# APPLIED DISCRIMINANT ANALYSIS IN ESTIMATION OF POTENTIAL EU MEMBERS

**Assist. Marija Trpkova M.Sc.**
**University "Ss. Cyril and Methodius"**
**Faculty of Economics**
**Assist. Prof. Dragan Tevdovski Ph.D.**
**University "Ss. Cyril and Methodius"**
**Faculty of Economics**

**Abstract:** The purpose of this research paper is to reveal which European countries are most suitable for EU membership using the multivariate method discriminant analysis.

Discriminant analysis is useful for building a model for separation of group membership based on observed characteristics of each country. This analysis is used to model the value of a dependent categorical variable EU membership based on its relationship to seven predictors as important variables for EU integration.

Final results confirm that all EU countries are correctly classified as members of the EU. On the other side, Croatia, Norway, Serbia, Switzerland, Turkey and Ukraine are non EU members, and according to the results, they should be part of the EU. Since Norway and Switzerland are not part of the EU due to non-economic reasons, the analysis points out Croatia, Serbia, Turkey and Ukraine as most suitable candidates for integration in the EU.

**JEL classification: F15, C19**

**Key words: EU integration, Canonical discriminant functions, pooled within-groups covariance matrices, Box's M statistic**

## 1. INTRODUCTION

Integration in the EU is one of the primary objectives of the government's policy from the southeast European countries. The purpose of this research paper is to reveal which European countries are most suitable for EU membership using the multivariate method discriminant analysis. Leaving the political issues aside, the analysis is only concerned with the economic and demographic variables that have potential influence on country's eligibility for EU entrance.

After the introduction, the second part of the paper analyses the remaining European countries that are applicants for EU accession. Except for the Iceland, all others are South-eastern European countries, and this is why this part is called EU enlargement for South-eastern European countries. It elaborates the economic structure of the accession countries, advantages and costs for joining the EU. The third part of the paper formulates the working hypothesis. The forth part of the paper is focused on the data and methodology used. It briefly explains the discriminant analysis as a multivariate analysis classification method, while it mainly describes the given output

of the analysis1. On one side, the paper explains the statistical procedure to see if all requirements and assumptions of the discriminant analysis are being met. On the other side, the paper interprets the results in terms of the considered problem, or which of the countries mostly resemble to the EU countries. The fourth part has the final conclusions and recommendations.

## 2. EU ENLARGEMENT FOR SOUTH – EASTERN EUROPEAN COUNTRIES

### *2.1 Economic structure*

In 2010 there are 9 countries - applicants for EU accession according to the latest data of the European Commission. Only 3 countries have the status "candidate country", while the other 6 countries are still "potential candidates". All countries, except for Turkey, Albania and Iceland, are countries from the former Yugoslavia (Table no. 1).

**Table no. 1 Applications for EU accession**

| Country | Status |
|---|---|
| Croatia | Candidate country |
| Macedonia | Candidate country |
| Turkey | Candidate country |
| Albania | Potential candidate |
| Bosnia and Herzegovina | Potential candidate |
| Kosovo | Potential candidate |
| Iceland | Potential candidate |
| Montenegro | Potential candidate |
| Serbia | Potential candidate |

The 8 accession countries (without Kosovo) could bring additional 94 millions of consumers to the European Union market.

The level of development represented by the GDP per capita (Table no. 2) proves that the accession countries are at a far lower level of development than the average GDP per capita for the Euro Area which in 2007 was 12.228 US dollars, especially countries like Albania, Macedonia, Bosnia and Herzegovina, Montenegro and Serbia. On the other side Iceland, Croatia and Turkey have high level of development, if indicated by the GDP per capita. Thus, there are three distinct groups of candidates for the European Union:

1. The first group represents very low developed countries with GDP per capita lower than 4.000 US dollars: Albania, Bosnia and Herzegovina and Macedonia.
2. The second group represents low developed countries: Serbia and Montenegro.
3. The third group represents relatively developed countries: Iceland, Croatia and Turkey.

[1] **The analysis uses the statistical software SPSS – Statistical Package for Social Sciences.**

**Table no. 2 Economic and demographic data for the applicant countries for 2007[2]**

| Applicant | Population | Annual GDP Growth (%) | GDP (in US dollars) | GDP per capita | Inflation (GDP deflator) (%) | Mobile subscriptions (per 100 people) |
|---|---|---|---|---|---|---|
| Croatia | 4.435.982 | 5 | 58.558.231.254 | 11.552 | 4 | 113 |
| Macedonia | 2.037.032 | 6 | 7.926.664.294 | 3.836 | 8 | 96 |
| Turkey | 73.003.736 | 5 | 655.881.426.190 | 9.390 | 8 | 85 |
| Albania | 3.132.458 | 6 | 10.831.224.735 | 3.383 | 3 | 73 |
| BIH | 3.778.410 | 7 | 15.144.156.753 | 3.985 | 6 | 65 |
| Iceland | 310.997 | 4 | 19.962.854.200 | 42.600 | 6 | 105 |
| Montenegro | 620.941 | 11 | 3.846.153.846 | 5.267 | 18 | |
| Serbia | 7.381.579 | 8 | 40.121.875.321 | 5.462 | 7 | 115 |
| Total | 94.701.135 | | 812.272.586.593 | | | |

Annual GDP growth for the Euro Area for 2007 is 2,7%, which is a stable growth rate. All the accession countries have higher GDP growth rate. This is probably because these are mostly developing countries. On the other side, the inflation rate in the Euro Area is 2,4%, which represents a stabile economy. The accession countries have significantly higher level of inflation, which is not a preferable indicator – it signals price instability, and it is incompatible with the Maastricht criteria.

Mobile cellular subscription can sometimes represent country's development level. The average rate per 100 people for the Euro Area is 116, and also countries like Croatia, Iceland and Serbia have high mobile cellular subscription rate. Except for Serbia, both Croatia and Iceland have high GDP per capita, which proves that mobile penetration rate can be a potential development indicator. According to this, it can be confirmed that most of the accession countries need further development.

**Table no. 3 GDP structure in the accession countries for 2007**

| Applicant | Agriculture (% of GDP) | Industry (% of GDP) | Services (% of GDP) |
|---|---|---|---|
| Croatia | 6 | 28 | 66 |
| Macedonia | 12 | 30 | 59 |
| Turkey | 9 | 28 | 63 |
| Albania | 21 | 20 | 59 |
| BIH | n/a | n/a | n/a |
| Iceland | n/a | n/a | n/a |
| Montenegro | 9 | 18 | 73 |
| Serbia | 13 | 28 | 59 |

Table no. 3 represents the GDP structure in three main economic sectors: Agriculture, Industry and Services. It is obvious that all of the accession countries have the sector Services as the dominant sector. The second sector after Services is Industry,

[2] **The data for 2007 were not available for Kosovo. Kosovo is not included in the further analysis due to the lack of data.**

and the last sector is the Agriculture. The structure in accession countries is compatible the EU structure.

### *2.2 Advantages and costs*

Both the accession countries and EU will have certain advantages and costs from the enlargement. Brief summary follows in Table no. 4.

The accession countries have three major advantages: the access to new and large market, great possibility for labour migration and access to significantly high EU funds. On the other hand, joining the European Union may mean great costs, since vast market may mean severe competition. Integration in the EU also means implementation of great number of EU regulations and legislative, as one aspect of major adjustment problems. Yet, the accession countries have more benefits from the enlargement than the European Union.

One advantage for the EU is securing its own values throughout the newly accepted countries. Yet, new countries may also mean significantly difficult union to govern. Also, increase in population is greater than the increase in the GDP. The disproportion may burden the EU economy. Another problem may be the large number of immigrant workers, yet this sometimes can be taken as advantage in terms of low – cost working force.

Another financial burden is financing the necessary adjustments of the accession countries. Also, the financial benefits that the EU will provide may mean potential loss of job and business in the "sensitive" manufacturing industries and in agriculture in the EU because of the penetration of goods from the east.

**Table no. 4 Advantages and costs from the EU enlargement[3]**

| | Accession countries | European Union |
|---|---|---|
| Advantages | ▪ Access to vast EU market<br>▪ Labour migration<br>▪ Access to EU funds | ▪ Securing the EU values |
| Costs | ▪ Fragile economies exposed to competition<br>▪ Adjustment problems<br>▪ Implementation of EU regulations | ▪ Complicated and ungovernable union<br>▪ Not proportional increase in population and GDP (if enlargement is realised)<br>▪ Large number of immigrant workers<br>▪ Financing the adjustments of the accession countries<br>▪ Potential loss of jobs and business |

## 3. FORMULATION OF HYPOTHESIS

As previously presented in section 2, a general conclusion can be drawn that most of the accession countries are not completely eligible for EU entrance, according to their economic performances: low level of gross domestic product, high inflation rate, low level of development and high unemployment. Yet, let us consider the Fifth EU Enlargement. Here eight Central and Easter European countries (Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia) and the Mediterranean islands of Malta and Cyprus joint the EU in 2004. This was the largest

[3] **'The economics of European integration – Limits and Prospects' (2005) from Miroslav N. Jovanović, pages 828 – 843.**

single enlargement in terms of people, landmass and number of countries, and not in terms of GDP. Romania and Bulgaria joined the EU in 2009, as part of the same enlargement. Many of current the member states were concerned that these are less developed countries and as a result of this placed number of travel and work rights restrictions.

The conclusion is that European Union has expanded with the countries that are less developed than the current member states and probably don't fulfil all the required Maastricht criteria for Economic and Monetary Union. This was the reason why this analysis was encouraged. Fact is certain countries included in the analysis are not yet prepared for the EU accession, according the economic criteria. Yet, EU integration is possible process, according to the past enlargements.

In order to define the working hypotheses, analysis of the economic structure previously explained was taken into account. The hypotheses are:

$H_1$: Very low developed countries Albania, Bosnia and Herzegovina and Macedonia should not be classified as EU members in the discriminant analysis.

$H_2$: Low developed countries Serbia and Montenegro could be classified as EU members in the discriminant analysis.

$H_3$: Relatively developed countries Iceland, Croatia and Turkey are most likely to be classified as EU members.

$H_4$: Countries that are not part of the EU due to political and not economical reasons are most likely to be classified as EU members.

The validity of the formulated hypothesis will be examined in the procedure of the discriminant analysis.

## 4. DATA AND METHODOLOGY

Discriminant analysis is useful for building a model for separation of group membership based on observed characteristics of each country. This analysis is used to model the value of a dependent categorical variable EU membership based on its relationship to seven predictors as important variables for EU integration. The procedure generates a discriminant function based on linear combinations of the predictor variables that provide the best discrimination between two groups of countries: EU members and non EU members. Analysis shows which countries should be part of their group and which are misclassified on the basis of the considered variables.

The main purpose of this paper is to perform discriminant analysis of the countries which are members of the European Union and countries which are not members of the European Union using the following independent and continuous variables: *foreign direct investments*, *gross domestic product*, *change in the gross domestic product (in percentage)*, *inflation*, *users of the mobile telephony*, *population* and *country's surface area*. Except for the last two variables, it is considered that all remaining variables are indicators of one country's growth and development, which is important prerequisite for European Union membership. The discriminant analysis should show which of the variables has the highest contribution in separation of the groups of EU member countries and non EU member countries. This variable would then be considered as the most important variable for membership in the European Union.

After the discrimination, on the basis of the calculated discriminant coefficients, classification of the countries is performed. This classification will show which of the countries are correctly classified, and which are misclassified. This will be used as indicator for the countries that have the largest potential for EU membership, or the countries that are not EU members, and according to the analysis are classified in the group of EU members.

The analysis includes 48 countries or observations, seven independent continuous variables and one dependent categorical variable – *EU membership*. The categorical variable has only two categories, or two groups: 1 – countries that are EU members and 0 – countries that are not EU members, which clearly reflect differences in the independent variables. When the number of groups is small, the complexity of the analysis is not increased.

The data source is the official statistics of the World Bank for 2007.

The discriminant analysis evaluates one discriminant function. The number of discriminant functions represents the number of groups minus one, or in this case 2-1=1. The sample size is also important, specially the number of observations per group. In the analysis each group has more than twenty observations, or the first group has 22 observations and the second group has 26 observations, and because the groups are approximately equal, there is no situation which involves disproportionate chances for classification. Due to the missing data, two observations are excluded from the analysis.

The next step is examination of the conditions of the discriminant analysis. First, the data are being scanned. For this Mahalanobis $D^2$ distance is used. The smallest value of the Mahalanobis distance is 0,65 for Macedonia and 43,89 for Russia. This is an indicator that in this data set Russia should be excluded from the analysis as nonstandard observation. Before the Russia is removed, normality test is conducted. Kolmogorov – Smirnov normality test is used for examination of the normality of the variables. The statistical software SPSS which is used for this analysis does not have a multivariate normality test, so only the separate normality tests for each independent variable is realized. Kolmogorov – Smirnov test proves that the independent variables do not have a normal distribution.

Since there are nonstandard observations and the variables do not have normal distribution, logarithmic transformation of the variables is performed. Only for the variable change in the gross domestic product has square root transformation. After the transformation, the Mahalanobis distance is 0,82 for Croatia and 19,14 for Luxemburg, which means that now there are no nonstandard operations. Also, the normality tests prove that after the transformation, all variables have normal distribution except inflation and users of mobile telephony. The data base for the discriminant analysis has been improved.

In order to save space, the independent variables are represented as: LOG_FDI – Logarithmic transformation of *Foreign direct investments in US dollars, net inflow*, LOG_GDP – Logarithmic transformation of *Gross domestic product in current prices, in US dollars*, SQRT_GDP% - Square root transformation of *Annual growth of the gross domestic product in percentage*, LOG_INF – Logarithmic transformation of *Inflation* or *gross domestic product deflator, annual and in percentage*, LOG_UMT – Logarithmic transformation of *Users of mobile telephony (on 100 residents)*, LOG_POP – Logarithmic transformation of *Total population*, LOG_SUR – Logarithmic transformation of *Surface area of the country in square kilometres*.

Further, the results from the SPSS software are presented.

There are several tables that assess the contribution of each variable to the model, including the tests of equality of group means, the discriminant function coefficients, and the structure matrix. The tests of equality of group means measure each independent variable's potential before the model is created. Each test displays the results of a one - way analysis of variance for the independent variable using the grouping variable *European Union membership* as the factor. If the significance value is greater than 0,10, the variable probably does not contribute to the model.

According to the results in this table, only variables *Population* and *Surface area* are not significant.

**Table no. 5. Test of equality of group means - Assessing the Contribution of Individual Predictors**

| Variable | Wilks' lambda | F | Degrees of freedom 1 | Degrees of freedom 2 | Significance |
|---|---|---|---|---|---|
| LOG_FDI | 0,73 | 16,67 | 1 | 44 | 0,00 |
| LOG_GDP | 0,75 | 14,73 | 1 | 44 | 0,00 |
| SQRT_GDP% | 0,78 | 12,75 | 1 | 44 | 0,00 |
| LOG_INF | 0,75 | 14,57 | 1 | 44 | 0,00 |
| LOG_UMT | 0,60 | 28,97 | 1 | 44 | 0,00 |
| LOG_POP | 0,99 | 0,27 | 1 | 44 | 0,60 |
| LOG_SUR | 0,96 | 2,04 | 1 | 44 | 0,16 |

Wilks' lambda is another measure of a variable's potential. Smaller values indicate the variable is better at discriminating between groups. The table suggests that *Mobile cellular subscriptions (per 100 people)* is best, followed by *Foreign direct investment*, *GDP (current US$)*, *Inflation*, *GDP deflator (annual %)*, and *GDP growth (annual %)*. (Table no. 5).

The pooled within-groups matrices are calculated and they display a covariance matrix and a correlation matrix. The pooled within-group matrices are very different from the usual matrices where all cases are treated as a single sample.

The within-groups correlation matrix shows the correlations between the predictors. Pooled within-groups correlations are computed from the covariances and variances. In the correlation matrix the coefficients on the main diagonal are always 1,0, because each variable has a perfect positive linear relationship with itself. Correlations above the main diagonal are a mirror image of those below. If there are several strong correlations (greater than say 0,75 or less than -0,75) there may be alternative subsets of variables that would perform equally well. The largest correlations occur between *GDP (current US$)* and the *Foreign direct investment, net inflows (BoP, current US$)*, but it is difficult to tell if they are large enough to be a concern. We shall look for differences between the structure matrix and discriminant function coefficients to be sure.

The pooled within-groups covariances are obtained by averaging the separate covariance matrices for all groups. The covariance matrix displays variances on the main diagonal and covariances elsewhere. Covariances above the main diagonal are a mirror image of those below.

One of the necessary assumptions for discriminant analysis is equality of group covariance matrices. In this example, the covariances and variances for *Mobile cellular subscriptions (per 100 people)* appear to differ most. There is no simple Levene test to test for equality of covariances.

One way to determine if the covariances are equal is to use the covariance matrices and the within-group scatter plots (setting all the plot scales the same).

**Table no. 6 Box's Test of Equality of Covariance Matrices - Checking Homogeneity of Covariance Matrices**

| | | |
|---|---|---|
| Box's M | | 52,06 |
| F | Approximately | 1,53 |
| | Degrees of freedom 1 | 28 |
| | Degrees of freedom 2 | 5824,99 |
| | Significance | 0,04 |

Box's test tests the null hypothesis of equal population covariance matrices or the assumption of equality of covariances across groups. The significance of Box's M statistic is based on an F transformation. The hypothesis of equal covariance matrices is rejected if the significance level is small (less than say 0,10). The hypothesis of equal covariance matrices is not rejected if the significance level is large (more than say 0,10). The test can be significant when within-group sample sizes are large or when the assumption of multivariate normality is violated.

Since Box's M is significant for level of significance 0,01 (if p-value=0,036>0,01 the null hypothesis is accepted), we look at the separate matrices to see if it gives radically different classification results. (Table no. 6).

The final comment about the assumptions of the discriminant analysis is that the most of the explanatory variables are independent, multivariate normality of the explanatory variables exists and there is homogeneity of the covariance matrices (at 0,01 level of significance). This leads to conclusion that the results of the discriminant analysis are valid.

The following section represents the assessment of the model fit trough the canonical discriminant functions. In addition to measures for checking the contribution of individual predictors to your discriminant model, the discriminant analysis procedure provides the eigenvalues and Wilks' lambda tables for seeing how well the discriminant model as a whole fits the data.

**Table no. 7 Eigenvalues [(a)]**

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 1,16 | 100,00 | 100,00 | 0,73 |

*(a) First 1 canonical discriminant functions were used in the analysis.*

The eigenvalues table provides information about the relative efficacy of each discriminant function. When there are two groups, the canonical correlation is the most useful measure in the table, and it is equivalent to Pearson's correlation between the discriminant scores and the groups. (Table no. 7).

**Table no. 8 Wilks' Lambda**

| Test of Function | Wilks' Lambda | Chi - square | Degrees of freedom | Significance |
|---|---|---|---|---|
| 1 | 0,46 | 31,15 | 7 | 0,00 |

The test of functions column tests the hypothesis that the means of the functions listed are equal across groups. The small significance value indicates that the discriminant function does better than chance at separating the groups, like in the analysis. (Table no. 8).

**Table no. 9 Standardized Canonical Discriminant Function Coefficients**

| Variable | Function |
|---|---|
| | 1 |
| LOG_FDI | -0,23 |
| LOG_GDP | 1,67 |
| SQRT_GDP% | 0,04 |
| LOG_INF | 0,44 |
| LOG_UMT | 0,38 |
| LOG_POP | -0,23 |
| LOG_SUR | -1,11 |

The standardized coefficients allow you to compare variables measured on different scales. Coefficients with large absolute values correspond to variables with greater discriminating ability, in this case *GDP (current US$)* is a variable with greatest significant discriminating ability, follows *Surface area (sq. km)* which is insignificant, then follows *Inflation, GDP deflator (annual %)*, *Mobile cellular subscriptions (per 100 people)*, *Foreign direct investment, net inflows (BoP, current US$)*, *Population, total* (which is insignificant) and the last variable *GDP growth (annual %)*. (Table no. 9).

The structure matrix contains within - group correlations of each predictor variable with the canonical function.

**Table no. 10 Structure Matrix**

| Variable | Function |
|---|---|
| | 1 |
| LOG_FDI | 0,75 |
| LOG_GDP | 0,57 |
| SQRT_GDP% | 0,54 |
| LOG_INF | -0,54 |
| LOG_UMT | -0,50 |
| LOG_POP | -0,20 |
| LOG_SUR | 0,07 |

Since the structure matrix is unaffected by collinearity, it's safe to say that the first several variables best discriminate between EU members and non EU members.

When there are more than two groups, the number of canonical variables is k-1 (where k is the number of groups) or p (the number of variables), whichever is smaller (Table 11).

**Table no. 51 Canonical Discriminant Function Coefficients**

| Variable | Function |
|---|---|
| | 1 |
| LOG_FDI | -0,33 |
| LOG_GDP | 2,33 |
| SQRT_GDP% | 0,07 |
| LOG_INF | 1,29 |
| LOG_UMT | 0,02 |
| LOG_POP | -0,39 |
| LOG_SUR | -1,73 |
| (Constant) | -13,42 |

The coefficients of the canonical variable are used to compute a canonical variable score for each case. The coefficients displayed in this table are the coefficients of the canonical variable. For example:

Canonical variable score = -0,33LOG_SDI + 2,33LOG_GDP + 0,07SQRT_GDP% + 1,29LOG_INF + 0,02LOG_MOB - 0,39LOG_POP - 1,73LOG_POV - 13,42.

In order to compute the canonical variable score it is necessery to substitute the values of LOG_SDI, LOG_GDP, SQRT_GDP%, LOG_INF, LOG_MOB, LOG_POP and LOG_POV for a specific case. When there are more than two groups, the number of canonical variables is $k-1$ (where k is the number of groups) or $p$ (the number of variables), whichever is smaller.

Standardized canonical discriminant function coefficients are used when variables are measured in different units and the magnitude of an unstandardized coefficient provides little indication of the relative contribution of the variable to the overall discrimination. Standardizing the coefficients allows examination of the relative standing of the measurements.

Unstandardized coefficients are used for calculation of the discriminant scores presented in the analysis.

**Table no. 62 Functions at Group Centroids**

| EU membership | Function |
|---|---|
| | 1 |
| 0 | -1,20 |
| 1 | 0,92 |

Unstandardized canonical discriminant functions evaluated at group means are presented in the table. Table no. 12 displays the canonical variable means by group. Within - group means are computed for each canonical variable. In this example for the first canonical variable, the average discriminant or canonical variable score for non EU countries is -1,2 and that for EU countries is 0,923.

Classification statistics are presented next.

**Table no. 73 Prior Probabilities for Groups**

| EU membership | Prior | Cases Used in Analysis | |
|---|---|---|---|
| | Unweighted | Weighted | Unweighted |
| 0 | 0,44 | 20 | 20 |
| 1 | 0,57 | 26 | 26 |
| Total | 1,00 | 46 | 46 |

Table no. 13 displays the prior probabilities for membership in groups. A prior probability is an estimate of the likelihood that a case belongs to a particular group when no other information about it is available.

The classification functions are used to assign cases to groups. There is a separate function for each group. For each case, a classification score is computed for each function. The discriminant model assigns the case to the group whose classification function obtained the highest score (Table no. 14).

Table that contains the casewise statistics which display the actual and the predicted group for all countries is not displayed in this paper due to its length.

The number and percentage of cases correctly classified and misclassified from the original sample are: 41 (15+26) or 85,42% of the cases were classified correctly, and only 7 cases or 14,58% from cases that belong to group 0, are classified in group 1.

There are no misclassified cases that originally belong to group 1 and are classified as cases from group 0.

**Table no. 84 Classification Function Coefficients**

| Variable | EU membership | |
|---|---|---|
| | 0 | 1 |
| LOG_FDI | -5,11 | -5,81 |
| LOG_GDP | 111,63 | 116,57 |
| SQRT_GDP% | 32,93 | 33,09 |
| LOG_INF | 98,05 | 100,79 |
| LOG_UMT | -0,37 | -0,33 |
| LOG_POP | -30,80 | -31,63 |
| LOG_SUR | -56,45 | -60,13 |
| (Constant) | -386,90 | -414,83 |

The original results may provide overly optimistic estimates. Cross-validation attempts to remedy this problem. With cross-validation, each case in the analysis is classified by the functions derived from all cases other than that case. In the cross-validate sample 37 (13+24) or 77,08% of the cases were classified correctly. Number of the misclassified cases is 11 (2+9) or 22,9%. Two cases that belong to group 1 are classified as members from group 0, and 9 cases that belong to group 0 are classified as members from group 1.

If the percentage for correct classification is significantly lower for the cross-validated cases than for the original cases, then there are too many predictors in the model. Here, the difference is 8,33% which does not appears to be significant in size.

Since Box's M is not significant, or variances of the independents among categories of the categorical dependent are not homogenous, it's no worth running a second analysis to see whether using a separate-groups covariance matrix changes the classification.

## 5. CONCLUSIONS

Final results confirm that all EU countries are correctly classified as members of the EU. On the other side, Croatia, Montenegro, Norway, Serbia, Switzerland, Turkey and Ukraine are non EU members, and according to the results, they should be part of the EU. Since Norway and Switzerland are not part of the EU due to non-economic reasons, the analysis points out Croatia, Serbia, Turkey and Ukraine as most suitable candidates for integration in the EU.

If we associate the findings of the analysis with the previously defined hypothesis, the final conclusions can be drawn:

$H_1$: Very low developed countries Albania, Bosnia and Herzegovina and Macedonia should not be classified as EU members in the discriminant analysis.

The first hypothesis is true. The analysis has confirmed that the very low developed countries are not ready yet for an EU accession.

$H_2$: Low developed countries Serbia and Montenegro could be classified as EU members in the discriminant analysis.

Low developed countries could be classified as EU members, since the discriminant analysis has confirmed this hypothesis. The difference between the very

low developed countries and low developed countries is crucial for membership in the EU.

$H_3$ : Relatively developed countries Iceland, Croatia and Turkey are most likely to be classified as EU members.

The analysis has confirmed that relatively developed countries are most likely candidates for EU membership, since Croatia and Turkey are classified as EU members. Iceland is not part of this group according to the analysis.

$H_4$ : Countries that are not part of the EU due to political and not economical reasons are most likely to be classified as EU members.

The last hypothesis also appears to be true. Norway, Switzerland and Ukraine are not part of the EU due to non economical factors. Since the discriminant analysis considers mostly economic and some demographic variables, these countries are classified as very likely EU members.

Variables that mostly contribute to the discrimination are mobile subscriber's penetration rate, foreign direct investments and size of the gross domestic product. High mobile penetration rate as indicator of country's development, high foreign direct investment inflow and large gross domestic product are key variables for EU membership and area of focus for potential candidate and applicant countries.

## REFERENCES


1. Anderson, T. E. An introduction to Multivariate Statistical Analysis, Third Edition, New York: John Wiley, 2003
2. Banićević, D. and Vasić, V. 'Business bank's discriminatory analysis: conservative or aggressive credit policy', in 'SymOgr 2006' , Proceedings, Zlatibor, 2006
3. European Commission: 'European Commission Enlargements', http://ec.europa.eu
4. Garson, D. 'Discriminant Analysis from Statnotes: Topics in Multivariate analysis', http://faculty.chass.ncsu.edu
5. Hair, J.F., W.C. Black, B.J. Babin, R.E. Anderson and R.L. Tatham Multivariate Data Analysis, Sixth Edition, Upper Saddle River, NJ: Prentice Hall, 2006
6. Johnsong, N., and D. Wichern Applied Multivariate Statistical Analysis, Fifth Edition, Upper Saddle River, NJ: Prentice Hall, 2002
7. Jovanović M. N. The economics of European integration – Limits and Prospects, Edward Elgar Publishing Limited, pages 828 – 843, 2005
8. Kovačić, Z. Multivarijaciona analiza, Ekonomski fakultet – Beograd, 1994
9. Porter, M.E. and Kramer, M.R. 'The Competitive Advantage of Corporate Philanthropy', Harvard Business Review (December 2002): 5-16, 2002
10. Vasić, V. and Banićević, D. 'Discriminant analysis algorithm in clients credit risks estimation', YU INFO 2006, Proceedings, Kopaonik, 2006
11. Welch, B. L. 'Note on Discriminant Functions', Biometrika, 31: 218 – 220, 1939
12. World Bank: 'Key Development Data and Statistics', http://web.worldbank.org