JOURNAL OF PREVENTIVE MEDICINE
2008; 16(1-2): 24-37

LIFE EXPECTANCY IMPACT ON GDP PER CAPITA IN ROMÂNIA

Anca Vitcu[1], Elena Lungu[2], Luminița Vitcu[3]

1. University "Al. I. Cuza", Faculty of Computer Science, Iaşi, România
2. Institute of Public Health, Iaşi, România
3. Institute of Public Health, Bucureşti, România

Abstract. Aim. The present survey was aimed at finding a statistical relationship between life expectancy at birth and Gross Domestic Products (GDP) per capita among a number of European countries and at national level, among Romanian counties. In the context of these models, the study tried to find the efficiency of policy plans which stimulate this connection. **Material and method.** The methods employed to find this connection between the two variables included multivariate statistical analysis. The efficiency was studied based on a common used method: ratios. **Results**. Regarding the analysis involving the European countries, there can be concluded that, the relationship between life expectancy at birth and GDP per capita can be defined by a linear regression model. Thus, higher levels of GDP per capita have been associated with high life expectancy at birth. Regarding efficiency of policy plans there can be noticed that the countries which explored in the best way the human capital in terms of life expectancy at birth *vs* GDP per capita, are countries with strong institutions. On the other hand, the analysis employed at Romanian national level, provides a model which envisages that there are other factors that boost or cripple its economic growth. In terms of policy plans there can be remarked that the Romanian counties which explored less the relationship between life expectancy at birth and GDP per capita, are counties which confronted with factors of a malfunction market. **Conclusions**. The analysis undertaken in this study proves that at European level there is a strong positive correlation between life expectancy at birth and economic development, and countries that can successfully manage this relationship possess a distinct advantage in front of economic and social challenges. For România the model developed is strongly influenced by complementary factors.
Key words: life expectancy at birth, GDP per capita, simple regression model, cluster analysis

Rezumat. Scop. Obiectivul studiului a fost determinarea unei legături statistice între speranța de viață la naştere şi Produsul Intern Brut (PIB) pe locuitor. Modelarea relației dintre cele două variabile a fost realizată atât le nivel European cât şi la nivel național, între județele României. In contextul acestor modele, studiul a definit eficiența politicilor care stimulează această conexiune. **Material si metodă**. Metodele utilizate pentru construirea modelelor au inclus elemente de statistică multivariată, iar studierea eficienței a avut la baza o metodă frecvent utilizată în econometrie pentru *Data Envelopment Analysis* (DEA): metoda proporțiilor. **Rezultate.** In ceea ce priveşte rezultatul analizelor la nivelul țărilor europene se poate afirma faptul că între speranța de viață la naştere şi PIB pe locuitor poate fi definit un model liniar, bazat pe o corelație pozitivă între cele două variabile. Pe de altă parte, analiza eficienței politicilor dezvoltate la nivel național este regăsită în modul în care aceste țări ştiu să exploateze relația dintre speranța de viață la naştere şi PIB per capita. In ceea ce priveşte România, corelația dintre cele două variabile a fost foarte mică, iar modelul a arătat implicarea altor factori care influențează creșterea economică. De asemenea, în termeni de eficiență, se poate observa că județele care nu știu să exploateze corect relația dintre cele două variabile se confruntă cu disfuncționalități ale pieței. **Concluzie.** Studiul arată că la nivel european există o puternică corelație pozitivă între speranța de viață la naștere și dezvoltarea economică, iar țările care dezvoltă un management adecvat al acesteia identifică soluții care răspund cu succes provocărilor economice și sociale. In cazul României, relația dintre cele două variabile este puternic influențată de factori complementari.

Cuvinte cheie: **sperața de viață la naștere, PIB pe locuitor, modelul liniar, analiza cluster**

INTRODUCTION

The economic performance depends on several factors which vary across countries and in countries across areas.

Long-run growth is the most important aspect of how the economy of a country functions. The disparities among countries are due to differences in growth policies regarding technology and human investment both influencing the efficiency of labor. In this context better technology produces a higher efficiency of labor and the more capital the worker has at the disposal to amplify productivity, the more prosperous the economy can be. In turn, the two principal determinants of capital intensity are: the investment effort made by economy (the share of real GDP saved and invested to increase the capital stock) and the requirements of economy (how much new investment is needed to equip new workers with the standard level of capital, to keep up with new technologies and replace damages). Good and bad policies can accelerate or slow all this growth.

Countries with high per capita GDP may have stronger institutions with a greater control over the resources and a greater capacity to change the flow of policy. We focused on GDP per capita to catch the effect of a country's wealth on policy volatility.

Life expectancy at birth is a weighted average of mortality rates at all ages, strongly influenced by infant and child mortality rates. It's value gives a great part of the picture describing community wellbeing (health and socioeconomic conditions). We focused on life expectancy at birth to capture the effect of a country's health on policy volatility.

The connection between these two variables can theoretical be explained by the fact that a population experiencing a rapid increase in life expectancy is likely to have more savings and contribute more to investments. The quantitative impact of life expectancy on economic growth is estimated to have an important magnitude.

MATERIAL AND METHODS

The present paper is based on the data provided by National Institute of Statistics România and WHO database for the years 2005 and 2006 (1, 2).

The purpose of the study was to characterize the relationship between two important variables, GDP per capita and life expectancy at birth employing components of multivariate statistical analysis, as well as some basic econometric elements.

The statistical analysis included a simple regression model to sense problem, and a cluster analysis to organize data into meaningful clusters and help select effective actions. The model is specified by the equation:

$$\ln(y_i) = \beta_{0i} + \beta_{1i} x_i + \varepsilon_i,$$

where i =1, 2,…, N is the number of countries, y_i is GDP per capita for country i, x_i is life expectancy at birth for country i, β_{0i} and β_{1i} are parameters to be estimated, and ε_i is a random error term associated with observation y_i. For statistical reasons the dependent variable, GDP per capita entered in the analysis as natural logarithm.

The econometric element incorporated into the analysis refers to the calculation of a performance ratio, more specific, the efficiency in converting the input (life expectancy at birth) into output (GDP per capita). This ratio is defined simply as:

$$Efficency = \frac{output}{input} \quad ,$$

under the hypothesis that between these two variables there is some connection.

We applied these methods to study the relationship between GDP per capita and life expectancy at birth for a number of countries in Europe and than for the Romanian counties looking for similarities between the variables behavior (3, 4).

RESULTS AND DISCUSSION

In the following lines we explore the statistical relationship between life expectancy at birth and GDP per capita in two study cases: 1) analysis among some European countries, and 2) analysis at national level among Romanian counties.

Hypothesis: Higher levels of GDP per capita are associated with high life expectancy at birth.

The case of European countries

The sample was composed of 29 European countries. The variables involved in the analysis are GDP per capita and life expectancy at birth. The data are valid for 2006.

Before accepting or rejecting the formulated hypothesis we have to study the relationship between ln(GDP per capita) and life expectancy at birth. We will try to answer the following issues: Is GDP per capita related to life expectancy at birth? If so, how are they related?

Firstly we generated the correlation matrix which enables us to test for a statistical relationship between the two variables. The outcome shows a value of 0.828 for the Pearson correlation coefficient, significant at 0.01 level. In other words, this means that GDP per capita appeared to have a positive correlation with life expectancy at birth.

Plotting the independent variable (life expectancy at birth) against the dependent variable [ln(GDP per capita)], the scatter diagram indicates that these variables are linearly related (fig. 1).

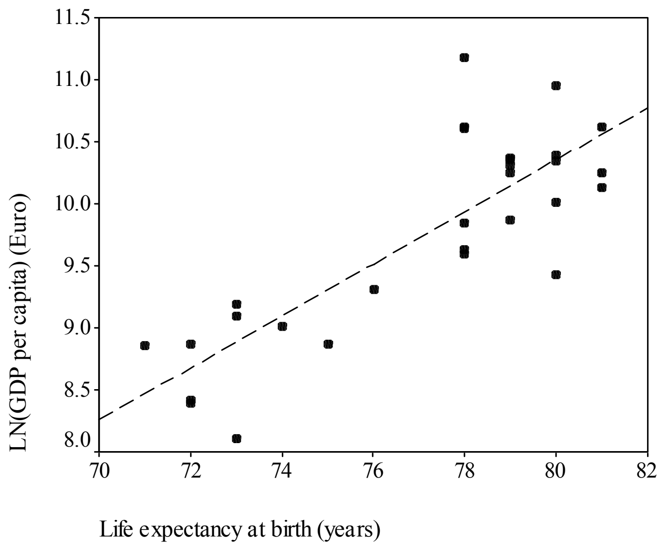

Fig 1. Scatter diagram of ln(GDP per capita) and life expectancy at birth, 2006

Considering the above information, we built a good linear regression model between the response variable ln(GDP per capita) and the independent variable (life expectancy at birth). The summary of the model is presented in table 1. According to this model, R-square, the percent of the GDP per capita, explained by life expectancy at birth is 0.686. This means that other variables, not included in the model, explain the rest of the variance. In other words, the one-independent model explains 68.6% of the variance. An important information presented in table 1 refers to the adjusted R-square, which value is 0.674, meaning that the penalty for the possibility that some of the variance may be due to chance is small. On the other hand, the calculated standard error is about 0.47 units, which means that we could be 95% confident that the actual value would be in the interval determined by the predicted value ± 1.96 * 0.47 (tab. 1).

Table 1. Model summary for the European countries

	R	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
Model summary	0.828	0.686	0.674	0.4658

In the ANOVA table we can see that the significance of the F value is below 0.05, so we can affirm that the model is significant (tab. 2).

The information regarding model coefficients, B and Beta, as well as the constant are presented in table 3.

Table 2. ANOVA table

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	12.789	1	12.789	58.937	0.000
Residual	5.859	27	0.217		
Total	18.647	28			

Table 3. Model Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	IC95% for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	-6.404	2.107		-3.040	0.005	-10.726	-2.082
Life expectancy at birth	0.209	0.027	.828	7.677	0.000	0.153	0.265

According to these values the prediction (regression) equation can be written in the following way:
predicted ln(GDP per capita)= 0.209 * (life expectancy at birth) – 6.404.
More precisely, reminding that the standard error of estimate is 0.47, this means that at the 0.05 significance level the estimate is the value given by this formula plus or minus 1.96*0.47.
Other important values mentioned in table 3 are: the value of Beta, the standardized regression coefficient (0.828), and the levels of significance of the t-test which show a significant B coefficient and also a significant constant.

The model described above is valid only if the framework of the following assumptions:

- the relationships between the independent and dependent variables is linear (which we already proved)
- the errors are normally distributed - technically normality is necessary only for the t-tests to be valid, estimation of the coefficients only requires that the errors be identically and independently distributed
- the error variance is constant (homogeneity of variance)
- the errors associated with one observation are not correlated with the errors of any other observation
- the model is properly specified

So, to answer to these issues we performed a residual analysis which summary is presented in table 4.

Table 4. Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	8.4687	10.5634	9.7544	.6758	29
Std. Predicted Value	-1.902	1.197	.000	1.000	29
Standard Error of Predicted Value	8.964E-02	.1885	.1193	2.773E-02	29
Adjusted Predicted Value	8.3933	10.6039	9.7563	.6796	29
Residual	-.9285	1.2439	1.041E-15	.4574	29
Std. Residual	-1.993	2.670	.000	.982	29
Stud. Residual	-2.059	2.721	-.002	1.013	29
Deleted Residual	-.9905	1.2917	-1.9307E-03	.4870	29
Stud. Deleted Residual	-2.200	3.134	.008	1.072	29
Mahal. Distance	.071	3.619	.966	.934	29
Cook's Distance	.000	.161	.032	.045	29
Centered Leverage Value	.003	.129	.034	.033	29

Studying the values of Mahalanobis distance and Cook’s distance as well as the centered leverage value we can conclude that it does not appear there are problem cases.

The regression standardized residual histogram provides a visual way of assessing if the assumption of normally distributed residual error is met (fig. 2).

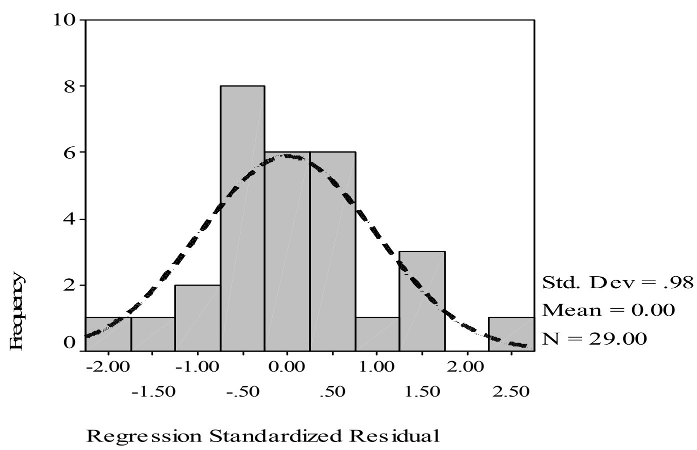

Fig. 2. The regression standardized residual histogram

When the sample size is small, the histogram, in general, shows irregularities from a normal curve. Studying Normal P – P Plot of Regression Standardized Residual we can notice that the plots have the trend of the 45-degree line, and they are close to it. The normality of the residuals can also be checked based on the tests of normality, in our case Shapiro-Wilk statistic is not significant (sig. =0.568). Based on these results, the residuals from this regression appear to conform to the assumption of being normally distributed.

The scatterplot of studentized deleted residuals *vs* standardized predicted values shows that the residuals fall between -2 and +2. Also, it reveals that all the points are positioned in a constant horizontal band (fig. 4).

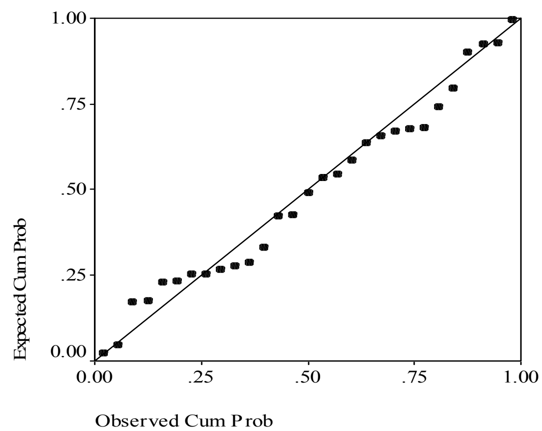

Fig. 3. Normal P – P Plot of Regression Standardized Residual

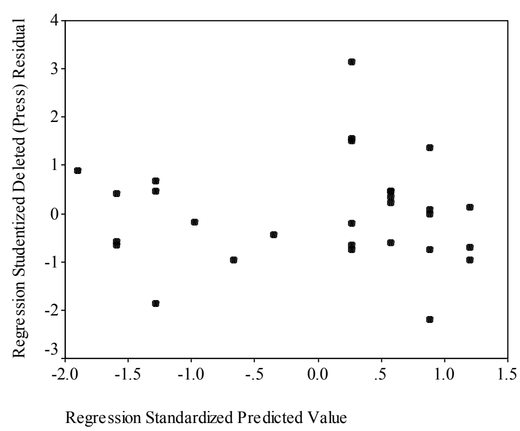

Fig. 4. The scatterplot of studentized deleted residuals *vs* standardized predicted values

In the scatterplot of standardized predicted values against observed values if 100% of the variance is explained in a linear relationship the points form a straight line. In our case the points are dispersed around a straight line (the trend) due to the standard error of estimate (fig. 5).

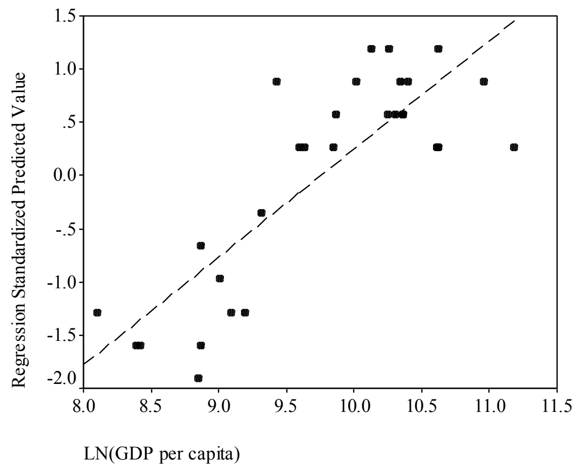

Fig. 5. The scatterplot of standardized predicted value against dependent variable

This reflects the fact that the model explains only 68.6% of the variance.

As we well know a model specification error can occur when one or more relevant variables are omitted from the model or one or more irrelevant variables are included in the model. If relevant variables are omitted from our model, the common variance they share with included variable may be wrongly attributed to this variable, and the error term can be inflated. Model specification errors can substantially affect the estimate of regression coefficients. To check if the model is properly specified we used the predicted value and the predicted value squared as predictors of the dependent variable. In this new approach, if the model is specified correctly the predicted value should be significant while the predicted value squared shouldn't be significant. The results of this analysis are presented in table 5 and we can see that the t-test envisages that the standardized predicted value square is not significant (the significance level is 0.411>0.05).

Thus, we can conclude that, the linear model is valid from statistical point of view, so we can accept the hypothesis that higher levels of GDP per capita are associated with high life expectancy at birth. In this context we can say that health is an important determinant of economic growth.

Table 5. Model coefficients (Standardized predicted value squared included)

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.	95% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	9.855	0.148		66.566	0.000	9.550	10.159
Standardized Predicted Value Squared	-0.104	0.124	-0.119	-0.836	0.411	-0.359	0.151
Standardized Predicted Value	0.613	0.116	0.752	5.293	0.000	0.375	0.851

Performing a cluster analysis based on GDP per capita and life expectancy at birth we found the following clusters to which components we associated the corresponding efficiency percentage:

Cluster 1: Luxembourg (100%);
Cluster 2: Norway (78%);
Cluster 3: Denmark (57%), Ireland (57%), Switzerland (55%);
Cluster 4: Netherlands (45%), Finland (44%), UK (43%), Austria (41%), Belgium (41%), Germany (39%), France (38%);
Cluster 5: Italy (34%), Spain (30%), Greece (27%), Cyprus (26%);
Cluster 6: Portugal (21%), Slovenia (21%), Malta (17%), Czech Republic (16%), Estonia (15%);
Cluster 7: Hungary (13%), Slovakia (12%), Latvia (11%), Lithuania (11%), Poland (10%), România (7%), Turkey (7%), Bulgaria (5%).

To calculate the relative efficiency scores (percentages) we used a classical econometric method. We considered the model in which the key performance outcome is GDP per capita (output variable), and one of the factors that can lead to this outcome, life expectancy at birth (input variable). For each country, we divided the output measure by the input measure; we selected the highest ratio and compare all other to it calculating their relative efficiency with respect to the best value. Then we converted the scores to percentages multiplying by 100.

This simple method was chosen as the analysis technique because variables can be measured in different units and it measure technical efficiency defined as the successful implementation of a policy plan, deviations from the plan being perceptible.

Studying the distribution of relative efficiency percentages among clusters, we can remark that the countries which explored in the best way the human capital in terms of life expectancy at birth *vs* GDP per capita, are countries with strong institutions.

The case of România

Preparing a similar model for Romanian counties we achieved that at the national level the outcomes are very different regarding the connection between the two variables (GDP per capita and life expectation at birth). For statistical reasons we excluded from the analysis Bucureşti and Satu Mare counties. The study uses the data available for 2005.

The correlation matrix shows for the Pearson correlation coefficient a low value (0.331) significant at 0.05 level.

Plotting the independent variable (life expectancy at birth) against the dependent variable [ln(GDP per capita)], the scatter diagram indicates that these variables can be linearly related (fig 6.).

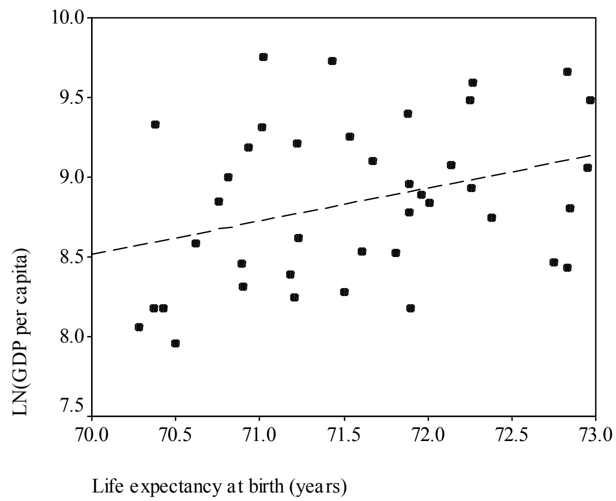

Fig 6. Scatter diagram of ln(GDP per capita) and life expectancy at birth, 2006

The R-square, the percent of the GDP per capita explained by life expectancy at birth has a very small value 0.110 (tab. 6). This means that other variables, not included in the model, explain the bulk of the variance. In other words, this one-independent model explains only 11% of the variance.

The adjusted R-square, calculated considering the variance due to chance is 0.086 (tab. 6).

The calculated standard error is about 0.48 units, which means that we could be 95% confident that the actual value would be in the interval determined by the predicted value ± 1.96 * 0.48 (tab. 6).

Table 6. Model Summary for România

	R	R Square	Adjusted R Square	Std. Error of the Estimate
Model summery	0.331	0.110	0.086	0.4783

In the ANOVA table we can notice that the significance of the F value is 0.037 (less than 0.05), so we can affirm that the model is significant (tab. 7).

The coefficients table envisaged that the independent value included in the model is significant (sig = 0.037) (tab. 8).

Table 7. ANOVA table

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.072	1	1.072	4.687	.037
Residual	8.694	38	.229		
Total	9.766	39			

Table 8. Model Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	IC95% for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	-6.016	6.865		-0.876	0.386	-19.912	7.881
Life expectancy at birth	0.208	0.096	0.331	2.165	0.037	0.013	0.402

The scatterplot of studentized deleted residuals *vs* standardized predicted values visualizes the fact that the residuals fall between -2 and +2 and that the all the points are positioned in a constant horizontal band (fig. 7). The test of normality, Shapiro-Wilk, is not significant (sig.=0.130), from which we can conclude, that the residuals from this regression appear to conform to the assumption of being normally distributed.

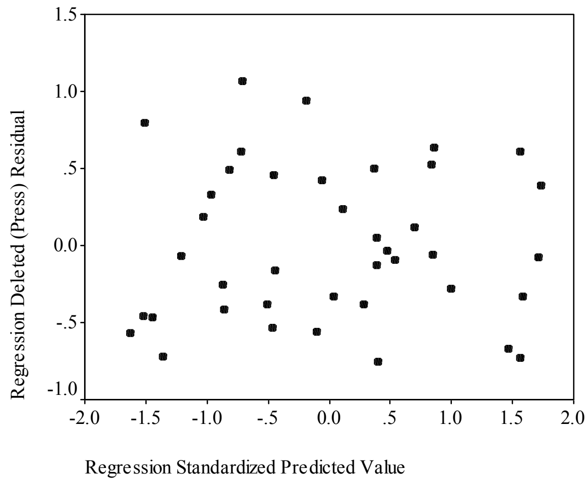

Fig. 7. The scatterplot of studentized deleted residuals *vs* standardized predicted values

In the scatter diagram of standardized predicted value against dependent variable the points are very spread around the trend, situation explicated by the fact that the model explains only 11% of the variance (fig.8).

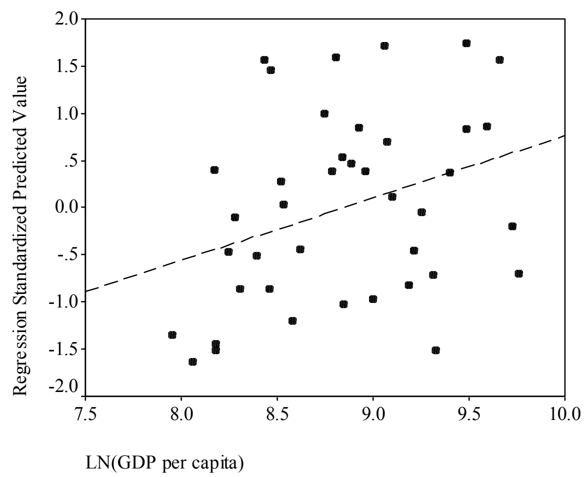

Fig. 8. The scatterplot of standardized predicted value against dependent variable

To verify if the model is correctly specified we computed the squared of predicted value and then included it in the model. The results show that the squared of predicted value is significant (sig.=0.306), suggesting that we have omitted important variables in our regression.

In this case, we don't have enough evidence to accept the hypothesis formulated at the beginning of this paper, that higher levels of GDP per capita are associated with high life expectancy at birth. The model shows that there are other factors that boost or cripple the economic growth.

Performing a cluster analysis based on GDP per capita and life expectancy at birth we found the following clusters to which components we associated the corresponding efficiency percentage:

Cluster 1:	Constanța (100%), Timiş (96%), Cluj (88%);
Cluster 2:	Prahova (84%), Argeş (75%), Braşov (74%);
Cluster 3:	Iaşi (69%), Bihor (66%), Bacău (64%), Mureş (60%), Dolj (58%);
Cluster 4:	Arad (56%), Galați (51%), Sibiu (50%), Suceava (49%), Hunedoara (47%), Ilfov (45%);
Cluster 5:	Dâmbovița (43%), Gorj (41%), Mureş (40%), Neamț (39%), Vâlcea (38%), Alba (37%), Buzău (36%);
Cluster 6:	Olt (32%), Caraş-Severin (31%), Harghita (29%), Brăila (29%), Bistrița-Năsăud (27%), Teleorman (27%), Vrancea (26%), Botoşani (25%);
Cluster 7:	Ialomița (24%), Vaslui (23%), Mehedinți (22%), Sălaj (21%), Tulcea (21%), Covasna (20%), Călăraşi (18%), Giurgiu (17%).

Studying the distribution of relative efficiency percentages among these clusters, we can remark that the counties which explored less the relationship between life expectancy at birth and GDP per capita are counties which confronted with factors of a malfunction market.

CONCLUSIONS

International literature proved that investment in public health produces enormous economic benefits both for the people and countries as a whole, being an important determinant of economic growth. The analysis undertaken in this study proves that at international level there is a strong positive correlation between life expectancy at birth and economic development, and countries that can successfully manage this relationship possess a distinct advantage in front of economic and social challenges.

The analysis provided at country level, for România envisaged a different model in terms of connection between life expectancy at birth and GDP per capita. The model developed is strongly influenced by complementary factors.

REFERENCES

1. WHO/Europe: *Europe health for all database (HFA –DB)*, July, 2008, http://www.euro.who.int/hfadb.
2. National Institute of Statistics*: Romanian Statistical Yearbook 2007*, Bucureşti.
3. Tandon Ajay: *Attaining Millennium Development Goals in Health: Its't Economic Growth Enough?*, March, 2005.
4. WHO, The Council of Europe Development Bank: *Health and Economic Development in South-Eastern Europe*, 2006.