KRISHNA MAZUMDAR

CLASSIFICATION OF COUNTRIES:
A SOCIO-ECONOMIC APPROACH

I. INTRODUCTION

Classification of countries on the basis of different economic indicators is a widely used phenomenon. The single indicator often used to classify the countries on the basis of level of development is the per capita gross domestic product expressed in U.S. dollars (PCGDP). But only PCGDP reflects distorted picture of the level of development of a country and in fact level of development a country has been achieved depends on a number of socio-economic aspects of life. This study attempts to classify the countries of the world on the basis of selected socio-economic indicators of life and test the following hypotheses:

(i)   Country clusters do not change a lot over the period 1960 to 1990 with any of the data set.

(ii)   The grouping pattern of countries does not change significantly with the inclusion of real per capita gross domestic product adjusted for purchasing power parity expressed in International Dollars (PCRGDP) in the variable list.

(iii)   Grouping pattern followed by the level of achievement of development differs with the grouping pattern of countries on the basis of PCGDP.

(iv)   Classification of countries on the basis of Physical Quality of Life Index (PQLI) variables differs a lot from classification of countries on the basis of achievement of development when some other development indicators are taken into account.

In order to test these hypotheses our classification scheme follows three steps: (a) In step one we use the core indicators of life to classify the countries, (b) in the second step we add a few back ground indicators for the purpose of classification and (c) the third step proceeds to include per capita real gross domestic product adjusted for purchasing power parity (PCRGDP) along with the nine variables used in step 2 to obtain the grouping pattern of the countries. This study incorporates four points of time, representing three decades of development, starting from 1960 enabling us to have an idea of the changing pattern of the grouping behaviour over the years.

The plan of the paper is as follows: section 2 describes the choice of variables, section 3 discusses the sources of data, section 4 illustrates the methodology, section 5 presents the findings and finally section 6 concludes.

II. CHOICE OF VARIABLES

As has been mentioned in the previous section our choice of variables proceeds in three steps. Our first set of variable, which we defined as the core set of variables used in the Physical Quality of Life Index (PQLI), the most widely used composite index to measure the level of development achieved. These variables are: life expectancy at birth, infant morality rate and adult literacy rate. In fact these three simple indicators taken together provide the best quantitative perspective on the quality of life or standard of living available in virtually all countries. These indicators are all output measures of development and they simultaneously reflect the achievement of several basic needs and give some idea of the distribution of welfare in the country.

The first indicator, life expectancy at birth directly reflects the levels of health, nutrition and income and thus indirectly links the employment and shelter. A low figure usually shows there is a sizable percentage of the population facing poor living conditions and there is a lack of proper health facilities in the country. The second indicator, infant mortality rate (in our study infant survival rate) not only helps to interpret life expectancy measures but also reflects nutrition, and sanitary conditions.

This indicator also reflects the prevalence of contagious disease in a country, as infants are most susceptible to these problems. It is a more sensitive indicator than life expectancy and shows a very rapid response to many health policies. It is obvious that where infant mortality is very high there are many people living in conditions under which basic health needs are not met. The third indicator used in the PQLI, adult literacy rate is a direct measure of the achievement of one basic right of the human beings, minimum education. This indicator is also well correlated with many other indices of quality of life such as measures of employment, income or health and therefore adult literacy can be considered as the most excellent overall quality of life indicator. With high values of these indicators such as, hundred percent of literacy, ten to twenty per thousand of infant morality and nearly eighty years of life expectancy prevailing in a country, it would be assumed that basic needs of life are adequately met in the country. In our study these three indicators together considered as our first set of data.

In our second set of data we widen our scope to include several back ground indicators essential to provide the context for interpreting the other measures of basic needs of life. We choose these additional indicators with a view that they also reflect the outcomes of development measures and policies adopted in a country. The additional indicators used in these sets are percentage of urban population to total population, physician per thousand of population, calorie supply per capita as percentage of requirement, teacher-pupil ratio in the primary level of education, passenger car per thousand of population, energy consumption per capita (kilograms of coal equivalent), per capita gross domestic product adjusted for purchasing power parity. These variables have been chosen for the following reasons.

We incorporate percentage of total population residing in the urban areas with a view that this indicates the percentage of total population with urban facilities such as better sanitation, medical and educational facilities, better communications, access to safe water etc. The second indicator chosen is the physician per thousand of population provides a general picture of the quantity of health care available in a country. The third additional variable chosen is the calorie intake per capita as percentage of requirement. For countries where available calorie per

person is well below the level of requirements there exists almost certainly a significant malnutrition problem for much of the population. The teacher-pupil ratio in the primary level of education gives an indication of quantity of educational care available in the primary level of education. The higher the number of pupil per teacher, lower is the quantity and quality of personal care available for each pupil in the school. Passenger car per thousand of population, the fifth additional indicator used in this study reflects the transportation system available in a country for the public. Better the transport facilities of a country more wide spread is the medical and educational facilities also. The last indicator incorporated in this set of data is the energy consumption per capita. In fact energy consumption of a country depends mostly on the availability of energy in the country or the possibility of importing. But it is obvious that modern amenities of life and industrialization, use of modern technology depends to a great extent on the availability of energy per capita in a country.

In the third set of data, in addition to these nine variables mentioned, we include PCRGDP. In fact, a single variable indicating the overall performance of a country is GDP per capita.

III. DATA

The data for 1960, 1970, 1980 and 1990 on all the social indicators considered in this study have been obtained from different issues of World Tables and World Development Reports. Figures not available for these 4 years have been imputed in some cases by figures from around these years, and in some cases by estimated figures using trends. The data for actual calorie supply for 1990 are available but percentage of requirements of calorie per capita for 1990 are not available. We approximate these figures using the ratio between percentage of requirement and actual supply for the year 1980.

A few variables required to be transformed to make all the variables move in the same direction along with the level of development. Data obtained from the official sources mentioned are mortality rate population per physician, pupil-teacher ratio. These variables are negatively

related with the level of development while other variables have positive correlation with the level of development. We transform these variables as follows:

(a)   Infant survival rate = 1000 − Infant mortality rate.

(b)   Physician per 1000 of population = 1000/population per Physician.

(c)   Teacher-pupil ratio = 1/Pupil-teacher ratio.

The data for per capita real gross domestic product (PCRGDP) have been taken from Summers and Heston (1984, 1988). We have all figures for 92 countries for 1960, 1970, 1980, 1990.

IV. METHODOLOGY

We use the technique of cluster to classify the countries on the basis of selected indicators of life. Cluster analysis is the modern statistical method of partitioning an observed sample population into relatively homogenous classes to produce an operational classification. The objective is to sort observations into groups called clusters so that the degree of statistical association is high among members of the same group and low between members of different groups.

Statistical association is expressed in terms of the squared difference between the values of the variables of cluster members and the cluster means. The partitioning into clusters is performed such that the sum of such differences over all variables, countries is minimized.

We use three data sets and attempt to form five clusters using 92 countries for 1960, 1970, 1980, 1990.

V. FINDINGS

The table attached presents groupings of the countries obtained from the cluster analysis using data set I, II, and III respectively for 4 years under study 1960, 1970, 1980 and 1990 such that we can have an idea

of the changing position of a country over the years with different data sets. We test the hypotheses mentioned in section I. The findings are as follows:

(i)   Grouping pattern of the countries using data sets I, II and III throws light to the fact that many countries are not changing their positions over the years under study. But some countries are changing their positions in favour of upward-direction and some others are moving downwards.

(ii)   Grouping on the basis of data set II and III almost similar for all the years. In other words, achievement of basic needs fullfilment and the level of PCRGDP achieved by a country is highly correlated.

(iii)   Comparing the grouping pattern of countries for all the data sets and the grouping pattern of countries by the World Bank on the level of PCGDP we find that (a) group I and II consist of countries mainly from high income and upper middle income group; (b) group III comprises countries from upper middle income and lower income group; (c) group IV and V consist of countries from the lower income bracket of the World Bank grouping. These findings are almost similar for four years under study.

(iv)   Almost all countries are found to be in the relatively higher groups when classified using PQLI variables (i.e. data set I) compared to their positions when grouped using more social indicators (i.e. when data set II is used). This indicates that the countries in general try to develop in the core indicators of life first.

The study also picturises the fact that the grouping pattern of the countries obtained on the basis of socio-economic aspects of life follows some regional pattern. Countries representing group I in our study are mainly from North America and West Europe. Countries in the second group belong to West Europe and Oceania. The third group comprises of countries from South America and Asia Pacific Rim mostly. Two

TABLE I

Grouping of countries on the basis of different data sets

Country	Years	Data Set I				Groups Data Set II				Data Set III			
		1960	1970	1980	1990	1960	1970	1980	1990	1960	1970	1980	1990
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Algeria		2	2	2	2	5	5	5	5	4	5	5	4
Argentina		2	1	1	2	2	2	2	3	2	2	2	3
Australia		1	1	1	1	1	1	1	2	1	1	1	2
Austria		1	1	1	1	2	2	2	2	2	2	2	2
Bangladesh		5	5	5	5	5	5	5	5	5	5	5	5
Barbados		2	1	1	1	4	4	4	3	3	4	3	3
Belgium		3	1	1	1	3	1	1	2	2	1	1	2
Benin		5	5	5	5	5	5	5	5	5	5	5	5
Bolivia		3	3	4	3	5	5	4	5	5	5	4	5
Brazil		4	4	2	4	4	4	3	3	3	4	3	3
Burma		3	3	3	3	4	4	4	4	3	4	4	4
Burundi		5	5	5	5	5	5	5	5	5	5	5	5
Cafrrep		5	5	5	3	5	5	5	5	5	5	5	5
Cameroon		5	5	3	3	5	5	5	5	5	5	5	5
Canada		1	1	1	1	1	1	1	1	1	1	1	1
Chile		2	2	2	1	4	3	3	3	3	3	3	3

Table I contd.

Country	Years	Data Set I				Data Set II				Data Set III			
		1960	1970	1980	1990	1960	1970	1980	1990	1960	1970	1980	1990
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Columbia		4	2	2	2	4	4	4	4	3	4	4	4
Congo		5	4	4	3	5	5	4	4	5	4	4	4
Costarica		2	2	1	1	4	4	4	4	3	4	3	4
Cyprus		2	2	1	1	4	4	3	3	3	3	3	3
Denmark		1	1	1	1	2	2	2	2	2	1	2	2
Dominican		4	4	4	4	4	4	4	4	3	4	4	4
Ecuador		4	4	2	4	4	4	4	4	3	4	4	4
Egypt		3	3	3	3	3	5	4	4	4	4	4	4
Elsalvador		3	3	3	3	3	4	4	4	4	4	4	4
Ethiopia		5	5	5	5	5	5	5	5	5	5	5	5
Finland		1	1	1	1	4	3	2	1	2	1	2	2
France		1	1	1	1	2	1	1	2	2	1	1	2
FRG		1	1	1	1	2	1	1	2	2	1	1	1
Gambia		5	5	5	5	5	5	5	5	5	5	5	5
Ghana		3	3	3	3	5	5	5	5	4	5	5	5
Greece		2	1	1	1	4	2	4	3	3	2	2	3

Groups

Country												
Guatemala	3	4	4	4	3	3	5	4	4	4	4	4
Guinea	5	5	5	5	5	5	5	5	5	5	5	5
Guyana	2	2	1	2	4	4	4	4	3	4	4	4
Haiti	5	3	3	3	5	5	5	5	5	5	5	5
Hondurus	3	4	4	4	3	3	4	4	4	4	4	4
Hongkong	2	2	1	1	2	3	3	3	2	3	3	3
Iceland	1	1	1	1	2	1	1	1	1	1	1	1
India	3	3	3	3	5	5	5	5	5	5	5	5
Iran	5	3	4	3	3	5	5	4	4	5	4	4
Iraq	5	3	1	2	3	3	5	4	4	5	4	3
Ireland	1	1	1	1	2	2	2	2	3	3	3	3
Israel	2	1	1	1	2	2	2	2	3	2	2	2
Italy	1	1	1	1	2	2	2	2	2	2	2	2
Japan	1	1	1	1	2	3	4	2	3	2	2	2
Jordan	3	4	4	4	3	5	4	4	4	4	4	4
Kenya	5	3	3	3	5	4	5	5	5	5	5	5
Korea	2	2	2	1	4	4	4	3	3	4	3	3
Liberia	5	5	5	5	5	5	5	5	5	5	5	5
Madagascar	3	3	3	3	5	5	5	5	5	5	5	5
Malawi	5	5	5	5	5	5	5	5	5	5	5	5
Malaysia	4	4	4	4	3	4	4	4	4	4	4	4

Table I contd.

Country	Years	Groups											
		Data Set I				Data Set II				Data Set III			
		1960	1970	1980	1990	1960	1970	1980	1990	1960	1970	1980	1990
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Maritius		4	2	2	2	4	4	4	4	3	4	4	4
Mexico		4	2	2	2	4	4	3	3	3	3	3	3
Morocco		5	3	3	3	5	5	5	5	5	5	5	5
Mozambique		5	5	5	5	5	5	5	5	5	5	5	5
Netherlands		1	1	1	1	2	1	1	2	2	1	1	2
New Zealand		1	1	1	1	1	1	1	2	1	1	1	2
Nicaragua		3	4	2	2	3	4	4	4	4	4	4	4
Niger		5	5	5	5	5	5	5	5	5	5	5	5
Nigeria		5	5	3	3	5	5	5	5	5	5	5	5
Norway		1	1	1	1	1	1	1	1	1	1	1	1
Pakistan		5	5	5	5	5	5	5	5	5	5	5	5
Panama		2	2	2	2	4	4	4	4	3	4	4	4
Papuang		5	3	3	3	5	5	5	5	5	5	5	5
Paraguay		4	2	2	2	4	4	4	4	3	5	5	5
Peru		4	4	4	4	4	4	4	4	3	4	4	4
Philippines		4	2	2	2	4	4	4	4	3	4	4	4

Country													
Portugal	4	2	2	2	2	4	4	2	3	3	3	4	3
Rwanda	5	5	3	3	5	5	5	5	5	5	5	5	5
Sierraleone	5	5	5	5	5	5	5	5	4	5	5	5	5
Singapore	4	2	2	2	3	3	3	2	2	3	3	3	3
Spain	2	1	1	2	2	2	2	2	2	3	3	2	2
Srilanka	2	2	2	4	4	4	4	3	4	4	4	4	4
Sudan	5	5	4	5	5	5	5	5	5	5	5	5	5
Surianme	2	2	2	4	4	4	4	3	4	3	4	3	4
Sweden	1	1	1	1	1	1	1	1	1	1	1	1	1
Syrianar	3	4	4	3	5	5	5	5	4	5	5	4	4
Tanzania	5	4	4	5	5	5	5	5	4	5	4	4	4
Thailand	4	2	2	4	4	4	4	3	4	3	4	4	4
Togo	5	5	3	5	5	5	5	5	5	5	5	5	5
Trin & Tobago	2	1	1	1	1	1	1	1	4	4	4	4	4
Turkey	3	4	4	3	3	4	4	4	4	4	4	4	4
UK	1	1	1	1	1	1	1	1	1	1	1	1	1
Upper Volta	5	5	5	5	5	5	5	5	5	5	5	5	5
Uruguay	2	1	1	2	3	3	3	2	3	2	3	3	3
USA	1	1	1	1	1	1	1	1	1	1	1	1	1
Venezuela	4	2	2	3	3	3	3	2	3	3	3	3	3

Table I contd.

Country		Data Set I				Data Set II				Data Set III			
Years	1960	1970	1980	1990	1960	1970	1980	1990	1960	1970	1980	1990
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Zaire	3	5	3	3	5	5	5	5	5	5	5	5
Zambia	5	4	3	4	5	5	5	4	5	5	5	5
Zimbabwe	3	4	4	4	3	4	4	4	4	4	4	4

Notes:

(1) Data Set I   includes variables life expectancy at birth, infant survival rate and adult literacy rate.

Data Set II   consists of physician per 1000 of population, urban population as percentage of total population, calorie intake as percentage of requirement, passenger car per 1000 of population, energy consumption per capita along with the variables included in Data Set I.

Data Set III   incorporates PCRGDP along with the variables in Data Set II.

(2) Figures in columns (2) to (13) denote in which groups the country belongs for the three Data Sets for four years.

(3) CAFRREP is Central African Republic
    DOM REP is Domican Republic
    FRG is Federal Republic of Germany
    KOREA is Republic of Korea
    PAPUANG is Papua New Guinea
    SYRINAR is Syrian Arab Republic

countries from West Europe. The fourth and fifth group mostly consist of countries from Asia, Africa and a few from South America.

VI. CONCLUSIONS

The study throws light to the fact that fullfilment of social attributes of life and achievement of PCRGDP goes hand in hand. In other words, if one country reaches certain level of PCRGDP it is almost sure to attain certain level of social standard.

The importance of the study lies in its attempt to classify the countries on the basis of socio-economic aspects of life instead of only economic aspects of life.

The study once again picturises the phenomenon of regional concentration of the level of socio-economic development i.e. the peripherial effect is present.

ACKNOWLEDGMENT

REFERENCES

Anderberg, M.: 1973, Cluster Analysis for Applications (Academic Press, New York).
Clague, Christopher: 1986, 'Short-cut estimates of real income', Review of Income and Wealth, 32(3), pp. 313–331.
Everitt, B.: 1974, Cluster Analysis (Heinemann Educational Books, London).
Havlik: 1991, 'East-west comparisons: problems, methods and results, no. 174', Vienna Institute For Comparative Economic Studies.
Janossay and Enrlich: Article in Economic Survey of Europe 1969 by the U.N.
Maser, P.: 1985, 'The dollar GNP of USSR and E. Europe', Johns Hopkins University Press.
Maser, P.: 1985, 'Alternative estimates of dollar GNP and growth rates of CMEA countries', Joint Economic Committee, U.S. Congress.
World Bank: World Development Report (World Bank, Washington, D.C., different issues).
World Bank: World Tables (World Bank, Washington, D.C., different issues).