SDU JOURNAL OF SCIENCE (E-JOURNAL). 2009, 4(2), 199-212

A COMPARISON OF MULTINOMIAL LOGISTIC AND MULTINOMIAL CONDITIONAL LOGISTIC REGRESSION MODELS IN ASSESSING EUROPEAN UNION MEMBERSHIP

Yüksel Akay ÜNVAN[*], Gamze ÖZEL[**]

[*]Export and Import Bank of Turkey, 06100, Bakanlıklar, Ankara, TURKEY
e-mail: aunvan@eximbank.gov.tr
[**] Hacettepe University, Department of Statistics, 06800, Beytepe, Ankara, TURKEY
e-mail: gamzeozl@hacettepe.edu.tr

Received: 26 June 2009, Accepted: 9 October 2009

Abstract: Multinomial logistic (ML) and multinomial conditional logistic (MCL) regression models are used for modeling the relationships between a polytomous response variable and a set of explanatory variables. In this study, key factors affecting the European Union (EU) membership process are determined using ML and MCL models. We compare the ML and MCL models and argue that MCL is more preferable than the more complex ML model. Then for each candidate or potential candidate country, the probability of the accession time for the EU membership is predicted. The findings indicate that human development index, gdp per capita, exports of goods and services are important factors in determining which of the countries will join the EU and when they will do so. Furthermore, the probabilities of the accession time for both candidate and potential candidate countries are predicted as more than six years.

Key words: Multinomial Logistic, Multinomial Conditional Logistic, ROC curve, European Union, Enlargement

ÇOKLU LOJİSTİK VE ÇOKLU KOŞULLU LOJİSTİK REGRESYON MODELLERİ: AVRUPA BİRLİĞİ ÜYELİĞİ İÇİN MODELLERİN KARŞILAŞTIRILMASI

Özet: Çoklu lojistik (ÇL) ve çoklu koşullu lojistik (ÇKL) regresyon çözümlemesi modelleri, çok düzeyli cevap değişkeni ile açıklayıcı değişkenler kümesi arasındaki ilişkileri modellemek için kullanılmaktadır. Bu çalışmada, Avrupa Birliği (AB) üyelik sürecini etkileyen temel faktörler ÇL ve ÇKL modelleri kullanılarak belirlenmiştir. ÇL ve ÇKL modelleri karşılaştırılmış ve ÇKL modelinin daha karmaşık olan ÇL modeline tercih edileceği sonucuna ulaşılmıştır. Daha sonra, her aday ve olası aday ülke için AB'ye üyelik zamanı olasılığı tahmin edilmiştir. Bulgular, insani gelişim indeksi, kişi başına düşen gayri safi milli hasıla, mal ve hizmet ihracatının hangi ülkelerin hangi tarihte AB'ye katılacağını belirlemede önemli faktörler olduğunu göstermiştir. Ayrıca, hem aday hem de olası aday ülkeler için giriş süresinin altı yıldan daha uzun olacağı tahmin edilmiştir.

Anahtar kelimler: Çoklu Lojistik, Çoklu Koşullu Lojistik, ROC eğrisi, Avrupa Birliği, Genişleme

1. INTRODUCTION

The EU was created by six founding states in 1957 and has grown to 27 member states (Austria, Belgium, Bulgaria, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom). There have been five enlargements, with the largest occurring in 2004, when 10 new member states joined, and the most recent in 2007, when Bulgaria and Romania joined (ARCHICK & KIM, 2008). The EU has tended to enlarge along regional lines, adding groups of nearby nations. In 2005, the EU agreed to open accession negotiations with Turkey, Croatia and the European Council granted the status of candidate country for the Former Yugoslav Republic of Macedonia. Currently, the EU is very interested in the integration of the Western Balkan countries (Albania, Bosnia and Herzegovina, Montenegro, Serbia, Kosovo[1]), which are potential candidates[2].

Frontier analyses are widely employed to investigate the effects of enlargements on the old and new members of the EU. The benefits and costs of the EU enlargement were explained by BALDWIN et al. (1997), BREUSS (2002), and DOYLE & FIDRMUC (2003). NICOLADIES (2003) noted that acceding countries can become effective members of the EU. It was reported that their ability to derive the maximum benefits from the EU membership has depended on their success or failure in influencing nascent EU rules and increasing their economies. The impact of the fifth enlargement on the EU was investigated by BARYSCH (2003). It clearly shows that this impact was marginally positive but negligible, because the economies of the acceding countries were so small. On the other hand, the result of the study supported that the accession of Central and Eastern European countries will contribute to enlarge the internal market and stimulate economies of scale as a result of the growth of the size of the businesses. DRAGAN (2007) mentioned that Bulgaria and Romania could be a real added value for the EU but proper assistance could be needed for small businesses in both countries in order to fully express their potential.

Although many studies have already been done to assess the enlargement of the EU, little attention has been focused on the effects of factors on the EU membership and the accession process of candidate countries. If the effective factors for the EU membership are determined, a candidate or potential candidate country can take into consideration these factors to become a member country. In this study, the EU data has been examined to determine the effects of factors on the probability of the EU membership. Then, for a candidate or potential candidate country, the probability of the accession process of the EU membership in less than eight years is predicted. For this aim, Multinomial Logistic (ML) and Multinomial Conditional Logistic (MCL) models have been used and the results of the models are compared. In this study, ML and MCL models are preferred for modeling of the EU data since these models have not been

[1] Kosovo is not included to the study due to lack of statistical data.
[2] European Commisson, Enlargement, Potential Candidate Countries, htp://ec.europa.eu/enlargement/potential-candidate-countries/index_en.htm (accessed on May 18, 2009).

performed before for the EU data. To our knowledge, this is the first study to examine effective factors in the context of the EU membership.

This paper is organized as follows. In Section II, the methodology of this study is given and in Section III the EU data set to be used in this study is briefly described. In Section IV, emprical investigation for the importance of various factors in determining the EU membership is given. These factors are used to find the probability of the accession process in less than eight years for each candidate and potential candidate country. Finally, a summary and conclusions are presented in Section V.

2. MATERIAL AND METHODS

Multinomial logistic regression analysis is used when the dependent variable in question is nominal and consists of more than two categories. When using multinomial logistic regression analysis, one category of the dependent variable is chosen as the comparison category. Multinomial logistic regression analysis has been used by SHABBIR (1993) to study the determinants of occupational choice, by WOJCIK (2000) to study the number of automobiles demanded, by COCKBURN & DOSTIE (2007) to study the structure of asset portfolios of households, and by MARTINEZ et al. (2009) to study choices of transportation modes.

Consider an individual choosing among alternatives in a choice set. Let P_{jk} denote the probability that individual j chooses alternative k, Let $\mathbf{X}_j$ be a (p+1) vector representing the characteristics of individual j, and let $\mathbf{Z}_{jk}$ be the characteristics of the k*th* alternative for individual j. For example, $\mathbf{X}_j$ may be the GDP ratio of the j*th* country and each $\mathbf{Z}_{jk}$ be a EU membership time of the j*th* country and k*th* membership time period.

Multinomial logistic regression analysis focuses on the individual as the unit of analysis and uses individual characteristics as explanatory variables. The explanatory variables, being characteristics of an individual, are constant over the alternatives. The probability that individual j chooses alternative k is

$$P_{jk} = \frac{\exp(\boldsymbol{\beta}_k' \mathbf{X}_j)}{\sum_{l=1}^{m} \exp(\boldsymbol{\beta}_l' \mathbf{X}_j)} = \frac{1}{\sum_{l=1}^{m} \exp[(\boldsymbol{\beta}_l - \boldsymbol{\beta}_k)' \mathbf{X}_j]} \qquad (1)$$

where $\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_m$ are m vectors of unknown regression parameters. Since $\sum_{k=1}^{m} P_{jk} = 1$, the m sets of parameters are not unique. By setting the last set of coefficients to null (that is, $\boldsymbol{\beta}_m = \mathbf{0}$), the coefficients $\boldsymbol{\beta}_k = (\beta_{0k},...,\beta_{mk})'$ represent the effects of the $\mathbf{X}$ variables on the probability of choosing the kth alternative over last alternative. In fitting such a model, $m-1$ sets of regression coefficients are estimated.

MCL models give the researchers control over the restrictions imposed on the dependent variable for each explanatory variable in the model. This makes it possible to

include models for square tables in a multinomial logistic framework. MCL model has been used by WANNER (2005) to investigate the occupation choices of Canadians, by KAUSHAL (2005) to define new immigrants' location choices, by TAYLOR (2009) to investigate if the quantity of food safety information publicly available impacts the consumers' decision to purchase fresh meat and poultry. A SAS macro (containing %mclgen, %design, %intrct and %mclest) is written by HENDRICKX (1995) for multinomial conditional logistic regression analysis.

Multinomial conditional logistic regression analysis entails using the conditional logistic (CL) model. A limitation of the ML model is that it allows only one response function for all explanatory variables in the model. If more flexibility is required in the specification of response functions, then a CL model can be used to estimate the MCL model. In a MCL model, the explanatory variable **Z** assume different values for each alternative and impact of a unit of **Z** is assumed to be constant across alternatives. The probability that the individual j chooses alternative k is

$$P_{jk} = \frac{\exp(\boldsymbol{\theta}'\mathbf{Z}_{jk})}{\sum_{l=1}^{m}\exp(\boldsymbol{\theta}'\mathbf{Z}_{jl})} = \frac{1}{\sum_{l=1}^{m}\exp[\boldsymbol{\theta}'(\mathbf{Z}_{jl}-\mathbf{Z}_{jk})']} \tag{2}$$

where $\boldsymbol{\theta}$ is a (p+1) vector of regression coefficients and $\mathbf{Z}_{jk}$ as follows:

$$\mathbf{Z}_{j1} = \begin{bmatrix} \mathbf{X}_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{Z}_{j2} = \begin{bmatrix} 0 \\ \mathbf{X}_j \\ \vdots \\ 0 \end{bmatrix} \quad , \ldots, \quad \mathbf{Z}_{j,m-1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ m-1 \end{bmatrix}, \quad \mathbf{Z}_{jm} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where the $\mathbf{0}$ is a $(p+1)$ vector of zeros. The impact of a variable on the choice probabilities derives from the difference of its values across the alternatives. For example, **Z** may be the GDP ratio of j*th* EU member country and $\mathbf{P}_{jk}$ is the k*th* EU membership time of j*th* country.

The MCL model can be used to estimate McFadden's choice model or matched case-control data. In McFadden's choice model, variables characterizing the choices (i.e. the categories of the dependent variable in the ML model) are included. With matching, cases are matched with respect to certain characteristics. When the model does not include choice characteristics or matched cases, the likelihood function of the MCL model is equivalent to that of the ML model. Under these circumstances, the MCL model will produce the same coefficients, standard errors and log likelihood values as the ML model. However, the MCL model is much more flexible in allowing restrictions on the choices (the dependent variable in ML). In the MCL model, the main effects of the choice variables correspond with the intercept term in an ML model. Interactions between these choice variables and the explanatory variables correspond with the effects of these variables. The MCL model can be estimated using programs for the Cox

proportional hazard model (event history) as well as programs for conditional logistic regression models.

For MCL and ML models, the probabilities are predicted with same method. To calculate P_j from β, the back transformation is

$$\hat{P}_{jk} = \frac{\exp(x'_j \boldsymbol{\beta}_k)}{1 + \sum_{l \neq j} \exp(x'_j \boldsymbol{\beta}_l)} \tag{3}$$

For the non-reference categories $k \neq k^*$, and the reference (also called baseline) category probability is

$$\hat{P}_{jk^*} = \frac{1}{1 + \sum_{l \neq j} \exp(x'_j \boldsymbol{\beta}_l)} . \tag{4}$$

where $0 \leq \hat{P}_{jk} \leq 1$.

In order to discriminate the groups based on the cut-off points, the ROC (receiver operating characteristic) curve analysis is conducted for the ML and MCL models. The ROC curve is defined as a plot of test sensitivity as the y coordinate versus its 1-specificity or false positive rate (FPR) as the x coordinate.

3. APPLICATION

The data for this study is annual and obtained from EUROSTAT (2008), UNDP (2008), and WORLD BANK (2008). Countries are classified as full members and candidates in the EU as of 2008. The list of the 27 full members and 3 candidates are given in Table 1.

As seen in Table 1, the EU welcomed 10 more member countries (Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia) on 1 May 2004. On 1 January 2007, this latest round of the enlargement came to its conclusion with the accession of two more countries, Bulgaria and Romania. Recently, Turkey, Croatia and the Former Yugoslav Republic of Macedonia have been recognized as candidate countries. However, potential candidate countries (Albania, Bosnia and Herzegovina, Montenegro, Serbia and Kosovo) are not included in the Table 1 since the accession negotiations have not been started.

Before the analysis, the dependent variable and explanatory variables affecting the dependent variable must be defined. To determine the dependent variable, the accession processes of the member countries are taken into consideration and the accession time is defined by difference between the full membership date and the full membership application date for a member country in Table 1. Then, the classification of EU member countries according to the accesssion time (year) is presented in Table 2.

Table 1. The List of Full Members and Candidates of the European Union (EUROSTAT 2009 Classification)

Country	Full Membership Application Date	Starting Date of Negotiation	Membership Date
Belgium	Founder	-	25.03.1957
Czech Republic	01.01.1996	01.03.1998	01.05.2004
Denmark	01.05.1967	01.06.1970	01.01.1973
Germany	Founder	-	25.03.1957
Estonia	01.11.1995	01.03.1998	01.05.2004
Greece	01.06.1975	01.07.1976	01.01.1981
Spain	01.07.1977	01.02.1979	01.01.1986
France	Founder	-	25.03.1957
Ireland	01.05.1967	01.06.1970	01.01.1973
Italy	Founder	-	25.03.1957
Cyprus	01.07.1990	01.03.1998	01.05.2004
Latvia	01.10.1995	01.02.2000	01.05.2004
Lithuania	01.10.1995	01.02.2000	01.05.2004
Luxembourg	Founder	-	25.03.1957
Hungary	01.03.1994	01.03.1998	01.05.2004
Malta	01.07.1990	01.02.2000	01.05.2004
The Netherlands	Founder	-	25.03.1957
Austria	17.07.1989	01.02.1993	01.01.1995
Poland	01.04.1994	01.03.1998	01.05.2004
Portugal	01.05.1978	01.11.1978	01.01.1986
Slovenia	01.06.1996	01.03.1998	01.05.2004
Slovakia	01.06.1995	01.02.2000	01.05.2004
Finland	18.03.1992	01.02.1993	01.01.1995
Sweden	01.07.1991	01.02.1993	01.01.1995
United Kingdom	01.05.1967	01.06.1970	01.01.1973
Bulgaria	14.12.1995	01.02.2000	01.01.2007
Romania	22.06.1995	01.02.2000	01.01.2007
Turkey	14.04.1987	03.10.2005	Candidate
Croatia	21.02.2003	03.10.2005	Candidate
The Former Yugoslav Rep. of Macedonia	22.03.2004	17.12.2006	Candidate

Table 2. Accession time and number of the EU members

Accession Time	Number of EU Member
≤ 3 Years	7
4-6 Years	6
7-9 Years	8
10-12 Years	7
≥ 12 Years	4

Values of the dependent variable (membership time) are given in Eq. (5) and the accession time of the EU members is given in Table 3. The dependent variable is defined as follows:

$$Y = \begin{cases} 1, & \text{if the EU membership period is less than } \leq 3 \text{ years} \\ 2, & \text{if the EU membership period is 4-6 years} \\ 3, & \text{if the EU membership period is 7-9 years} \\ 4, & \text{if the EU membership period is 10-12 years} \\ 5, & \text{if the EU membership period is more than 12 years.} \end{cases} \quad (5)$$

In this study, the last category is the reference for ML and MCL models. After determining the dependent variable, explanatory variables, believed to have an effect on the dependent variable, are chosen suitable to Copenhagen criteria[3] and Maastricht Treaty[4].

Table 3. Values of dependent variable and accession time of EU members

Country	Accession Time (Year)	Y (Membership Time)
Belgium	0	1
Czech Republic	8	3
Denmark	6	2
Germany	0	1
Estonia	9	4
Greece	6	2
Spain	9	4
France	0	1
Ireland	6	2
Italy	0	1
Cyprus	14	5
Latvia	9	4
Lithuania	9	4
Luxembourg	0	1
Hungary	10	4
Malta	14	5
The Netherlands	0	1
Austria	6	2
Poland	10	4
Portugal	8	3
Slovenia	8	3
Slovakia	9	4
Finland	3	1

[3] The Copenhagen criteria are the rules that define whether a country is eligible to join the EU. This criteria require that a state have the institutions to preserve democratic governance and human rights, a functioning market economy, and that the state accept the obligations and intent of the EU.

[4] The Maastricht Treaty was signed on February 7, 1992 in Maastricht, between the members of the European Community. It led to the creation of the EU and was the result of separate negotiations on monetary union and on political union.

Sweden	4	2
United Kingdom	6	2
Bulgaria	12	5
Romania	12	5

Since multicollinearity may result in incorrect signs and magnitudes of the regression parameter estimates, correlations between explanatory variables were examined. Because of multicollinearity problem, high correlated explanatory variables were removed and 38 explanatory variables were used. These explanatory variables are collected into ten main groups as agriculture, economy and finance, education, health, labour market, life quality, population, priorities in public spending, status of major international human rights instruments, and trade. Variable definitions and descriptive statistics to be used in this study are listed in Table 4.

Table 4. Explanatory variable definitions and descriptive statistics (EUROSTAT 2008, Worldbank 2008, UNDP 2008)

Variable Group	Variable Name	Variable Definition	Mean	Standart Derivation
Agriculture	X_1	Share of population dependent on agriculture in total population (%)	10.448	9.545
	X_2	Share of agricultural GDP in total GDP (%)	2.396	1.377
	X_3	Share of agricultural land in total area (%)	42.346	20.645
Economy and Finance	X_4	GDP per capita (US $)	29,900	11,773
	X_5	National income per capita (US $)	16,811	11,553
	X_6	GDP per capita in Purchasing Power Standards (PPS), (EU-27 = 100)[5]	89	41
	X_7	Inflation rate (%)	1.984	5.223
	X_8	Tax rate (% of GDP)	35.884	8.841
	X_9	General government debt (% of GDP)	50.333	26.824
	X_{10}	Public balance (% of GDP)	-2.544	3.481
	X_{11}	Total investment (% of GDP)	21.637	3.343
Education	X_{12}	Pupil/teacher ratio in primary education (pupils per teacher)	15.090	3.269
	X_{13}	Average education level (years)	17	1
	X_{14}	Adult literacy rate (% ages 15 and above)	97.525	3.523
	X_{15}	Women literacy rate (%)	96.851	4.148
	X_{16}	Number of doctors (per 10,000 population)	320	88

[5] The volume index of GDP per capita in PPS is expressed in relation to the EU (EU-27) average set to equal 100. If the index of a country is higher than 100, this country's level of GDP per head is higher than the EU average and vice versa.

Health	X_{17}	One-year-olds fully immunized against measles (%)	91.745	7.294
	X_{18}	Life expectancy at birth (years)	76	3
	X_{19}	Under-five mortality rate (per 1,000 live births)	12.371	19.963
Labour Market	X_{20}	Share of agricultural labour force in total labour force (%)	10.454	9.545
	X_{21}	Female economic activity rate (% of ages 15 and above)	55.673	8.788
	X_{22}	Long-term unemployment rate (% of labour force)	3.781	2.787
	X_{23}	Total unemployement rate (% of labour force)	8.270	3.666
Life Quality	X_{24}	Human development index	89.213	4.857
	X_{25}	Share of income or consumption in poorest 10%	7.833	1.185
Population	X_{26}	Population under age 15 (% of total)	17.300	1.988
	X_{27}	Population age 65 and above (% of total)	15.048	2.018
	X_{28}	Total fertility rate (births per woman)	1.470	0.255
Priorities in Public Spending	X_{29}	Public education expenditure (% of GDP)	5.287	1.200
	X_{30}	Public health expenditure (% of GDP)	5.600	1.175
	X_{31}	Public human resource expenditure (% of GDP)	5.350	1.186
Status of Major International Human Rights Instruments	X_{32}	International covenant on civil and political rights (1: if the convenant signed, 0: otherwise)	-	-
	X_{33}	International covenant on economic, social and cultural rights (1:if the convenant signed,0: otherwise)	-	-
	X_{34}	International convention on the elimination of all forms of racial discrimination (1: if the convenant signed, 0: otherwise)	-	-
	X_{35}	Convention on the rights of the child (1: if the convenant signed, 0: otherwise)	-	-
Trade	X_{36}	Imports of goods and services (% of GDP)	0.531	0.245
	X_{37}	Exports of goods and services (% of GDP)	0.525	0.269
	X_{38}	Balance of the current account (% of GDP)	-0.848	0.310

4. RESULTS

The explanatory variables in Table 4 were categorized for ML and MCL models. Then, MCL model was performed with a SAS macro written by HENDRICKX (1995) and ML model was performed with procedure PHREG in SAS. To determinate the factors affecting the EU membership, statistically significant ML and MCL models were obtained. These models were compared to reach the best model explaning the EU data with the help of $-2\log(L)$ and AIC. Model fit statistics of these models are given in Table 5.

Table 5. Model fit statistics

Model	df	-2log(L)	AIC
ML	8	13.941	15.941
MCL	12	6.254	8.254

Based on the AIC and $-2\log(L)$ values, it can be said that MCL model is more efficient than ML model. All coefficients of ML and MCL models were tested and the estimation results for the ML model are reported in Table 6.

Table 6. Results of multinomial logistic regression analysis

Variable	Parameter Estimate	Standart Error	z	P>|z|
Intercept	-3.405	1.211	-2.812	0.045*
INF(1)	2.914	1.378	2.121	0.034*
INF(2)	1.843	1.381	1.334	0.182
INF(3)	-3.138	1.475	-2.127	0.033*
INF(4)	-6.131	1.441	-4.254	0.000*
INF(1)*DEBT	1.305	1.043	1.251	0.211
INF(2)*DEBT	0.628	1.055	0.595	0.552
INF(3)*DEBT	0.332	1.103	0.301	0.763
INF(4)*DEBT	-0.226	1.093	-0.207	0.836

Notes: 1. * represents statistical significance at the 5%.

2. INF=Inflation Rate (X_9), DEBT=General Government Debt (X_{11}).

3. Inflation rate is categorized as INF(1)= $\leq 0.9\%$, INF(2)=$1\% - 1.9\%$, INF(3)=$2\% - 2.9\%$, INF(4)=3%-3.9%, and INF(5)=$\geq 4\%$.

4. General government debt is categorized as ($0 = \leq 50\%, 1 = \geq 50\%$).

According to Table 6, the ML model equation is given by

$$\ln \frac{P_{jk}}{P_{j5}} = -3.405 + 2.914\text{INF}(1) - 3.138\text{INF3} - 6.131\text{INF}(4).$$

This study takes into consideration parameter estimation results of MCL model since AIC value of MCL model is lower than ML model. The estimation results of MCL model is presented in Table 7.

Table 7. Results of multinomial conditional logistic regression analysis

Variable	Parameter Estimate	Standart error	χ^2	P-value
GDP(1)	-132.673	66.513	3.979	0.046*
GDP(2)	-133.163	69.080	3.716	0.053**
GDP(3)	-26.382	40.824	0.418	0.518
GDP(4)	44.136	37.398	1.393	0.238
GDP(1)*HDI	0.768	2.768	0.077	0.781
GDP(2)*HDI	-0.557	2.908	0.037	0.848
GDP(3)*HDI	1.403	2.247	0.390	0.532
GDP(4)*HDI	-0.614	2.061	0.089	0.766
GDP(1)*EXP	-146.495	72.955	4.032	0.045*
GDP(2)*EXP	-14.261	75.741	3.780	0.052**
GDP(3)*EXP	29.350	45.411	0.418	0.518
GDP(4)*EXP	-49.114	42.208	1.354	0.245

Notes: 1. * and ** represent statistical significance at the 5% and 10% level, respectively.

2. GDP= GDP per capita (US $), ($X_5$); EXP=Exports of goods and services (% of GDP), (X_{40});
HDI= Human development index, (X_{27}).

3. GDP per capita is categorized as GDP(1)=$\leq 5000\$$, GDP(2)=5000-14999, GDP(3)=15000-24999, GDP(4)=25000-34999, and GDP(5)=$\geq 35000\$$.

From Table 7, the MCL model equation is given by

$$\ln\frac{P_{jk}}{P_{j5}} = -132.673GDP(1) - 133.163GDP(2) - 146.495GDP(1)*EXP - 14.261GDP(2)*EXP.$$

The results show that the MCL model containing GDP per capita (X_5), human development index (X_{27}), export of goods and services (X_{40}) are significant at a 95% confidence level in explaining the probability of the EU membership time. On the basis of statistical tests and the explanatory variable groups, it can be concluded that the explanatory variables related with economy and finance, life quality and trade are effective on the EU membership. These results can be informative for the candidate countries how to get a place in the EU.

The results in Table 7 indicate that relative to countries with GDP per capita more than $35000\$$, the likelihood of the EU membership is lower for countries with GDP per capita $\leq 5000\$$ and $5000\$ - 14999\$$. Moreover, if exports of goods and services is low and GDP per capita is lower than 5000$ for a candidate country, the likelihood of the EU membership decreases. Similarly the likelihood of the EU membership decreases

when GDP per capita between 5000$ and 14999$ and exports of goods and services is low.

After the determination of the best models explaining the EU data, now the probabilities of the accession time are obtained for the candidate countries, i.e., Croatia, the Former Yugoslav Republic of Macedonia, and Turkey. The EU's relations with Western Balkan states (Albania, Bosnia and Herzegovina, Montenegro and Serbia) were moved from external relations to the enlargement policy segment. These countries currently are not recognized as candidate countries, but only as potential candidate coutries. In this study, the probabilities of the EU membership for these potential candidate countries are also predicted. According to MCL model, whose parameter estimates are given in Table 7, the values of GDP(1), GDP(2), GDP(1)*EXP, GDP(2)*EXP of candidate and potential candidate countries are used to predict probabilities and the predicted probabilities are given in Table 8.

Table 8. The predicted probabilities of accession time for candidate and potential candidate countries

Candidate Country	**Probability**	**Potential Candidate Country**	**Probability**
Croatia	0.4965	Albenia	0.3249
		Bosnia and Herzegovina	0.2436
FYRM	0.4968	Montenegro	0.0800
		Serbia	0.2591

Note: The probability of accession time for Turkey is not predicted since it is not logical that the full membership application date of Turkey was 14.04.1987.

In this study, the EU membership time is interpreted according to a five-point scale and one can choose from four different cut off levels. Therefore, there are four pairs of sensitivity and specifity values. The emprical ROC curve is shown in Figure 1.

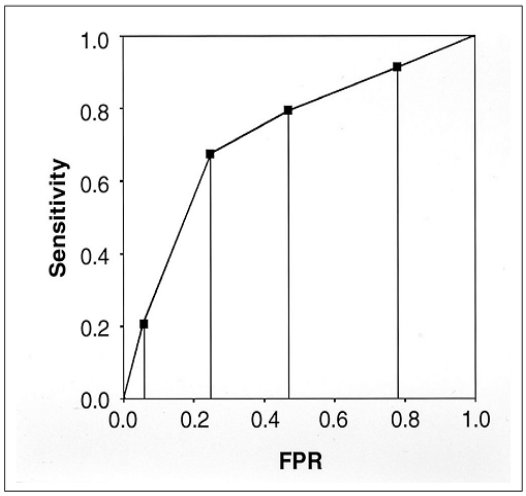

Figure 1. The emprical ROC curve.

As seen in Table 8 and Figure 1, FYRM and Croatia become an EU member between 7 and 9 years with the probability 0.49. The results in Table 8 also suggest that the accession time of the Albania, Bosnia and Herzegovina and Serbia is predicted as between 10 and 12 years while Montenegros' is more than 12 years. These results can be helpful to both candidate and potential candidate countries to become a member of the EU. The results are consistent to some extent with the statements of the EU officials. The expected accession of Croatia would likely happen between 2016 and 2017 according to the EU officials, and Gruevski has suggested for the Former Yugoslav Republic of Macedonia that the country could join in 2016 or 2017.

5. DISCUSSION

The EU is a political and economic community of 27 member states, located primarily in Europe. The EU comprises a single market created by a system of laws, which apply in all member states, guaranteeing the freedom of movement of people, goods, services, and capital. Previous studies on the EU have normally been concerned with the assessment the enlargement and impact of enlargement on the EU. However, little attention has focused on the effects of factors on the EU membership and the accession process of candidate countries.

In this study the EU data has been examined for the emprical investigation of various important factors in determining the EU membership. Then, these factors are used for a candidate or potential candidate country to find the probability that the accession time for the EU membership. For this aim, ML and MCL models have been used to compare the results for more reliable results. The results of this study suggest that the human development index, GDP per capita, exports of goods and services are effective factors for the EU membership. This is the outcome of human development index and GDP per capita are concerned with many important factors such as prosperity level, marketing economy and competing pressure. This study also suggests that if a candidate country can increase its GDP per capita, value of human development index, these improvements can accelerate the accession process. When the current candidate countries are considered together, according to our findings, the Former Yugoslav Republic of Macedonia has the longest timeline to joining the EU, while Croatia has the shortest timeline for the EU membership. Similarly, among the potential candidate countries, Albenia is closer than others for the EU membership. However, the accession time is found as more than six years for both candidate and potential candidate countries.

REFERENCES

ARCHICK K, KIM J, 2008. European Union Enlargement, *Crs Report for Congress*, Belgium, 1-6.

BALDWIN RE, FRANCOIS JF, PORTES R, 1997. The costs and benefits of eastern enlargement: the impact on the EU and Central Europe, *Economic Policy*, 24, 127-176.

BARYSCH K, 2003. Does Enlargement Matter for the EU Economy?, *Policy Brief*, Centre For European Reform, 1-6.

BREUSS F, 2002. Benefits and dangers of EU enlargement, *Empirica*, 29(3), 245-274.

COCKBURN J, DOSTIE B, 2007. Child work and schooling: the role of household asset profiles and poverty in rural Ethiopia, *Journal of African Economies*, 16(4), 519-563.

DOYLE O, FIDRMUC J, 2003. Who is in favor of enlargement? Determinants of Support for EU Membership in the Candidate Countries' Referenda, *ZEI/WDI Conference on Political Economy of Transition: Job Creation and Job Destruction,* Bonn, 1-30.

DRAGAN G, 2007. Romania's accession to EU: Challenges and opportunities, EUIJ Workshop: EU Enlargement and Its Economic Impacts on Transitional Countries,Romania, 1- 24.

EUROPEAN COMMISSION, 2009. Enlargement, Potential Candidate Countries, http://ec.europa.eu/enlargement/potential-candidate-countries/index_en.htm.

EUROSTAT, 2008. Eurostat Indicators, Statistical Office of the European Communities, http://epp.eurostat.ec.europa.eu/portal/page?

HENDRICKX J, 1995. Multinomial conditional logit models for the analysis of status attainment and mobility, *The Institute of China Studies (ICS) Working Papers*, 1, 75-96.

KAUSHAL N, 2005. New immigrants' location choices: magnets without welfare, *Journal of Labor Economics*, 23 (1), 59-80.

MARTINEZ F, AGUILA F, HURTUBIA R, 2009. The constrained multinomial logit: A semi-compensatory choice model, *Transportation Research Part B: Methodological*
43 (3), 365-377.

NICOLAIDES P, 2003. *The Application of EU competition rules to services of general economic interest: How to reduce competitive distortions*. Springer Verlag, Berlin, pp. 88-96.

SHABBIR T, 1993. Multinomial logit model of occupational choice: A latent variable approach, *The Pakistan Development Review*, 32 (4), 687-698.

TAYLOR MR, 2009. Does Food Safety Information Affect Consumers' Decision to Purchase Meat and Poultry? Evidence from U.S. Household Level Data, *School of Economic Sciences Seminar*, 27 February, Washington State University, USA, 1-33.

UNDP, 2008. Human Development Report, http://hdr.undp.org/en/reports/.

WANNER RA, 2005. Twentieth-century trends in occupational attainment in Canada, *The Canadian Journal of Sociology*, 30 (4), 441-469.

WOJCIK C, 2000. Alternative models of demand for automobiles, *Economics Letters,* 68, 113–118.

WORLDBANK, 2008. World Development Indicators (Washington, D.C.: World Bank), http://www.worldbank.org/data/wdi2003/index.htm