CLASSIFYING AND PREDICTING COUNTRY TYPES THROUGH DEVELOPMENT FACTORS THAT INFLUENCE ECONOMIC, SOCIAL, EDUCATIONAL AND HEALTH ENVIRONMENTS OF COUNTRIES

Sidika Gulfem Ozturk
The University of Texas-Pan American College of Business Administration
1201 West University Dr. Room 120J , Edinburg, TX 78539, USA
Phone: 1-956-381-2590
E-mail address: sozturk@panam.edu

Abstract

The purpose of this study is to classify and predict country types through some of the development indicators by using multiple discriminant analysis as the statistical technique. Human Development Index scores published each year by the United Nations have been considered as the development classification directory of the countries. The results of this study indicate that health development factors are good indicators for the distinction between developed and underdeveloped countries. On the other hand, women's involvement in governments and national parliaments, economic growth, and trade development factors are good indicators of development for a developing country.

Keywords: HDI, country development, multiple discriminant analysis, economic indicators, women participation in national parliament, health development, educational development, trade.

JEL classification: N01, N70, O15, O19

1. Introduction

Development of countries had been long discussed in and outside of academia and it remains a popular topic in the literature. Definition of country development has been controversial because of the fact that every discipline that has studied country development has a different perspective over the definition. Sociologists, economists, medical experts, public administers, scientists and psychologists have differing opinions on country development and its definition. Since the definitions are significantly different from each other and each discipline tends to measure development with different indicators, it has been difficult to come up with a single composite measure that summarizes the level of overall development of a country.

The purpose of this study is to predict the classification of countries using their various characteristics. These characteristics primarily are health, education, social environment, trade and economics-related factors. Classifying countries correctly based on these factors will identify the shortcomings of undeveloped and developing countries and will enable these countries to improve in those areas. It will also help foreign investors in their decision making processes where economic indicators of the country development are not sufficient to identify a country's development level. The main contribution of this study is the usage of multiple discriminant analysis as the statistical technique for the study. Also, in the past, level of country development was assumed to be a function of economic development for the most part. However, this approach was unable to explain some of the discrepancies such as a country with a high Gross Domestic Product ("GDP") but a poor level of life standard. Therefore, it was clear that economic indicators needed to be combined with social, educational, health, cultural, and global indicators of development. Even then, it might have not been able to give a deeper insight into any country's development.

There has been an ongoing attempt to measure the development level of countries by using different indicators of development by many researchers. The reason for the researchers to strive for an objective measure of development was to identify the indicators that determine the level of development. If this were possible, then it would also be possible for countries to sustain or improve their development levels by using the necessary tools to improve on the areas that caused the development shortcomings. The United Nations and the World Bank have been the pioneers in measuring countries' development levels in a single composite measure. As a result, the United Nations developed an index called the Human Development Index ("HDI"). In this study, it will be investigated whether variables other than GDP are good indicators of country classification.

2. Review of Literature

The United Nations ("UN") defines human development as the enlargement of the range of people's choices. According to the UN's definition, human development is an extension, enlargement and deepening of the currently unpopular basic needs approach. Streeten (2001) argues the basic needs approach and states: "Human development goes beyond basic needs in that it is concerned with all human beings, not only poor, not only poor countries, not only basic needs. Human development applies to the advanced, industrial countries, as much as to middle-income and low-income countries."

Moreover, Hicks (1979) explores the relationship between economic growth and basic needs and concludes that "the development of critical minimum level of basic human capital may be an important prerequisite for accelerating the growth of output". Taking this statement into consideration, GDP per capita itself is not used as an independent variable in this study. However, it makes more sense to use GDP growth per capita as an independent variable and

investigate its explanatory power to distinguish among country types. Self and Grabowski (2003) find that public health expenditures face diminishing returns in developed countries, life expectancy is used to measure health. They also posit that "it is in the least developed countries that public health expenditures have significant positive impact on life expectancy, yet in these countries public health expenditures contributed far less towards funding of health care than the developed countries". Hence, health development indicators can be good indicators of country types. In this study, health expenditure per capita (in US dollars) and health expenditure as a percentage of GDP are used as indicators of health development for all countries. Anand and Sen (1993) have suggested the classification of all countries into three groups: low, medium and high levels of human development. This is one of the reasons that in this study the level of country development have been classified into three categories: developed, developing, and undeveloped.

Sharma (1997) argues that in every society, women play vital roles. Yet, most women do not have an equal share of land, credit, education, employment, and political power, in comparison to the men in their society. Fukuda-Parr (2001) also discusses social development from a gender perspective. In his study, he emphasizes that gender empowerment measure is a measure of women's participation and empowerment in the society, i.e., in key areas of national political decision making, professional activities and income earning activities. Therefore, as a measure of women's involvement in national politics, proportion of seats held by women in national parliaments is used as one the factors in this study to investigate whether this factor is a good indicator of county development levels. Anderson and Morrissey (2006) conduct a study for a set of countries that can be classified as poor performers. They utilize economic indicators and infant mortality over two periods, 1980-1990 and 1990-2000, and use four different statistical criteria to identify poor performance. Their main finding is that very few countries appear as poor performers; those that perform poorly on one indicator, or in one period, typically do not perform poorly on other indicators or in the other time period. In order to utilize their study, Immunization DPT percentage of children and mortality rate under 5 per 1000 are used in this study to test whether these indicators are useful in identifying country types.

Pillay (2006) states: "Some high performing countries have combined rapid economic and social progress and now have high-performing economies (e.g. Republic of Korea, Malaysia, Mauritius-World Bank, 2004). They achieved social progress early in their development processes, when national incomes were still low- suggesting a certain sequence for investments. In other high-achieving countries economic growth has been slower and less consistent. Nevertheless, all these high performers show that with the right government priorities and policies high social development is possible, even without a thriving economy. This suggests that for countries to become high achievers investments in education and health need to be the highest priority. In countries where growth has been historically low, such investments can provide the foundation for stimulating growth; in countries which are already growing, further investment in human capital development can lead to these countries embracing a growth path characterized by increasingly sophisticated technology, high value-added processes and rapidly rising standards of living".

Consequently, it is clear that there are some countries with low growth rates but higher social and educational development levels. When only economic indicators of development are used to identify countries, it is likely that the result might be misleading. This study particularly avoids such one-dimensional approach and includes indicators from different areas of development.

Storm (2005) emphasizes the importance of international trade in countries' development processes by saying that: "International trade plays a crucial role in the economies of the least

developed countries. In most least developed countries, trade accounts for more than half of GDP- a larger share than in the high-income OECD countries. Least developed countries have undertaken greater trade liberalization than most other developing countries (often compelled by World Bank and IMF loan conditionalities) and, by all indicators, have to be regarded as "open". …But least developed countries imports remain larger than exports, as reflected in a structural average trade deficit of the least developed countries as a whole,…" This concludes that international trade retains its importance as a driving force in economies. Merchandise exports and imports are considered in this study as a measure of trade development of countries.

3. Data

Data for this study have been collected from the World Bank and the United Nation databases. The data consist of 90 countries and developed, developing and undeveloped countries are represented in equal numbers in the study. Therefore, 30 developed, 30 developing and 30 undeveloped countries are analyzed in the study. The country type classifications were made based on the countries' development scores on the Human Development Index that was provided by the United Nations for the year 2005. According to the HDI, the top 30 countries are classified as developed countries. Countries that were in the middle 30 of the HDI ranking are considered as developing countries, and the countries that were ranked in the bottom 30 of the HDI are designated as the undeveloped countries.

As mentioned above, the initial sample size consisted of 90 countries. However, because of some missing data complications, there were only 77 countries valid for the analysis. Yet, usage of multiple discriminant analysis only requires group sample sizes to be equal or relatively equal. There is no concern due to the group sample sizes used in the analysis since relatively equal group sample sizes assumption of multiple discriminant analysis is met.

Another important point regarding the data used in the study is the time interval that the data were taken from. All of the data used in the analysis are from year 2005 figures.

3.1 Variables

Country type is the dependent variable for this study. It is a categorical variable with three possible outcomes: "0= Developed", "1= Developing", and "2= Undeveloped". The country development as the dependent variable will be denoted by "CT" in further sections. Thus, CT measures the extent to which the level of a country's economic, social, educational, and health development is perceived.

There are 10 independent variables used to distinguish among the characteristics of the country types in the study. These are School Enrollment Percentage ("SER"), GDP Growth Percentage Annual ("GDPGR"), Health Expenditure Per Capita USD ("HEPC"), Health Expenditure Percentage of GDP ("HEGDP"), Immunization DPT Percentage of Children ("IMM"), Mortality Rate Under 5 per 1000 ("MR"), Trade Percentage of GDP ("TRDR"), Proportion of Seats Held By Women in National Parliament Percentage ("WMN"), Merchandise Exports ("ME"), Merchandise Imports ("MI"). All of these independent variables are continuous variables. Thus, SE, GDPG, HEGDP, IMM, TR, MR, and WMN are measured by percentages. HEPC, ME, and MI are represented with real dollar values. Table 1 shows the descriptive statistics and the measurement of variables in the study.

Table 1 is about here.

As mentioned above, the HDI scores are used to classify the level of development of countries into three categories. There is no variable in the study that is directly related to the nominal dollar values of GDP, since GDP has already been included in the calculation of HDI. Including independent variables that are directly related to nominal GDP values would cause

high multicollinearity among the dependent and independent variables and may lead to misleading results or interpretations in the analysis.

3.2 The Model

This study investigates whether countries that have different level of development can be distinguished through factors other than just economic indicators. It is clearly established in the literature that countries show differences in their levels of economic development. However, the reason for this study is not just to classify the countries correctly, but also to predict the classification of countries in the future through other indicators that are used in the study. Given these primary purposes of the study, multiple discriminant analysis ("MDA") is the appropriate statistical technique for this study. Using MDA will also fill a gap in the literature because the issue of country development classification has never been studied using MDA.

MDA requires a categorical dependent variable. Thus, the dependent variable, *CT*, in the study is country type which has three possible outcomes. Independent variables of the study can be grouped under five different headings: health-related development indicators, social development indicators, trade development indicators, education-related development indicators, and growth-related development indicators.

Formally, the model used for this study is as below:

$$CT_{jk} = a + W_1 SER_{1k} + W_2 GDPGR_{2k} + W_3 HEPC_{3k} + W_4 HEGDP_{4k} + W_5 IMM_{5k} + W_6 MR_{6k} + W_7 ME_{7k} + W_8 MI_{8k} + W_9 TR_{9k} + W_{10} WMN_{10k}$$

CT_{jk} is the discriminant CT score of discriminant function *j* for object *k*, a is the intercept, W_i is the discriminant weight for independent variable *i*, and X_{ik} is the independent variable *i* for object *k*.

4. Results

The multiple discriminant analysis was run for the sample with SPSS 15.0. Since the country type had three possible outcomes, the results of the MDA provided two discriminant functions. These functions are:

Function 1: *CT*= -0.133 *SER*+ 0.033 *GDPGR*- 0.457 *HEPC*- 0.286 *HEGDP* – 0.276 *IMM*+ 0.817 *MR* – 0.171 *ME* – 0.148 *MI* – 0.03 *TR* – 0.111 *WMN*

Function 2: *CT*= -0.161 *SER*- 0.232 *GDPGR*+ 0.778 *HEPC*+ 0.436 *HEGDP* – 0.119 *IMM*+ 0.369 *MR* + 0.2 *ME* + 0.193 *MI* – 0.071 *TR* + 0.352 *WMN*

MDA has three important assumptions. The first one is the normality and equality of covariance matrices of the independent variables. Kolmogorov-Smirnov test of normality is performed for the sample and there were some violations of this assumption in some categories of *MR, ME,* and *MI*. However, in many instances remedies for data not meeting the multivariate normality is ineffectual. The equality of covariance matrices of the independent variables is checked from Box's M test results. For the data in this study, the Box's M test is significant at .000 significance level and it is concluded that the country type groups differ in their covariance matrices, violating an assumption of MDA. Yet, discriminant function analysis is robust even when the equality of variances assumption is not met.

Standardized discriminant function coefficients and Structure correlation coefficients should be interpreted together as shown in Table 2. The standardized discriminant function coefficients indicate the partial contribution of each variable to the discriminant functions, controlling for other independent variables in the equation whereas the structure coefficients indicate the simple

correlations between the variables and the discriminant functions in the study. Therefore, usage of the structure coefficients enables us to assign meaningful labels to the discriminant functions. As shown in Table 3, *IMM* and *MR* are contributing significantly to Discriminant Function 1 in the study. On the other hand, *SER, GDPGR, HEPC, HEGDP, ME, MI, TR,* and *WMN* are contributing significantly to Discriminant Function 2 in the study.

Eigenvalues show how much of the variance in the country type is accounted for by each of the discriminant functions. Table 2 shows that in this study Function 1 and Function 2 respectively account for 83.9 % and 16.1 % of the explained variance of the discriminant scores, which is explained by the differences among the country types. They together account for 94.8 % of the total variance in the discriminant scores that is explained by the differences among the country types.

The Wilk's Lambda (4.01 %) indicates the proportion of the variance in the discriminant scores unexplained by the differences among the three groups of country type. The Wilk's Lambda significance level for both functions in this study is .000. Therefore, it is concluded that the means of the three groups of country type on the discriminant functions are equal and the functions are well discriminating.

Group centroids also need to be viewed for a complete interpretation of the models in this study. Table 2 shows the corresponding group centroid for each group of the country type. Function 1 distinguishes between developed countries and undeveloped countries. Whereas, Function 2 primarily distinguishes developing countries from the other two groups, developed and undeveloped countries.

Hit ratio for the study is 94.8 %. This means that 94.8 % of the classifications of the country type were correctly predicted by the model in this study. Therefore, the model has a high predictive ability. The internal validity of the model is established through the usage of leaving-one-out case estimation. Utilization of the cross validation approach resulted in the hit ratio of 93.5 %. When compared to the original hit ratio of 94.8 %, it is obvious that the internal validity of the model is accomplished.

Table 2 is about here.

5. Conclusion

Knowledge of country types has been very critical for many different reasons. There are institutions such as the World Bank, United Nations, governments, international credit unions, private companies and people who seek international investment opportunities and entrepreneurship. All of these institutions and people would benefit a great deal if they can classify a country's type based on some factors other than economic factors. Some countries may have small GDPs but they can still be classified as developing or developed countries. The opposite of this situation might also occur as it did in the past with the example of Iraq. Therefore, when judging the development levels of countries, more factors that may play important roles in a country's development need to be considered carefully.

The model in this study gives us an opportunity to identify the development levels of countries based on some factors related to health, education, trade and women's involvement in political decisions. Mortality rate for children under 5 and Immunization DPT percentage of children can easily distinguish between developed and undeveloped countries. Proportion of seats held by women in national parliament, annual GDP growth rates, Merchandise Exports and Imports, percentage of School enrollment, and percentage of Trade can distinguish developing countries from the others. Therefore, it can be easily generalized that health development is a

discriminant factor between developed and undeveloped countries whereas women's involvement in governments and national parliaments, economic growth, and trade development are good indicators for judging the level of development of a developing country.

References:

Anand, S., Sen, A., 1993. Human Development Index: Methodology and Measurement. UNDP Occasional Papers No.8.

Anderson, E., Morrissey, O., 2006. A Statistical Approach to Identifying Poorly Performing Countries. Journal of Development Studies Vol.42, No 3., 469—489

Fukuda-Parr, S., 2001. Indicators of Human Development and Human Rights -- Overlaps, Differences ... and What about the Human Development Index? Statistical Journal of the United Nations Economic Commission for Europe Vol. 18., 239—248

Hicks, N. L., 1979. Is There a Tradeoff between Growth and Basic Needs? Finance and Development June Vol. 17, No 2., 17--20.

Pillay, P., 2006. Human Resource Development and Growth: Improving Access to and Equity in the Provision of Education and Health Services in South Africa. Development Southern Africa Vol. 23, No 1., 64--83

Self, S., Grabowski, R., 2003. How effective is public health expenditure in improving overall health? A cross-country analysis. Applied Economics Vol. 35, Issue 7, 835—845

Sharma, S. D., 1997. Making the Human Development Index (HDI) gender-sensitive. Gender and Development Vol. 5, No 1., 60—61

Streeten, P., 2001. Reflections on Social and Antisocial Capital. Journal of Human Development Vol. 3, Issue 1, 7 –22

Storm, S., 2005. Development, Trade or Aid? UN Views on Trade, Growth and Poverty. Development and Change Vol. 36, Issue 6, 1239–1261.

World Bank web site www.worldbank.org

United Nations web site www.un.org

Table 1

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Country Type	90	2	0	2	1	0.821	0.674
School Enrollment %	85	106	39.1	145.1	99.424	17.888	319.977
GDP Growth % Annual	90	33.3	-7.1	26.2	4.254	3.950	15.600
Health Expenditure per capita USD	88	5708	3	5711	898.034	1394.361	1944241.941
Health Expenditure(% of GDP)	88	12.4	2.8	15.2	6.478	2.457	6.039
Immunization DPT, % of children	87	74	25	99	84.483	15.071	227.136
Mortality Rate Under 5 (per 1000)	88	280	3	283	71.489	78.264	6125.264
Merchandise Exports	89	970679	9	970688	93807.596	191272.760	36585268587.630
Merchandise Import	89	1732586	120	1732706	98573.865	233980.187	54746728108.232
Trade(% of GDP)	83	361	22	383	90.663	61.300	3757.690
Proportion of seats held by women in national parliament(%)	88	48.8	0	48.8	16.785	10.699	114.466
Valid N (listwise)	77						

Table 2

Multiple Discriminant Analysis Results

Group Centroids:		**Discriminant Function 1**		**Discriminant Function 2**	
Developed Countries:		-3.1067		1.1035	
Developing Countries		-0.4987		-1.6889	
Undeveloped Countries		3.9266		0.7045	
Predictor Variables		**Std.Discriminant Coeffs.(STAN)**	**Structure Correlation Coeffs.(SC)**	**Std.Discriminant Coeffs.(STAN)**	**Structure Correlation Coeffs.(SC)**
SER		-0.217	-0.133	-0.131	-0.161*
GDPGR		-0.270	0.033	-0.279	-0.232*
HEPC		-0.541	-0.457	0.700	0.778*
HEGDP		-0.005	-0.286	0.280	0.436*
IMM		0.116	-0.276*	-0.018	-0.119
MR		0.931	0.817*	0.490	0.369
ME		-0.362	-0.171	0.480	0.200*
MI		0.395	-0.148	-0.608	0.193*
TR		0.055	-0.030	-0.115	-0.071*
WMN		-0.014	-0.111	0.221	0.352*
Discriminant Functions Summaries:					
Canonical Correlations	0.9599		0.8945		0.0654
Wilk's Lambda (1 through 2)	0.0401				
Eigenvalues			8.482		1.630
Chi-Sq Sig.(1 through 2)	223.53				

* indicates statistical significance at 5 %.