08 RO

• Between the value of GDP per capita and population employment rate recorded in Romania in the period 2002 - 2011 there is a significant direct relationship. Thus, we can say that a one percent increase in employment will lead to an increase with 11,150.08 currency units for the capita GDP.

• The high value of the constant term reflects that the influence of the unspecified factors in the model on resultative variable evolution (GDP per capita) is significant, which leads us to conclude that the model used (although is correct) can be developed to ensure even better outcomes for activity prediction.

Between the employment rate and GDP / capita there is a direct linear whose trend can be evidenced by the equation: $\hat{y} = -633.663,4 + 11.150,08 \bullet x$, which gives the following results:

Correlation coefficient $r_{y/x}$ = 0.7804 indicates a strong link between the two variables:

$$r = \sqrt{R^2} = \sqrt{0.609} = 0,7804;$$

The determination report confirms that the employment rate is a determinant factor ($R^2 > 50\%$) for GDP growth/capita, its variation influencing the rate by 61%.

To verify the significance of the linear correlation coefficient t test (Student) is applied, by calculating the variable t_{calc} by the relation:

$$t_{calc} = \frac{r_{y/x}}{\sqrt{1 - r_{y/x}^2}} \times \sqrt{n-2}\,;$$ where: $r_{y/x}$ = linear correlation coefficient.

n = number of pairs of observed values =10

$$t_{calc} = \frac{0,7804}{\sqrt{1 - 0,609}} \times \sqrt{10-2} = 3,531$$

t_{calc} value is compared with the critical value, the table, $t_{tabelat}$, that is probabilistic set to a level of significance α and n-2 degrees of freedom. If $|t_{calc}| > |t_{tabelat}|$ the hypothesis significance of correlation is checked, and if $|t_{calc}| < |t_{tabelat}|$ the relationship is insignificant, so a key determinant will have to be found to apply the correlation method.

With a 95% probability and 8 degrees of freedom $t_{tabelat}$ has a value of 2.306. Because $|t_{calc}| > |t_{tabelat}|$, $|3,531| > |2,306|$, we can say that the hypothesis for the significance of correlation between variables investigated is verify and and there is a connection between investigated variables significant, so $r_{y/x}$ is statistically significant and analysis model is correctly specified.

Bibliografie selectivă:

[1] *** Institutul Naţional de Statistică – Anuarul Statistic al României, Ediţiile 2010, 2011, 2012

- Anghelache C., (2008) – „Tratat de statistică teoretică şi economică", Editura Economica, Bucureşti

- Anghelache, C (coordonator) şi alţii (2012) – „Modele statistico – econometrice de analiză economică – utilizarea modelelor în studiul economiei României", Revista Română de Statistică (Supliment)

- Turdean, M.S., Prodan L., (2012) – „Statistică pentru afaceri", Editura ProUniversitaria, Bucuresti, ISBN 978-606-647-312-5