USING THE LINEAR REGRESSION MODEL IN ORDER TO ANALYSE THE CORRELATION BETWEEN THE GROSS DOMESTIC PRODUCT AND THE HOUSEHOLD EFFECTIVE INDIVIDUAL FINAL CONSUMPTION

Lecturer Marian SFETCU PhD.
"ARTIFEX" University of Bucharest
Andreea-Ioana MARINESCU, PhD. Student
Bucharest University of Economic Studies

Abstract

This article aims to review the changes in gross domestic product compared to the progress of the household effective individual final consumption for the 1995-2015 period. By using a series of online data published by the National Institute of Statistics, we will analyse if there is a dependency relationship between the two variables and if so, we will also study the dependence type. By using the statistical-econometric model of simple linear regression, we will notice how a variable y, called dependent variable, can be described using another variable x, called independent variable.

Key-words: *simple regression, GDP, dependent variable, final consumption*
JEL Classification: *B22, N10*

Introduction

The linear regression model is a statistical-econometric instrument used for proving the existence of a dependency relationship between two variables. The use of the linear regression model implies, first of all, identifying the variables for the analysis, namely a variable y, called dependent variable or endogenous, and another variable x, called independent variable or exogenous. After defining the two variables, we will proceed to the definition of the residual variable, denoted by ε. Given the fact that any modelling process or analysis of data series involves certain measurement errors which could influence the estimation of the parameters, it is necessary to include the residual variable into the model, in order to encapsulate all the phenomena which can influence the dependent variable, but is not necessarily connected to the independent variable.

Let us assume that we are discussing about an analysis of two variables y and x and we are looking to describe the variable y depending on the variable x. If there is a dependency relationship between the two variables and their graphic representation (Scatter chart) is a line, then it means we are dealing with a **simple linear regression model**. If the dependent variable y can be described depending on two or multiple independent variables, it means we are dealing with a **multiple linear regression model**.

The simple linear regression model can be represented by using the following function:

$$y = a + bx + \varepsilon$$

where,

y = dependent variable (explained)

x = independent variable (explanatory)

ε = residual variable

a,b = parameters estimated using the data series used for the definition of the two variables

In addition to linear models, there are also nonlinear models which can be transformed in linear models by using the logarithm function.

When the points resulting from the graphic representation of the two variables are not on the line, but around it, we are dealing with a nonlinear regression model. For instance, we have the following function:

$$y_i = a \bullet x_i^b$$

Using the logarithm function, we obtain:

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

Thus, a nonlinear model was transformed, by using the logarithm function, in a linear model, having as dependent variable $\log y_i$ and as independent variable $\log x_i$.

The linear regression model is an instrument often used for the analysis of the correlation and dependences between the main macroeconomic aggregates, because it provides an overview of the economy, emphasising the intensity of the influence exercised by certain indicators upon other indicators and also representing the basis for future developments.

In this article we will focus on studying the dependences between the gross domestic product per capita (GDP) as dependent variable (y) and the household effective individual final consumption, as independent variable (x).

According to the National Institute of Statistics, the gross domestic product represents the main macroeconomic aggregate, which reflects the final result of the production of resident production units. This indicator reflects the economic activity within a country, region or municipality for a specified period of time (usually a year).

The gross domestic product represents the result of the summation of gross added value for different economic sectors and product taxes and the reduction of the subsidies on products.

The household effective individual final consumption reflects the spending on goods and services made by households in order to meet their own needs.

Literature review

Anghelache (2015) studied Romania's macroeconomic indicators in depth, focusing on the study of gross domestic product. Anghelache, C., Manole, A., Anghel, M. G. (2015) use the regression model for the analysis of a macroeconomic indicator's dependence on other macroeconomic indicators. The work of Anghelache, C., Pagliacci, M. G. R., Prodan, Ligia (2013) presents the linear regression model and its applicability in studying the correlation between two variables. Anghelache C. (2008) deepens the statistical-econometric instruments used in the economic analysis and Ladiray, Dominique (2006) focuses on statistical data analysis using time series.

The research methodology

In order to build the simple linear regression model for the correlation between the gross domestic product and the household effective individual final consumption, we will analyse the variation of these two indicators for the 1995-2015 period, by using a series of online data published by the National Institute of Statistics.

Nr. Crt.	Year	Gross Domestic Product per capita	GDP per capita growth index compared to the previous year (%)	Household effective individual final consumption	Household effective individual final consumption growth index compared to the previous year (%)
1	1995	337,6	0,00	254,2	0,00
2	1996	507,1	50,21	401,6	57,99
3	1997	1.139,4	124,69	903,6	125,00
4	1998	1.655,7	45,31	1.386,8	53,47
5	1999	2.470,4	49,21	2.005,2	44,59
6	2000	3.622,7	46,64	2.784,0	38,84
7	2001	5.280,5	45,76	4.116,4	47,86
8	2002	7.041,5	33,35	5.396,3	31,09
9	2003	9.212,8	30,84	6.928,7	28,40
10	2004	11.595,6	25,86	8.927,4	28,85
11	2005	13.625,4	17,50	10.647,5	19,27
12	2006	16.373,0	20,17	12.670,3	19,00
13	2007	20.028,7	22,33	15.135,6	19,46
14	2008	25.532,8	27,48	18.482,1	22,11
15	2009	25.065,6	-1,83	17.760,4	-3,90
16	2010	26.368,7	5,20	19.071,1	7,38
17	2011	28.047,8	6,37	20.049,1	5,13
18	2012	29.679,1	5,82	21.320,9	6,34
19	2013	31.895,4	7,47	21.538,5	1,02
20	2014	33.524,4	5,11	21.984,2	2,07
21	2015	35.873,6	7,01	22.851,5	3,95

Source: http://statistici.insse.ro

The evolution of gross domestic product and household effective individual final consumption for the 1995-2015 period

Between 1995 and 2015, in Romania, the GDP experienced a generally positive variation. Thus, the biggest growth of GDP volume for the 1995-2015 period was in 1997, when the GDP increased by 124,69% compared to the previous year.

By analysing the data from the table above, we note that, in 2009, the GDP decreased by 1,83% compared to the previous year, which can be explained by the worsening of the economic crisis erupted in 2008, which had repercussions on the entire national economy.

Starting from 2010 and till 2015, the GDP amount increased continuously, but the values of the GDP growth index remain pretty low compared to the 1995-2008 period.

As regards the household effective individual final consumption, such as GDP variation, we notice a progress, the biggest growth being the one from 1997 (125% compared to 1996). The economic crisis which erupted in 2008 has negatively influenced household effective individual final consumption and this explains the decrease in 2009 by 3,90% in terms of final consumption, 2,07% more than the GDP decrease.

In the 2010-2015 period, household effective individual final consumption continued to increase, but, as we have seen regarding GDP, the rhythm of these increases was pretty low, compared to the 1995-2008 period.

Considering the data series presented in the table above, we can notice that the two indicators used for the analysis have varied in the same way, which is why we can state that there is a dependency relationship between the two variables.

In order to analyse the type of the two variables' regression model, we will proceed to their graphic representation (Scatter chart):

By analysing the chart above, we can notice that the graphic representation of the two indicators is a line. Thus, we can appreciate that there is a direct and linear connection between the two variables corresponding to the simple linear regression, described as $y = a + bx + \varepsilon$.

By using the regression function described above, we obtain the following values of parametres a and b:

$a = 620,4$ and $b = 0,6738$

The linear regression function becomes:

$y = 620,4 + 0,6738x$

An important element in determining the connection's intensity between the two variables is represented by the coefficient of determination (R^2). This coefficient shows to what extent is a dependent variable y explained by an independent variable x. Therefore, the closer to 1 is the value of the coefficient of determination, the bigger is the proportion in which the variable x explains the variable y.

In our analysis, $R^2 = 0,9918$, a value very close to 1, indicating that the variable y, namely the gross domestic product, is explained at a rate of 99,18% by the variable x, namely the household effective individual final consumption.

Conclusions

The gross domestic product and the household effective individual final consumption are two macroeconomic aggregates which have evolved between 1995 and 2015, their growth's annual index being alert during 1997-2005 compared to 2010-2015.

The use of the regression model for the two variables has reflected a direct and linear connection between the two indicators, while the main coefficients' values, which were close to 1, have strengthened the correlation between them, showing that the chosen model is valid. The dependent variable x has explained the independent variable y at a rate of 99%.

Bibliography

1. Anghelache, C., Manole, A., Diaconu, A., Evoluția economică și socială a României în ultima perioadă – analiză multidimensională, Revista Română de Statistică nr. 8/2016, pp. 12-18;
2. Anghelache, C., Manole, A., Anghel, M. G. (2015), *Analysis of the final consumption and gross investment influence on GDP – multiple linear regression model*, Theoretical and Applied Economics, no. 3/2015 (604), Autumn, pp. 137-142;
3. Anghelache, C., Pagliacci, M. G. R., Prodan, Ligia (2013), *Model de analiză macroeconomică bazat pe funcţia de regresie*, Revista Română de Statistică nr. 1/2013;
4. Anghelache, C., Bugudui, Elena, Deatcu, C. (2012), *Econometrie*, Editura Artifex, Bucureşti, 2012, pp. 77-130;
5. Anghelache, C. (2008), *Tratat de statistică teoretică şi economică*, Editura Economică, Bucureşti;
6. Ladiray, Dominique (2006), Économétrie, prévision et analyse des données, Revista Modulad nr. 35/2006;
7. Anuarul statistic al României, 2015.