Current Research Journal of Social Sciences 2(3): 174-180, 2010
ISSN: 2041-3246

Submitetd Date: February 28, 2010 Accepted Date: March 26, 2010 Published Date: June 15, 2010

Education Levels and Household Location Preferences: A Case Study of Turkey

[1]Nüket Kırcı Çevik, [2]Emel Kuruoğlu and [3]Şenay Üçdoğruk
[1]Department of Econometrics, Institute of Social Sciences,
[2]Department of Statistics, Faculty of Arts and Sciences,
[3]Department of Econometrics, Faculty of Economics and Administrative Sciences,
Dokuz Eylul University, Izmir, Turkey

Abstract: Metropolitan cities are attractive household locations for individuals due to the social and job opportunities they provide. This study investigates the effect of university education on individuals' preference of metropolitan cities as their household location and intends to suggest the internal relationship between university education and individuals' preferences of household location. The bivariate probit model and the multinomial logit model are used so as to describe the relationship in question and the findings obtained demonstrate that university education is effective on individuals' preference of metropolitan cities as their household location in Turkey.

Key words: Bivariate probit model, education level, household location

INTRODUCTION

Many studies have been carried out about the fact that the individuals with a high education level prefer to live in metropolitan cities. There is considerable literature on the influence of settlement in metropolitan cities on income (Muth, 1969; Wheaton, 1977). Margo (1992) showed that half of the households living in metropolitan cities in England between 1950 and 1980 had high incomes. In their study, Costa and Kahn (2000) demonstrated that metropolitan cities created more considerable job opportunities and were more attractive for employees with a university education. Glaeser and Mare (2001) stated that although employees earned more in metropolitan cities, this was not more than their investments in their education. Glaeser (1998) suggested that the high income in metropolitan cities have a complete balance with the expensiveness of life; Asiyanbola (2006) suggested that education level, occupation, marital status and cultural structure were important components in the decisions by Nigerian women about choosing metropolitan cities as their household location and Sander (2004) examined the positive relationship between the education level of individuals in England and households' preferences of household location by the bivariate probit model and displayed the existence of the relationship between education level and household location preference. In addition, while estimating the relationship between households' preferences of household location in England and individuals' having a university education, Sander (2005) also considered whether the household belonged to any ethnic groups or not.

This study has investigated the factors that affect the household location preferences of individuals at the ages ranging between 25 and 47 in Turkey and the effect of individual's education level on household location preference has been investigated in particular and it has focused on whether there is a causal relationship between this preference and individual's having a university education. In order to support the findings obtained from the bivariate probit model, the multinomial logit model has been used and political comments have been made concerning the current socioeconomic structure of our country depending on the results found.

MATERIALS AND METHODS

This study was conducted with the 2003 data obtained from Turkish Statistical Institute (TSI) in Turkey. In the study, the definitions of regional variables have been made according to "the Nomenclature of Territorial Units for Statistics" (NUTS) performed by the TSI. The NUTS consists of territorial definitions in five steps. However, only the first three steps of it have been created in Turkey. According to this classification, each of 81 provinces is named as level-3 unit in the lowest step; some provinces are combined and 26 level-2 units and, at the highest step, 12 level-1 units have been created. The dependent and independent variables used in

Corresponding Author: Nüket Kırcı Çevik, Department of Econometrics, Institute of Social Sciences, Dokuz Eylul University, Izmir, Turkey

Fig. 1: The first regional definitions of units on level-1 (4 regions)

Fig. 2: The second regional definitions of units on level-2 (5 regions)

our study have been performed with two different definitions of units via NUTS. The first regional definitions of units are made by dividing level-1 units into four distinct regions. These regions have been referred to as the first and second western regions and the first and second eastern regions by us and they are independent variables in our models. The individuals in the western regions have a higher welfare level.

Figure 1 elaborates the first regional definitions of units, the First Western Region is composed of Istanbul, Western Anatolian and Eastern Marmara Regions; the Second Western Region is composed of the Aegean, Western Marmara and Mediterranean Regions; the First Eastern Region is composed of Western Black Sea, Central Anatolian and Eastern Black Sea Regions; and the Second Eastern Region is composed of the Southeastern Anatolian, middle-eastern Anatolian and Northeastern Anatolian Regions. This scope is composed by us.

In the second regional definitions of units, five distinct regions of preference have been composed from level-2 units. These regions range between the first to the fifth regions of preference and they have been used as dependent variables in the multinomial logit model.

Figure 2 shows that the socioeconomic development level of regions decreases as it is progressed from the first region of preference towards the fifth region of preference. The western regions are the centers for migration from the eastern regions in Turkey due to the social and economic opportunities they provide.

"The Research on Socioeconomic Development Level Ranking of Provinces and Regions" (RSDLR), prepared in 2003 by the State Planning Organization (SPO), has been taken as the basis while defining a metropolitan city. According to this research, Istanbul, Ankara and Izmir Subregions, which are socioeconomically the most developed and Bursa and Kocaeli Subregions (all level-2 units) have been named as the First Region of Preference (metropolitan city) respectively by us and, more comprehensively, all provinces ranking in the first ten have been named as the Comprehensive Region of Preference (metropolitan city) again by us. Both variables are dependent variables in the bivariate model. Furthermore, both units have been used as metropolitan cities in the two different models (Table 1).

Table 1: Scope of regions of preference (Dependent variable-model 1 and model 2)

	First regional definitions of units on level-2	The first 10 subregions sequentially according to RSDLR	Provinces covered
Comprehensive region of preference (It covers the first region of preference and the second region of preference together (Metropolitan city))	The first region of preference (Metropolitan city)	Istanbul Subregion	Istanbul
		Ankara Subregion	Ankara
		Izmir Subregion	Izmir
		Bursa Subregion	Bursa, Eskişehir, Bilecik
		Kocaeli Subregion	Kocaeli, Yalova, Bolu, Sakarya, Düzce
	The second region of preference(Metropolitan city)	Tekirdağ Subregion	Tekirdağ, Kırklareli, Edirne
		Adana Subregion	Adana, Mersin
		Aydın Subregion	Denizli, Muğla, Aydın
		Antalya Subregion	Antalya, Isparta, Burdur
		Balıkesir Subregion	Balıkesir, Çanakkale

Table 2: Descriptive statistics table

Variables	Frequency (%)
Being a university graduate	1278 (0.061)
Household location in the first western region	5806 (0.277)
Household location in the second western region	7231 (0.345)
Household location in the first eastern region	4261 (0.203)
Household location in the second eastern region	3675 (0.175)
Household location in the first region of preference	6098 (0.291)
Household location in the second region of preference	4446 (0.212)
Household location in the third region of preference	3691 (0.176)
Household location in the fourth region of preference	3439 (0.164)
Household location in the fifth region of preference	3313 (0.158)
Household location in the comprehensive region of preference	10542 (0.503)
Being insured	13129 (0.626)
Married	18248 (0.870)
Male	10115 (0.482)

According to the data of the TSI, Istanbul is found out to be the greatest receiving region whereas the Western Black Sea is found out to be the greatest sending region. Intensive migrations resulting from the obvious difference among the interprovincial repulsive and attractive factors create great population pressure and cause settlement/urbanization problems that affect all spatial units nationwide (Pazarlıoglu, 2007). The settlement problems that are becoming widespread at country scale change quality in time and influence economic, social and political structures as well and deepen. Various insufficiencies of service occur in metropolitan cities that grow by migration waves and the differences in interregional socioeconomic development become a problem required to be corrected not only in terms of backward regions but also in terms of developed regions.

In the application, the data of "2003 Household Income Survey" were used and the units of level-1 and level-2 NUTS by the TSI were taken as basis. Moreover, the". The reason why the data of 2003 Income Survey are used in our study is that the RSDLR prepared by the SPO in 2003 was used in defining the subregions as "metropolitan cities information on the household location of individuals is obtained on the basis of level-1 and level-2 units only in the survey of the year concerned. The TSI did not ask any questions about the units concerned in the surveys it carried out during 2004 and 2006. This is a shortcoming of our study. Although the TSI has not conducted any activities on level-3 basis yet due to high costs, it has announced that such an activity will be conducted in 2009. We aim at comparing the results of this study of ours with a second study by using the data of this survey to be performed in 2009. The sample size in 2003 Household Income Survey was determined as 107.615 people, sample ages were limited to 25, the age of graduation from university, to 47, the mean retirement age according to the retirement law in our country, since the preferences by university graduates of living in metropolitan cities are investigated in the study. The descriptive statistics of the respondents of the survey are given in Table 2. According to the results in Table 2, the ratio of university graduates is 6% in our case study with a volume of 20.973 and 42% of them live in the First Region of Preference while 62% live in the Comprehensive Region of Preference. The ratio of university graduates, who live in the remaining 16 subregions, is only 38%. Furthermore, 94% of the university graduates are insured employees/employers. Besides this, 29% of the respondents live in the first region of preference, 21% live in the second region of preference, 17% live in the third region of preference, 16% live in the fourth region of preference and finally 15% live in the fifth region of preference. The ratio of the respondents living in the first and second regions of preference or, in other words, in the comprehensive region of preference is 50%. However,

when the regional definitions of eastern and western regions referred to as the first regional definition of units is taken into consideration, 62% of the respondents live in western regions whereas 38% live in eastern regions.

The other variables used in the models are defined as follows: If the first region of preference, the dependent variable of Model 1, is one of the regions of the household location of an individual, it is the dummy variable which takes on the value of one. If it is not, it is the dummy variable, which takes on the value of zero. If the comprehensive region of preference, the dependent variable of Model 2, is one of the regions of the household location of an individual, it is defined as the dummy variable, which takes on the value of one. If it is not, it is the dummy variable which takes on the value of zero: University: It is the dummy variable that takes on the value of one for the individuals who are graduates of a four-year faculty, graduate degree and doctorate degree, and it takes on the value of zero for the individuals having other educational positions. Insurance: The variable of insurance is the dummy variable that takes on the value of one for compulsorily insured individuals, voluntarily insured individuals and individuals having both while taking on the value of zero for uninsured individuals with a green card. Marital status: Marital status is the dummy variable that takes on the value of one for married individuals and that takes on the value of zero for the individuals who have never got married, who live together, whose spouse has died, who have divorced, and who have separated.

The bivariate probit model has been used as the estimation method in the study. Like the seemingly unrelated regression model, the bivariate probit models allow two equation systems with related error terms. A system is organized here so as to express binary decisions. The simultaneous determinants of binary dependent variables and the causal relationship between them can be suggested with this technique. For instance, if we assume that education level and household location preference are simultaneously related, the general specifications for a two-equation model can be written as follows (Greene, 2003):

$$\begin{aligned} y_1^* &= x_1'\beta_1 + \varepsilon_1 \\ y_2^* &= x_2'\beta_2 + \varepsilon_2 \end{aligned} \qquad (1)$$

Where if $y_1^* > 0$ $y_1=1$;otherwise, it takes on the value of 0. Likewise, if $y_2^* > 0$ $y_2=1$; otherwise, it takes on the value of 0.

In addition,

$$E[\varepsilon_1 | X_1, x_2] = E[\varepsilon_2 | X_1, x_2] = 0$$

$$Var[\varepsilon_1 | X_1, x_2] = Var[\varepsilon_2 | X_1, x_2] = 1 \text{ and}$$

$$Cov[\varepsilon_1, \varepsilon_2 | X_1, x_2] = \rho \text{ in the model}$$

The Lagrange Multiplier (LM) test is used in the model in order to test the existence of the correlation between two variables. In the test method, where the null hypothesis is established as ρ = 0 (there are no relations between education level and household location preference), the model including the independent probit equations is estimated separately and the LM test statistics is obtained in the following way:

$$LM = \frac{\left[\sum_{i=1}^{n} q_{i1} q_{i2} \frac{\phi(w_{i1})\phi(w_{i2})}{\Phi(w_{i1})\Phi(w_{i2})}\right]^2}{\sum_{i=1}^{n} \frac{\left[\phi(w_{i1})\phi(w_{i2})\right]^2}{\Phi(w_{i1})\Phi(-w_{i1})\Phi(w_{i2})\Phi(-w_{i2})}} \qquad (2)$$

Both dependent variables used in the application, namely, the graduation of an individual from the university and whether he/she lives in either of the first region of preference or the comprehensive region of preference, are the bivariate dependent variables. The reason why the above-mentioned bivariate probit model is used as the method is to find out the simultaneous equations of these binary dependent variables and to suggest the relationship between them.

RESULTS AND DISCUSSION

The results obtained from the bivariate model applied for the model of household location preference are given in Table 3. The dependent variables are the individuals living in the first region of preference in Model 1 and the individuals living in the comprehensive region of preference in Model 2 as a second scenario. The independent variables used in the study are gender, age, age*age, marital status, whether the individual is insured or not while the dummy variables are composed of the region composed of the first and second western regions and the first and second eastern regions that we have been divided depending on development level. In Model 1, where the individuals, whose household locations are in the first region of preference and who are university graduates in the Eq. (2), are the dependent variables, it is observed that the factors that influence both dependent variables are the same. The state of being a university graduate has been used as an independent variable in the Eq. (1) whereas it has been used as a dependent variable in the Eq. (2). In addition, the fact that the individual is insured has been used as the instrumental variable in the Eq. (2). Evans and

Table 3: Region of preference and university graduation results of bivariate probit model

	Model 1 Dependent variables		Model 2 Dependent variables	
Independent Variables	First Region of preference	University graduation	Comprehensive Region of preference	University graduation
Constant	-0.279	-2.021[a]	-0.592[c]	-1.952[a]
Being a University Graduate	1.300		1.307[a]	
Gender (Male)	-0.037[c]	0.377[a]	-0.707[a]	0.370[a]
Age	0.071[a]	0.003[c]	0.084[a]	-0.0001[c]
Age*Age	-0.0001	-0.0001[c]	-0.001[a]	-0.000[c]
Marital Status (Married)	-0.067[b]	-0.561[a]	-0.052[c]	-0.548[a]
Second Western Region	-1.933[a]	-0.122[a]	-0.182[a]	-0.116[a]
First Eastern Region	-8.070[a]	-0.294[a]	-8.523[a]	-0.294[a]
Second Eastern Region	-8.095[a]	-0.077[b]	-8.497[a]	-0.078[b]
Insurance		1.126[a]		1.106[a]
The Wald Test H_0: $\rho=0$	$\chi2(1) = 35.633$ [0.000]		$\chi^2(1) = 77.078$ [0.000]	
N	20973		20973	

[a], [b] and [c] show 5%, 10% and 15% statistical significance, respectively. Base category: the first western region.

Table 4: Results of the Multinomial Logit Model

	Dependent Variables			
Independent Variables	Second region of preference R.R.R	Third region of preference R.R.R	Fourth region of preference R.R.R	Fifth region of preference R.R.R
Being a University Graduate	0.662[a]	0.527[a]	0.657[a]	0.870
Gender (Male)	1.010[c]	1.018[c]	0.986[c]	1.000[c]
Age	1.012[c]	0.860[a]	0.906[a]	0.853[a]
Age*Age	0.999[c]	1.002[a]	1.001[a]	1.002[a]
Marital Status (Married)	1.067[c]	1.399[a]	1.487[a]	1.253[a]
Insurance	0.810[a]	0.692[a]	0.604[a]	0.252[a]

[a], [b] and [c] show 5%, 10% and 15% statistical significance, respectively
Base category: It is the first region of preference.
R.R.R.: Relative Risk Ratio

Table 5: Results of the hausman test

Categories	χ^2 (degree of freedom)	p-value	Result
2nd Region of Preference	-5.548 (18)	1.000	H_0 cannot reject
3rd Region of Preference	0.903 (18)	1.000	H_0 cannot reject
4thRegion of Preference	1.143 (18)	1.000	H_0 cannot reject
5th Region of Preference	1.703 (18)	1.000	H_0 cannot reject

Montgomery (1994) stated that some health habits could be used as the instrumental variable for education level in their study where they investigated the relationship between individuals' education level and health condition. Fairlie (2005) simultaneously estimated the school enrollment of the child and his probability of owning a computer, and used the bivariate probit model as the estimation model. Sander (2004) chose smoking as the instrumental variable in his study where he investigated the relationship between the household location preferences of households and the education level of individuals. Bivariate probit models are synonyms of instrumental variables or two-step least-squares model. Thus, the choice of instrumental variable is of extreme importance in bivariate probit models. The instrumental variable to be used in the models should be highly correlated with the dependent variable (Fairlie, 2005). Since whether the individual is insured or not is considered to be highly correlated with education level in this study as well, this variable is chosen as the instrumental variable. In the Eq. (1), the factors influencing the individual's settlement in the first region of preference, namely one of the metropolitan cities, are being a university graduate, gender, age, marital status and the settlement of the individual in the second western region, the first and second eastern regions. According to the data in Table 3, it is observed in the Eq. (1) that being a university graduate and age have a significant and positive effect on household location in the metropolitan city while marital status has a significant but negative effect. A negative and significant relationship has been found between household location preference in the first region of preference and the household location in the second western and eastern regions. Depending on the first regional definitions of units, the first western region is the region with the highest socioeconomic development ranking whereas the second eastern region is the region with the lowest socioeconomic development ranking. Therefore, when the coefficient estimates of these regions are examined in Model 1, the likelihood of individuals' settlement in the first region of preference, namely the metropolitan city, decreases when progressed from the west to the east and as the socioeconomic development ranking of the region

decreases. In other words, the individuals, whose household location is in regions with a low socioeconomic development ranking, are not in the first region of preference since the settlement has not been able to be qualified as a metropolitan city yet and is not ready to employ qualified employees. Actually, as it is also observed in the Eq. (2), the settlers in the second western and eastern regions are observed to have a low education level. In other words, we can state that the individuals, whose household location is in the first western region, have a more qualified education. In this equation, the coefficient estimates for the variables of age and age*age have been found to be statistically insignificant. It is estimated that the coefficient estimates are found to be insignificant since the study has been carried out with the active population in the models. Males have a positive but being married has a negative effect on receiving university education. The positive and significant effect of being insured on being a university graduate has been put forth in the Eq. (2). The common result to be obtained from both equations is that household location in the second western and eastern regions has a simultaneously negative effect on both the household location preference in metropolitan cities and on university education. Our indirect result is that the individuals with a university education generally prefer to live in metropolitan cities. The Wald test indicates a significant relationship that is different from zero between the error terms of both equations and displays the existence of the simultaneous relationship. The dependent variable in Model 2 is composed of the individuals, whose household location is in the comprehensive region of preference, and the results obtained from Model 2 are quite similar to the results obtained from Model 1. The coefficient estimates obtained from Model 2 are not interpreted in order to save space since they have been found to be very close to the coefficient estimates obtained from Model 1. With a view to supporting the results of the bivariate probit model, the multinomial logit model has been estimated for the household location preference and the results are shown in Table 4. The regions of preference created in the second regional definitions of units have been used as the dependent variables of this model. The definitions of region used here are the same with the definitions used previously in Model 1. First of all, the Hausman test has been applied for in order to find out independence among categories. According to the results of the Hausman test in Table 5, the null hypothesis that puts forward independence among categories has been not rejected (Greene, 2003). Thus, the multinomial logit model can be applied.

According to the results of the multinomial logit model in Table 4, a significant relationship has been found between individuals' preferences of household location in the second, third, fourth and fifth regions of preference and being university graduates. In addition, the existence of a similar relationship has been put forth between household location preference in these regions and being insured. In terms of all regions, it has been found out that gender does not have any effects on household location preference in these regions. Additionally, it has been found out that age and marital status do not have any effects on household location preference in these regions either in the second region of preference. Being married has a significant effect on household location preference in the third, fourth and fifth regions of preference. The results obtained from this model support the results of the bivariate probit model shown in Model 1. The results of the bivariate probit model have demonstrated that individuals with a university education prefer the regions that rank among the top in socioeconomic development ranking, namely, the "metropolitan cities" as household location and that there is a relationship between university education and household location preferences.

The factors affecting university education and household location preferences of individuals at the ages between 25 and 47 in general in Turkey have been examined by means of the bivariate probit model and the multinomial logit model in this study. University education, gender, age, marital status, individual's insurance status and the regional dummy variables have been considered as the base category. In the two-variable probit model, it has been found that university graduation has a significant and positive effect on preferring the comprehensive region of preference as the household location and, the negative effect of living in eastern regions on university education has been found out. The metropolitan cities in Turkey are located in the first and second western regions and continue to grow as the center of migration from the eastern regions. It is possible to state that the individuals, who prefer metropolitan cities as household location, have a high education level and that these regions attract the educated population not only from eastern regions but also from western regions. Individuals with a high education level migrate in order to find a suitable job for their education. It should be taken into account that university education is a determinant in individuals' household location preference and that education increases the phenomenon of urbanization. It gains vital importance to plan and encourage regional investments in such a way that they will eliminate regional differences and that balance will be established in parallel to the proportional increase of individuals. Otherwise, it will not be possible to prevent irregular urbanization, environmental problems and the increase in crime rate, which are the basic urban problems, and unemployment and socio-political instabilities will be unavoidable.

REFERENCES

Asiyanbola, A.R., 2006. Residential location decision-making and gender in Nigeria. Geoforum, 37: 1059-1065.

Costa, D.L. and M.E. Kahn, 2000. Power couples: changes in the locational choice of the college educated 1940-1990. Q. J. Econ., 115: 1287-1315.

Evans, W.N. and E. Montgomery, 1994. Education and health: where there's smoke there's an instrument. Working Paper No: 4949. National Bureau of Economic Research, University of Maryland.

Fairlie, R.W., 2005. The effects of home computers on school enrollment. Econ. Edu. Rev., 24: 533-547.

Glaeser, E.L., 1998. Are cities dying. J. Econ. Perspect., 12: 139-160.

Glaeser, E.L. and D.C. Mare, 2001. Cities and skills. J. Labor Econ., 19: 316-342.

Greene, W., 2003. Econometric Analysis. Prentice Hall, New Jersey.

Margo, R.A., 1992. Postwar suburbanization of population in the United States: The role of income. J. Urban Econ., 31: 301-310.

Muth, R., 1969. Cities and Housing. University of Chicago Press, Chicago.

Pazarlioglu, M.V., 2007. Izmir orneginde ic gocun ekonometrik analizi. Yönetim ve Ekonomi, 14(1): 121-136. (In Turkish)

Sander, W., 2005. On the demand for city living. J. Econ. Geography., 5: 351-364.

Sander, W., 2004. Education and urban household location. Econ. Lett., 82: 327-331.

SPO (Devlet Planlama Teşkilatı), 2003. İllerin ve bölgelerin sosyo-ekonomik gelişmişlik sıralaması araştırması. Yayın No: 2671. (In Turkish)

Wheaton, W., 1977. Income and urban residence: An analysis of consumer demand for location. Am. Econ. Rev., 67: 620-631.