

DETERMINATION OF THE KEY FACTORS THAT INFLUENCE POVERTY THROUGH ECONOMETRIC MODELS

Evgjeni Xhafaj, Msc.

Department of Mathematics, Faculty of Information Technology,
University Alexandër Moisiu (Durrës, Albania)

Ines Nurja, PhD

University of New York Tirana (Tirana, Albania)

Abstract

The purpose of this study is to determinate the most important factors that influence poverty through econometric models that are the logistic regression and the linear log regression. The data are obtained from the Living Standards Measurement Study (LSMS) for 2008 that includes 3600 household interviewed in Albania. At the linear log regression model as dependent variable are the expenditure of consumption of the household per capita and as independent variable (explanatory) are the demographic, educational, zone variables. The logistic regression model has as depended variable the economic status (poor and non poor) and as independent variables are the same mentioned previously. The result of both econometric models confirm that the variables that are strong connected with the expenditure of consumption per capita and with the economic status are: household size, the educational level and gender of the head of household, the zone. This study recommends a carefully review on the reforms to be taken in relation to education in Albania.

Keywords: Poverty, linear regression, logistic regression, household size, Albania

Introduction

Poverty is a complex occurrence that includes different dimensions of deprivation, such as incomes or expenditure of consumption, the insufficiency of goods and services. The evaluation of poverty in Albania is based on a definition multidimensional of poverty. Except deprivation of goods, poverty is defined also in the relation with the inadequacy of a series of provisions of social care that has no relation with the incomes, as in

education, health care, authority, utilization of the base services and infrastructure.

The literature and from the different poverty study it is shown how the different methodology are related with individual or social nature of the above phenomenon. That is to say, can a person considered poor just evaluating his personal condition or evaluating in relation with other social characteristics in which he lives? Two answers can give two definitions in relation to poverty. An individual is considered poor if his expenditure level per capita is under the minimal level that is needed to fulfill his/her basic need for food and non food goods. This minimal level of consumption is differently called as the 'poverty line' and is a margin which represent the breaking point among poor and non poor. This is called the absolute margin of poverty. The absolute margin of poverty is 4.891 Leke per capita per month (Poverty Assessment` Report No. 26213-AL).

The relative margins of poverty reflect the absence of which a household or an individual suffers in relation with the incomes of the other part of population, so it is a social phenomenon. Generally, in poor countries, the absolute margins of poverty are used. So, in Albania the absolute margin of poverty is used.

In terms of poverty measures, the standard decomposable indicators in the Foster-Greer-Thorbecke (FGT) family are used. These include: a) the incidence of poverty, or headcount ratio, reflecting the percentage of population whose per capita consumption falls below the poverty line; b) the depth of poverty, measured by the poverty gap index, quantifying the average income shortfall of the poor in relation to the poverty line; and c) the poverty gap squared, a measure of severity of poverty.

In relation with the techniques that are used for the evaluation of poverty in Albania, in a study by Carlo Azzari, the results and trends of poverty obtained from the linear log regression model and from the methods of principal components are compared.

Instat has published in 2006 an article in which the linear regression model is used and the results are compared with those of 2002 and 2005.

One recent study related with the factors that influence poverty in Albania published by Camilla Mastromarco *et al.* in 2010, explains the factors that are strongly related to the deprivation condition for 2005: the geographic zone, kind of work, level of education, household size and the absence of migration experience.

A study in relation with the determining of poverty using logistic regression is conducted by Thomas N O Achia *et al.* In 2010, the results of which shows that the determining poverty factors in Kenya are the educational level of the head of household, rural zone, the age of the head of the household and the ethnicity.

From the recent study of Tshediso Joseph Sekhampu in 2010 for South African Township confirms that the age of the household head, the employment status, household size, has a strong influence, whereas the gender of the household, his/her marital status are not important in the explanation of economic status of household.

Methods

The survey includes a sample of 3600 household that make the observation unit representative for 4 zones: Center, Coast, Mountain, Tirana.

The poverty threshold is calculated using the basket of goods that is consumed from the individuals from the second to the fourth deciles. Keeping in mind the FAO recommendations for the minimal consumption of calories according to age and gender, and fitting these requirements per calories according to the distribution of population in Albania in 2001, the calories intake necessary per capita is estimated 2,288 calories/day. The non-food component of poverty is estimated taking into consideration the percentage of expenditure for non-food articles of those household that spend for food articles a value near the poverty threshold for food. Estimated this way, the poverty threshold for food (or the extreme poverty threshold) is set to 3,047 Lekë per capita/month, whereas the total poverty threshold, together with the needs of non food articles is set to 4,891 Lekë per capita/month with constant costs (2002). For the estimation of the complete poverty threshold, in the beginning it has been estimated a poverty threshold for food (or for the cost of a minimal calories intake) and then is added in order to include non food base articles that are considered minimally needed to survive. The non food part is calculated as an average of the non food expenditure of those household that spend almost the same amount for food, as defined in the poverty threshold.

Linear regression model

A preliminary analysis for searching the relation of expenditure of consumption with a series of other variables is conducted. Positively is observed the multiple regression model, the relation of logarithm of expenditure of consumption with other explanatory variable that are continuously or categorical. The model shows how the demographic data of the household, the educational level of the household head influence in the dependent variable.

$$\log y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i \quad (1)$$

y_i is the household expenditure of consumption. For every household there are different expenditure of consumptions y_i ,

x_1, x_2, x_3, x_4 are the independent variables

β_0 is a constant

β_j are the coefficients of x variables, the coefficients that are to be estimated.

ε_i being a random normally distributed error term with an average of 0 (J. Haughton et.al 2009)

These independent variables mentioned above, are due to the selection procedure stepwise and the variables which don't have an explanatory power for their statistic strength are deleted. All the calculations are made using the statistical software SPSS.

Logistic regression

Logistic regression is part of a category of statistical models called generalized linear models. Logistic regression allows the prediction of a discrete outcome such as group membership from a set of explanatory variables that may be continuous, discrete, and dummy or a mixture of these.

Logistic regression based on what connection there is a logistic (sigmoid dependence) between the probability of group membership and independent variables. It is assumed that observations are independent. Generally when we two groups we use binary logistic regression, when we have three more sets we use multiple logistic regression, nominal or ordinal.

Generally, dependent variable or the result is a dichotomy, for example presence / absence or success / failure.

Logistic regression analysis does not require that the data have a normal distribution with multi-dimensional dispersion and covariance equal for all variables.

For logistic regression model used in this study, household expenditures per capita are considered dependent variables. These were codified in the poor (1) and non-poor (0). Families with per capita consumption expenditure less than the poverty line are considered poor and those with costs greater than the poverty threshold are considered non-poor. For the determination of economic status y^* is the dependent variable defined from the regress:

$$y^* = \sum x_i \beta_i + u_i \tag{2}$$

Where $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ $x_i = (1, x_{i2}, x_{i3}, \dots, x_{ik})$

y^* is expressed by a variable dummy

$$y = 1 \text{ if } y^* > 0, \text{ and} \tag{3}$$

$$y = 0 \text{ otherwise}$$

From the equations (1) and (2) derive that :

$$P(y_i = 1) = P(u_i > -\sum x_i \beta) = 1 - F(-\sum x_i \beta) \tag{4}$$

Where F is the cumulative distribution function for u_i

$$P(y_i = 0 | \beta, x_i) = F(-\sum x_i \beta) \tag{5}$$

The variable y_i with the probability according to the respective formula (4) and (5) takes just two values 0 and 1 and is a case variable with Bernuli distribution. Variable y_i are independent because the observations are independent, so, the likelihood function is given with the below formula:

$$L = \prod_{i=1}^n [1 - F(-\sum x_i \beta)]^{y_i} F[-\sum x_i \beta]^{1-y_i} \tag{6}$$

The functional form imposed on F in equation (6) depends on the assumptions made about in equation (2). It is supposed that the cumulative distribution of u_i is logistic we have the logistic model. In this case is obtained:

$$F(-\sum x_i \beta) = \frac{e^{-\sum x_i \beta}}{1 + e^{-\sum x_i \beta}} \tag{7}$$

$$1 - F(-\sum x_i \beta) = \frac{e^{\sum x_i \beta}}{1 + e^{\sum x_i \beta}} \tag{8}$$

Where :

x_i are the characteristics of household (educational level of the household head, household size, the area, the gender of the household head).

β , the respective coefficients of these variables in the logistic regression.

Equation (7) gives the probability for a household being non poor ($P(y_i=0)$).

Equation (8) gives the probability for a household to be poor. ($P(y_i=1)$)(Maddala, G.S., 1983).

Results

Explanatory variables (independent):

1. The family size. Number of household members
2. Educational level of the household head.

Education that household head has committed (elementary, secondary school, middle school, professional, university).

3. Male or female household head: male = 1, female = 0.
4. Living area: urban = 1

Table 1. The coefficients of the logistic regression

	β	Wald	p
-Zone	0,094	100,909	0,000
-Education of household head	-0,350	6338,956	0,000
-Female head	-0,381	560,004	0,000
-Household size	0,548	41092,929	0,000
-Constant	-4,146	61346,523	0,000

If we look at the table we see that household size is a statistically significant variable (B = 0.548 Wald = 41092,929 p =0.000) and its positive coefficient indicates that with increasing household size increases the probability that the household be poor.

Education is one of the determinants of human capital in every country of the world. The quality of education depends on the number of persons with higher level of education and training. In this present study is the education level of the household head from primary education to higher education. If we look at the education level of the household head say is a statistically significant variable (B = -0.350 Wald = 6338,956 p =0.000) and its negative coefficient indicates that increased education has a significant impact in reducing the probability of being poor. Female-headed households variable is an important factor in explaining the economic status of the family (B=-0,381 Wald=560,004, p=0.000) but the negative coefficient indicates that households headed by female have lower probability of being poor than male-headed households.

Logistic regression model presented above reflects that families living in rural areas have a higher probability of being poor than families urbane. The variable rural/ urban areas is statistically significant (B =0.094, Wald = 100.909, p =0.000).

Values of statistics Cox Snell R-square and Nagerlelke R square parameters are 0.11 and 0.22 which indicate that the model explains 11% to 22% of the variance in the data. This low value is explained primarily by the fact that the main variable affecting the poor are household income, this variable did not have entered in the model and secondly the number of variables is not too large.

Table 2. The coefficients of the regression model

Depend variable: log of consumption		
	Coefficients (β)	t- statistics
Highest level	0,038	215,083
Household size	-0,003	-10,862
Male head	-0.021	-25,691
Zone	-0,022	40,943
R- squared		0,072

All the variables of which t value >2 , are statistically significant and t value <2 , are not statistically significant

From the results, the educational level of the household head is positive, which means that this variable influence in the increasing of the expenditure of consumption. Whereas the households headed by males and the household size are negative coefficients in the expenditure of consumption. So, with the increasing of household size, there is a decreasing in the expenditure of consumption. Also, the last variable that influence is the region, which is characterized by a negative coefficient. The urban zone has higher expenditure of consumptions compared to the rural zones.

Conclusion

The results shows that the household size, the educational level and gender of the household head, the zone (urban or rural) are variables statistically significantly important for the explanation of economic status as well as expenditure of consumption. So, the educational level and the household headed by females, reduce the probability of being poor, whereas the household size increases it.

The same results are confirmed by the linear log regression model in which the dependent variables are the expenditure of consumption per capita. It is observed a decrease in the expenditure of consumption with an increase of the household size. The increase of the educational level of the household head reflects in a increasing of the expenditure of consumption. The households that live in the rural zones have lower expenditure of consumption than those in urban zones. The households headed by females have higher expenditure of consumption than those headed by males. This study can be a reference when different social policies of increasing the educational level for decreasing poverty are undertaken in the Republic of Albania.

References:

- Albanian Institute of Statistics www.instat.gov.al
- Achia T, Wangombe A, Khadioli N 2010. A logistic regression model to identify key determinants of poverty using demographic and health survey data. *European Journal of Social Sciences*, 13(1): 38-45
- Ahmed, A., H. Bouis., 2003. Weighing what's practical: Proxy means test for targeting food subsidies in Egypt, Disc. Pap. 213, International Food Policy Research Institute, Washington, D.C
- Azzarri, C., G. Carletto, and A. Zezza (2006), "Monitoring Poverty Without Consumption Data. An Application Using the Albania Panel Survey". *Eastern European Economics*, 44, pp. 59-82
- Foster, J., J. Freer, and E. Thorbecke (1984), "A Class of Decomposable Poverty Measures". *Econometrica*, 52, pp. 761-766.
- J. Haughton, Sh.R. Khandeker, 'Handbook on poverty and inequality', ISBN: 978-0-8213-7613-3
- Maddala, G.S., 1983, *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, pp 9-23 ISBN 0 521 33825 5
- Mastromarco C, Peragine V, Russo F, Serlenga L, (2010). *Poverty, Inequality and Growth in Albania: 2002 - 2005 Evidence*, February 2010
- Mwabu, Germano, Wafula Masai, Rachel Gesami, Jane Kirimi, Godfrey Ndeng'e, Tabitha Kiriti, Francis Munene, Margaret Chemngich and Jane Mariara, 2000, *Poverty in Kenya: Profile and Determinant*', Nairobi: University of Nairobi and Ministry of Finance and Planning.
- Rodriguez, A.G., and S.M. Smith (1994). A comparison of determinants of urban, rural and farm poverty in Costa Rica. *World Development*, 22 (3), 381-397.
- Sekhampu Tshediso Joseph Determinants of poverty in South African Township, *Journal of Social Science* 34(2):145-152(2013)
- World Bank ``Albania trends in poverty and inequality 2002-2005
- World Bank ``Poverty Assessment`` Report No. 26213-AL