

Identifying chronically deprived countries: results from cluster analysis

Edward Anderson, January 2007



**Poverty and Public Policy Group
Overseas Development Institute
111 Westminster Bridge Road, London
SE1 7JD, UK**

e.anderson@odi.org.uk

CPRC Working Paper 70

**Chronic Poverty Research Centre
ISBN 1-904049-69-9**

Abstract

This paper provides a classification of non-OECD countries on the basis of the level of, and the change in, their average welfare over recent decades. The method used is that of cluster analysis. The main finding is that four welfare indicators – GDP per capita, child mortality, fertility and under-nourishment – all show strong evidence of three distinct country clusters: a ‘chronically deprived’ group, characterised by low initial levels of welfare and relatively slow rates of progress; a ‘good performers’ group, consisting of countries with similarly low initial levels of welfare but faster rates of progress; and a group of ‘others’, consisting of countries with higher initial levels of welfare and moderate rates of progress.

On the basis of these results, approximately 30 countries are defined as chronically deprived ‘across-the-board’, i.e. with all four welfare indicators showing the three-cluster pattern (or at least all of those four for which data are available). These account for approximately 10% of the total population of non-OECD countries, but close to 20% of the total number of people living in less than \$1-a-day poverty, and more than 30% of the total number of child deaths. A further 45 or so countries are defined as ‘partially’ chronically deprived, i.e. with one but not all of the four indicators showing the three-cluster pattern. They account for over 50% of total \$1-a-day poverty and 55% of total child deaths, compared with approximately 40% of total population.

Various factors increase the probability of being a chronically deprived country. Although it is difficult to separate out the effects of these different influences, the results of multivariate regression analysis suggest that geography (e.g. location in tropics, distance from major markets) and external conditions (e.g. shifts in the terms of trade) play the largest role. Furthermore, the most recent rates of progress in chronically deprived countries continue to lag significantly behind rates of progress in all other non-OECD countries. This suggests that whatever has been driving the differences in performance between chronically deprived and other countries, it has been persisting in recent years.

As regards policy implications, one of the main issues is the amount of aid the different groups receive. The across-the-board chronically deprived countries currently receive around 30% of total aid. Although this is close to their shares of total poverty and total child mortality, one could make a case for their share of aid being even larger, given the natural disadvantages they appear to face. The partially chronically deprived countries are more clearly under-aided: their share of total aid is also currently around 30%, compared with their shares of total \$1-a-day poverty and child deaths of over 50%. Whether donors and multilateral agencies are pursuing the right aid strategies in the different country groups is something which needs to be considered further.

Acknowledgements

This paper was commissioned as a background paper to the 2nd Chronic Poverty Report and was funded by AusAID. The author is grateful for comments and suggestions from Andrew Shepherd, Ursula Grant, Martin Prowse, Tony Addison and Bob Baulch.

Edward Anderson is a Research Fellow at the Overseas Development Institute in London. His main research interests are the links between inequality, poverty and growth, the effects of globalisation on inequality, and international aid allocation. Prior to joining ODI he was a Lecturer in Economics at Keele University, having completed his PhD in Economics at the University of Sussex.

Annexes

Four Excel spreadsheets containing Annexes 1-4 are available on request from the author, or from the CPRC website: <http://www.chronicpoverty.org/resources/cp70.htm>.

Table of contents

	<i>Abstract and acknowledgements</i>	<i>(i)</i>
1.	Introduction	1
2.	Method	1
3.	Data	3
4.	Results of cluster analysis	5
4.1	Existence and number of clusters	5
4.2	Location of clusters	5
4.3	Identifying country groups	6
4.4	Summary of results	8
5.	Characteristics of country groups	8
5.1	Share of total non—OECD aggregates	8
5.2	Most recent growth rates	9
5.3	Characterisation of country groups	10
5.4	Growth trajectories	12
6.	Conclusion	15
	References	17
	Tables 1 – 17	18
	Figures 1 - 14	38

1. Introduction

This paper asks whether distinct groups of developing countries can be identified, using statistical methods, on the basis of both the level of average welfare and the change in average welfare witnessed in recent decades. Potential groups include:

- ‘chronically deprived’ countries: characterised by low initial levels of welfare and stagnant or negative improvements in welfare in recent decades;
- ‘good performers’: countries with initially low levels of welfare but strong improvements in welfare in recent decades; and
- ‘poor performers’: countries with initially moderate levels of welfare or deprivation, but slow, stagnant or negative improvements in welfare in recent decades.

These groups are shown graphically in Figures 1 and 2. Further categories could of course be defined: for example, the ‘good performers’ group could be split into ‘moderate’ and ‘top’ performers, as in Figure 3.

Various country classification schemes exist in the literature already. Of these, the approach used in this paper shares most in common with the definition of ‘top-priority’ and ‘medium-priority’ countries provided by UNDP (2003). This also groups countries according to both their initial starting points and their progress in recent decades (see Figure 4). The main difference is that, in the UNDP (2003) approach, the threshold values separating the country groups are determined in advance. For child mortality for example, top priority countries are defined as those with an initial mortality rate greater than 100 deaths per 1,000 live births, and a rate of reduction in mortality which is less than half the rate required to meet the MDG target of reducing child mortality by two thirds by 2015.

One limitation with the UNDP approach (2003), however, is that many countries may be quite close to the particular threshold values chosen, and that small differences between countries on either side of the thresholds will as a result be given too much emphasis. By contrast, in the approach used in this paper the threshold values are determined by the data, in a way which seeks to keep to a minimum the number of countries which are just either side of the thresholds. The approach in this paper also allows for the possibility that no threshold levels, and therefore no country groupings, can be identified. This may appear a disadvantage, but it is actually a strength. It would after all be misleading to provide a country classification analysis if there was a continuum of different country outcomes, rather than sets of reasonably distinct country groups.

2. Method

The method used to classify countries is the technique of cluster analysis.¹ The basic principle is to combine countries which are similar to one another into subsets or clusters. The first step is to decide how the similarity or dissimilarity between any two countries is measured. There are various measures available, but in this paper one of the most common measures is used, that of Euclidian distance. This is denoted D_{ij} and is given by:

$$D_{ij} = \sqrt{(y_i^0 - y_j^0)^2 + (\Delta y_i - \Delta y_j)^2}, \quad (1)$$

¹ For a detailed exposition of the technique of cluster analysis see Bartholomew *et al* (2002).

where y_i^0 and y_j^0 are the initial values of some welfare indicator in countries i and j at the start of the relevant period, and Δy_i and Δy_j are the changes in the indicator in countries i and j over the relevant period. This measure of distance can be represented graphically as the shortest distance between two countries plotted on a graph with, as in Figures 1-3, the initial level of the indicator shown on the horizontal axis and the change in the indicator on the vertical axis.²

The second step is to decide how the distance between any two clusters is measured. Again there are various options, but in this paper the Ward's linkage approach is used, because of its similarity with the well-known analysis of variance (ANOVA) method. The Ward's linkage method begins by locating each cluster's central value. This is denoted $(y_r^0, \Delta y_r)$ and is given by:

$$(y_r^0, \Delta y_r) = \left(\frac{1}{n_r} \sum_{i=1}^n y_{ri}^0, \frac{1}{n_r} \sum_{i=1}^n \Delta y_{ri} \right), \quad (2)$$

where y_{ri}^0 is the initial value of the welfare indicator for country i in cluster r , Δy_{ri} is the change in the welfare indicator over the relevant period for country i in cluster r , and n_r is the number of countries in cluster r . The approach then measures each cluster's 'within-cluster' sum of squares. This is defined as the sum of the squares of the distances between all countries in the cluster and the central value. The approach then measures the distance between two clusters by the extent to which their within-cluster sum of squares when combined exceeds the sum of their within-cluster sums of squares when kept separate.

The clustering process itself can be illustrated using a diagram known as a dendrogram (Figure 5). It proceeds either by initially treating all countries as representing one cluster and then dividing that cluster into successively smaller clusters (divisive or top-down clustering), or by initially treating each country as a separate cluster and then merging into successively larger clusters (agglomerative or bottom-up clustering).

The final step is to determine how many clusters should be defined. Here there is no hard and fast rule, only a rule of thumb known as the 'elbow criterion'. This can be assessed by plotting the proportion of total variation in the data explained by the clustering process against the number of clusters (see Figure 6).³ This will generate a line stretching from 0 (when one cluster contains all countries) to 100% (when the number of clusters equals the number of countries). If at some point along the x-axis there is a marked kink or deceleration in this line, this would indicate that adding another cluster does not add very much in terms of explaining the patterns displayed by the data. This would in turn indicate the appropriate number of clusters (in Figure 6, this number would be 4). If, by contrast, there was no

² To use this measure of distance both y^0 and Δy must be measured on the same scale. This is achieved through normalisation, i.e. expressing y^0 and Δy in terms of standard deviations from their mean.

³ The proportion of total variance in the data explained by the clusters is defined as (one minus) the ratio of the sum of the within-cluster sums of squares to the sum of squares among the dataset as a whole.

marked kink in the line, this would indicate that the data do not fit into clearly identifiable and distinct groups.

We use two measures to determine whether there is a marked kink in the relationship between the number of clusters and the proportion of total variation in the data explained by the clusters. The first is the absolute difference between the additional amount of variation in the data explained by one cluster and the additional amount explained by the next cluster. This is denoted $d(abs)_k$ and is given by:

$$d(abs)_k = dp_k - dp_{k+1}, \quad (3)$$

where p_k is the proportion of total variance in the data explained by all clusters up to and including cluster number k , and $dp_k = p_k - p_{k-1}$. The second measure is the relative difference between the additional amount of variation in the data explained by one cluster and the additional amount explained by the next cluster. This is denoted $d(rel)_k$ and is given by:

$$d(rel)_k = dp_k / dp_{k+1}. \quad (4)$$

The value of k at which the values of $d(abs)_k$ and $d(rel)_k$ are highest indicates the point in the relationship between the number of clusters and the proportion of total variation in the data explained by the clusters at which there is the largest kink; and the higher their values, the larger the kink. (For an illustration based on the hypothetical example shown in Figure 6, see Table 1). The two measures of difference do not always tally, however, nor is there a natural cut-off point indicating whether the maximum values of $d(abs)_k$ or $d(rel)_k$ are in some sense significant. In this paper, strong evidence of clustering is defined as being indicated by: a) the maximum value of $d(abs)_k$ exceeding 20 percentage points; b) the maximum value of $d(rel)_k$ exceeding 4; and c) both maxima occurring at the same value of k . However, it should be recognised that these choices are somewhat arbitrary, and there may as a result often be an unavoidable amount of uncertainty about the appropriate number of country clusters, and indeed whether clusters of countries really do exist.

3. Data

The method described in the previous section is applied to seven commonly used and widely available indicators of welfare: GDP per capita, life expectancy, adult literacy, child mortality, infant mortality, the total fertility rate, and the prevalence of under-nourishment. The GDP per capita data are measured in PPP US\$, taken from the most recent Penn World Tables.⁴ The under-nourishment data are obtained from the FAO Statistics Division.⁵ The data on life expectancy, adult literacy, child and infant mortality and fertility are from World Development Indicators (World Bank, 2005). The main reason for using more than one indicator is that the ordinal and cardinal ranking of welfare levels across countries, and the comparison of

⁴ Available at: <http://pwt.econ.upenn.edu>.

⁵ Available at: <http://www.fao.org/faostat/foodsecurity>.

welfare changes over time, typically vary according to the welfare measure used. We restrict the analysis to 130 non-OECD countries with populations exceeding one million.⁶

We begin by calculating the average rate of growth of each indicator in each country over the periods 1960-1993 and 1970-2003. We do this for all countries which have data for periods covering at least 20 years in each period. The countries which do not have data for the indicators covering a period of at least 20 years, or which lack data for the indicators altogether, are listed in Annex 1.

We then calculate both the proportional and the absolute average annual growth rate of each indicator, by estimating an ordinary least squares regression of the form:

$$y_{it} = \alpha_i + \beta_i \cdot t + \varepsilon_{it}, \quad (5)$$

where t is the year, y_{it} is the level of the welfare indicator in country i in year t , and ε_{it} is an error term. The slope coefficients β_i give the average proportional annual growth rate of each country when y_{it} is expressed in logarithms, and the average absolute annual growth rate of each country when y_{it} is expressed in standard units (e.g. US\$ or deaths per 1,000 births).

For each indicator we measure growth in either proportional or absolute terms, and initial levels in either standard or log units, depending on which of the four possible combinations generates the patterns conforming most closely to the stylised groups shown in Figures 1-3. Prior analysis suggested that the most appropriate combinations according to this criterion were: proportional growth rates and standard initial levels (GDP per capita); absolute growth rates and standard initial levels (life expectancy and adult literacy); and proportional growth rates and log initial levels (child and infant mortality, fertility and under-nourishment).

We repeat the analysis separately for each indicator, and then compare the results. We also compare the results across the two time periods (1960-93 and 1970-2003). Finally, we carry out three robustness tests. First, we repeat the calculations using a longer time period, 1960-2003 (although maintaining a minimum period length of 20 years in any one country). Second, we restrict the sample to countries with data for periods of at least 30 years during 1970-2003 and 1960-1993, and for at least 40 years during 1960-2003. In this case the data for each country refer to periods of similar length, although the number of countries which can be included in the analysis is smaller. We also restrict the sample to the countries with data for all seven of our indicators over a period of at least 30 years (this can only be done for the period 1970-2003). In this case direct comparisons of the results across the different indicators can be made, as both the sample of countries and the approximate length of period are held constant. Finally, we repeat the analysis using an alternative measure of distance between countries, the squared Euclidian distance, and an alternative method for measuring distance between clusters, the between-groups linkage method.⁷

⁶ There are 184 non-OECD countries in the World Bank (2005) database (when using the pre-1996 OECD membership). Of these, 130 have populations greater than 1 million.

⁷ The between-groups linkage method measures the distance between two clusters as the average distance between all inter-cluster pairs.

4. Results of cluster analysis

4.1 Existence and number of clusters

Figures 7 and 8 plot the proportion of total variation explained by the clusters against the number of clusters, for each of the seven indicators and when using the full sample of countries, for 1970-2003 and 1960-1993, respectively. Recall that large kinks in this relationship are evidence of distinct country clusters, with the point along the x-axis at which the kinks occur indicating the number of clusters. Visual inspection of Figures 7 and 8 in this light suggests that clustering is arguably a phenomenon in four of our seven indicators, namely child mortality, fertility, under-nourishment and adult literacy. For child mortality, fertility, under-nourishment, the suggested number of clusters indicated is three, whereas for adult literacy it is two.

These results are confirmed by Table 2, which shows the values of $d(abs)_k$ and $d(rel)_k$ for each indicator, again when using the full sample of countries and for the periods 1960-1993 and 1970-2003, respectively. The results for child mortality, fertility and under-nourishment all show strong evidence of three clusters, while those for adult literacy show strong evidence of two clusters. The results for GDP per capita also show strong evidence of three clusters. By contrast, the results for life expectancy and infant mortality do not show strong evidence of clustering.

A summary of the results for other time periods and country samples is shown in Table 3. On the whole, the evidence of clustering is less strong when considering shorter time periods and smaller country samples. The main exception is GDP per capita, which does not show strong evidence of clustering over the period 1960-2003, but does show strong evidence for the periods 1970-2003 and 1960-1993. The results are also very similar when using the squared Euclidian distance as the measure of similarity between any two countries, except that in this case GDP per capita does not show strong evidence of clustering in any of the three periods (details available on request). The results are more different when using the between-groups linkage method, however, which tends to place more weight on country outliers in the aggregation process. Nevertheless, there continues to be strong evidence of two clusters for the adult literacy rate during 1960-1993 and 1970-2003, and three clusters during 1960-2003 and 1970-2003 for the fertility rate. For child mortality and under-nourishment, there is now strong evidence of four clusters, during 1960-1993 and 1970-2003, respectively (details available on request).

Overall, five of the seven indicators show strong evidence of clustering for at least one of the seven combinations of country samples and time periods. Of these, the indicator showing the most consistently strong evidence of clustering is the fertility rate. For four of these indicators (GDP per capita, fertility, child mortality and under-nourishment), the number of clusters is generally three, while for the other indicator (adult literacy) the number of clusters is two.

4.2 Location of clusters

We now ask what sorts of country clusters emerge from the clustering process. To shed light on this question, Figures 8-14 show the results in graphical form, while Table 4 shows the range of growth rates and initial starting levels contained in each cluster. In each case the focus is on the five indicators showing most evidence of clustering (GDP per capita, child mortality, fertility, under-nourishment and adult literacy). Also, the results shown are those for the full samples of countries for both 1960-1993 and 1970-2003. The number of clusters, as determined by the elbow criterion, in these cases is three for GDP per capita, child mortality, fertility and under-nourishment, and two for adult literacy. Lists of countries contained in each cluster, together with each country's growth rate and starting level, are contained in Annex 2.

The main finding from Figures 8-14 and Table 4 is that GDP per capita, child mortality, fertility and under-nourishment all show the same basic clustering pattern, consisting of:

- 'chronically deprived countries', characterised by relatively low initial levels of welfare (relatively low GDP per capita, and relatively high mortality, fertility and under-nourishment) and relatively slow rates of progress over time;
- 'good performers': countries also with relatively low initial levels of welfare, but faster rates of progress over time; and
- 'others': countries with higher initial levels of welfare, and with rates of progress somewhere in between the other two clusters.

For adult literacy however, we see a different cluster pattern. Here the clusters have both different initial levels and different rates of progress over time; one with low initial levels but progressing relatively rapidly, the other with higher initial levels progressing less rapidly. Clearly, there is no 'chronically deprived' cluster in this case.

These results refer only to the full country sample, but equivalent results for other country samples and for 1960-2003 are very similar. Although the ranges of initial starting levels and growth rates in each cluster vary somewhat, the same three broad country clusters emerge. When child mortality shows more evidence of a two-cluster rather than three-cluster pattern (e.g. over the period 1970-2003 when using the common country sample), the only difference is that the 'good performers' cluster merges with 'others', with the 'chronically deprived' cluster remaining separate. When child mortality and under-nourishment show evidence of a four-cluster pattern (as when using the between-group linkages method), the difference is that the 'good performers' cluster divides into one cluster of 'top performers' and another cluster of 'moderate performers', as in Figure 3.

4.3 Identifying country groups

The next question is which countries feature in each of the clusters identified in the previous sub-section. Of particular interest is whether the same countries feature in each cluster when considering the four different indicators showing strong evidence of the three-cluster pattern (we omit adult literacy from now on, on the basis that it shows an altogether different cluster pattern to the other indicators). Table 5 sheds light on this question, by dividing countries into five groups:

- across-the-board chronically deprived countries: countries which are chronically deprived in all four indicators;
- across-the-board good performers: countries which are good performers in all four indicators;
- across-the-board others: countries which are neither chronically deprived nor good performers in any of the four indicators;
- partially chronically deprived countries: countries which are chronically deprived in one but not all of the four indicators; and
- partial good performers: countries which are good performers in one but not all of the four indicators, and which are not chronically deprived in any of the four indicators.

The results are shown for the whole country sample, the two periods 1960-1993 and 1970-2003. The first row in each panel shows the results when considering the 73 countries with data for all four indicators. The second row then adds the 16 countries with data for child mortality, fertility and under-nourishment but not GDP per capita. Of these, those which are chronically deprived, good performers or others in all of the three indicators are included in the across-the-board groups; those which are chronically deprived in one but not all of the three indicators are placed in the partially chronically deprived group, and the remainder are placed in the partial good performers group. The third row then adds the 36 countries with

data for child mortality and fertility but not GDP per capita and/or under-nourishment. Of these, those which are chronically deprived, good performers or others in both indicators are included in the across-the-board groups; those which are chronically deprived in one but not both indicators are placed in the partially chronically deprived group, and the remainder are placed in the partial good performers group. This leaves only five countries which remain unclassified.

The results show a fairly large number of across-the-board chronically deprived countries: 32 during 1970-2003 and 27 during 1960-1993. There are fewer across-the-board good performers: 11 during both 1970-2003 and 1960-1993. However, there are more partially chronically deprived countries: 44 countries during 1970-2003 and 37 during 1960-1993. Thus, the number of both across-the-board and partially chronically deprived countries has increased over time, while the number of across-the-board good performers has remained stable.

Lists of the countries making up each of the groups shown in Table 5 for 1970-2003 are shown in Tables 6-11. These also show whether the countries were also in the same category during 1960-1993. For example, of the 32 across-the-board chronically deprived countries listed in Table 6, 22 were also chronically deprived across-the-board during 1960-1993 (a further four have no data for that period). Similarly, of the 11 across-the-board good performers listed in Table 7, six were also across-the-board good performers during 1960-1993 (one other country has no data for that period), while of the 18 across-the-board others in Table 8, 15 were also across-the-board others during 1960-1993. This is to be expected given the large amount of overlap between the two periods.

For the two 'partial' groups, Tables 9-10 make a further distinction between which indicator(s) each country is chronically deprived or performing well in. Thus, for the 44 partially chronically deprived countries in Table 9, a distinction can be made between countries which are:

- chronically deprived in GDP per capita, child mortality and fertility, but not in under-nourishment (Benin, Burkina Faso, Ghana, Lesotho, Mali, Mauritania and Nigeria);
- chronically deprived in child mortality, fertility and under-nourishment, but not in GDP per capita (Botswana, Guinea-Bissau, Haiti, Namibia, Pakistan and Uganda);
- chronically deprived in GDP per capita, fertility and under-nourishment, but not in child mortality (Bolivia, Guatemala and Honduras);
- chronically deprived in child mortality and fertility, but not in GDP per capita or under-nourishment (Gabon, Nepal);
- chronically deprived in child mortality and under-nourishment, but not in GDP per capita or fertility (Bangladesh, Panama); and
- chronically deprived in under-nourishment, but none of the other three indicators (Dominican Republic, Jamaica, Sri Lanka, Thailand, Trinidad and Tobago and Venezuela).

The remaining five countries in Table 9 with data for all four indicators follow a more idiosyncratic pattern. Of the other countries, a further eight are chronically deprived in child mortality but not in fertility, while the other five follow a more idiosyncratic pattern.

Similarly, for the 20 partial good performers in Table 10, a distinction can be made between countries which are:

- good performers in child mortality, fertility and under-nourishment, but others in GDP per capita (Algeria, Brazil, Colombia, Costa Rica, El Salvador, Iran, Mexico and Peru);

- good performers in child mortality and fertility but others in under-nourishment (South Korea, Malaysia, United Arab Emirates);
- good performers in fertility and under-nourishment, but others in child mortality (Kuwait, Lebanon);
- good performers in GDP per capita, but others in child mortality and fertility (Romania, Singapore); and
- good performers in child mortality but others in fertility (Bosnia and Herzegovina, Macedonia, Serbia and Montenegro, Mauritius and Chile).

4.4 Summary of results

To summarise, the analysis in this section shows that four welfare indicators (GDP per capita, child mortality, fertility and under-nourishment) all show strong evidence of three distinct country clusters, all of which fit the same broad pattern: one 'chronically deprived' cluster, one 'good performer' cluster, and one 'all others' cluster.

On the basis of these results, one can define five main groups of countries. The first includes countries which are chronically deprived across-the-board; the second includes countries which are partially chronically deprived; the third includes countries which are good performers across-the-board; the fourth includes countries which are partial good performers, while the fifth includes countries which are neither chronically deprived nor good performers in any of the indicators.

We cannot say for certain how many countries are in each of these groups, because we don't have data on all of the four relevant indicators for all 130 countries. The results suggest, however, that approximately 30 countries are chronically deprived across-the-board, a further 45 countries are partially chronically deprived, but only approximately 10 countries are across-the-board good performers. The remaining countries are either partial good performers or across-the-board others (approximately 20 countries in each case).

Finally, we can sometimes differentiate further between countries within these broad groups. Among the partially chronically deprived countries for instance, one can identify countries which are chronically deprived in a particular welfare indicator or set of indicators, but not in others.

5. Characteristics of country groups

This section provides some descriptive information about the groups of countries identified in Section 4.3.

5.1 Share of total non-OECD aggregates

Table 11 shows estimates of the share of the groups in certain key aggregates for all non-OECD countries, namely mortality (infant and child), income poverty (\$1-a-day and \$2-a-day), under-nourishment and international assistance. The population and aid data are taken from World Bank (2005), while the income poverty estimates are from underlying World Bank data gathered from household surveys. The other variables are taken from the same sources outlined in Section 3.⁸

⁸ We extend the sample of countries with estimates of \$1-a-day by assuming that the income distributions of countries lacking household survey data can be approximated by those of a representative country in their respective region. Nevertheless, we do not have poverty or under-

The across-the-board chronically deprived countries account for a much larger share of total mortality, income poverty and under-nourishment than their share of total population. For the period 1970-2003, this share is as high as 36% when considering child mortality, compared with 10% of population. The shares of mortality, income poverty and under-nourishment in this group have also been rising over time. There can be little doubt, therefore, that the proportion of poverty and deprivation contained in chronically deprived countries is rising.⁹

The across-the-board good performers, by contrast, account for a much smaller share of mortality, income poverty and under-nourishment than their share of total population, and this share has been falling over time. The largest shares of mortality, income poverty and under-nourishment are found in the partially chronically deprived countries: up to 57% for \$1-a-day poverty. This share has remained fairly stable over time, rising for some indicators but falling for others.

In terms of international assistance, the shares of aid received by the across-the-board chronically deprived countries and the across-the-board good performers are broadly in line with their shares of mortality, income poverty and under-nourishment. For example, the across-the-board chronically deprived countries during 1970-2003 received 29% of total aid in 2002, which is lower than their shares of both infant and child mortality in 2002, but higher than their shares of income poverty and under-nourishment. Similarly, the across-the-board good performers during 1970-2003 received 14% of total aid in 2002, which is lower than their shares of income poverty and under-nourishment, but higher than their shares of infant and child mortality.

By contrast, the shares of aid received by the partially chronically deprived countries, the partial good performers and the across-the-board others are very different from their shares of mortality and poverty. Considering again the results for 1970-2003, the partially chronically deprived countries received 29% of total aid in 2002, compared with their shares of mortality, income poverty and under-nourishment of between 49% and 57%. By contrast, the partial good performers and the across-the-board others both received 12% of total aid, compared with their shares of mortality, income poverty and under-nourishment of less than 5%. As a group therefore, the partially chronically deprived countries appear to be significantly under-aided, while the partial good performers and the across-the-board others appear significantly over-aided.

5.2 Most recent growth rates

The growth rates used in the cluster analysis typically refer to periods of around 30 years in length. What about the more recent growth rates of countries in each cluster? Are there, for example, any signs of improvement in chronically deprived countries? Are the good performers continuing their progress? To answer these questions we calculate each country's most recent trend in each indicator. We then compare these most recent trends against a) the range of growth rates of their cluster for the period 1970-2003, and b) most recent trends in other countries. Depending on data availability, the starting year of the period over which the most recent trend is measured lies between 1991 and 1998, and the length of period varies between five and nine years.

Table 12 summarises the results (figures for each individual country are shown in Annex 3). The first panel shows the results for the across-the-board chronically deprived countries. Here, there is evidence of improvement in GDP per capita and under-nourishment: in 11 out

nutrition data for all countries which can be classified. This means that the figures in the table are approximations rather than exact figures.

⁹ DFID estimate that around one third of the world's poor live in 'fragile states', and that by 2010 this figure will have risen to around half (DFID White Paper 2006: 7-8). The figures arrived at in this paper are clearly similar to the DFID estimates – although not all chronically deprived countries are necessarily fragile states, and vice versa.

of 23 countries (GDP per capita) and 14 out of 26 countries (under-nourishment), the most recent rates of progress exceed the fastest rate of progress of chronically deprived countries during 1970-2003. For child mortality and fertility, however, there are much fewer signs of improvement: in only one out of 31 countries (child mortality) and two out of 31 countries (fertility) do the most recent rates of progress exceed the fastest rate of progress of chronically deprived countries during 1970-2003. Furthermore, current rates of progress among the across-the-board chronically deprived countries lag significantly behind rates of progress in other non-OECD countries, in all four indicators.

The results for partially chronically deprived countries are shown in the second panel of Table 12. (These figures refer to those countries which were chronically deprived in each indicator, and not to all countries in this group.) The results are similar to those for the across-the-board chronically deprived countries: signs of improvement in GDP per capita and under-nourishment, but not in child mortality and fertility. Most recent rates of progress are also significantly lower than among other non-OECD countries, with the exception of GDP per capita, where the difference is not statistically significant. The third and fourth panels of Table 12 show the results for the across-the-board and partial good performers, respectively. Here the majority of countries are continuing to grow at the rates of progress shown by their clusters during 1970-2003. Moreover, they continue to out-perform other non-OECD countries, with only one or two exceptions.

In summary therefore, relative to the period 1970-2003 as a whole, the most recent years have seen only limited improvement in chronically deprived countries (both across-the-board and partial), and only a small amount of deceleration among the good performers (both across the board and partial). At the same time, there are large and statistically significant differences in most recent rates of progress across these main groups, in all four indicators. Whatever is driving the differences in their performance, it appears to be persisting in recent years.

5.3 Characteristics of country groups

This section asks whether there are observable characteristics which make countries more or less likely to be found in any one particular country group. We look at eight sets of characteristics, namely:

- *geography*, measured by dummy variables for location in the tropics and being land-locked, and by distance in kilometres to major OECD markets (all from Gallup and Sachs, 1999);
- *history*, measured by the ethno-fractionalisation index in 1960 (from Easterly and Levine, 1997), and a proxy for European settlement (the proportion of the population speaking a Western European language in 1960, taken from Hall and Jones, 1999);
- *conflict*, measured by dummy variables for whether a country participated in an inter- or intra-state war over the relevant period (from Sarkees, 2000);
- *resources*, measured by average years of schooling in the adult (15+) population at the start of the relevant period (from Barro and Lee, 2001), and by arable and permanent crop area, annual renewable water resources, and value of known metal, gas and oil reserves, all in or around 1990 (from Wood and Mayer, 2001);
- *inequality*, measured by the average Gini coefficient over the relevant period (from Dollar and Kraay, 2002), and by the average ratio of male to female years of schooling, again over the relevant period (from Barro and Lee, 2001);
- *institutions*, measured by a rule of law index in 2001 (from Kaufmann *et al.*, 2003), and by the average score of the Freedom House political rights index over the relevant period;

- *policy*, measured by the average Sachs-Warner trade policy index over the relevant period (from Easterly *et al.*, 2003), and the average fiscal balance and rate of inflation over the relevant period (from World Bank, 2005);
- *external conditions*, measured by the average rate of change in a country's terms of trade index over the relevant period (from World Bank, 2005).

We look at the effects of these variables on group membership in both 1970-2003 and 1960-1993.

We first show, in Tables 13-14, the results of simple bivariate analysis. Column (1) shows the results for all chronically deprived countries, both across-the-board and partial. For the dummy variables (e.g. being located in tropics, being land-locked), the column shows the difference between the proportion of countries with a score of one which are chronically deprived, and the proportion of countries with a score of zero which are chronically deprived. For the continuous variables, the column shows the difference between the proportion of countries with above the median score for that indicator (across all countries) which are chronically deprived, and the proportion of countries with below the median score which are chronically deprived. Column (2) then shows the equivalent results for just the across-the-board chronically deprived countries, while column (3) shows the results for the across-the-board good performers. Each column also shows the p-values associated with a hypothesis test that the difference between the two proportions equals zero.

The main finding from Tables 13-14 is that almost all of the explanatory variables considered have a significant effect on the groups countries are found in. Considering a 5% level of statistical significance, the only variables which do not are mineral resources, water resources, inflation and population. Moreover, the effects are almost always in the ways one would expect. In particular, the following variables either increase the probability of being an across-the-board and/or a partially chronically deprived country, and/or reduce the probability of being an across-the-board good performer: location in the tropics, being land-locked, being both in the tropics and land-locked, distance from major OECD markets, ethnic fractionalisation, intra-state conflict, and inequality. The following variables have the opposite effect: European settlement, education, arable land, a higher Sachs-Warner index, greater fiscal surpluses (or smaller deficits), favourable terms of trade changes, and higher institutional quality.

The one surprising finding is that inter-state conflict reduces the probability of being an across-the-board or partially chronically deprived country, and increases the probability of being an across-the-board good performer. Reverse causation is the most likely explanation here.

Two further points are evident from Tables 13-14. First, each variable does not always affect group membership in exactly the same way. For example, a higher rule of law index significantly reduces the probability of being either an across-the-board or a partially chronically deprived country, but it does not significantly increase the probability of being an across-the-board good performer. Second, in some cases the effect of a variable is so strong as to completely determine group membership. For instance, all countries which are both located in the tropics and land-locked (12 in total) were either across-the-board or partially chronically deprived countries in both 1970-2003 and 1960-1993. Similarly, no land-locked countries have been across-the-board good performers in 1970-2003, while only one land-locked country (Botswana) achieved this status during 1960-93.

Tables 15-16 show the results of the multivariate analysis. In this case, the explanatory variables are entered simultaneously rather than separately, and the continuous explanatory variables (e.g. distance from major markets) are entered as they are rather than as dichotomous (above/below the median) variables. The main advantage in this case is that the effect of each variable can be measured holding others constant. However, entering all variables simultaneously would reduce the sample size considerably, and so to avoid this problem the different variables are entered in stages, with a common group of variables

included throughout. The estimation procedure is standard probit analysis, with the coefficients in Tables 15-16 showing the 'marginal effects', i.e. the change in the probability of being in a particular country group resulting from a one unit change in each explanatory variable.

Columns (1)-(7) show the effects on the probability of being either an across-the-board or partially chronically deprived country. In all columns, location in the tropics, being land-locked, and distance from major markets increase the probability, while favourable terms of trade shifts reduce it. In the majority of cases, these effects are statistically significant at the 5% level or lower. Among the remaining variables, the direction of effect is generally as expected, although the effects are not always statistically significant. Nonetheless, the effects of the rule of law index (column 6) and the Sachs-Warner index (column 7) are statistically significant at the 5% level, while the effects of the budget balance (column 7) and gender inequality (column 5) are statistically significant at the 10% level. The one surprising finding is the negative and statistically significant effect of inflation (column 7), for which the most likely explanation is reverse causation.

Columns (8)-(14) show the effects on the probability of being an across-the-board chronically deprived country. The results are similar to those in columns (1)-(7), except that being land-locked (all columns), the rule of law index (column 13), and all three policy indicators (column 14) are no longer statistically significant at the 5% level. However, the effect of initial education is now statistically significant, and negative as expected. Finally, the results for the across-the-board good performers are shown in columns (15)-(21). In this case, neither of the two geographical variables is statistically significant at the 5% or 10% levels, although the effect of favourable terms of trade shifts remains positive and statistically significant in the majority of cases. Of the remaining variables, the rule of law index and the rate of inflation both have statistically significant impacts in the directions one would expect (positive and negative, respectively). However, inter-state conflict and political rights also have statistically significant impacts (positive and negative, respectively) which are not in the directions one would expect.

A final point to note concerns the explanatory power of the regressions. A comparison of columns (1) and (7) in Table 15 shows that adding the three policy variables to the base regression (which includes just the three geographical variables and one external conditions variable) increases the adjusted R-squared from 0.47 to 0.63, an approximate increase of one third. Similarly, a comparison of columns (1) and (6) shows that adding the two institutional variables to the base regression increases the adjusted R-squared from 0.47 to 0.55, an approximate increase of one fifth. The magnitude of these increases is even smaller when considering the across-the-board chronically deprived countries. This suggests that characteristics which are beyond the control of government policy – in particular, geography and external conditions – play the largest role in determining whether or not a country is chronically deprived.

5.4 Growth trajectories

Section 5.2 showed that at least some chronically deprived countries have improved their performance, in at least some indicators, in more recent years. This raises the broader question of what types of trajectories chronically deprived countries have followed in recent decades. For example, does relatively slow progress on average reflect persistent decline or stagnation, or instead a period of reasonable progress followed by an adverse shock and subsequent stagnation or even collapse? If the latter, when did the transition occur?

To answer these questions we use the approach used in previous work by Pritchett (2000). The basic idea is to identify structural breaks in the time-path of a given welfare indicator in each country using statistical methods. One can then classify countries according to the size and direction of the trend in the welfare indicator before and after the structural break. Categories of countries defined by Pritchett (2000) are:

- a) *plateaus*: countries which had growth rates higher than 1.5 percent before their structural break, but less than 1.5 percent (but still positive) afterward;
- b) *mountains*: countries which had growth rates higher than 1.5 percent before their trend break, but negative rates afterward;
- c) *plains*: countries with growth rates less than 1.5 percent (but still positive) before and after their structural break;
- d) *valleys*: countries with negative growth rates before and after their structural break;
- e) *accelerators*: countries with growth rates less than 1.5 percent before their structural break, but more than 1.5 percent afterward;
- f) *hills*: countries with growth rates between 1.5 and 3 percent before and after their structural break; and
- g) *steep hills*: countries with growth rates greater than 3 percent before and after their structural break.

We apply this procedure to trends in GDP per capita, child mortality, fertility and under-nourishment over the period 1970-2003.¹⁰ For the three non-income welfare indicators, we use the median growth rate over the whole period across all non-OECD countries in place of the 1.5 percent threshold used by Pritchett (2000).¹¹ We also multiply the growth rates for the three non-income indicators by minus one to ensure compatibility with the GDP per capita growth figures.

Also, two groups are added to ensure that all countries are classified: 'vales', defined as countries with growth rates less than zero before their structural break, and with growth rates greater than zero but less than 1.5 percent (or equivalent threshold) after their structural break, and 'glaciers', defined as countries with growth rates less than 1.5 percent (or equivalent), but still positive, before their structural break, and with growth rates less than zero afterward. Finally, we rename the accelerators 'beacons' (as in Ditchling Beacon) to ensure geo-morphological terminology throughout.

Table 16 summarises the results (results for each individual country are shown in Annex 4). Among the across-the-board chronically deprived countries, there are not surprisingly no instances of persistently good progress (hills or steep hills). There are, however, several instances of growth starting and/or accelerating during the period (vales and beacons). The number is highest for fertility and under-nourishment, with 23 (out of 32) and 17 (out of 26) countries respectively, but is still significant for the GDP per capita and child mortality.¹² Among the remaining cases, there are generally more instances of growth which subsequently stagnated or was reversed (the plateaus, mountains and glaciers) as opposed to persistent decline or stagnation (the valleys and the plains). The exception is the fertility

¹⁰ The original analysis by Pritchett (2000) referred to the period 1960-1992. This section, therefore, updates this analysis and extends it to other welfare indicators.

¹¹ 1.5 percent was the median growth rate in GDP per capita across all non-OECD countries in the sample of countries used by Pritchett. In the sample of countries used in this paper, and for the period 1970-2003, the equivalent figure is 1.3%, but we round this up to 1.5% and therefore use the same figure as Pritchett. The median rates of growth for the other indicators during 1970-2003 (rounded to the nearest half percentage point) are: -3% (child mortality), -2% (fertility), and -1% (under-nourishment).

¹² These shifts are generally small in size however, since the countries remain chronically deprived when considering 1970-2003 as a whole. Also, despite these improvements few countries are progressing in child mortality or fertility at rates beyond their cluster range in the most recent years, as we saw in Table 12.

rate, where there have been very few instances of progress stagnating or reversing once it has been set in motion.

For the partial chronically deprived countries, the results are very similar: typically no instances of persistently good progress (the hills and the steep hills),¹³ a significant number of positive transitions (the vales and beacons), and with the exception of fertility, relatively more instances of growth which subsequently stagnated or was reversed (the plateaus, mountains and glaciers) than instances of persistent decline or stagnation (the valleys and the plains).

Among the good performers (across-the-board and partial), there are not surprisingly no instances of persistent decline or stagnation, but more surprisingly there are a number of instances of progress which subsequently stagnated or was reversed (the plateaus, mountains and glaciers). In these cases growth prior to the slowdown was high enough to ensure that average growth over the period remained high: Indonesia being a classic example. The remaining countries typically experienced persistently good progress rather than bursts of growth following initial stagnation. The exception is again the fertility rate, where a significant number of countries did experience such a shift. In most cases this corresponded with the onset or acceleration of the demographic transition.¹⁴

The years in which the positive and negative transitions occur are shown in Annex 4. For the chronically deprived countries, most of the negative transitions (the mountains, plateaus and glaciers) in GDP per capita occur in the late 1970s and early 1980s, while most of the positive transitions (the beacons and the vales) occur in the early to mid-1990s. For child mortality, however, most of the negative transitions occur in the early to mid-1990s, while most of the positive transitions occur in the 1980s. In these cases however, we cannot be sure about the precise transition years since the data are typically not available on an annual basis. For under-nourishment, like GDP per capita, most of the positive transitions occur in the early to mid-1990s, and most of the negative transitions occur in the early 1980s, although there is also a small cluster of negative transitions in the mid-1990s. Again, however, we cannot be sure about the precise transition years in this case. For fertility, the positive transitions are evenly spread throughout the period.

For the good performers, the instances of negative transitions are spread fairly evenly across the period, although with again a group of such transitions in the early to mid-1990s in the case of child mortality. The positive transitions witnessed in the case of the fertility rate occur mainly in the late 1970s and early 1980s.

To summarise, there is evidence for three types of trajectories within chronically deprived countries: persistent worsening or stagnation, growth which was subsequently reversed or at least came to a halt, and (on a more positive note) stagnation followed by the onset or acceleration of growth. This provides a way of disaggregating further among the chronically deprived countries. Similarly, within the good performers we find three types of trajectories: persistently good progress, bursts of growth following initial stagnation, and (on a more negative note) strong growth which was subsequently reversed or at least came to a halt. Research into the causes of large performance transitions has been expanding in recent years, particularly the causes of economic growth transitions (e.g. Hausmann *et al.*, 2005, 2006). Extending this analysis to non-income indicators, where the transitions do not appear to always occur at the same points in time, would be a useful task for further work.

¹³ The one exception is Nepal, which is a 'hill' for child mortality despite being in the child mortality chronically deprived cluster. The average rate of reduction in child mortality in Nepal over the period 1970-2003 was 4.1% per year, which also happens to be the maximum rate of reduction in the cluster; this explains the apparent anomaly.

¹⁴ Of these, the countries experiencing the largest transitions were all from North African and the Middle East: Algeria, Iran, Libya, Oman, Syria and Tunisia.

6. Conclusion

This paper provides a classification of developing countries on the basis of the level of, and the change in, their average welfare over recent decades. The aim is to improve on existing classifications, and to provide useful and policy relevant analysis for the second Chronic Poverty Report. The method used is the statistical technique of cluster analysis. This is applied to seven welfare indicators (GDP per capita, life expectancy, adult literacy, child mortality, infant mortality, fertility and under-nourishment) and to all non-OECD countries with populations exceeding one million (130 in total).

The main finding from the cluster analysis is that four welfare indicators (GDP per capita, child mortality, fertility and under-nourishment) show evidence of three distinct country clusters. These clusters all fit the same basic pattern, namely a 'chronically deprived' group, characterised by low initial levels of welfare and relatively slow rates of progress over time; a 'good performers' group, consisting of countries with similarly low initial levels of welfare, but faster rates of progress over time; and a group of 'others', consisting of countries with higher initial levels of welfare and with rates of progress somewhere in between the other two groups.

On the basis of these results, countries are classified into five groups. The first includes approximately 30 countries which are chronically deprived 'across-the-board', i.e. in all four welfare indicators showing the three-cluster pattern, or at least all of those four for which data are available. The second includes approximately 45 countries which are 'partially' chronically deprived, i.e. in one but not all of the four indicators. The third includes around 10 countries which are good performers across-the-board, while the fourth includes around 20 countries which are 'partial' good performers, i.e. not chronically deprived in any of the four indicators, but not good performers in all. The fifth and final group consists of around 20 countries which are others across-the-board. This leaves around 5 countries which cannot be classified due to lack of data.

The across-the-board chronically deprived countries account for 17% of the total number of people living on less than \$1-a-day in the countries considered in the analysis, and 36% of the total number of child deaths, compared with approximately 10% of total population. These shares have also been rising over time. However, the largest shares of total poverty and child mortality are found in the partially chronically deprived countries: 57% of \$1-a-day poverty and 55% of total child deaths, compared with 39% of total population. These shares have remained fairly stable over time.

In recent years there has been only limited improvement in chronically deprived countries, both across-the-board and partial. At the same time, rates of progress in chronically deprived countries (both across-the-board and partial) in recent years continue to lag significantly behind rates of progress in all other non-OECD countries (including, but not only, the good performers), in all four indicators. Whatever it is that has been driving the differences in performance between the groups, it appears to be persisting in recent years.

Various factors increase the probability of being a chronically deprived country. These include location in the tropics, being land-locked, greater distance from major OECD markets, ethnic fractionalisation, intra-state conflict and inequality. At the same time, the probability is reduced by European settlement, higher average years of schooling, greater stocks of arable land, a higher Sachs-Warner index, greater fiscal surpluses (or smaller deficits), favourable terms of trade changes, and higher institutional quality. Although it is difficult to separate out the effects of these different influences, the results of multivariate regression analysis suggest that geography (e.g. location in tropics, distance from major markets) and external conditions (e.g. shifts in the terms of trade) play the largest role in determining group membership.

Countries within each group show different types of performance trajectories. Among the chronically deprived countries for example (across-the-board and partial), some countries

experienced persistent worsening or stagnation, others experienced growth which was subsequently reversed or at least came to a halt, while some experienced stagnation followed by the onset or acceleration of growth. The precise numbers following each trajectory varies according to the welfare indicator. Similarly, within the good performers, some experienced persistently good progress, while others experienced bursts of growth following initial stagnation.

As regards policy implications, one of the main issues is the amount of aid the different groups receive. The partially chronically deprived countries, for example, appear significantly under-aided: their share of total aid was around 30% (in 2002), compared with shares of over 50% of total \$1-a-day poverty and total child deaths. By contrast, the partial good performers and the across-the-board others appear significantly over-aided. Furthermore, even though the share of total aid received by the across-the-board chronically deprived countries (in 2002) was close to their shares of total poverty and total child mortality, one could make a case for their share of aid being even larger, given the natural disadvantages they appear to face (e.g. being land-locked or located in the tropics). Whether donors are pursuing the right aid strategies in different groups of countries is another issue which needs to be considered.

In terms of further research, one of the most promising avenues is research into the causes of performance transitions, be they positive (e.g. growth accelerations) or negative (e.g. growth collapses). Research into this issue has been expanding in recent years, particularly in terms of economic growth transitions (e.g. Hausmann *et al.*, 2005, 2006). Extending this analysis to non-income indicators, where the transitions do not appear to always occur at the same points in time, would be a useful task for further research.

References

- Bartholomew, D., Steele, F., Moustaki, I and Galbraith, J. (2002) *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Boca Raton: Chapman and Hall.
- Barro, R. and Lee, J. (2001) International data on educational attainment: updates and implications. *Oxford Economic Papers* **53**(3): 541-563.
- DFID (2006) *Eliminating World Poverty. Making Governance Work for the Poor*. A White Paper. UK Department for International Development; London, UK.
- Dollar, D. and Kraay, A. (2002) Growth is good for the poor. *Journal of Economic Growth* **7**(3): 195-225.
- Easterly, W. and Levine, R. (1997) Africa's growth tragedy: policies and ethnic divisions. *Quarterly Journal of Economics* **107**: 1203-1250.
- Easterly, W., Levine, R. and Roodman, D. (2003) New data, new doubts: revisiting "Aid, policies and growth". Working Paper No.26, Center for Global Development, Washington D.C.
- Gallup, J. and Sachs, J. (1999) Geography and economic development. Working Paper No.1, Center for International Development, Harvard University.
- Hall, R. and Jones, C. (1999) Why do some countries produce so much more output than others? *Quarterly Journal of Economics* **114**: 83-116.
- Hausmann, R., Pritchett, L. and Rodrik, D. (2005) Growth accelerations. *Journal of Economic Growth*, **10**(4): 303-329.
- Hausmann, R., Rodriguez, F. and Wagner, R. (2006) Growth collapses. CID Working Paper No. 136, Center for International Development, Harvard University.
- Kaufmann, D., Kraay, A. and Mastruzzi, M. (2003) Governance matters III: Governance indicators for 1996-2002. Policy Research Working Paper 3106, World Bank.
- Pritchett, L. (2000) Understanding patterns of economic growth: searching for hills among plateaus, mountains and plains. *World Bank Economic Review* **14**(2): 221-250.
- Sarkees, M. (2000) The correlates of war data on war: an update to 1997. *Conflict Management and Peace Science* **18**(1): 123-144.
- UNDP (2003) *Human Development Report 2003*, pp. 347-349.
- Wood, A. and Mayer, J. (2001) Africa's export structure in comparative perspective. *Cambridge Journal of Economics* **25**: 369-394.
- World Bank (2005) *World Development Indicators, 2005 Edition*. CD-ROM, World Bank, Washington D.C.

Table 1 Determining the number of clusters: a hypothetical example

Cluster number	Proportion of variance explained (%)	Additional amount of variance explained (%)	Absolute difference in additional amount of variance explained (%)	Proportional difference in additional amount of variance explained (%)
1	0	-	-	-
2	30	30	5	1.2
3	55	25	5	1.3
4	75	20	17	6.7
5	78	3	0	1.0
6	81	3	0	1.0
7	84	3	0	1.0
8	87	3	1	1.5
9	89	2	-	-

Notes: The data in the data are the same as those underlying the example shown in Figure 6. Figures highlighted in bold show the maximum values of the absolute/proportional differences in the additional amount of variance explained by any one cluster.

Table 2 Determining the number of clusters: 7 welfare indicators, 1970-2003 and 1960-1993

Cluster number	GDP per capita	Life expectancy	Adult literacy	Child mortality	Infant mortality	Fertility	Under-nourishment
<i>Absolute difference in additional amount of variance explained, 1970-2003 (%)</i>							
2	3.6	12.2	57.4	14.0	<u>35.0</u>	9.8	-1.7
3	21.7	<u>16.1</u>	8.4	22.4	9.8	26.0	27.0
4	-0.6	5.9	0.0	2.4	0.8	0.4	2.8
5	1.2	2.8	1.5	1.6	2.4	1.2	0.6
6	2.4	0.4	1.0	-0.4	0.8	1.6	-1.7
7	0.6	1.2	1.0	0.8	-0.4	-0.4	3.9
8	1.8	0.4	-1.0	0.8	1.2	1.2	0.6
9	-0.6	0.0	1.0	0.4	0.4	0.0	0.0
<i>Proportional difference in additional amount of variance explained, 1970-2003 (%)</i>							
2	1.1	1.4	5.6	1.5	<u>3.2</u>	1.3	1.0
3	4.3	<u>2.4</u>	3.1	4.5	2.6	6.1	4.7
4	0.9	2.1	1.0	1.6	1.1	1.1	1.6
5	1.2	2.0	1.6	1.7	1.8	1.3	1.1
6	1.7	1.2	1.7	0.9	1.3	1.8	0.7
7	1.2	2.0	3.0	1.4	0.9	0.8	3.3
8	2.5	1.5	0.3	1.7	1.7	2.0	1.5
9	0.7	1.0	3.0	1.5	1.3	1.0	1.0
<i>Absolute difference in additional amount of variance explained, 1960-1993 (%)</i>							
2	1.2	27.3	57.4	18.3	28.8	-8.0	7.1
3	25.3	2.9	6.4	<u>18.3</u>	15.0	35.2	<u>19.8</u>
4	-1.2	10.9	-2.5	2.7	-1.8	0.4	2.7
5	0.6	0.8	5.4	3.1	2.2	1.6	3.8
6	1.2	0.0	0.0	0.4	1.3	2.4	0.5
7	3.0	2.1	1.0	0.4	0.4	0.8	0.0
8	0.0	0.4	-1.0	0.4	1.3	0.4	1.1
9	-0.6	-0.4	1.0	0.4	-0.4	-0.8	0.0
<i>Proportional difference in additional amount of variance explained, 1960-1993 (%)</i>							
2	1.0	2.5	6.3	1.7	2.5	0.8	1.2
3	5.7	1.2	2.4	<u>3.3</u>	4.4	6.9	<u>3.1</u>
4	0.8	3.6	0.6	1.5	0.7	1.1	1.4
5	1.1	1.3	4.7	2.4	1.6	1.4	2.4
6	1.3	1.0	1.0	1.3	1.5	2.5	1.3
7	2.7	2.7	3.0	1.3	1.2	2.0	1.0
8	1.0	1.5	0.3	1.5	2.5	2.0	2.0
9	0.7	0.7	3.0	2.0	0.7	0.3	1.0

Notes: In each case the full sample of countries is used (for sample sizes see Section 2). Figures highlighted in bold show the instances of strong evidence of clustering, as defined in this paper. Figures underlined show the instances in which the maximum values of $d(abs)_k$ and $d(rel)_k$ both occur at the same cluster number, but are not high enough to be classified as strong clustering.

Table 3 Maximum values of $d(abs)_k$ and $d(rel)_k$: robustness tests

Period	Difference measure	GDP per capita	Life expectancy	Adult literacy	Child mortality	Infant mortality	Fertility	Under-nourishment
<i>Full sample of countries</i>								
1970-2003	$d(abs)$	21.7	<u>16.1</u>	57.4	22.4	<u>35.0</u>	26.0	27.0
		3	3	2	3	2	3	3
	$d(rel)$	4.3	<u>2.4</u>	5.6	4.5	<u>3.2</u>	6.1	4.7
		3	3	2	3	2	3	3
1960-1993	N	84	128	102	126	128	128	90
	$d(abs)$	25.3	27.3	57.4	<u>18.3</u>	28.8	35.2	<u>19.8</u>
		3	2	2	3	2	3	3
	$d(rel)$	5.7	3.6	6.3	<u>3.3</u>	4.4	6.9	<u>3.1</u>
		3	4	2	3	3	3	3
1960-2003	N	84	120	102	113	114	126	92
	$d(abs)$	14.5	19.7		30.0	27.6	34.0	
		3	2	as	3	2	3	as
	$d(rel)$	2.3	3.7	1970-2003	13.5	4.2	9.7	1970-2003
		7	5		3	3	3	
	N	84	128		126	128	129	
<i>Restricted sample of countries (A)</i>								
1970-2003	$d(abs)$		<u>17.2</u>	57.8	18.0	32.0	26.0	
		as full sample	3	2	3	2	3	as full sample
	$d(rel)$		<u>3.1</u>	5.8	3.8	2.8	5.9	
			3	2	4	2	3	
1960-1993	N		117	97	112	112	124	
	$d(abs)$	<u>20.0</u>	26.1	-	19.5	24.8	32.2	-
		3	2	-	2	2	3	-
	$d(rel)$	<u>3.5</u>	3.0	-	3.2	3.0	6.1	-
		3	4	-	3	3	3	-
1960-2003	N	81	118		111	112	119	
	$d(abs)$	13.0	<u>16.1</u>	-	27.1	27.1	36.4	-
		3	3	-	3	2	3	-
	$d(rel)$	2.7	<u>3.5</u>	-	6.9	3.7	10.6	-
		7	3	-	3	3	3	-
	N	82	116	-	110	110	119	-
<i>Restricted sample of countries (B)</i>								
1970-2003	$d(abs)$	25.4	39.5	39.5	<u>40.4</u>	<u>40.4</u>	15.8	<u>19.3</u>
		3	2	2	2	2	3	3
	$d(rel)$	6.8	4.7	3.1	<u>3.4</u>	<u>3.9</u>	3.5	<u>3.2</u>
		3	4	3	2	2	5	3
	N	58	58	58	58	58	58	58

Notes: Figures in italics show the cluster number at which the maximum values occur. Figures in bold show the instances in which the evidence of clustering is 'strong' (as defined in Section 2). Figures underlined show the instances in which the maximum values of $d(abs)_k$ and $d(rel)_k$ both occur at the same cluster number, but are not high enough to be classified as strong clustering.

Table 4 Size and location of clusters, 1970-2003 and 1960-1993

Cluster*	Number of countries			Mean growth rate			Mean starting level			Range of growth rates			Range of starting levels		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1970-2003															
GDP	36	25	23	-0.7	3.2	0.8	1,613	1,983	5,466	-5.0, 1.3	1.5, 5.9	-1.2, 2.7	445, 4008	334, 6540	2896, 10342
CM	62	38	26	-1.7	-7.5	-3.7	77	46	7	0.7, -4.4	-4.6, -12.0	-0.5, -7.4	13, 175	11, 96	3, 19
FERT	53	42	33	-1.0	-2.8	-2.0	7	6	3	0.2, -2.4	-1.7, -4.3	-0.9, -3.5	4, 8	4, 9	2, 4
NOUR	49	34	7	0.0	-4.0	-0.9	31	32	3	2.9, -1.5	-1.5, -7.3	1.6, -3.6	11, 64	11, 58	1, 6
ALIT	59	43	-	0.9	0.2	-	46	89	-	0.4, 1.6	-0.0, 0.5	-	11, 72	69, 100	-
1960-1993															
GDP	43	29	12	0.1	3.5	1.5	1,547	1,659	5,071	-2.7, 1.9	1.7, 6.5	-1.0, 3.2	383, 3306	450, 3084	3480, 7751
CM	48	31	34	-2.0	-7.6	-4.2	105	64	18	-0.3, -3.9	-4.5, -12.9	-1.3, -7.1	62, 215	28, 117	4, 47
FERT	48	52	26	-0.2	-2.4	-1.0	7	6	3	0.8, -0.9	-1.2, -4.8	0.1, -2.8	4, 8	3, 8	2, 5
NOUR	43	38	11	1.0	-5.4	1.5	28	35	4	3.7, -1.9	-1.8, -14.6	9.2, -5.1	12, 58	11, 64	1, 11
ALIT	59	43	-	0.9	0.2	-	44	89	-	0.4, 1.7	-0.1, 0.5	-	11,72	69, 100	-

Notes:*For child mortality, fertility, under-nourishment, and GDP per capita: 1=chronically deprived country, 2=good performer, 3=other. For adult literacy, 1=low initial level and high progress, 2=high initial level and low progress. Growth rates shown are proportional (% per year) for child mortality, fertility, under-nourishment and GDP per capita, and absolute (years or percentage points per year) for adult literacy. Starting levels are in absolute (not log) terms for all variables.

Table 5 Cluster membership across indicators, 1970-2003 and 1960-1993

	CDC in all*	GP in all*	Other in all*	CDC in one but not all	GP in one but not all**	Missing data
<i>1970-2003</i>						
Data for 4 indicators (GDP, CM, FERT, NOUR)	21	7	2	31	12	57
Data for 3 indicators (CM, FERT, NOUR)	26	9	3	36	15	41
Data for 2 indicators (CM, FERT)	32	11	18	44	20	5
<i>1960-1993</i>						
Data for 4 indicators (GDP, CM, FERT, NOUR)	20	9	3	31	8	59
Data for 3 indicators (CM, FERT, NOUR)	24	10	3	35	13	45
Data for 2 indicators (CM, FERT)	27	11	17	37	19	19

Notes: *Refers to all indicators which are available. **Not including countries which were chronically deprived in one indicator. Results are based on full sample of countries.

Table 6 Across-the-board chronically deprived countries, 1970-2003

CDCs in all 4 indicators (GDP, CM, FERT, NOUR)	Additional CDCs in all of 3 indicators* (CM, FERT, NOUR)	Additional CDCs in all of 2 indicators** (CM, FERT)
Angola Burundi Cameroon Central African Republic Chad ^{##} Congo, Dem. Rep. Cote d'Ivoire Gambia, The [#] Guinea Kenya [#] Madagascar Malawi Mozambique Niger Rwanda Senegal Sierra Leone Tanzania [#] Togo Zambia Zimbabwe [#] (21)	Cambodia ^{##} Liberia Sudan Swaziland Yemen, Rep. [#] (5)	Afghanistan Eritrea ^{##} Ethiopia Iraq [#] Papua New Guinea Somalia ^{##} (6)

Notes: Countries are listed in alphabetical order. Each country's starting level, start year, growth rate and period over which growth rate is measured is shown in Annex 2. *Countries lack data for GDP per capita; **countries lack data for GDP per capita and under-nourishment. [#]Not an across-the-board chronically deprived country during 1960-1993; ^{##}not classified during 1960-1993 owing to lack of data.

Table 7 Across-the-board good performers, 1970-2003

GPs in all 4 indicators (GDP, CM, FERT, NOUR)	Additional GPs in all of 3 indicators* (CM, FERT, NOUR)	Additional GPs in all of 2 indicators** (CM, FERT)
China Egypt, Arab Rep. [#] Indonesia [#] Jordan ^{##} Morocco Syrian Arab Republic Tunisia (7)	Libya Vietnam [#] (2)	Albania Oman [#] (2)

Notes: Countries are listed in alphabetical order. Each country's starting level, start year, growth rate and period over which growth rate is measured is shown in Annex 2. *Countries lack data for GDP per capita; **countries lack data for GDP per capita and under-nourishment. [#]Not an across-the-board good performer during 1960-1993; ^{##}not classified during 1960-1993 owing to lack of data.

Table 9 Partially chronically deprived countries, 1970-2003

Country	GDP	CM	FERT	NOUR
Benin	1	1	1	2
Burkina Faso	1	1	1	2
Ghana [#]	1	1	1	2
Lesotho	1	1	1	2
Mali	1	1	1	2
Mauritania	1	1	1	2
Nigeria	1	1	1	2
Botswana [#]	2	1	1	1
Guinea-Bissau ^{##}	2	1	1	1
Haiti [#]	2	1	1	1
Namibia [#]	3	1	1	1
Pakistan	2	1	1	1
Uganda [#]	2	1	1	1
Dominican Rep. [#]	2	2	2	1
Jamaica	3	3	2	1
Sri Lanka	2	2	3	1
Thailand	2	2	2	1
Trinidad and Tob.	3	3	3	1
Venezuela	3	3	2	1
Bolivia	1	2	1	1
Guatemala	1	2	1	1
Honduras	1	2	1	1
Gabon ^{##}	3	1	1	2
Nepal	2	1	1	2
Bangladesh	2	1	2	1
Panama	3	1	2	1
Ecuador [#]	1	2	2	2
India	2	1	2	2
Nicaragua	1	2	2	1
Paraguay [#]	3	3	1	1
Philippines	1	1	2	2
Korea, Dem. Rep. [#]	.	1	3	1
Lao PDR [#]	.	2	1	1
Mongolia ^{##}	.	2	2	1
Myanmar	.	1	2	2
Saudi Arabia	.	2	1	2
Azerbaijan ^{##}	.	1	2	.
Kazakhstan ^{##}	.	1	3	.
Kyrgyz Republic [#]	.	1	2	.
Moldova [#]	.	1	3	.
South Africa ^{##}	3	1	2	.
Tajikistan ^{##}	.	1	2	.
Turkmenistan ^{##}	.	1	2	.
Uzbekistan [#]	.	1	2	.

Notes: 44 countries in total. Cluster numbers are: 1=chronically deprived country, 2=good performer, 3=other. [#]Not partially chronically deprived during 1960-1993; ^{##}not classified during 1960-1993 owing to lack of data.

Table 10 Partial good performers, 1970-2003

Country	GDP	CM	FERT	NOUR
Algeria [#]	3	2	2	2
Brazil [#]	3	2	2	2
Colombia [#]	3	2	2	2
Costa Rica	3	2	2	2
El Salvador [#]	3	2	2	2
Iran, Islamic Rep. [#]	3	2	2	2
Mexico	3	2	2	2
Peru [#]	3	2	2	2
Korea, Rep.	2	2	2	3
Kuwait	.	3	2	2
Lebanon	.	3	2	2
Malaysia	2	2	2	3
Mauritius	2	2	3	2
United Arab Emirates	.	2	2	3
Bosnia and Herzegovina	.	2	3	.
Macedonia, FYR	.	2	3	.
Serbia and Montenegro	.	2	3	.
Chile	3	2	3	3
Romania	2	3	3	.
Singapore	2	3	3	.

Notes: 20 countries in total. Cluster numbers are: 2=good performer, 3=other. [#]Not partial good performer during 1960-1993.

Table 11 Share of non-OECD totals in each country group

	CDC in all*	GP in all*	Others in all*	CDC in one but not all	GP in one not all**	Missing data
<i>1970-2003</i>						
Number of countries	32	11	18	44	20	5
Population (% of total)	10	33	7	39	11	<1
Child mortality (% of total)	36	6	<1	55	1	<1
Infant mortality (% of total)	30	11	1	54	4	<1
\$1-a-day poverty (% of total)	17	22	1	57	3	<1
\$2-a-day poverty (% of total)	13	28	2	53	4	<1
Under-nourishment (% of total)	23	23	1	49	4	<1
Aid (% of total)	29	14	12	29	12	5
<i>1960-1993</i>						
Number of countries	27	11	17	37	19	19
Population (% of total)	7	32	8	44	6	3
Child mortality (% of total)	24	9	<1	64	1	2
Infant mortality (% of total)	21	13	1	61	2	3
\$1-a-day poverty (% of total)	9	37	1	51	1	1
\$2-a-day poverty (% of total)	8	35	2	53	2	1
Under-nourishment (% of total)	11	28	<1	57	2	1
Aid (% of total)	24	10	8	50	4	4

Notes: Results refer to sample of countries with data for child mortality and fertility. *Refers to all indicators which are available. **Not including countries which were chronically deprived in one indicator. Results are based on full sample of countries.

Table 12 Most recent growth rates

	Min growth rate	Max growth rate	Mean growth rate	No. >/< cluster range*	Relative growth rate**	p-value	No. of countries
<i>Across-the-board chronically deprived countries</i>							
GDP per capita	-4.4	5.5	0.6	11	-1.3	0.05	23
Child mortality	-6.4	5.9	-0.4	1	7.3	0.00	31
Fertility	-2.9	0.4	-1.3	2	2.0	0.00	31
Under-nourishment	-6.2	3.0	-1.7	14	3.2	0.00	26
<i>Partially chronically deprived countries***</i>							
GDP per capita	-2.0	5.0	1.1	6	-0.7	0.34	13
Child mortality	-7.2	7.9	-1.5	5	4.9	0.00	29
Fertility	-2.5	-0.4	-1.5	2	1.7	0.00	21
Under-nourishment	-7.3	3.3	-1.3	10	2.8	0.00	22
<i>Across-the-board good performers</i>							
GDP per capita	-0.4	6.4	2.1	4	1.0	0.25	7
Child mortality	-15.0	-2.3	-7.9	3	-3.6	0.01	11
Fertility	-4.2	0.3	-2.1	2	-1.0	0.10	11
Under-nourishment	-5.2	4.1	-0.4	7	0.1	0.92	9
<i>Partial good performers***</i>							
GDP per capita	0.3	3.8	2.1	2	1.2	0.29	5
Child mortality	-15.7	0.0	-5.4	6	-2.9	0.01	15
Fertility	-5.6	-1.0	-2.6	3	-1.5	0.00	13
Under-nourishment	-6.6	4.1	-1.8	6	-1.0	0.36	11

Notes: *No. of countries exceeding maximum growth rate of cluster during 1970-2003 for chronically deprived countries (across the board and partial); number of countries below minimum growth rate of cluster during 1970-2003 for good performers (both across-the-board and partial). **Results show estimated coefficient on a dummy variable for each respective country group in a linear regression of each country's most recent growth rate on the initial level of the indicator at the start of the period over which growth is measured. *** Figures refer only to countries which are chronically deprived/good performers in each specific indicator, and not to all countries in these groups.

Table 13 Effects of country characteristics on group membership, bivariate analysis: 1970-2003 results

	(1)		(2)		(3)		Sample sizes	
	All CDCs		Across-the-board CDCs		Across-the-board GPs			
	p1-p0	p-value	p1-p0	p-value	p1-p0	p-value	n1	n2
TROPICS	0.42	0.00	0.35	0.00	-0.10	0.04	66	59
LLOCK	0.30	0.00	0.09	0.24	-0.12	0.00	91	34
TROPICS-LLOCK	0.43	0.00	0.45	0.00	-0.10	0.00	113	12
DISTANCE	0.43	0.00	0.36	0.00	-0.11	0.03	63	62
TTRADE	-0.37	0.00	-0.27	0.00	0.14	0.01	58	62
ETHNO	0.34	0.00	0.34	0.00	-0.11	0.09	38	38
ESETTLEMENT	-0.19	0.07	-0.31	0.00	-0.09	0.08	59	36
CONFLICT1	-0.23	0.04	-0.08	0.27	0.28	0.00	99	26
CONFLICT2	0.11	0.18	0.17	0.06	0.00	0.40	81	44
SCHOOLING	-0.40	0.00	-0.50	0.00	0.00	0.40	36	35
LAND	-0.25	0.01	-0.30	0.00	0.16	0.02	44	44
WATER	-0.07	0.31	0.02	0.39	0.07	0.23	44	44
MINERALS	0.07	0.31	-0.02	0.39	-0.11	0.08	44	44
POPULATION	-0.05	0.33	0.00	0.40	0.02	0.38	63	62
GINI	0.29	0.00	0.15	0.09	-0.10	0.05	48	48
GGAP	0.13	0.21	0.42	0.00	0.10	0.04	29	42
RLAW	-0.34	0.00	-0.15	0.07	0.02	0.38	62	62
PRIGHTS	-0.26	0.00	-0.32	0.00	-0.11	0.03	62	62
SW	-0.30	0.00	-0.37	0.00	0.06	0.24	49	51
BBAL	-0.07	0.30	-0.18	0.03	-0.08	0.15	52	55
INF	-0.10	0.20	-0.06	0.30	-0.08	0.12	62	61

Notes: TROPICS: dummy variable for location in the tropics; LLOCK: dummy variable for being land-locked; TROPICS-LLOCK: dummy variable for being both located in the tropics and land-locked; DISTANCE: distance to major OECD markets (log km); TTRADE: rate of change in country's terms of trade index; ETHNO: ethno-fractionalisation index; ESETTLEMENT: dummy variable for whether any of the population spoke a Western European language as a mother tongue in 1960; CONFLICT1: dummy variable for whether a country participated in an inter-state war; CONFLICT2: dummy variable for whether a country participated in an intra-state war; SCHOOLING: average years of schooling in the adult (15+) population; LAND: arable and permanent crop area; WATER: annual renewable water resources; MINERALS: value of known metal, gas and oil reserves; POPULATION=total population; GINI: Gini coefficient; GGAP: average ratio of male to female years of schooling; RLAW: rule of law index; PRIGHTS: political rights index; SW: Sachs-Warner trade policy index; BBAL: fiscal balance (% of GDP); INF: rate of change in GDP deflator. For dummy variables, p1-p0 = difference between proportion of countries with score of 1 and proportion of countries with score of 0. For continuous variables, p1-p0 = difference between the proportion of countries with above the median score and the proportion of countries with below the median score. n1=number of countries with score of 0 (dummy variables) or below median score (continuous variables) of each explanatory variable; n2=number of countries with score of 1 or above median score.

Table 14 Effects of country characteristics on group membership, bivariate analysis: 1960-1993 results

	(1)		(2)		(3)		Sample sizes	
	All CDCs		Across-the-board CDCs		Across-the-board GPs			
	p1-p0	p-value	p1-p0	p-value	p1-p0	p-value	n1	n2
TROPICS	0.54	0.00	0.28	0.00	-0.12	0.03	57	54
LLOCK	0.07	0.32	0.22	0.04	-0.08	0.11	84	27
TROPICS-LLOCK	0.47	0.00	0.54	0.00	-0.11	0.00	100	11
DISTANCE	0.41	0.00	0.35	0.00	-0.12	0.03	56	55
TTRADE	-0.16	0.12	-0.18	0.05	0.17	0.01	48	47
ETHNO	0.44	0.00	0.36	0.00	-0.22	0.00	36	36
ESETTLEMENT	-0.30	0.00	-0.23	0.01	0.07	0.23	55	34
CONFLICT1	0.03	0.38	-0.21	0.01	0.08	0.24	86	25
CONFLICT2	0.21	0.03	0.04	0.36	0.04	0.33	70	41
SCHOOLING	-0.36	0.00	-0.44	0.00	0.01	0.39	21	46
LAND	-0.20	0.05	-0.28	0.01	0.20	0.01	42	41
WATER	-0.01	0.40	-0.04	0.37	0.15	0.04	42	41
MINERALS	0.01	0.40	-0.06	0.34	-0.05	0.31	41	42
POPULATION	0.01	0.40	-0.09	0.23	0.06	0.25	56	55
GINI	0.23	0.03	0.00	0.40	0.05	0.30	43	43
GGAP	0.25	0.04	0.35	0.00	-0.06	0.30	29	38
RLAW	-0.27	0.00	-0.11	0.16	-0.02	0.38	55	55
PRIGHTS	-0.22	0.03	-0.30	0.00	-0.01	0.39	56	52
SW	-0.17	0.10	-0.31	0.00	0.10	0.13	47	43
BBAL	-0.15	0.13	-0.12	0.18	0.00	0.40	44	45
INF	-0.20	0.04	-0.05	0.33	-0.06	0.26	53	52

Notes: See Table 13.

Table 15 Effects of country characteristics on group membership, multivariate analysis: 1970-2003 results

	1	2	3	4	5	6	7
TROPICS	0.534 0.00	0.557 0.00	0.545 0.00	0.737 0.01	0.558 0.00	0.452 0.00	0.679 0.00
LLOCK	0.464 0.00	-	0.468 0.00	-	0.419 0.04	0.447 0.01	0.479 0.00
DISTANCE	0.433 0.00	0.135 0.46	0.421 0.00	0.291 0.28	0.235 0.16	0.501 0.00	0.469 0.00
TTRADE	-3.167 0.02	-8.165 0.00	-3.608 0.02	-4.497 0.17	-3.641 0.12	-2.089 0.15	-2.200 0.04
ETHNO	-	0.438 0.14	-	-	-	-	-
ESETTLEMENT	-	-0.199 0.24	-	-	-	-	-
CONFLICT1	-	-	-0.070 0.75	-	-	-	-
CONFLICT2	-	-	0.209 0.19	-	-	-	-
SCHOOLING	-	-	-	-0.340 0.18	-	-	-
LAND	-	-	-	0.005 0.80	-	-	-
WATER	-	-	-	0.019 0.92	-	-	-
MINERALS	-	-	-	-0.083 0.32	-	-	-
POPULATION	-	-	-	0.153 0.46	-	-	-
GINI	-	-	-	-	0.324 0.47	-	-
GGAP	-	-	-	-	0.213 0.06	-	-
RLAW	-	-	-	-	-	-0.367 0.02	-
PRIGHTS	-	-	-	-	-	0.186 0.62	-
SW	-	-	-	-	-	-	-0.721 0.00
BBAL	-	-	-	-	-	-	-0.046 0.07
INF	-	-	-	-	-	-	-0.599 0.01
N	86	52	86	43	49	86	86
Pseudo R-squared	0.47	0.48	0.49	0.39	0.42	0.55	0.63
Sample probability	0.63	0.63	0.63	0.58	0.63	0.63	0.63

Notes: For variable definitions, see Table 13. Dependent variable is a dummy variable indicating whether a country is either an across-the-board or partially chronically deprived country. P-values shown below each coefficient.

Table 15 cont.

	8	9	10	11	12	13	14
TROPICS	0.164 0.00	0.007 0.02	0.156 0.00	0.105 0.05	-	0.115 0.01	0.151 0.00
LLOCK	0.040 0.37	0.007 0.23	0.041 0.36	0.028 0.58	-	0.009 0.79	0.032 0.50
DISTANCE	0.191 0.00	0.006 0.06	0.182 0.00	0.163 0.02	-	0.162 0.01	0.195 0.00
TTRADE	-1.452 0.01	-0.059 0.01	-1.369 0.01	-1.050 0.09	-	-0.911 0.05	-1.120 0.07
ETHNO	-	-0.004 0.21	-	-	-	-	-
ESETTLEMENT	-	-0.028 0.11	-	-	-	-	-
CONFLICT1	-	-	-0.018 0.71	-	-	-	-
CONFLICT2	-	-	0.002 0.96	-	-	-	-
SCHOOLING	-	-	-	-0.067 0.01	-	-	-
LAND	-	-	-	0.003 0.19	-	-	-
WATER	-	-	-	-0.008 0.77	-	-	-
MINERALS	-	-	-	0.001 0.87	-	-	-
POPULATION	-	-	-	0.000 0.99	-	-	-
GINI	-	-	-	-	-	-	-
GGAP	-	-	-	-	-	-	-
RLAW	-	-	-	-	-	-0.026 0.45	-
PRIGHTS	-	-	-	-	-	-0.119 0.17	-
SW	-	-	-	-	-	-	-0.085 0.17
BBAL	-	-	-	-	-	-	-0.002 0.59
INF	-	-	-	-	-	-	-0.105 0.13
N	86	65	86	53	-	86	86
Pseudo R-squared	0.43	0.59	0.43	0.63	-	0.49	0.47
Sample probability	0.24	0.32	0.24	0.23	-	0.24	0.24

Notes: For variable definitions, see Table 13. Dependent variable is a dummy variable indicating whether a country is an across-the-board chronically deprived country. P-values shown below each coefficient. Column (12) is not reported because of multi-collinearity (all coefficients insignificant).

Table 15 cont.

	15	16	17	18	19	20	21
TROPICS	-0.085	-	-0.029	-0.045	-0.026	-0.009	-0.028
	0.17		0.54	0.55	0.73	0.58	0.14
LLOCK	-	-	-	-	-	-	-
DISTANCE	-0.022	-	-0.014	-0.075	-0.030	-0.004	-0.003
	0.60		0.68	0.24	0.61	0.68	0.82
TTRADE	1.253	-	0.823	0.286	0.302	0.196	0.402
	0.01		0.03	0.78	0.75	0.04	0.01
ETHNO	-	-	-	-	-	-	-
ESETTLEMENT	-	-	-	-	-	-	-
CONFLICT1	-	-	0.285	-	-	-	-
			0.01				
CONFLICT2	-	-	-0.053	-	-	-	-
			0.23				
SCHOOLING	-	-	-	0.001	-	-	-
				0.99			
LAND	-	-	-	0.000	-	-	-
				0.96			
WATER	-	-	-	0.010	-	-	-
				0.85			
MINERALS	-	-	-	-0.009	-	-	-
				0.50			
POPULATION	-	-	-	0.009	-	-	-
				0.87			
GINI	-	-	-	-	-0.069	-	-
					0.66		
GGAP	-	-	-	-	0.011	-	-
					0.75		
RLAW	-	-	-	-	-	0.026	-
						0.03	
PRIGHTS	-	-	-	-	-	-0.147	-
						0.01	
SW	-	-	-	-	-	-	0.004
							0.82
BBAL	-	-	-	-	-	-	-0.003
							0.12
INF	-	-	-	-	-	-	-0.149
							0.07
N	86	-	86	53	49	86	86
Pseudo R-squared	0.25		0.39	0.18	0.09	0.54	0.39
Sample probability	0.10	-	0.10	0.08	0.06	0.10	0.10

Notes: For variable definitions, see Table 13. Dependent variable is a dummy variable indicating whether a country is an across-the-board good performer. P-values shown below each coefficient. Column (16) is not reported because of multi-collinearity (all coefficients insignificant); LLOCK is omitted as it predicts the dependent variable perfectly.

Table 16 Effects of country characteristics on group membership, multivariate analysis: 1960-1993 results

	1	2	3	4	5	6	7
TROPICS	0.501 0.00	0.508 0.00	0.569 0.00	0.853 0.09	0.801 0.01	0.494 0.00	0.691 0.00
LLOCK	0.034 0.83	0.075 0.67	0.132 0.44	0.048 0.70	0.181 0.24	-0.244 0.25	0.182 0.22
DISTANCE	0.251 0.08	0.000 1.00	0.279 0.06	0.111 0.37	0.237 0.15	0.274 0.09	0.430 0.01
TTRADE	-1.287 0.12	-2.327 0.04	-1.353 0.10	-4.174 0.15	-8.382 0.04	-0.913 0.31	-1.832 0.12
ETHNO	-	0.226 0.33	-	-	-	-	-
ESETTLEMENT	-	-0.346 0.01	-	-	-	-	-
CONFLICT1	-	-	0.236 0.17	-	-	-	-
CONFLICT2	-	-	0.116 0.43	-	-	-	-
SCHOOLING	-	-	-	-0.131 0.18	-	-	-
LAND	-	-	-	-0.004 0.53	-	-	-
WATER	-	-	-	-0.083 0.18	-	-	-
MINERALS	-	-	-	-0.030 0.29	-	-	-
POPULATION	-	-	-	0.169 0.08	-	-	-
GINI	-	-	-	-	-0.549 0.23	-	-
GGAP	-	-	-	-	0.193 0.04	-	-
RLAW	-	-	-	-	-	-0.262 0.03	-
PRIGHTS	-	-	-	-	-	-0.794 0.02	-
SW	-	-	-	-	-	-	-0.703 0.00
BBAL	-	-	-	-	-	-	-0.065 0.02
INF	-	-	-	-	-	-	-0.912 2.00
N	67	58	67	46	43	67	67
Pseudo R-squared	0.27	0.42	0.32	0.73	0.62	0.50	0.56
Sample probability	0.67	0.71	0.67	0.58	0.60	0.67	0.67

Notes: For variable definitions, see Table 13. Dependent variable is a dummy variable indicating whether a country is either an across-the-board or partially chronically deprived country. P-values shown below each coefficient.

Table 16 cont.

	8	9	10	11	12	13	14
TROPICS	0.297 0.01	0.001 0.03	0.305 0.01	-	0.194 0.22	0.255 0.02	0.249 0.01
LLOCK	0.135 0.31	0.006 0.13	0.157 0.23	-	0.212 0.16	0.038 0.75	0.110 0.38
DISTANCE	0.314 0.04	0.000 0.76	0.297 0.05	-	0.016 0.91	0.258 0.09	0.266 0.07
TTRADE	-1.681 0.04	-0.006 0.01	-1.615 0.03	-	-1.294 0.12	-1.435 0.05	-2.051 0.08
ETHNO	-	0.000 0.34	-	-	-	-	-
ESETTLEMENT	-	-0.003 0.19	-	-	-	-	-
CONFLICT1	-	-	-0.017 0.91	-	-	-	-
CONFLICT2	-	-	-0.134 0.22	-	-	-	-
SCHOOLING	-	-	-	-	-	-	-
LAND	-	-	-	-	-	-	-
WATER	-	-	-	-	-	-	-
MINERALS	-	-	-	-	-	-	-
POPULATION	-	-	-	-	-	-	-
GINI	-	-	-	-	0.082 0.69	-	-
GGAP	-	-	-	-	0.038 0.20	-	-
RLAW	-	-	-	-	-	-0.009 0.93	-
PRIGHTS	-	-	-	-	-	-0.745 0.01	-
SW	-	-	-	-	-	-	-0.466 0.08
BBAL	-	-	-	-	-	-	-0.004 0.81
INF	-	-	-	-	-	-	-0.351 0.19
N	67	58	67	-	43	67	67
Pseudo R-squared	0.25	0.58	0.28	-	0.46	0.40	0.37
Sample probability	0.27	0.31	0.27	-	0.19	0.27	0.27

Notes: For variable definitions, see Table 13. Dependent variable is a dummy variable indicating whether a country is an across-the-board chronically deprived country. P-values shown below each coefficient. Column (11) is not reported because of multi-collinearity (all coefficients insignificant).

Table 16 cont.

	15	16	17	18	19	20	21
TROPICS	-0.143 0.07	-0.203 0.04	-0.147 0.09	0.000 0.98	-0.094 0.28	-0.167 0.06	-0.178 0.04
LLOCK	-	-	-	-	-	-	-
DISTANCE	-0.022 0.71	-0.066 0.34	-0.023 0.70	-0.002 0.30	-0.035 0.62	-0.036 0.48	-0.013 0.80
TTRADE	0.541 0.30	0.264 0.71	0.541 0.30	0.033 0.20	1.380 0.20	0.356 0.38	0.046 0.93
ETHNO	-	-0.05 0.72	-	-	-	-	-
ESETTLEMENT	-	0.072 0.37	-	-	-	-	-
CONFLICT1	-	-	-0.007 0.94	-	-	-	-
CONFLICT2	-	-	0.009 0.91	-	-	-	-
SCHOOLING	-	-	-	-0.001 0.62	-	-	-
LAND	-	-	-	0.001 0.19	-	-	-
WATER	-	-	-	0.003 0.10	-	-	-
MINERALS	-	-	-	0.000 0.93	-	-	-
POPULATION	-	-	-	-0.004 0.09	-	-	-
GINI	-	-	-	-	0.438 0.08	-	-
GGAP	-	-	-	-	0.004 0.93	-	-
RLAW	-	-	-	-	-	-0.018 0.73	-
PRIGHTS	-	-	-	-	-	0.248 0.11	-
SW	-	-	-	-	-	-	0.161 0.15
BBAL	-	-	-	-	-	-	0.010 0.35
INF	-	-	-	-	-	-	-0.048 0.71
N	67	58	67	46	43	67	67
Pseudo R-squared	0.11	0.20	0.11	0.44	0.24	0.18	0.19
Sample probability	0.10	0.12	0.10	0.13	0.12	0.10	0.10

Notes: For variable definitions, see Table 13. Dependent variable is a dummy variable indicating whether a country is an across-the-board good performer. P-values shown below each coefficient. Column (16) is not reported because of multi-collinearity (all coefficients insignificant); LLOCK is omitted as it predicts the dependent variable perfectly.

Table 17 Performance transitions, 1970-2003

	Plateaus	Mountains	Glaciers	Plains	Valleys	Beacons	Vales	Hills	Steep hills	Persistently good*	Persistently bad*	Negative transitions*	Positive transitions*
Across-the-board chronically deprived countries													
GDP	2	5	3	0	6	5	2	0	0	0	6	10	7
CM	5	6	8	7	0	4	0	0	0	0	7	19	4
FERT	1	0	1	7	0	8	15	0	0	0	7	2	23
NOUR	0	6	0	0	3	14	3	0	0	0	3	6	17
Partially chronically deprived countries**													
GDP	4	2	0	0	3	2	2	0	0	0	3	6	4
CM	7	3	3	7	0	6	2	1	0	1	7	13	8
FERT	0	0	0	8	0	9	4	0	0	0	8	0	13
NOUR	1	9	0	0	1	10	1	0	0	0	1	10	11
Across-the-board good performers													
GDP	2	2	0	0	0	1	0	2	0	2	0	4	1
CM	1	0	0	0	0	0	0	8	2	10	0	1	0
FERT	2	0	0	0	0	5	0	4	0	4	0	2	5
NOUR	3	5	0	0	0	1	0	0	0	0	0	8	1
Partial good performers**													
GDP	0	1	0	0	0	0	0	1	3	4	0	1	0
CM	3	2	0	0	0	0	0	3	2	5	0	5	0
NOUR	3	0	0	0	0	4	0	6	0	6	0	3	4
FERT	0	3	0	0	0	2	0	5	1	6	0	3	2

Notes: *Persistently good=hills+ steep hills; persistently bad=plains+valleys; positive transitions=beacons+vales; negative transitions=plateaus+mountains+glaciers. For definition of each category, see main text. ** Figures refer only to countries which are chronically deprived/good performers in each specific indicator, and not to all countries in these groups. .

Figure 1 Stylised country groupings based on indicators of welfare (e.g. literacy, GDP per capita)

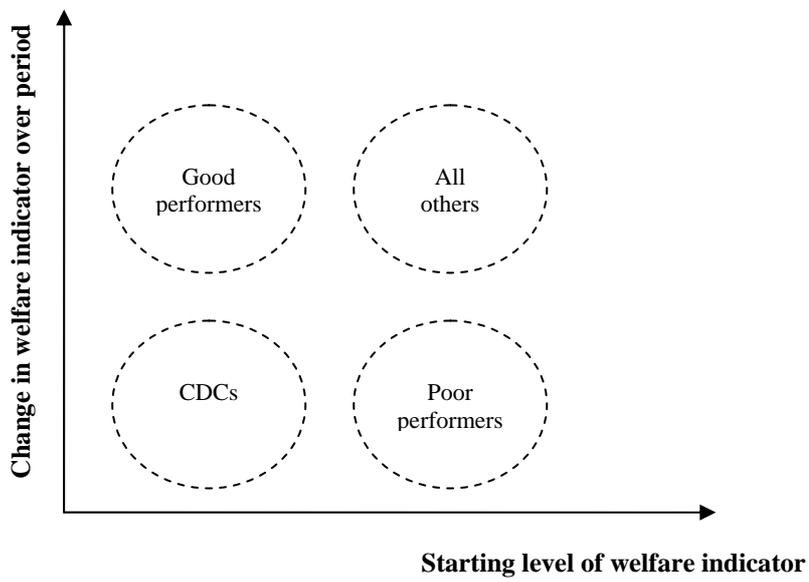


Figure 2 Stylised country groupings based on indicators of deprivation (e.g. child mortality)

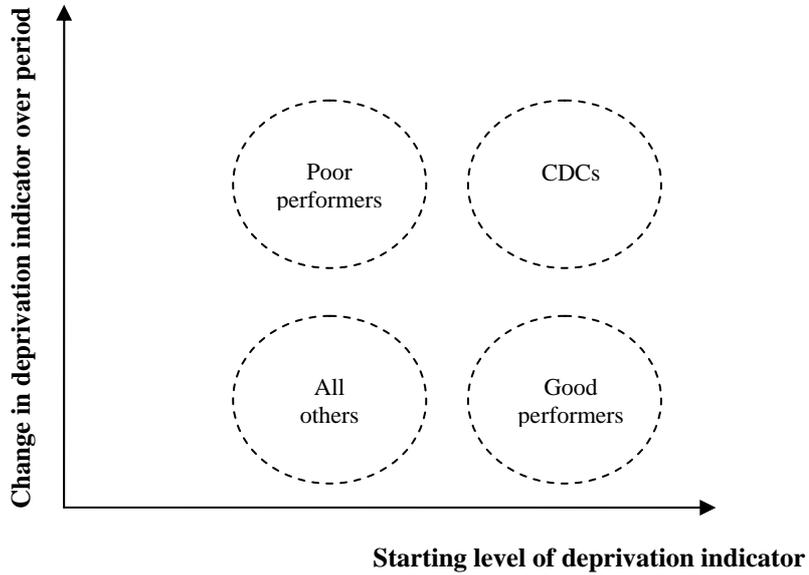


Figure 3 Alternative stylised country groupings based on indicators of welfare (e.g. literacy, GDP per capita)

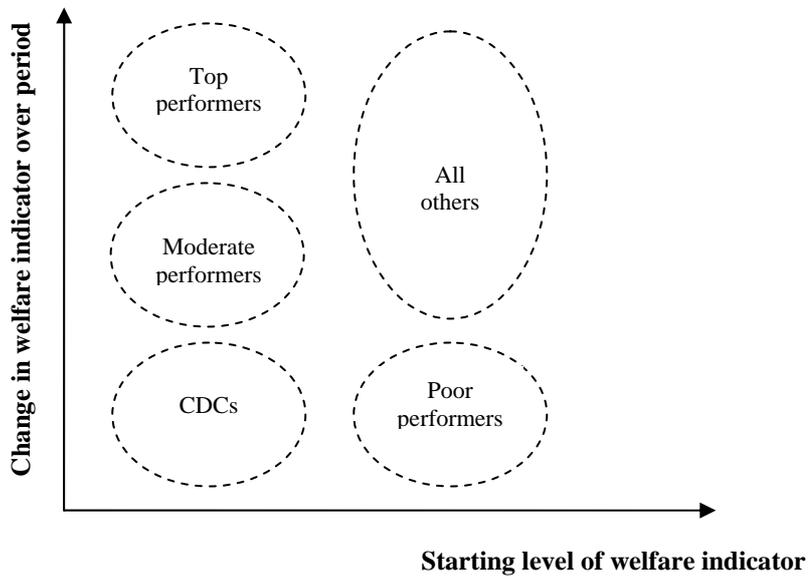
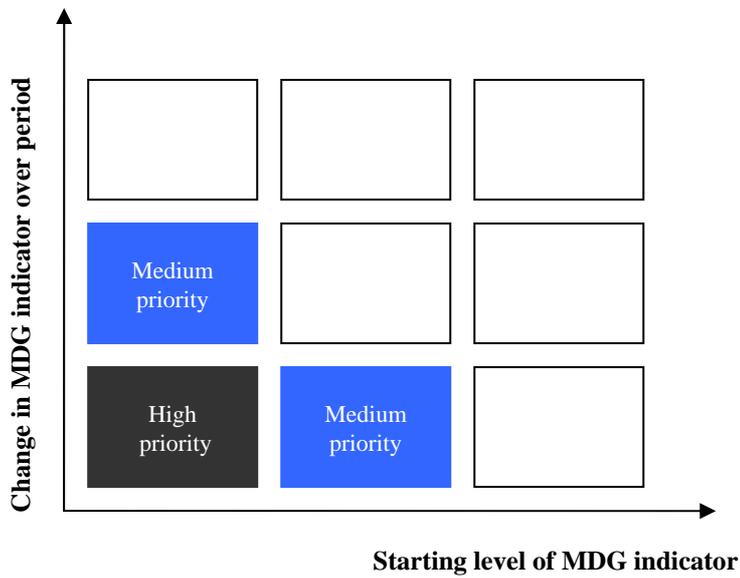
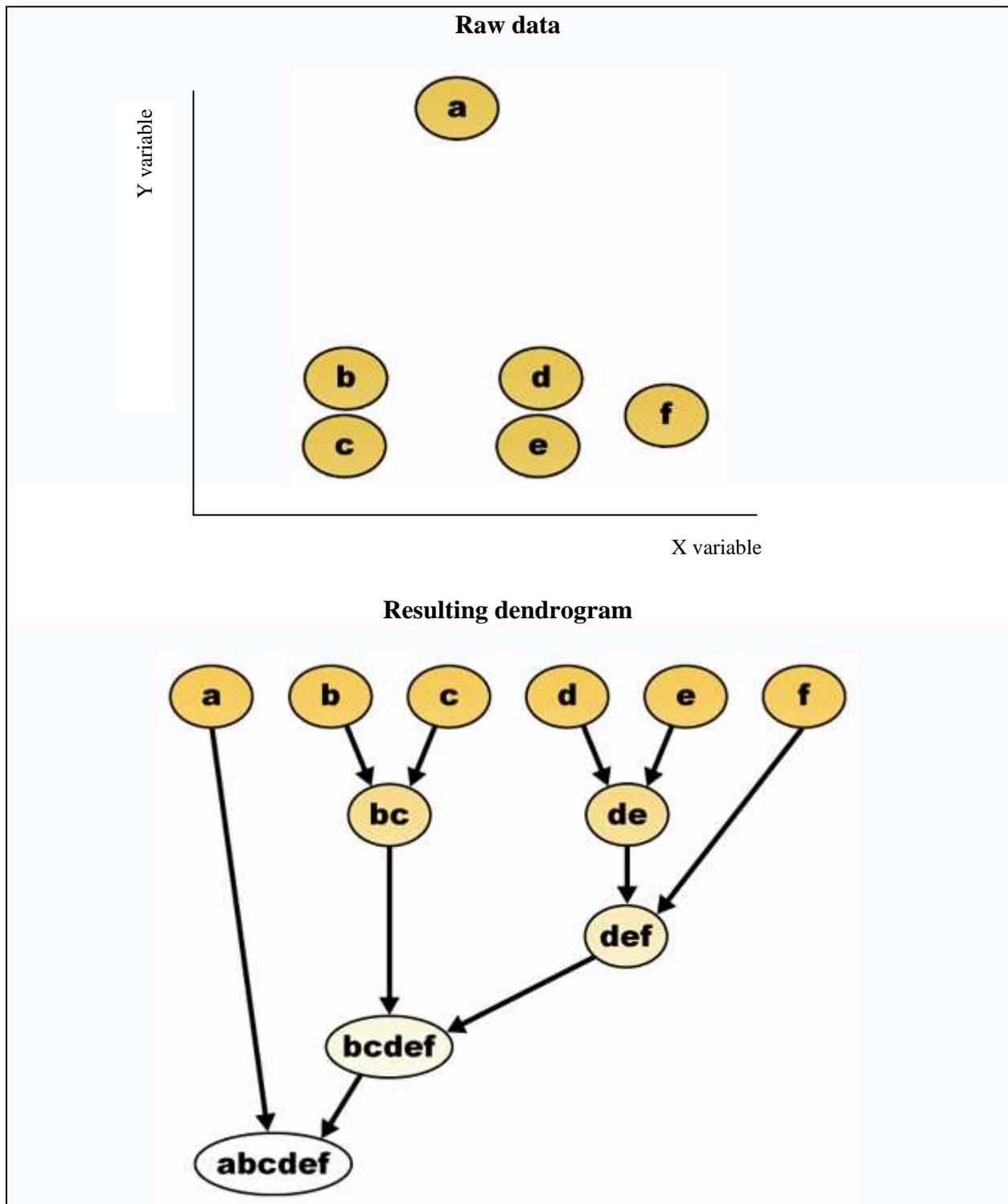


Figure 4 The UNDP (2003) approach



Source: UNDP (2003), Human Development Report 2003, pp. 347-349.

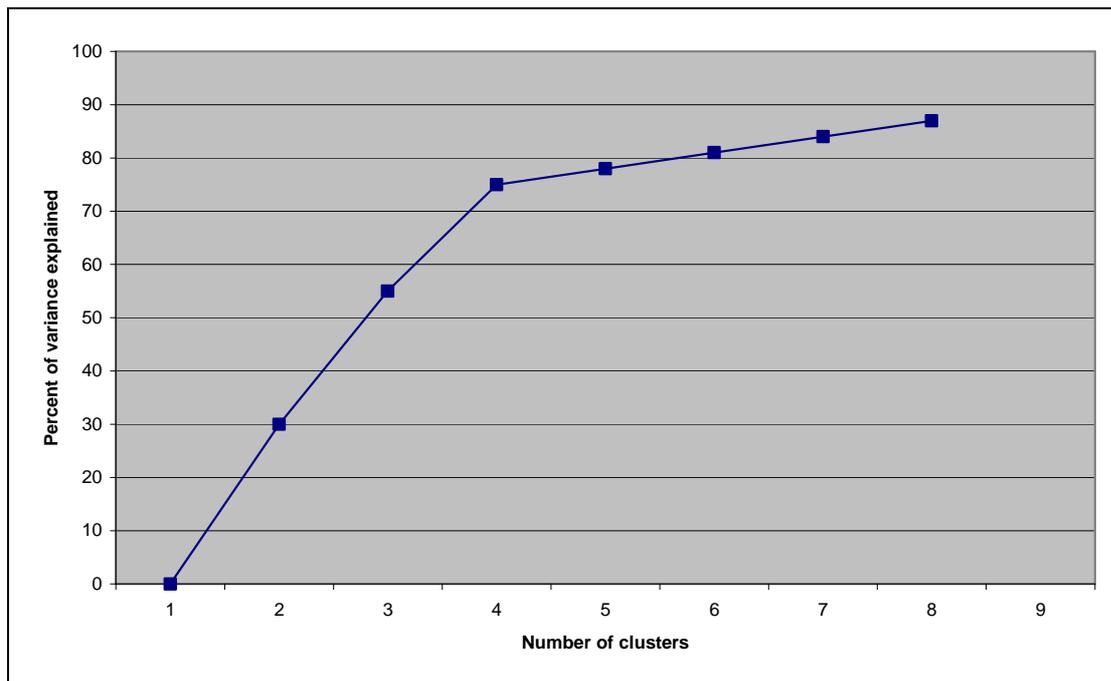
Figure 5 A hypothetical example of clustering and the dendrogram



Notes: The upper panel shows the raw data on two variables x and y for 6 hypothetical observations (e.g. countries) a , b , c , d , e and f . The lower panel shows the dendrogram summarising the process of agglomerative (bottom-up) clustering. This hypothetical example uses the Euclidian distance as the measure of distance (or dis-similarity) between any two observations (see main text).

Source: http://en.wikipedia.org/wiki/Cluster_analysis

Figure 6 The elbow criterion: a hypothetical illustration



Notes: In this hypothetical example, the appropriate number of clusters would be four.

Source: http://en.wikipedia.org/wiki/Cluster_analysis

Figure 7 Assessing the elbow criterion for 7 welfare indicators, 1970-2003

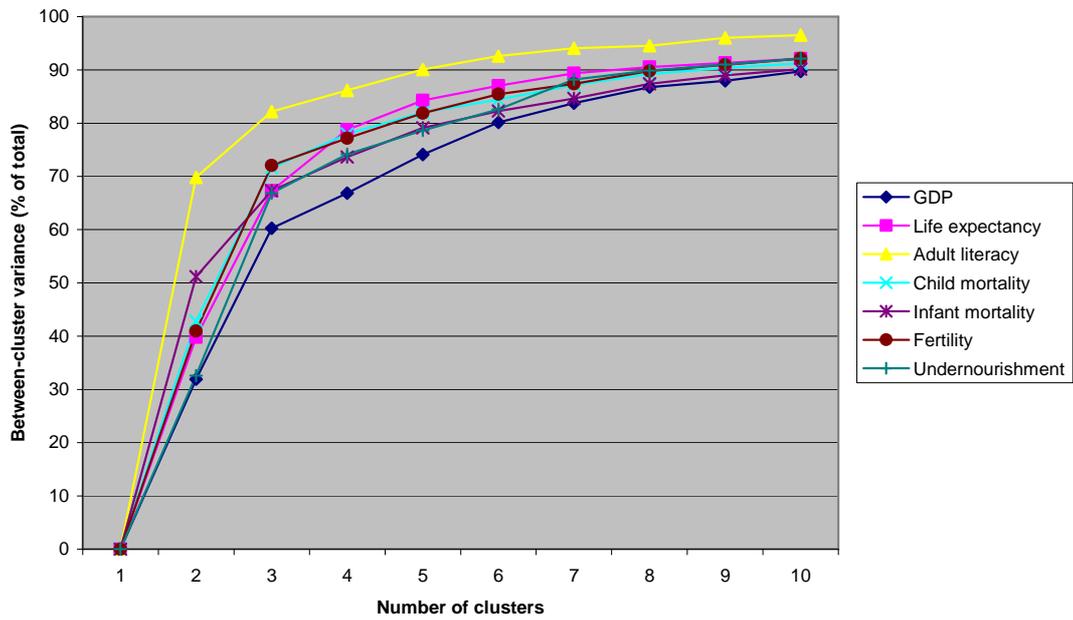


Figure 8 Assessing the elbow criterion for 7 welfare indicators, 1960-1993

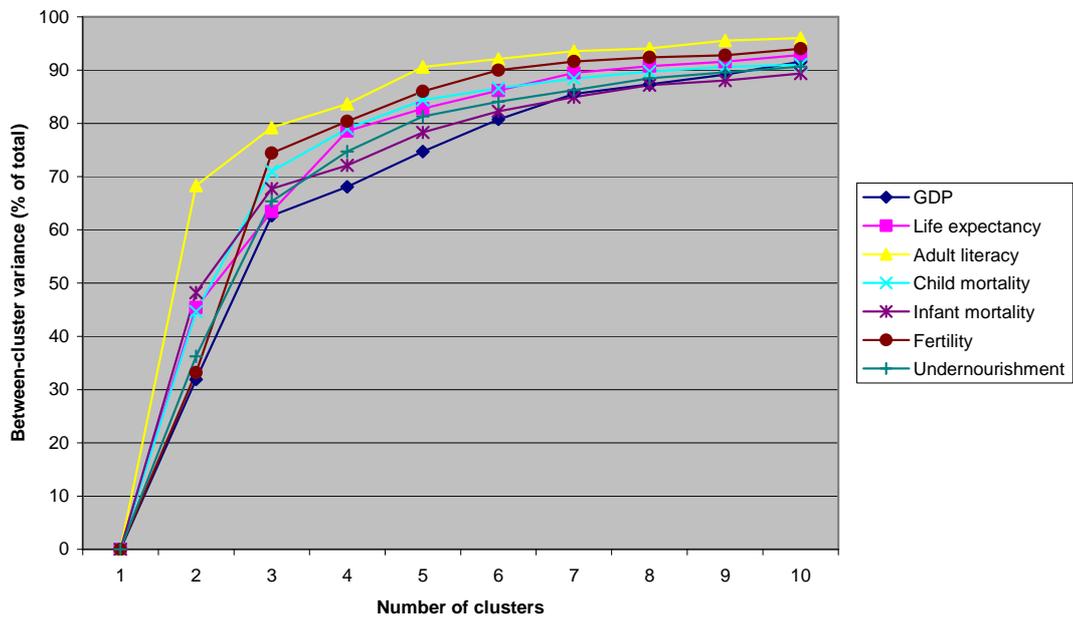


Figure 8 Scatter plot, GDP per capita (1970-2003)

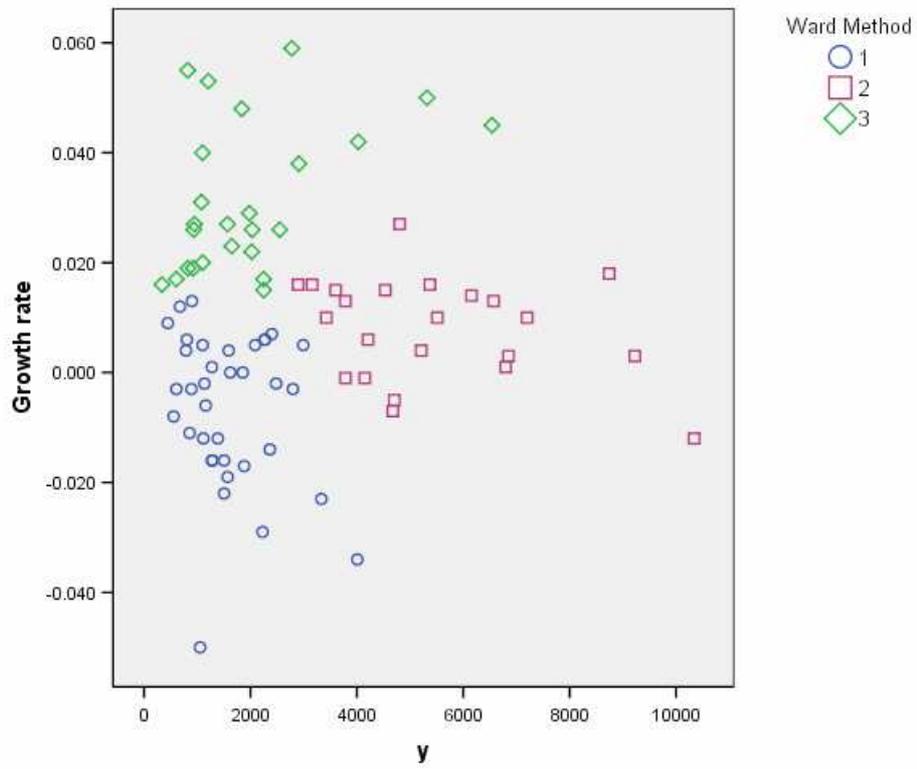


Figure 8a Scatter plot, GDP per capita (1960-1993)

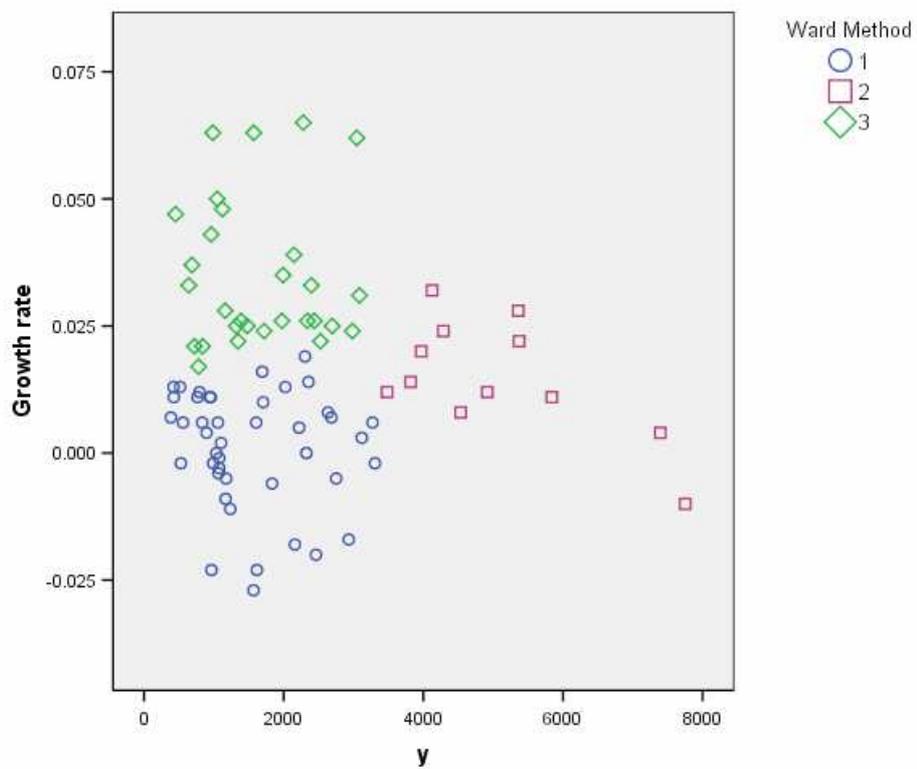


Figure 9 Scatter plot, child mortality (1970-2003)

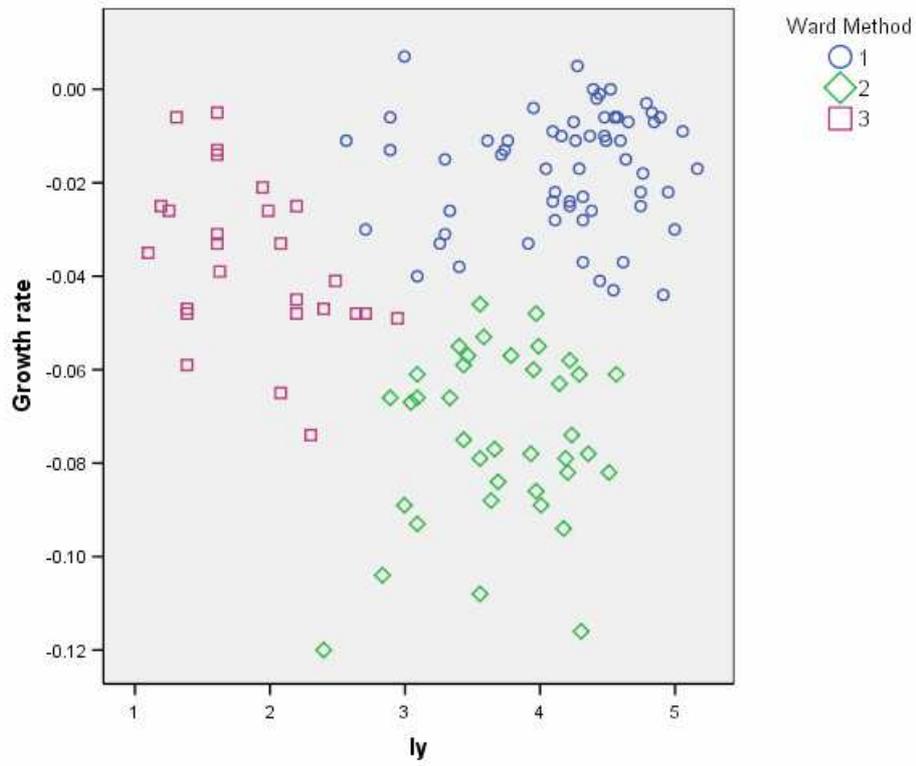


Figure 9a Scatter plot, child mortality (1960-1993)

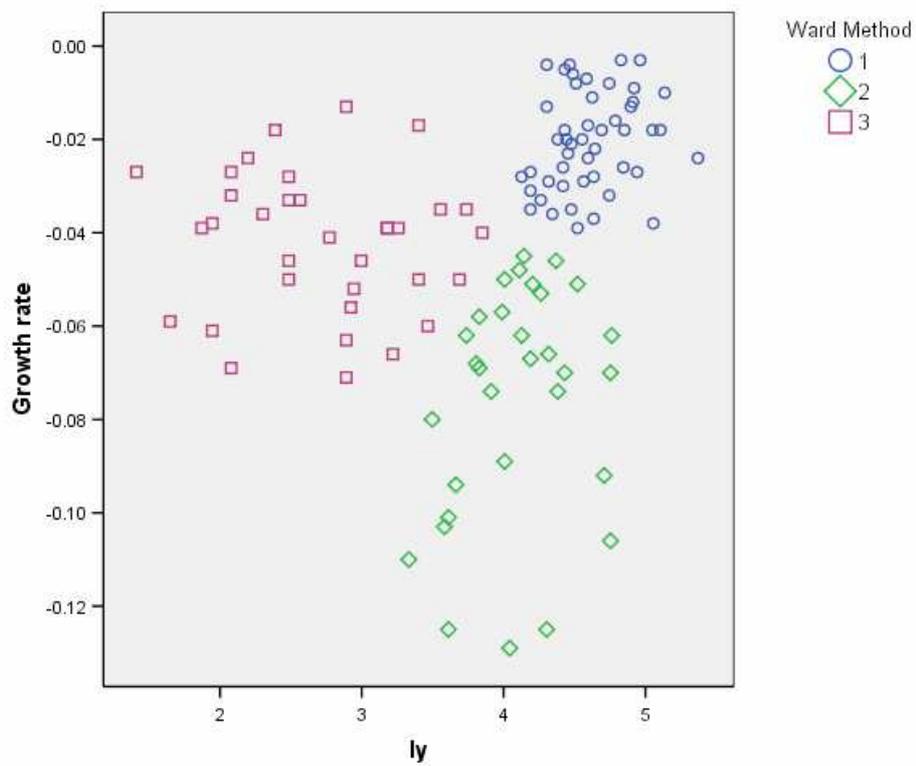


Figure 10 Scatter plot, fertility (1970-2003)

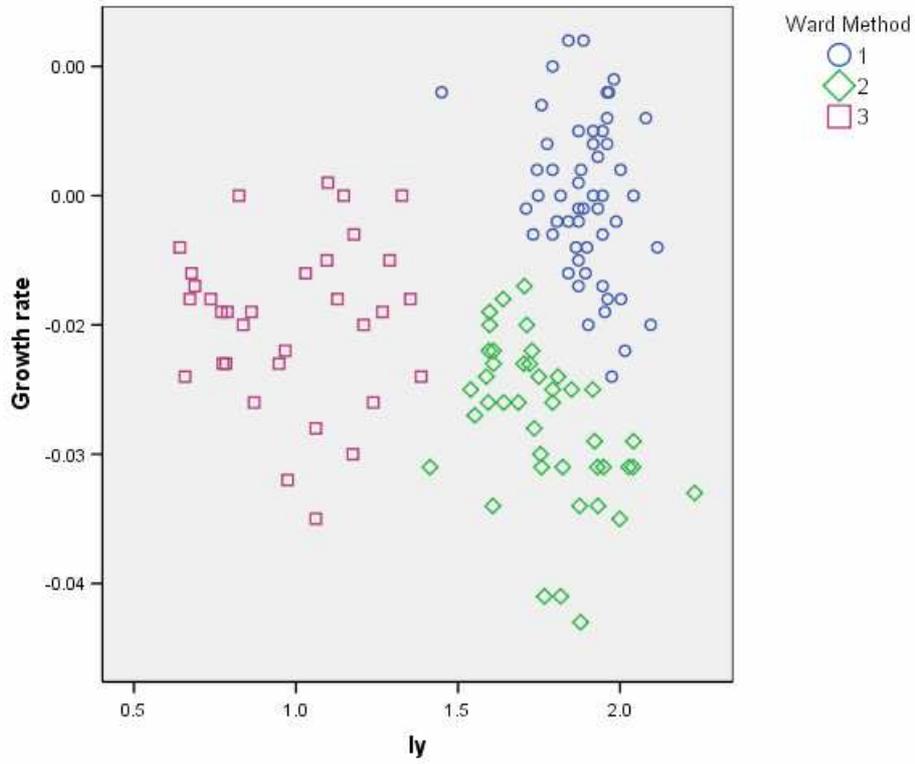


Figure 10a Scatter plot, fertility (1960-1993)

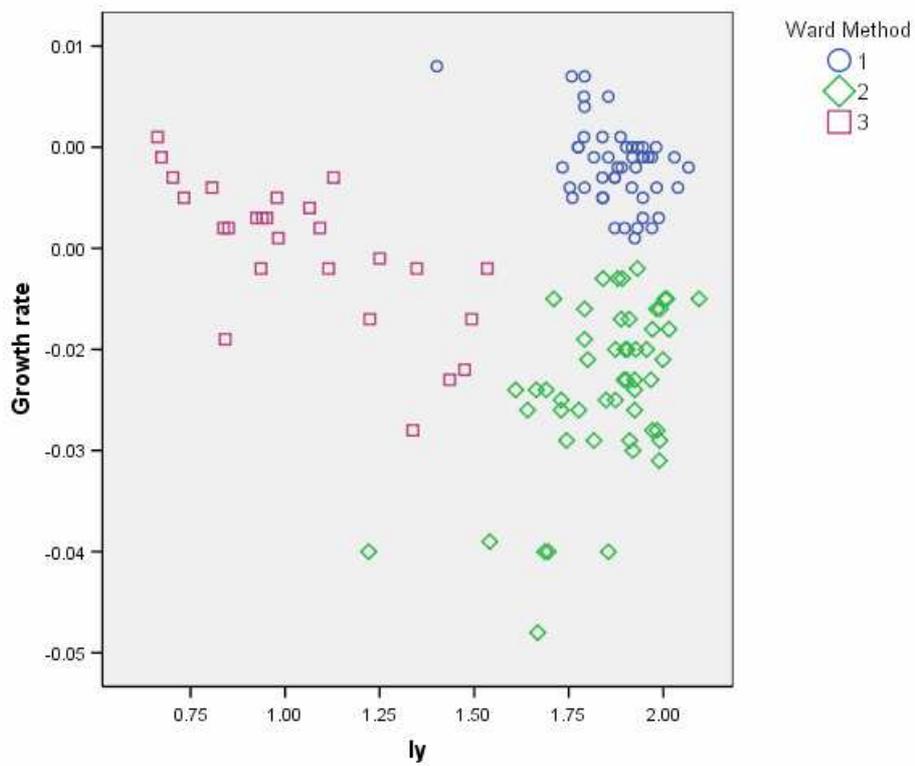


Figure 11 Scatter plot, under-nourishment (1970-2003)

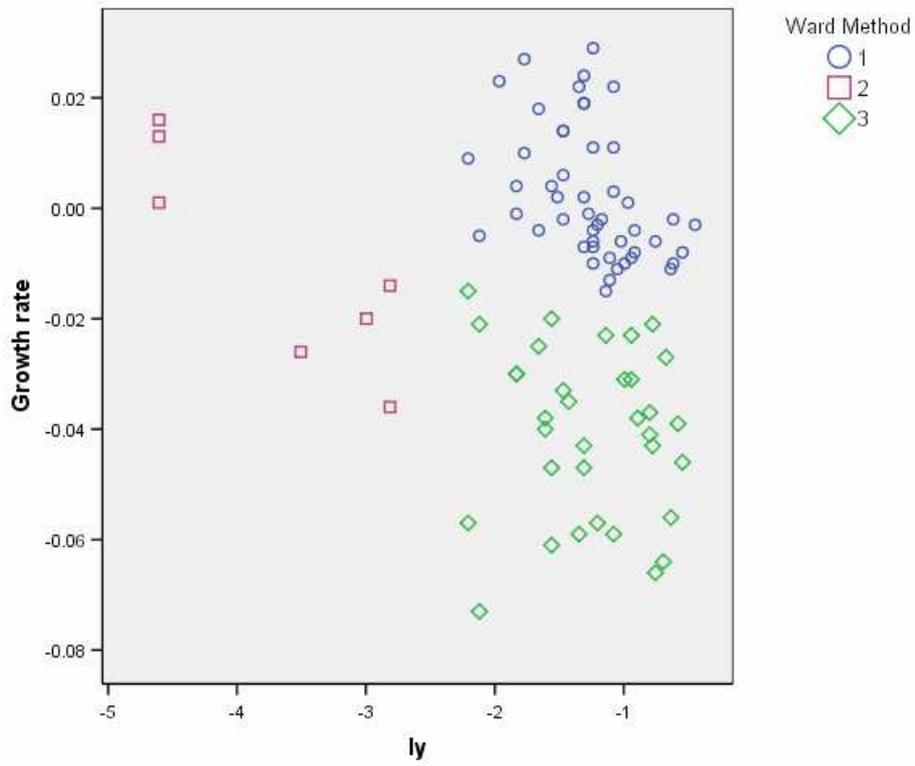


Figure 11a Scatter plot, under-nourishment (1960-1993)

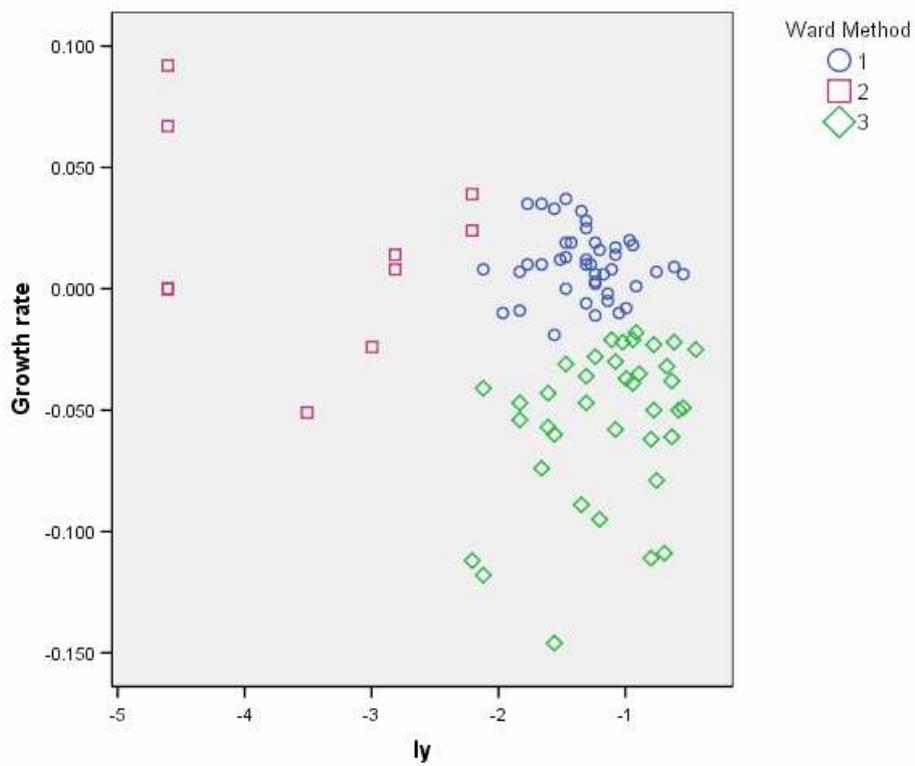


Figure 14 Scatter plot, adult literacy (1970-2003)

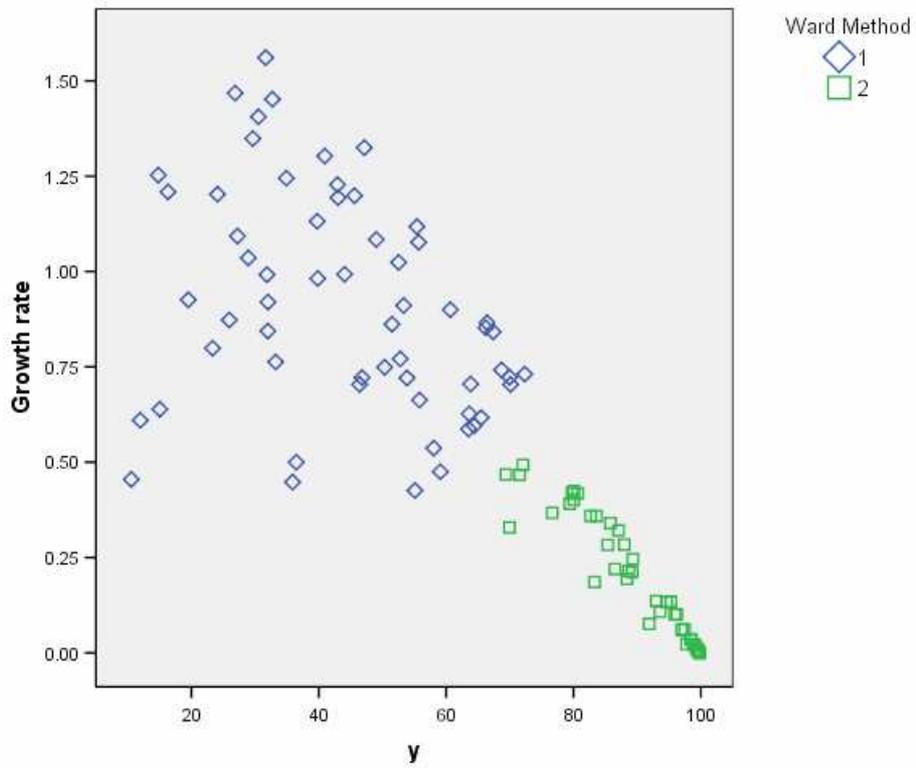


Figure 14a Scatter plot, adult literacy (1960-1993)

