STATISTICAL EVALUATION OF THE INFLUENCE OF DETERMINING FACTORS OF LIFE EXPECTANCY

Elisabeta JABA
Faculty of Economics and Business Administration
Alexandru Ioan Cuza University
Iasi, Romania
ejaba@uaic.ro

Christiana BALAN
Faculty of Economics and Business Administration
Alexandru Ioan Cuza University
Iasi, Romania
christiana.balan@uaic.ro

Silvia PALAŞCĂ
Faculty of Economics and Business Administration
Alexandru Ioan Cuza University
Iasi, Romania
silvia_palasca@yahoo.com

Abstract

The aim of this paper is to reach the following objectives: to identify the determinants that have significant influence on life expectancy for European countries, in 2007; to measure the significance level of the influence of the determinants on life expectancy; to classify the European countries according to life expectancy and its determining factors.

The research is based on the regression method and cluster analysis.

The results show that the determining factors of life expectancy, in 2007, for European countries, are: education, gender, health, earnings, and marital status. The influence of these factors is different for life expectancy of male population and for life expectancy of female population. The similarities of factors' influence create clusters of countries.

Keywords: Life expectancy, education, health, regression analysis, cluster analysis
JEL classification: C10, C20, C82, I15, I25

1. INTRODUCTION

Each century has a determining feature, and wellness is the dominant of the twenty first century. The current level of knowledge, technology, and comfort guide the human being towards improving the factors that increase life expectancy. Life expectancy is the

expected (in the statistical sense) number of years of life remaining at a given age [Tsai, 2010].

In the knowledge society, the economical growth and development are directly influenced by the investment in human capital, thus resulting in a development of well-being, quantifiable by the improvement of life expectancy.

Most common factors associated with life expectancy are: gender, infant mortality rate, level of education, health state of the population, earnings, economic status and marital status. A number of studies were conducted for each one of these factors.

Gender was the first variable of study as seen in Monden [2010], Jaba [2007], Corsini [2010], Leung [2010] and the conclusion was that women have a higher life expectancy.

The second variable was education, a great deal of studies indicate that tertiary education tends to increase life expectancy [Cooney, 2008], [Mazumdar, 2001], [Balan, 2011], [Spoerri, 2006] in both genders and even secondary education gives a better chance to women [Corsini, 2010], [Rogot, 1992], [Stack, 1998].

Closely related to education is health. Life expectancy is influenced by health in a number of ways such as the expenses for health [Stubich, 2008], the number of doctors available to 1000 persons [Motivans, 2005], the healthier way of life of educated persons reflected also on their children's life expectancy [Balan, 2011], [Stubich, 2008], [Lind, 2010].

Another important factor is constituted by earnings as can be seen in [Leung, 2010], [Rogot, 1992], [Monden, 2010]. It has an overwhelming influence on life expectancy because it interacts with the other factors, earlier mentioned, facilitating or obstructing access to education and health services. [Lind, 2010] notes: "Social inequality and chances in life represent the educational gradient on health."

The last factor to be considered is marital status. Married persons are likely to live longer than single persons [Rossman, 1994]. However, a surprising result occurs in this paper: single persons tend to have greater life expectancy, probably due to lack of stress and responsibilities.

Results of the previous papers can be summarized as a spiral: education growth leads to technological development that can be seen in economic expansion, a result of which is an increase in well-being that triggers a new impulse for a need of raising the level of education.

The paper aims to reach the following objectives: to identify the determinants that have significant influence on life expectancy for European countries, in 2007; to measure the significance level of the influence of the determinants on life expectancy; to classify the European countries according to life expectancy and its determining factors.

2. DATA AND METHOD

a. Population. Sample. Variables

Data on life expectancy and its determining factors are observed for a sample of 29 European countries, in 2007, from the EUROSTAT and World Bank databases.

Considering the main results presented in the literature of speciality on the factors that influence life expectancy, the following variables were considered in the paper: •Health expenditure as % of GDP; •Total public expenditure on education as % of GDP, for all levels of education combined; •Percent of persons with pre-primary, primary and lower secondary education (total and by gender); •Percent of persons with upper secondary and post-

secondary non-tertiary education (total and by gender); •Percent of persons with tertiary education (total and by gender); •Gross domestic product at market prices (Euro per inhabitant); •Infant mortality rate; •Annual net earnings (Euro per person); •Unemployment rate (%);•Percent of single/married/widowed/divorced persons;

In the first stage of research we used the 13 variables recorded on a sample of countries in 2007. On the variables considered in the literature as determining factors on life expectancy, we applied an analysis of their significance, using the F test.

From the observed set of variables, there were identified, using the significance level for the F test, the following determining factors for life expectancy: •Percent of persons with tertiary education; •Annual net earnings; •Infant mortality rate; •Percent of single persons.

Life expectancy at certain ages represents the mean number of years still to be lived by a person who has reached a certain exact age, if subjected throughout the rest of his or her life to the current mortality conditions (age-specific probabilities of dying).

Infant mortality rate represents the ratio of the number of deaths of children under one year of age during the year to the number of live births in that year. The value is expressed per 1000 live births. It reflects the health level of a population and the concern for health in one economy.

The classification of the educational activities is based on ISCED (International Standard Classification of Education – UNESCO 1997). The categories are groups as it follows: pre-primary, primary and lower secondary (levels 0-2); upper secondary and post-secondary non-tertiary education (levels 3-4); tertiary education (levels 5-6).

Marital status reflects the condition of a person to be in one of the following categories: ingle, married, widowed, divorced or separated.

Annual net earnings are calculated from gross earnings by deducting the employee's social security contributions and income taxes, and adding family allowances in the case of households with children. Income tax and social security contributions parameters refer to the beginning of the fiscal year, i.e. changes occurring later during the year are ignored. Family allowances are cash transfers paid in respect of dependent children aged between 6 and 11 years. If the amounts vary within this age range, the most generous are used. The case of twins is disregarded.

b. Method

In order to highlight the influence of the determining factors and its level of significance, we used the analysis of regression and correlation.

The choice of the best regression model is based on the analysis of residuals (calculated as a difference between the observed values of the dependent variable and the fitted values from the regression line of the influence of each of the selected factor).

For each determining factor of life expectancy we retained the regression model for which the R square is the highest.

In order to identify homogenous groups of countries based on the observed variables, we used the hierarchical cluster analysis.

The method starts with clusters formed by one case that merge together until they form one big cluster. The first step in the cluster analysis is the proximity matrix (similarity matrix) or the distance matrix. The distance between two cases is measured using the Euclidean distance. It represents the geometrical distance in a multidimensional space (the square root

of the sum of the squared differences between all the "i" variables on two cases): $d_{Euc}(x,y) = \sqrt{\sum_i (x_i - y_i)^2}$.

The distances are calculated in order to group the cases into clusters. The rule used for clustering the cases is the Ward's method: the cluster membership is based on the sum of the squared deviations from the cluster mean. The optimal number of clusters is decided on subjective reasons: as the cases form groups, the clusters that merge are bigger and bigger and they are composed of more and more dissimilar cases.

3. RESULTS

The relationship between each factor and life expectancy is analysed by a simple linear regression model.

The graph for the analysis of residuals is presented for each model, considering the residuals on the horizontal axis and the explanatory variable on the vertical axis (see Figure no. 1). There can be noticed that the points in the graphs are randomly spread around the line for $\bar{e} = 0$, therefore the linear regression model is appropriate.

For each model we estimated the regression equation. Table 1 presents the estimates for the regression coefficients, standard errors within brackets, and the significance level noted with asterisk.

Table no. 1 Life expectancy regressions for pooled population

Explanatory variable	**Regression model**
X_1 – Secondary education (3-4) (%)	$\hat{y} = 81.766^{***} - 0.09^{**} x_1$ (2.001) (0.041)
X_2 – Tertiary education (5-6) (%)	$\hat{y} = 73.341^{***} + 0.208^{**} x_2$ (1.725) (0.080)
X_3 – Infant mortality rate	$\hat{y} = 81.891^{***} - 0.832^{***} x_3$ (0.853) (0.143)
X_4 – Annual net earnings	$\hat{y} = 73.206^{***} + 9.896E-5^{***} x_4$ (0.749) (0.000)
X_5 – Single persons (%)	$\hat{y} = 63.119^{***} + 0.349^{**} x_5$ (4.400) (0.105)

*Sig.<0.10, ** Sig.<0.05, *** Sig.<0.001

There can be noticed that educational level is an important determinant of life expectancy. However the impact is differentiated between the secondary and tertiary levels. Upper secondary and post-secondary non-tertiary education has a small negative influence on life expectancy (–0.09) while tertiary education has a positive influence. This fact can be explained in correlation with the different activities on labour market and the different income of the two groups. Persons with secondary education work mostly in industry (construction, mining, etc.) and perform mostly physical work in difficult conditions; moreover they have

a low income. On the other side, persons highly educated work in services sector, in a better work environment and have higher incomes, so that they can afford to spend more on health services and therefore improve their life expectancy.

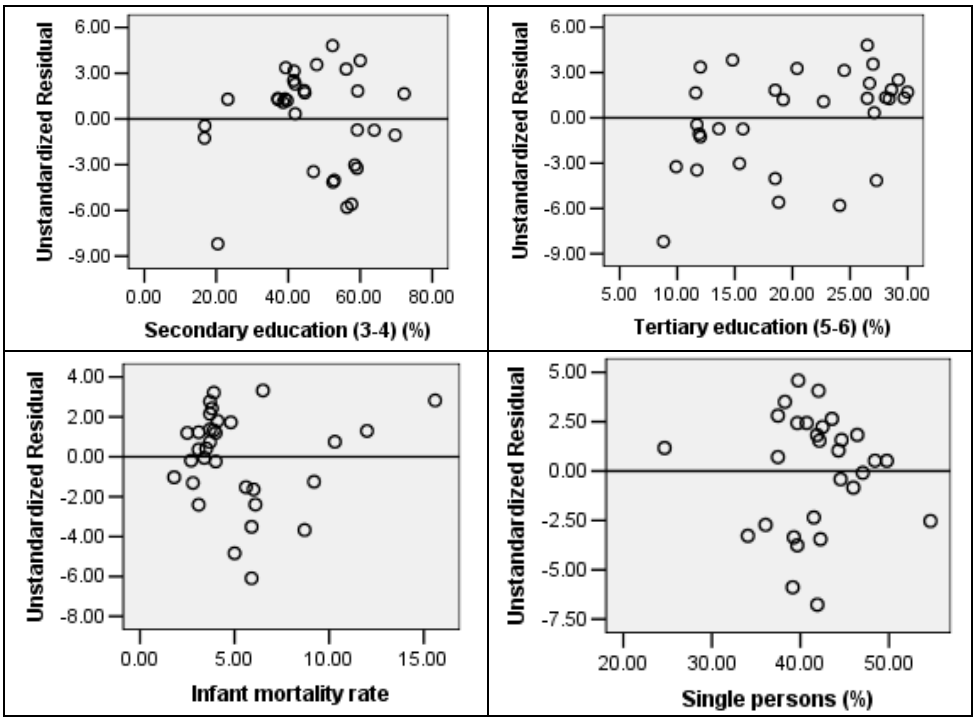

Figure no. 1 Scatter plot of Residuals for each regression model

Infant mortality rate has a negative significant impact on life expectancy. Indeed, mortality rate at an early age reflects the quality of health care system in one country which is an essential determinant on life expectancy. Therefore, a lower infant mortality rate implies a higher life expectancy.

As shown in literature income is a determining factor of life expectancy due to its correlation with access to education and health services. The positive dependence between annual net earnings and life expectancy underlines this argument. Higher earnings imply higher access to health care services and better education, hence a higher life expectancy.

Also, marital status proves to influence life expectancy. This hypothesis is verified for one category of persons, namely single persons. There is a positive impact of single persons on life expectancy. One reason for this finding may be the work

There are differences in the explanatory factors of life expectancy according to gender.

For male population, secondary education has a negative impact, as also seen for total population, while tertiary education has a positive impact (see Table 2).

Table no. 2 Life expectancy regressions for male population

Explanatory variable	Regression model
X_1 – Secondary education (3-4) (%)	$\hat{y} = 80.728^{***} - 0.130^{**} x_1$ (2.355) (0.047)
X_2 – Tertiary education (5-6) (%)	$\hat{y} = 68.158^{***} + 0.327^{**} x_2$ (1.941) (0.095)

*Sig.<0.10, ** Sig.<0.05, *** Sig.<0.001

For female population, the only category of education that influences life expectancy is the tertiary education (see Table 3).

Table no. 3 Life expectancy regressions for female population

Explanatory variable	Regression model
X_1 – Tertiary education (5-6) (%)	$\hat{y} = 78.131^{***} + 0.127^{**} x_1$ (1.344) (0.059)

*Sig.<0.10, ** Sig.<0.05, *** Sig.<0.001

After identifying the determining factors on life expectancy in 2007, we clustered the countries in homogenous groups according to the selected variables. We applied the hierarchical cluster analysis for the standardized variables using the Ward's cluster method and the Squared Euclidian distance measure.

Dendrogram graphically describes the formation of clusters. In this graph, the horizontal axis scales the distance of links between cases. These links are represented by lines forming an inverted U. Therefore the height of U letter indicates the distance between cases. For each node on the graph we read the horizontal distance criterion on which these cases were grouped into a single cluster. When the data are clearly divided into clusters of similar cases, different branches will be represented on the dendrogram.

According to the distance values, we kept the solution with 4 clusters (see Figure no. 2). The distribution of countries into the 4 clusters is presented in the cluster membership table (see Table 4).

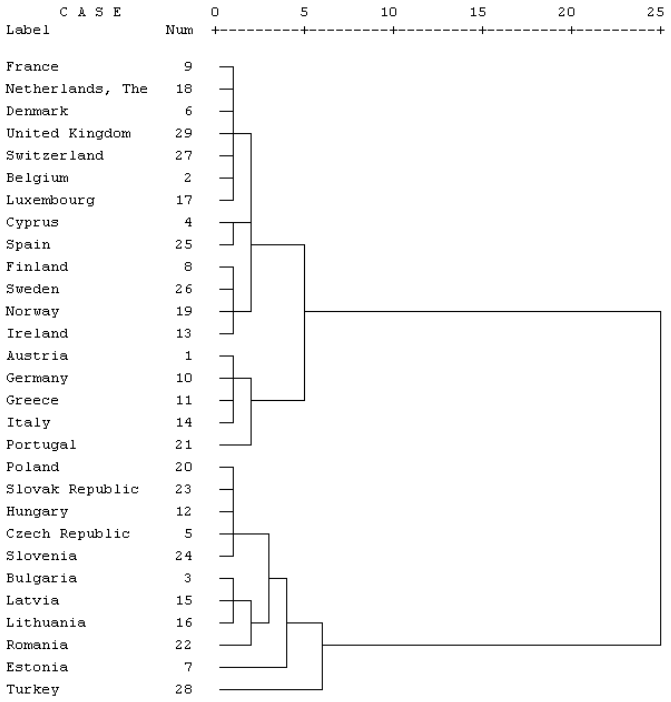

Figure no. 2 Dendrogram for hierarchical cluster analysis

Table no. 4 Cluster membership

Case	Clusters	Case	Clusters
1: Austria	1	16: Lithuania	3
2: Belgium	2	17: Luxembourg	2
3: Bulgaria	3	18: The Netherlands	2
4: Cyprus	2	19: Norway	2
5: Czech Republic	3	20: Poland	3
6: Denmark	2	21: Portugal	1
7: Estonia	3	22: Romania	3
8: Finland	2	23: Slovak Republic	3
9: France	2	24: Slovenia	3
10: Germany	1	25: Spain	2
11: Greece	1	26: Sweden	2
12: Hungary	3	27: Switzerland	2
13: Ireland	2	28:Turkey	4
14: Italy	1	29:United Kingdom	2
15: Latvia	3		

We characterized the 4 clusters according to the values observed for the selected variables. We computed the mean and the std. dev. for each variable by clusters (see Table 5).

Table no. 5 Statistics by cluster of countries

Cluster	Stat.	Life expectancy	Secondary education (%)	Tertiary education (%)	Infant mortality rate	Single persons (%)	Annual net earnings
1	N	5	5	5	5	5	5
	Mean	80.03	42.36	15.68	3.64	39.53	56769.76
	Std. Dev.	.99	17.14	3.95	.19	1.92	19139.87
2	N	13	13	13	13	13	13
	Mean	80.12	40.87	27.31	3.48	45.51	68832.10
	Std. Dev.	.95	6.82	2.06	.80	3.87	14556.18
3	N	10	10	10	10	10	10
	Mean	73.92	60.18	17.17	6.47	38.65	14724.32
	Std. Dev.	2.36	6.59	5.52	2.81	5.50	5712.732
4	N	1	1	1	1	1	1
	Mean	71.73	20.40	8.80	15.60	34.07	17585.37

There can be noticed that the 2nd cluster (Belgium, Denmark, Finland, France, Ireland, Luxembourg, The Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom, and Cyprus) is characterised by the highest level for: life expectancy, percent of population with tertiary education, percent of single persons and annual net earnings. In the same time it is characterised by the lowest level of infant mortality rate.

The 5 countries grouped in the 1st cluster are characterised by high performance on the life expectancy and its determinants. This cluster is closed to the 2nd cluster.

Romania belongs to the 3rd cluster of countries that is characterised by: low life expectancy (92.3% of the 2nd cluster mean); high infant mortality rate (185% of 2nd cluster mean); low tertiary education proportion (67% of 2nd cluster mean). The level of annual net earnings is the lowest one compared to all the other clusters (21% of 2nd cluster mean).

Turkey is analysed as an independent cluster as it is the country with the poorest performance on the observed variables: the lowest level for life expectancy along with the lowest level of tertiary education and the highest level for infant mortality rate.

4. CONCLUSION

In the paper we identified the determining factors of life expectancy, in 2007, for European countries. Then, we measured the influence of determining factors of life expectancy. We found that these factors differ depending on the period under research. For the modern context, we discovered which factors have a major influence. We used statistical methods and we identified important relations between life expectancy and education (secondary and tertiary), infant mortality rate, earnings, and some very interestingly the marital status, more exactly, the status of being single has an important significance.

In the end, we found the relativity of the determining factors on life expectancy.

Our findings correspond to the main ideas in the literature. The relevance of this study is essentially to integrate Romania in the European environment form the economic, social and demographic point of view. The cluster analysis used highlighted the long-lasting effects of communist period on the Eastern European countries. Even after almost 20 years the heritage of that period affects both life expectancy and its determining factors.

Acknowledgement

This work was co-financed from the European Social Fund through Sectoral Operational Programme Human Resources Development 200732013, project number POSDRU/89/1.5/S/59184 Performance and excellence in postdoctoral research in Romanian economics science domain.

References

[1] Balan, C., Jaba, E., Statistical analysis of the determinants of life expectancy in Romania, *Romanian Journal of Regional Science*, Vol. 5, No. 2, Winter: 25-38, 2011;
[2] Cooney, E., Better education translates into longer life expectancy, *The Boston Globe*, 2008;
[3] Corsini, V., Highly educated men and women likely to live longer, *Eurostat*, 24, 2010;
[4] Healy, A., Women beat men on life expectancy, education, *The Irish Times*, 2010;
[5] Jaba, E., Iaţu, C., Pintilescu, C., Dynamique de la structure économique en Roumanie et l'impact sur le marché, *Revue d'Economie Régionale et Urbaine*, Armand Collin, France, Vol. Novembre, Issue 4: 609-623, 2007;
[6] Leung, L., Effects of Internet Connectedness and Information Literacy on Quality of Life, *Social Indicators Research* 98:273–290, 2010;
[7] Lind, N.A., Calibrated Index of Human Development, *Social Indicators Research* 98:301–319, 2010;
[8] Mazumdar, K., Improvements in life expectancy: 1960-1995, an exploratory analysis, *Social Indicators Research* 55: 303–328, 2001;
[9] Monden, C., Do Measured and Unmeasured Family Factors Bias the Association Between Education and Self-Assessed Health?, *Social Indicators Research* 98:321–336, 2010;
[10] Motivans, A., Using educational indicators for policy: school life expectancy, *Prospects*, vol. XXXV, no. 1, March, 2005;
[11] Rogot, E., Sorlie, P., Johhnson, N., Life expectancy by employment status, income, and education in the National Longitudinal Mortality Study, *Public Health Rep*. Jul–Aug; 107(4): 457–461, 1992;
[12] Rossman, J., College, D., Televisions, Physicians, and Life Expectancy, *Journal of Statistics Education*, 1994;
[13] Spoerri, A., Zwahlen, M., et. al., Educational inequalities in life expectancy in the German speaking part of Switzerland between 1990 and 1997: Swiss National Cohort, *Swiss Med Wkly*, 2006;
[14] Stack, S., Marriage, family and loneliness: A cross-national study, *Sociological Perspectives* 41 (2), pp. 415-432, 1998;
[15] Stubich, M., Education, A Longevity Secret, *Scandinavian Journal of Public Health*, 2008;
[16] Tsai, C.M., Human Capital Composition and Economic Growth**,** *Social Indicators Research* 99:41–59, 2010;
[17] Udell, J., *Two implementations of US health care cost vs. life expectancy*, blog.jonurdell.net, 2010;
[18] Sullivan, A., Sheffrin, M., *Economics: Principles in action*, Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. (2003).