IDS WORKING PAPER
Volume **2012** No **404**

Beyond Low and Middle Income Countries: What if There Were Five Clusters of Developing Countries?

Sergio Tezanos Vázquez and Andy Sumner
September 2012

IDS Institute of Development Studies

Poverty and Inequality Research Cluster

The Poverty and Inequality research cluster, part of the Vulnerability and Poverty Reduction team at IDS, produces research on poverty, inequality and wellbeing. Our research challenges orthodox views on the nature of poverty, how poverty is understood and how policy can best accelerate poverty reduction. Our work focuses on poverty and wellbeing through the lens of equity and inequality. Poverty is not only about 'poor' people but also about the social and economic inequalities that compound and reproduce poverty.
Email: poverty@ids.ac.uk
Web: www.ids.ac.uk/research-teams/vulnerability-and-poverty-reduction-team/research-themes/poverty-inequality-and-wellbeing

PI WP1

The Vulnerability and Poverty Reduction (VPR) Team aims to construct dynamic and multi-dimensional perspectives on vulnerability and poverty in order to transform thinking, policy and practice.

The VPR team produces working papers on social protection; conflict, violence and development; and poverty and inequality. Follow this link to view a full list of publications: www.ids.ac.uk/research-teams/vulnerability-and-poverty-reduction-team/publications/vpr-working-paper-series

Beyond Low Income and Middle Income Countries: What if There Were Five Clusters of Developing Countries?
Sergio Tezanos Vázquez and Andy Sumner
IDS Working Paper 404
First published by the Institute of Development Studies in September 2012

ISSN: 2040-0209 ISBN: 978-1-78118-082-2

A catalogue record for this publication is available from the British Library.

Available from:
Central Communications, Institute of Development Studies, Brighton BN1 9RE, UK
Tel: +44 (0) 1273 915637 Fax: +44 (0) 1273 621202
E-mail: bookshop@ids.ac.uk
Web: www.ids.ac.uk/ids/bookshop
IDS is a charitable company limited by guarantee and registered in England (No. 877338)

Beyond Low and Middle Income Countries: What if There Were Five Clusters of Developing Countries?

Sergio Tezanos Vázquez and Andy Sumner

Summary

Many have challenged the use of income per capita as the primary proxy for development. This paper continues this tradition with a twist. The paper challenges the continuing use of income per capita to classify developing countries as low income or middle income now that most of the world's poor no longer live in low income countries (LICs) and ambiguity over the usefulness of the middle income country (MIC) classification given the diversity in the group of over 100 MICs. We use a cluster analysis to identify five types of developing countries using a set of indicators covering definitions of development based on the history of thinking about 'development' over the last 50 years from four conceptual frames: development as structural transformation; development as human development; development as democratic participation and good governance; and development as sustainability. We find that the cluster analysis produces five types of developing country using data for the period 2005-2010. Our development taxonomy differs notably from the usual income classification of GNI per capita (Atlas method) used to classify LICs and MICs. Notably many countries commonly labelled "emerging economies" are not in the two clusters related to emerging economies because they retain characteristics of poorer countries.

We find that there is no simple "linear" representation of development levels (from low to high development countries). We find that each development cluster has its own and characteristic development issues. There is no group of countries with the best (or worst) indicators in all development dimensions. It thus would be more appropriate to build "complex" development taxonomies on a five-year basis than ranking and grouping countries in terms of per capita incomes, as this will offer a more nuanced image of the diversity of challenges of the developing world and policy responses appropriate to different kinds of countries.

Keywords: Poverty; Low-income countries; Middle-income countries; Country Taxonomy

Andy Sumner is a Research Fellow at the Institute of Development Studies at the University of Sussex.

Sergio Tezanos Vázquez is an associate professor at the Economics Department and a research fellow at the Iberoamerican Research Office on International Development & Co-operation at the University of Cantabria

Contents

Tables and Graphs

Acknowledgements

We would like to thank Richard Manning, Ainoa Quiñones, José María Larrú, Duncan Green and Rafael Domínguez for their comments and suggestions.

Abstract

Many have challenged the use of income per capita as the primary proxy for development. This paper continues this tradition with a twist. The paper challenges the continuing use of income per capita to classify developing countries as low income or middle income now that most of the world's poor no longer live in low income countries (LICs) and ambiguity over the usefulness of the middle income country (MIC) classification given the diversity in the group of over 100 MICs. We use a cluster analysis to identify five types of developing countries using a set of indicators covering definitions of development based on the history of thinking about 'development' over the last 50 years from four conceptual frames: development as structural transformation; development as human development; development as democratic participation and good governance; and development as sustainability. We find that the cluster analysis produces five types of developing country using data for the period 2005-2010. Our development taxonomy differs notably from the usual income classification of GNI per capita (Atlas method) used to classify LICs and MICs. Notably many countries commonly labelled "emerging economies" are not in the emerging economies clusters because they retain characteristics of poorer countries.

Our clusters are as follows:

- *Cluster 1*: ***High poverty rate countries with largely traditional economies.*** Those countries with the highest poverty and malnutrition headcounts, who are also countries with low productivity and innovation and mainly agricultural economies, with severely constrained political freedoms.
- *Cluster 2*: ***Natural resource dependent countries with little political freedom****.* Those countries with high dependency on natural resources, who are also countries with severely constrained political freedom and moderate inequality (relative to the average for all developing countries).
- *Cluster 3*: ***External flow dependent countries with high inequality.*** *Those coun*tries with high dependency on external flows, who are also countries with high inequality, and moderate poverty incidence (relative to the average for all developing countries).
- *Cluster 4*: ***Economically egalitarian emerging economies with serious challenges of environmental sustainability and limited political freedoms.*** Those countries with most equal societies, who are also countries with moderate poverty and malnutrition but serious challenges of environmental sustainability and –perhaps surprisingly– limited political freedoms.
- *Cluster 5*: ***Unequal emerging economies with low dependence on external finance****.* Those countries with the lowest dependency on external finance and who are also countries with the highest inequality.

Two-thirds of the world's poor – not surprisingly given the characteristics noted above - live in *Cluster 1* countries though this is largely due to the inclusion of four populous countries (Bangladesh, India, Pakistan and Nigeria and one should remember a third of world poverty is accounted for by India). About a quarter of world poverty is situated in *Cluster 3* and *Cluster 4* countries and the remaining 5% live in *Cluster 2* and *Cluster 5.*

We find that there is no simple "linear" representation of development levels (from low to high development countries). We find that each development cluster has its own and characteristic development issues. There is no group of countries with the best (or worst) indicators in all development dimensions. It thus would be more appropriate to build "complex" development taxonomies on a five-year basis than ranking and grouping countries in terms of per capita incomes, as this will offer a more nuanced image of the diversity of challenges of the developing world and policy responses appropriate to different kinds of countries.

1. Introduction

In 1963 Dudley Seers wrote –in *The Limitations of the Special Case*– of developing countries:

> *[t]he typical case is a largely unindustrialised economy, the foreign trade of which consists essentially in selling primary products for manufactures. There are about 100 identifiable economies of this sort, covering the great majority of the world's population (Seers, 1963, p. 80).*

And perhaps most famously, Seers wrote in *The Meaning of Development*:

> *The questions to ask about a country's development are therefore: What has been happening to poverty? What has been happening to unemployment? What has been happening to inequality? If all of these three have become less severe, then beyond doubt this has been a period of development for the country concerned […] If one or two of these central problems have been growing worse, especially if all three have, it would be strange to call the result 'development', even if per capita income has soared (Seers, 1969:24).*

Since then many have challenged the use of income per capita as the primary proxy for development. This paper continues this tradition with a twist. The paper challenges the continuing use of income per capita to classify developing countries as low income countries (LICs) or middle income countries (MICs), given that most of the world's poor live in the later group (Alkire *et al*. 2011; Chandy and Gertz 2011; Glassman *et al*. 2011; Kanbur and Sumner 2011; Sumner, 2012a; 2012b). Further, the ambiguity over the usefulness of the MIC classification given the diversity in the group of over 100 countries that includes Ghana and Zambia, as well as India, China and Brazil.

We use a cluster analysis to identify five types of developing country using a set of indicators covering definitions of development based on the history of thinking about 'development' over the last 50 years from four conceptual frames: development as structural transformation; development as human development; development as democratic participation and good governance; and development as sustainability.

The paper is structured as follows: Section 2 presents the four frames on 'development' emerging from the last five decades. Section 3 discusses the most used existing international classifications for countries. Section 4 presents our methodology and analysis. Section 5 concludes.

2.What is development? 50 years of thinking

In considering the history of thinking about 'development' over the last 50 years, four conceptual frames can be unambiguously identified (and many more so we just choose most influential ones here).[1] First, development as structural change–an idea particularly prevalent in the 1960s/70s. Second, development as human development –an approach that emerged from *basic needs* work in the 1970s and strongly in the 1990s in the UNDP *Human Development Report.* Third, development as democratic participation and good governance –a frame that arose strongly in the late 1990s and 2000s. Fourth, development as sustainable development – an idea that has steadily risen in prominence since the 1970s.[2]

All of these approaches to 'development' are boundless –there is not an end point of achieving 'final' development–; rather 'development' is positive if it moves in the direction of more structural change, or progress in human development, or better governance or a more sustainable trajectory. Here we label these frames thus:

> Development as structural transformation.
> Development as human development.
> Development as democratic participation and good governance.
> Development as sustainability.

The first frame –*development as structural transformation* – can be traced to thinking around the time of decolonisation of many countries in the 1950s and 1960s and the work of Arthur Lewis, Hans Singer, Raul Prebisch, Gunnar Myrdal and Dudley Seers.

Seers (1963) provided the seminal discussion of development as structural transformation. In which he discusses developed country characteristics, and their divergence from the characteristics of developing countries. Seers referred to the developed, or industrialised, countries 'a special case' of 'a few countries with highly unusual, not to say peculiar, characteristics' (p. 80). Furthermore, Seers (1963: 81-83) identified the characteristic features of the 'special case' or advanced economies in 'note form' including, for example, factors of production (e.g. literacy and the mobility of labour), sectors of the economy (e.g. manufacturing much larger than either agriculture or mining), public finance (e.g. reliance on direct taxes), households (e.g. very few below subsistence level and a moderately equal distribution of income), savings and investment (e.g. well-developed financial intermediaries), and 'dynamic influences' (e.g. slow population growth and high urbanisation).

This eludes to an underlying notion of 'development' as transformation from rural to urban, non-agricultural society as the work of Lewis (1954), Structuralists and others. Such a transformative view of societal change dominated the 1950s and 1960s if not beyond. Industrialization, structural societal change and economic development were defining aspects of development. However, such a perspective 'slipped from view' in the 1970s and since. As Gore (2000: 794-5) notes:

> The dynamics of long-term transformations of economies and societies [has] slipped from view and attention was placed on short-term [indicators] [...] The shift to ahistorical

[1] This section draws upon and develops Sumner (2004), Sumner and Tribe (2009), Sumner and Tiwari (2009) and work since.

[2] Other recently emerging 'candidates' might be "subjective wellbeing" and/or "state fragility". We did not include these in this paper as both are still evolving conceptually and empirically and remain contented in meaning and measurement. Some aspects of "state fragility" are captured by the governance and democracy measures. The indicators chosen were done so because they established conceptually and with long standing datasets necessary for the analysis.

> performance assessment can be interpreted as a form of the post-modernization of development policy analysis.

The shift Gore refers to is a shift towards annual economic growth rather than transformation of the economy and the emergence of tracking of poverty indicators. This leads us to a second frame –*development as human development.*

This second frame – can also be linked to another seminal work by Seers –*The Meaning of Development* (1969)– which led to the questioning of development as growth in Gross Domestic product (GDP) per capita alone (see Seer quote in introduction).

Seers's (1969) sought to push understandings of development beyond GDP per capita and into 'basic needs'. Further major contributions on 'basic needs' were made by other development economists, notably Paul Streeten (see Hicks and Streeten, 1979; Streeten, 1984) and staff at the ILO (1976; 1977). These 'basic needs' included not only income and employment but also the physical necessities for a basic standard of living such as food, shelter and public goods. This coincided with the emergence in the 1960s and 1970s of 'levels of living indicators' due to dissatisfaction. The culmination of efforts was the first composite measure of standards of living –Morris's (1979) physical quality of life index (PQLI).

The research of ILO, Morris, Baster (1979), McGranahan *et al.* (1985) and UNRISD (1970) set the foundations for Amartya Sen's work with the United Nations Development Programme (UNDP) on the 'human development'.

In the 1990s the meaning of development was more fundamentally reshaped by the work of Sen and the new annual *Human Development Report*, launched in 1990 by the UNDP. The new approach was known as 'human development' (or Sen's 'capabilities approach') and a related set of composite indicators were developed led by the *Human Development Index*.

Sen (see in particular 1999), Nussbaum (see in particular 2000) and UNDP (1990-2012) have argued consistently that development should focus on expanding capabilities –means, opportunities or substantive freedoms– which permit the achievement of a set of 'functionings' – things which human beings value in terms of 'being' and 'doing'. Development is not, as previously conceived, based on desire fulfilment (utility or consumption measured by a proxy for income, such as the GDP per capita) as this does not take sufficient evaluative account of the physical condition of the individual and of a person's capabilities. In short, income is *only* an instrumental freedom which can help to achieve other constitutive freedoms. Sen does not ignore income; rather he argues that too much emphasis can be placed on this dimension of development. Instead:

> Development consists of the removal of various types of unfreedom that leave people with little opportunity of exercising their reasoned agency […] Development can be seen […] as a process of expanding the real freedoms that people enjoy,[…] the expansion of the 'capabilities' of persons to lead the kind of lives they value –and have reason to value (Sen, 1999: xii,1,18).

Sen argued that there is a set of conditions including being fed, healthy, clothed and educated that together constitute a good life.. Individuals have a set of entitlements (command over commodities) which are created through a set of endowments (assets owned –physical and self– financial, human, natural, social and productive) and exchange (production and trade by the individual). These entitlements are traded for a set of opportunities (capabilities) in order to achieve a set of functionings (outcomes of well-being). There have been numerous attempts at constructing sets of capabilities (see review in Alkire, 2002). The most recently work in this area is that of the work of the Oxford Poverty and Human Development Initiative (OPHI) which has extended thinking with a set of multi-dimensional poverty measures (Alkire *et al*., 2011).

The third frame is that of *democratic participation and good governance*, which grew to prominence in the late 1990s and 2000s in part building on the work of Sen and others.

Access to governance structures and 'voice' of citizens in policy processes has both intrinsic and instrumental value to citizens and to the poor in particular (as Sen noted). Earlier, Myrdal (1956, *An international economy: problems and prospects*, p.180) was a pioneering development thinker who stressed the idea that it is necessary to promote social and political changes in order to improve the wellbeing; thus, neither the industrialization process nor the economic growth are possible without distributive reforms. In fact when one takes the governance and poverty literatures several areas are common to both: not only democratic participation and 'voice', but also human rights and freedoms and access to/delivery of quality public services for example.

Graham *et al*. (2003: 1-2) define governance as 'the traditions, institutions and processes that determine how power is exercised, how citizens are given a voice, and how decisions are made on issues of public concern'. Hyden *et al*. (2004: 5) define governance as 'the formation and stewardship of the rules that regulate the public realm –the space where state as well as economic and societal actors interact to make decisions'. Thus, governance is about the relationship(s) between governments and society. However, governance is not the same as government and the solutions to poor governance are not solely in the domain of governments and governance is about more than just corruption. In sum, governance is about who decides – who sets what rules, when and how. Such rules are no-longer the preserve of the state alone. The trend has been from representative or formal democracy (i.e. indirect participation) towards more mechanisms for ensuring citizens voice in the decision-making processes.

Finally, the fourth frame chosen is that of *sustainable development* (SD). Concern with the impacts of economic growth and development on depletion and degradation of the natural environment emerged in the 1970s. Attention culminated in the World Commission on Environment and Development (WCED) and the later Rio Earth Summits.

Over the last few years such concerns have started to take on a new impetus in light of climate change discussions. However, attempts to bring sustainability dimensions into policy making have been plagued by one basic question –what to sustain? The most often cited definition of SD is still that of the WCED (1987: 43) that identifies SD as meeting the 'needs' of the present without undermining the ability to meet the 'needs' of the future. However, this definition – although the most often used– is of little practical use. What exactly it means in policy is not clear: what is to be sustained, consumption at current levels (for future generations' 'needs' to be met) or sustain the environmental resources themselves (for future generations to meet their self-defined 'needs')?

There have been several attempts at composite measures that aimed to challenge the primacy of GDP (incorporating social, economic and environment components) such as the GPI –the *Genuine Progress Indicator* and the ISEW–, and the *Index of Sustainable Economic Welfare.*[3]

However, Neumayer (2003: 2-3) argues that the measurement of well-being (social and economic progress) and sustainability should remain conceptually separated because what affects the former is not necessarily the same as what affects the later and *vice versa*. Additionally, the first refers to total *current* capital stock whilst the later to the total *future* capital stock. For this reason, Neumayer rejected attempts at amalgamating social development and sustainable development completely. This is not to say the concepts are not linked –poverty related to current well-being and SD to future well-being (hence the New Economic Foundation's Happy Planet Index)–; but the first is not about the future flow of well-being and the latter is not concerned with the current stock of well-being.

[3] For greater detail on these see, for example, Moffatt et al. (1996).

Attempts at reconciling the conceptual discussion led to back-and-forth debates –notably in the pages of *Environment Values* (see for details Dower, 1994; Beckerman, 1995; Daly, 1995; El Serafy, 1996; Common, 1996). Initially, Pearce *et al.* (1989) temporarily addressed the question of what to sustain with the concepts of *strong sustainability* and *weak sustainability* based on the infinite substitutability (weak) or non-substitutability (strong) of natural capital. *Strong sustainability* –also known as *utility-based sustainability* or *Solow sustainability*– is based on the work of neo-classical economists such as Hicks (1939). Within this framework of inter-generational equity, Solow (1974; 1986) and Hartwick (1977) argued that SD is sustaining the utility of future generations and that this is possible if the *Solow-Hardwick rule* is followed: to maintain constant consumption over time, dependent countries must reinvest all rents from natural resource extraction in productive capital (assuming an initial endowment of natural resources adequate for a certain standard of living). Economic growth is sustained as the scarcity of one resource (natural capital) can be compensated (substitution) by the availability of another (man-made or produced capital).

In contrast, Daly (2002: 1) argued that *weak sustainability* or *throughput-based sustainability* is based on sustaining the entropic physical flow from nature and back to nature as non-declining. In order to do this there has to be zero economic growth and zero population growth –known as a *steady-state* because natural capital is infinitely non-substitutable. Daly (2002: 7) went so far as to argue that GDP growth might be *un*economic growth, because environmental depletion and GDP growth generate not only wealth but *illith* –John Ruskin's term for the opposite of wealth.

However, Beckerman's (1994) devastating critique undermined both *weak* and *strong sustainability.* He argued that *weak sustainability* took debates not much further than the current model of economic welfare maximisation and *strong sustainability* was morally unacceptable and totally impractical.

The weaknesses of *weak sustainability* were largely agreed, but no such consensus was reached on the 'moral injunction' on *strong sustainability* –whether considering something to be absolutely or relatively more or less sustainable than other options confers a moral imperative or not.

In the forthcoming discussion we take an understanding of SD that is entirely focused on climate change as the pressing development issues and use CO_2 emissions per capita.

3. Main international classifications of development

It is not easy to classify countries according to their levels of development, to begin with because any definition of "development" is complex and multidimensional. Added to this difficulty is the fact that the socio-economic realities of the so-called "developing countries" are becoming more diverse and heterogeneous, which makes universally valid analysis even more difficult and unreliable. In fact, as stated by Nielsen (2011: 3), "when it comes to classifying countries according to their levels of development, there is no criterion (either grounded in theory or based on an objective benchmark) that is *generally accepted*".

Despite these difficulties, there are several international classifications of development that use different criteria to draw some kind of "global development threshold" that separates the "developed" and the "developing" countries. Four particularly influential classifications are those developed and used by the World Bank, the OECD, the UNDP and UNCTAD.

On the one hand, the World Bank provides, since 1978, a ranking of countries according to their corresponding levels of per capita income (proxied by the per capita GNI based on the Atlas method, largely an exchange rate conversion). Although the World Bank recognizes that development is not only a matter of income, it believes that the per capita GNI is "the best single indicator of economic capacity and progress" (World Bank, 2012a). Thus, the successive *World Development Reports* (and their corresponding statistical appendixes, the *World Development Indicators*) classify countries into four income groups. According to the latest edition (World Bank, 2011) these groups are:

- The "low income countries" (LIC), with less than $1,005 per capita GNI in 2010.
- The "lower middle income countries" (LMIC), with per capita incomes between $1,006 and $3,975.
- The "upper middle income countries" (UMIC), with incomes between $3,976 and $12,275
- The "high income countries" (HIC), with more than $12,276 per capita income.

On the other hand, the OECD's Development Assistance Committee (DAC) uses the World Bank's income classification in order to distinguish two groups of countries (DAC, 2011): the "developing countries" (LIC, LMIC and UMIC, according to the World Bank)[4], and the "developed countries" (basically high-income countries). The former are potential recipients of Official Development Assistance (ODA).

Further, the UNDP ranks countries by levels of "human development" by means of a composite index –the *Human Development Index*, HDI– that tries to capture the multidimensionality of the development process (see earlier discussion). The HDI was first developed by Mahbub ul Haq with the collaboration of the Nobel laureate Amartya Sen and other leading development thinkers for the first *Human Development Report* in 1990. Specifically, the index includes 3 dimensions of development: health, education and living standards. Thus the HDI breaks the conventional classification of countries according to per capita income levels, and, instead, classifies countries into four relative groups of human development (UNDP, 2011):

Very high human development countries, with HDI greater than 0.79 in 2011.
High human development countries, with HDI between 0.698 and 0.79
Medium human development countries, with HDI between 0.52 and 0.698.

[4] Specifically, the DAC classification divides the LIC group into 'Least Developed Countries' (LDC, as we will explain later) and 'other low income countries'.

Low human development countries, with HDI less than 0.52.

There is also the UN category of 'Least Developed Countries' (LDC), which utilises a sophisticated methodology that combines human assets (including nutrition, child mortality, school enrolment and adult literacy), economic vulnerability (including measures of the instability of agricultural production, population displaced by natural disasters, instability in exports, the share of agriculture in GDP and exports), and proxies for economic 'smallness' (less than 75 million people), 'remoteness' and GNI per capita. However, the graduation criteria make it very difficult to leave the category (see Guillaumont, 2010) and a third of the 48 LDCs are MICs.

Curiously enough, the most internationally widespread development classification is just the simplest one, based solely on a per capita income indicator. Thus, according to the World Bank's classification, the majority of the developing countries' population (and the world's poor) live in the generically called "middle income countries" (LMIC and UMIC)[5]. However –as we will show later–, the simplicity of this criterion and the large interval width of the middle income group (with a range greater than 12) masks significant differences in terms of the development challenges faced by these countries. It is worth remembering that any system for thinking about countries has to retain some level of simplicity to be taken up.

4. An alternative development classification: taxonomy of developing countries using cluster analysis

There are different procedures for classifying countries in "development groups" –once the development indicators have been chosen–. In the case of those groupings listed above such as the LICs and MICs, the groupings are made by means of an ordinal criterion. However, this procedure fails to determine both the appropriate number of groups of countries, and the "development thresholds" that separate the groups.[6] As we will explain below, cluster analysis offers a more nuanced and objective statistical technique for the composition of groups of countries than the mere ordering of a given development indicator, and it allows to include a more complete set of indicators that better reflects the multidimensionality of the development process.

In the following pages we propose an alternative development taxonomy for 139 developing countries –almost all of the LICs and MICs according to the current World Bank's list. We first identify the main development dimensions and proxies that we use for the classification. Secondly, we justify the convenience of cluster analysis in building a development taxonomy. Thirdly, we analyze the cluster solutions and characterize each group of countries.

[5] See Sumner (2012a) for an update of the world's poor distribution.

[6] See Nielsen (2011) for a detailed explanation on how the World Bank, the IMF and the UNDP determine the number of countries that makes up each income group. Nielsen also proposed an alternative "data-driven" methodology for grouping countries according to a single development indicator, which overcomes the arbitrary definition of the income intervals.

4.1. Identification of the Development dimensions

We propose a set of proxies for the development conceptions and dimensions using the four frames for 'development' (**Table 1**).[7] To measure development as economic independence and structural transformation we use data for structural change (GDP in non-agriculture), natural resource dependency (exports of primary commodities), labour productivity (GDP per worker), innovation capacities (production of scientific articles) and external finance (the sum of Official Development Assistance, Foreign Direct Investment, portfolio investment and remittances).

To measure development as human development we use health (malnutrition prevalence of children under five), purchasing power (GDP per capita PPP), income poverty rates (at $2/day as the average poverty line for all developing countries –Chen and Ravallion, 2008), and inequality (Gini coefficient).[8]

To measure development as democratic participation and good governance we use good governance (World Governance Indicators, WGI) and quality of democracy (POLITY 2).

And, finally, to measure development as sustainability we use CO_2 emissions (metric tons per capita).

Table 1. Development dimensions/concepts and data used. 2005-2010

Development dimensions/conceptions	Sub-dimensions	Proxies	Sources	Methods of construction
I. Development as human development	1.1. Poverty	Poverty headcount (2$ PPP a day)	World Bank (2012c)	Closest available years
	1.2. Inequality	Gini coefficient	Solt (2009)	Closest available years
	1.3. Health	Malnutrition prevalence, weight for age (% of children under 5)	World Bank (2012b)	5-years averages
	1.4. Purchasing power	GDP per capita, PPP (constant 2005 $)	World Bank (2012b)	5-years averages
II. Development as economic autonomy	2.1. Structural change	GDP in non-agricultural sectors (% of GDP)	World Bank (2012b)	5-years averages
	2.2. Dependency on natural resources	Exports of primary commodities (% of GDP)	UNCTAD (2012) and World Bank (2012b)	5-years averages
	2.3. Labour productivity	GDP per worker, PPP (constant 2005 $)	Heston et al. (2011)	5-years averages
	2.4. Innovation capacities	Scientific articles (per million inhabitants)	World Bank (2012b)	5-years averages
	2.5. External finance	(ODA+FDI+portfolio investment+remittances)/GDP	DAC (2012) and World Bank (2012b)	5-years averages
III. Development as political freedom	3.1. Good governance	World Governance Indicators	Kaufmann et al. (2011)	2-years averages of 6 governance indicators
	3.2. Quality of democracy	POLITY 2	Marshall and Jaggers (2011)	5-years averages
IV. Development as sustainability	4.1. Environmental sustainability	CO2 emissions (metric tons per capita)	World Bank (2012b)	5-years averages

4.2. Statistical procedure: cluster analysis of developing countries

Cluster analysis is a numerical technique that is suitable for classifying a sample of heterogeneous countries in a limited number of groups, each of which is internally

[7] See **Appendix 1** for descriptive statistics of the dataset.

[8] We initially considered (un)employment but we finally ruled it out due to comparability problems of the unemployment rates across countries.

homogeneous in terms of the similarities between the countries that comprise it. Ultimately, the goal of cluster analysis is to provide classifications that are reasonably "objective" and "stable" (Everitt *et al.*, 2011): objective in the sense that the analysis of the same set of countries by the same numerical methods produces similar classification; and stable in that the classification remains similar when new countries –or new characteristics describing them– are added. According to Everitt *et al.* (2011: 13):

> Cluster analysis techniques are concerned with exploring data sets to assess whether or not they can be summarized meaningfully in terms of a relatively small number of groups or clusters of objects or individuals which resemble each other and which are different in some respects from individuals in other clusters.

Specifically, hierarchical cluster analysis allows to build a "taxonomy" of countries with heterogeneous levels of development in order to divide them into a number of groups so that: *i*) each country belongs to one –and only one– group; *ii*) all countries are classified; *iii*) countries of the same group are, to some extent, internally "homogeneous"; and *iv*) countries of different groups are noticeably dissimilar.[9] In the end, this type of analysis allows to discern the association structure between countries, which –in our analysis– facilitates the identification of the key development challenges that characterize each cluster.

Furthermore, cluster analysis deals with two intrinsic problems of the design of a development taxonomy. On the one hand, it facilitates the determination of the appropriate number of groups in which to divide the sample of countries. On the other hand –given that each country has different values for the set of development indicators–, cluster analysis allows bringing together the different indicators by building a synthetic distribution that makes easier comparison of the development indicators across countries.

Nevertheless, cluster analysis also poses some particular difficulties for the classification of countries. Nielsen (2011) pointed out two main difficulties: Firstly, if the values of the development indicators are evenly distributed across countries, the analysis is not able to distinguish groups, even though there may be important differences between the indicators for each country. However, as we will discuss below, this limitation does not affect our case of study, as the analysis clearly discern the association structure across developing countries and thus allow us to identify a small number of developing groups. And secondly, Nielsen criticizes that clustering techniques involve a large degree of freedom in choosing among alternative distance measures and cluster algorithms, which in turn complicates the selection of time-invariant variables that can be used in periodic updates of the classification. However, this difficulty only applies in the case of restricting the classification over-time to the same exact number of groups (regardless of what the cluster analysis suggests); cluster analysis can be used to replicate the analysis in different periods, to compare the groups built in each period (not necessarily the same number of groups), and to analyze the dynamics of the development process of each country in comparative terms (i.e. in terms of their movements across development groups).

In the following pages we carry out a hierarchical cluster analysis using the Ward's method, computing the squared Euclidean distances between each element and standardizing the variables in order to correct their differences in scale.[10] The analysis includes 101 of the 139

[9] In this way, Tezanos and Quiñones (2012) applied cluster analysis techniques for classifying the middle-income countries of Latin America and the Caribbean.

[10] See **Appendix 2** for an explanation of the clustering method used in this piece of research. Regarding the standardization method, we use the "range -1 to 1" as it has been proven to work better than other methods "in most situations" (Mooi and Sarstedt, 2011: 247).

LICs and MICs (i.e. 72.7% of the targeted countries, and 95.3% of the population of the developing world).[11]

Before starting the clustering process, we first examine the variables for substantial collinearity. As the initial dataset includes 12 variables that proxy different development dimensions, it is reasonable to suspect that there may be some highly correlated variables in our dataset.[12] The correlation matrix (**Table 2**) confirms this suspicion, as there are two variables, *GDPpc* and *productivity*, that have a high correlation coefficient (close to 0.9), indicating possible collinearity issues. Thus we opt to omit *GDPpc* from the subsequent cluster analyses[13]. The remaining variables still provide a sound basis for carrying out cluster analysis.

[11] The countries not included in the analysis are either insular States with less than one million inhabitations (Antigua and Barbuda, Dominica, Fiji, Grenada, Kiribati, Maldives, Marshall Islands, Mauritius, Mayotte, Palau, Samoa, Sao Tome and Principe, Seychelles, Solomon Islands, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Tonga, Tuvalu and Vanuatu), or countries with limited statistical information (Afghanistan, Bosnia and Herzegovina, Cuba, Eritrea, Kosovo, Lebanon, Libya, Mongolia, Myanmar, North Korea, Somalia, Sudan, Timor-Leste, Uzbekistan, West Bank and Gaza, and Zimbabwe).

[12] If highly correlated variables are used for cluster analysis, specific aspects covered by these variables will be overrepresented in the solution. In practical terms, Everitt *et al.* (2011) and Mooi and Sarstedt (2011) argue that absolute correlations above 0.9 are problematic.

[13] In the end we choose to drop *GDPpc* instead of *productivity* because the later is both a determinant of the income level and a determinant of the development level; whereas high per capita income levels do not necessary lead to higher labour productivity.

Table 2. Correlation matrix

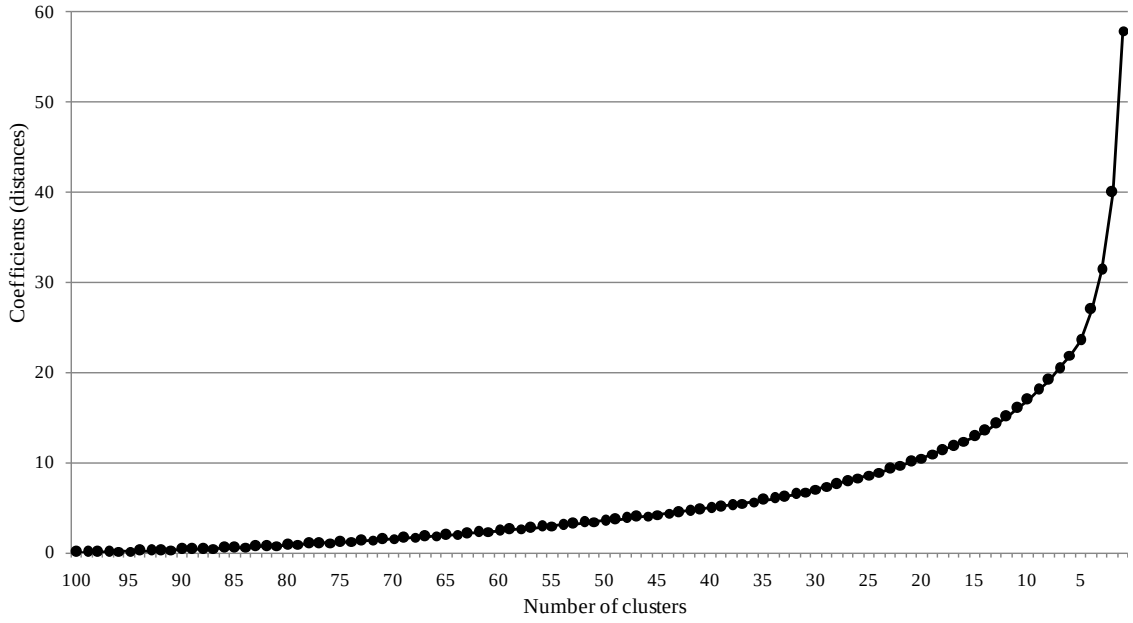
		Poverty	Gini	Malnutrition	GDPpc	Non-agriculture GDP	Primary exports	Productivity	Articles	External finance	WGI	POLITY	CO2pc
Poverty	Pearson	1											
	Sig. (2-tailed)												
	N	111											
Gini	Pearson	,086	1										
	Sig. (2-tailed)	,381											
	N	105	112										
Malnutrition	Pearson	,746	-,046	1									
	Sig. (2-tailed)	,000	,635										
	N	107	109	124									
GDPpc	Pearson	-,749	,000	-,630	1								
	Sig. (2-tailed)	,000	,999	,000									
	N	111	112	120	132								
Non-agriculture GDP	Pearson	-,708	,086	-,565	,705	1							
	Sig. (2-tailed)	,000	,370	,000	,000								
	N	108	111	119	129	131							
Primary exports	Pearson	-,048	-,027	-,124	,081	,140	1						
	Sig. (2-tailed)	,621	,776	,177	,358	,116							
	N	110	111	120	130	128	133						
Productivity	Pearson	-,759	,096	-,594	,899	,675	,181	1					
	Sig. (2-tailed)	,000	,320	,000	,000	,000	,047						
	N	108	110	120	121	121	121	123					
Articles	Pearson	-,511	-,074	-,465	,561	,398	-,131	,495	1				
	Sig. (2-tailed)	,000	,438	,000	,000	,000	,139	,000					
	N	109	111	121	130	130	130	122	134				
External finance	Pearson	,018	-,091	-,027	-,190	-,391	-,105	-,365	-,197	1			
	Sig. (2-tailed)	,848	,339	,770	,029	,000	,228	,000	,024				
	N	111	112	121	131	130	133	122	131	135			
WGI	Pearson	-,380	,114	-,437	,564	,452	-,334	,438	,374	-,072	1		
	Sig. (2-tailed)	,000	,233	,000	,000	,000	,000	,000	,000	,413			
	N	111	112	123	130	129	131	122	132	133	136		
POLITY	Pearson	-,064	,304	-,124	,132	,038	-,299	,081	,133	,186	,518	1	
	Sig. (2-tailed)	,517	,001	,188	,163	,686	,001	,392	,156	,047	,000		
	N	104	107	115	113	114	113	114	115	114	117	117	
CO2pc	Pearson	-,650	-,169	-,514	,702	,547	,233	,641	,483	-,181	,201	-,165	1
	Sig. (2-tailed)	,000	,075	,000	,000	,000	,007	,000	,000	,037	,020	,077	
	N	111	112	123	132	130	132	123	134	133	134	116	136

4.3. Main results[14]

A key aspect of the development classification is deciding on the number of developing country groups (i.e. the number of clusters to retain from the data). For guiding this decision we will use 3 different tools: the agglomeration schedule, the dendrogram and the variance ratio criterion.

The agglomeration schedule (see **Appendix 3**) displays the clusters combined at each stage (second and third column) and the distances at which the mergers take place.[15] The agglomeration schedule can be used to determine the optimum number of groups of countries: by plotting the distances ("coefficients" column) against the number of clusters, we can identify a distinct break ("elbow") in the number of cluster (i.e. where an additional combination of two clusters occurs at a greatly increased distance). Thus, the number of clusters prior to this merger is the most probable solution. In this way –and despite the high number of countries included in the graph–, the scree plot shows a distinct break due to the increase in distance when switching from a five to a six-cluster solution (**Graph 1**).

Graph 1. Scree plot: distances against number of clusters



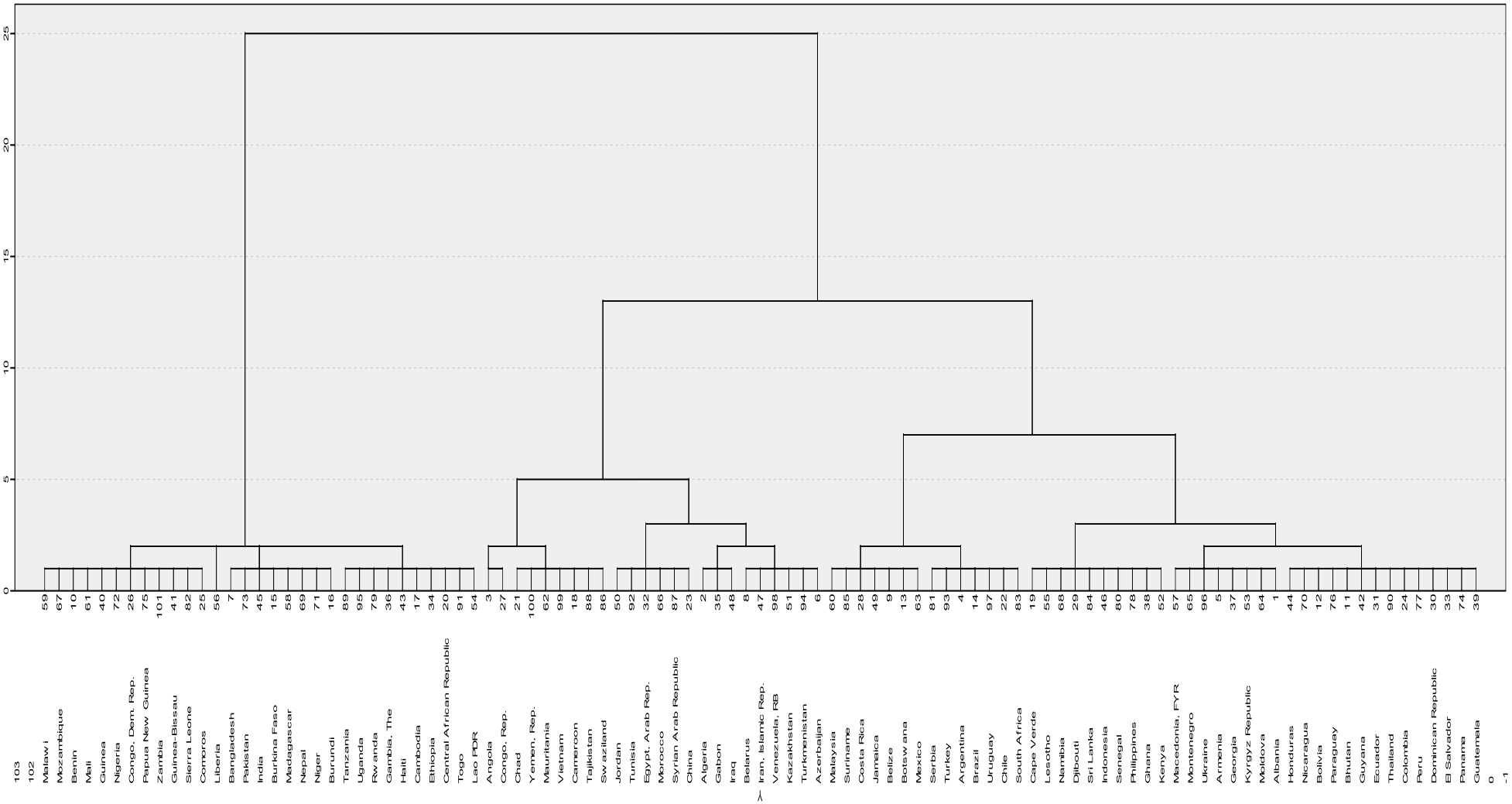
The dendrogram (**Graph 2**) graphically displays the distances at which countries (and clusters of countries) are joined. The dendrogram is read from left to right; vertical lines are countries joined together –their position indicates the distance at which the mergers take place–[16]. This graph provides a rough guidance regarding the number of groups to retain, suggesting that between 4 and 6-clusters solutions may be appropriate.

[14] The analysis was carried out using *IBM SPSS Statistics*.

[15] For example, in the first stage, Malawi (country 59) and Mozambique (67) are merged at a distance of 0.149. From here onward, the resulting cluster is labelled as indicated by the first country involved in this merger (in this case, country 59).

[16] *SPSS* re-scales the distances to a range of 0 to 25. Therefore, the last merging step to a 1-cluster solution takes place at a (re-scaled) distance of 25.

Graph 2. Dendrogram



A more precise and objective method for determining the optimum number of clusters was proposed by Calinski and Harabasz (1974), which has proven to work well in many situations (Milligan and Cooper, 1985). The so-called "variance ratio criterion" (VRC) recommends choosing the number of clusters that maximizes the ratio between the overall between-cluster variation and the overall within-cluster variation with regard to all clustering variables (i.e. a good clustering yields groups of countries with small within-cluster variation but high between-cluster variation). In our case, this suggests that the optimum number of clusters is five (**Table 3**).

Table 3. Variance Ratio Criterion (VRC)

# clusters	VRC_k	w_k
2	650.47	..
3	542.11	233.52
4	667.26	-50.97
5	741.45	-127.31
6	688.33	259.60
7	894.81	..

Note: VRC implies choosing the cluster with minimum *w*. See Mooi and Sarstedt (2011, appendix of chap. 9) for a practical explanation of this criterion.

Therefore, the 3 procedures (the distances scree plot, the dendrogram and the VCR) confirm that the optimum number of clusters is five. The resulting development classification in five groups is shown in **Table 4**.

Table 4. Cluster membership

Country	GNI per capita	Development cluster	Income group*	Income rank
Burundi	170	*C1*	LIC	1
Congo, Dem. Rep.	180	*C1*	LIC	2
Liberia	200	*C1*	LIC	3
Malawi	330	*C1*	LIC	4
Sierra Leone	340	*C1*	LIC	5
Niger	370	*C1*	LIC	6
Ethiopia	390	*C1*	LIC	7
Guinea	400	*C1*	LIC	8
Madagascar	430	*C1*	LIC	9
Mozambique	440	*C1*	LIC	10
Gambia, The	450	*C1*	LIC	11
Central African Republic	470	*C1*	LIC	12
Nepal	490	*C1*	LIC	13
Togo	490	*C1*	LIC	14
Uganda	500	*C1*	LIC	15
Rwanda	520	*C1*	LIC	16
Tanzania	540	*C1*	LIC	17
Burkina Faso	550	*C1*	LIC	18
Guinea-Bissau	590	*C1*	LIC	19
Mali	600	*C1*	LIC	20
Haiti	650	*C1*	LIC	22
Bangladesh	700	*C1*	LIC	23
Cambodia	750	*C1*	LIC	24
Comoros	750	*C1*	LIC	25
Benin	780	*C1*	LIC	26
Lao PDR	1,040	*C1*	LMIC	31
Pakistan	1,050	*C1*	LMIC	32
Zambia	1,070	*C1*	LMIC	33
Nigeria	1,230	*C1*	LMIC	40

India	1,270	*C1*	LMIC	42
Papua New Guinea	1,300	*C1*	LMIC	43
Chad	620	*C2*	LIC	21
Tajikistan	800	*C2*	LIC	27
Mauritania	1,000	*C2*	LMIC	30
Vietnam	1,160	*C2*	LMIC	37
Yemen, Rep.	1,170	*C2*	LMIC	38
Cameroon	1,200	*C2*	LMIC	39
Congo, Rep.	2,240	*C2*	LMIC	51
Swaziland	2,930	*C2*	LMIC	61
Angola	3,960	*C2*	LMIC	70
Kenya	810	*C3*	LIC	28
Kyrgyz Republic	830	*C3*	LIC	29
Senegal	1,080	*C3*	LMIC	34
Lesotho	1,090	*C3*	LMIC	35
Nicaragua	1,110	*C3*	LMIC	36
Ghana	1,250	*C3*	LMIC	41
Djibouti	1,300	*C3*	LMIC	44
Bolivia	1,810	*C3*	LMIC	45
Moldova	1,810	*C3*	LMIC	46
Bhutan	1,870	*C3*	LMIC	47
Honduras	1,870	*C3*	LMIC	48
Philippines	2,060	*C3*	LMIC	49
Sri Lanka	2,240	*C3*	LMIC	50
Indonesia	2,500	*C3*	LMIC	54
Georgia	2,690	*C3*	LMIC	55
Paraguay	2,720	*C3*	LMIC	56
Guatemala	2,740	*C3*	LMIC	57
Guyana	2,870	*C3*	LMIC	60
Ukraine	3,000	*C3*	LMIC	62
Armenia	3,200	*C3*	LMIC	63
Cape Verde	3,270	*C3*	LMIC	64
El Salvador	3,380	*C3*	LMIC	65
Ecuador	3,850	*C3*	UMIC	68
Albania	3,960	*C3*	UMIC	69
Thailand	4,150	*C3*	UMIC	71
Namibia	4,510	*C3*	UMIC	76
Macedonia, FYR	4,570	*C3*	UMIC	77
Peru	4,700	*C3*	UMIC	79
Dominican Republic	5,030	*C3*	UMIC	81
Colombia	5,510	*C3*	UMIC	83
Montenegro	6,740	*C3*	UMIC	88
Panama	6,970	*C3*	UMIC	91
Iraq	2,340	*C4*	LMIC	52
Egypt, Arab Rep.	2,420	*C4*	LMIC	53
Syrian Arab Republic	2,750	*C4*	LMIC	58
Morocco	2,850	*C4*	LMIC	59
Turkmenistan	3,790	*C4*	LMIC	66
Tunisia	4,160	*C4*	UMIC	72
China	4,270	*C4*	UMIC	73
Jordan	4,340	*C4*	UMIC	74
Algeria	4,390	*C4*	UMIC	75
Iran, Islamic Rep.	4,600	*C4*	UMIC	78
Azerbaijan	5,330	*C4*	UMIC	82
Belarus	5,950	*C4*	UMIC	85
Kazakhstan	7,580	*C4*	UMIC	92
Gabon	7,650	*C4*	UMIC	93
Venezuela, RB	11,590	*C4*	UMIC	101
Belize	3,810	*C5*	LMIC	67
Jamaica	4,800	*C5*	UMIC	80
Serbia	5,630	*C5*	UMIC	84
Suriname	6,000	*C5*	UMIC	86

South Africa	6,090	*C5*	UMIC	87
Botswana	6,740	*C5*	UMIC	89
Costa Rica	6,810	*C5*	UMIC	90
Malaysia	7,760	*C5*	UMIC	94
Argentina	8,620	*C5*	UMIC	95
Mexico	8,930	*C5*	UMIC	96
Brazil	9,390	*C5*	UMIC	97
Turkey	9,890	*C5*	UMIC	98
Chile	10,120	*C5*	UMIC	99
Uruguay	10,230	*C5*	UMIC	100

Before comparing the characteristics of the five clusters obtained in the analysis, it is convenient to distinguish which variables are more influential in discriminating these five groups of countries. This step is particularly important as cluster analysis sheds light on whether the groups of developing countries are statistically distinguishable (i.e. the clusters exhibit significantly different means in the development indicators). In order to verify if there are significant differences between clusters, we perform a one-way ANOVA analysis to calculate the cluster centroids and compare the differences formally. According to this analysis, the 11 variables include in the classification are statistically significant (**Table 5**). The size of the *F* statistics shows the relation between the overall between-cluster variation and the overall within-cluster variation and, therefore, it is a good indicator of the relevance of each variable for identifying groups of countries. According to this criterion, the variable with the greatest discriminating power is *poverty*, followed by *productivity* and *quality of democracy*. By contrast, the variables with lowest relative importance in the classification are *external finance*, *inequality* and *primary exports*.

Table 5. ANOVA output

		Sum of squares	Df.	Mean square	F	Sig.
Poverty	Between	73,653.18	4	18,413.30	98.39	0.00
	Within	17,965.57	96	187.14		
	Total	91,618.75	100			
Gini	Between	972.33	4	243.08	4.54	0.00
	Within	5,135.75	96	53.50		
	Total	6,108.08	100			
Malnutrition	Between	7,321.58	4	1,830.40	33.96	0.00
	Within	5,174.59	96	53.90		
	Total	12,496.17	100			
Non-agriculture GDP	Between	12,098.98	4	3,024.75	47.34	0.00
	Within	6,133.92	96	63.90		
	Total	18,232.90	100			
Primary exp	Between	7,714.92	4	1,928.73	11.01	0.00
	Within	16,818.07	96	175.19		
	Total	24,532.99	100			
Productivity	Between	4,304,000,000	4	1,076,000,000	78.33	0.00
	Within	1,319,000,000	96	13,737,304		
	Total	5,623,000,000	100			
Articles	Between	30,536.43	4	7,634.11	20.58	0.00
	Within	35,603.23	96	370.87		
	Total	66,139.66	100			
External finance	Between	4,250.50	4	1,062.63	4.28	0.00
	Within	23,834.78	96	248.28		
	Total	28,085.29	100			
WGI	Between	12.69	4	3.17	22.77	0.00
	Within	13.37	96	0.14		
	Total	26.06	100			
POLITY	Between	2,091.16	4	522.79	54.39	0.00
	Within	922.78	96	9.61		

	Total	3,013.94	100			
CO2pc	Between	304.75	4	76.19	25.95	0.00
	Within	281.86	96	2.94		
	Total	586.61	100			

In summary, the first cluster (***C1***) includes 31 countries (25 of them are LICs and the remaining six are LMIC); the second (***C2***) is composed of nine countries (two LICs and seven LMICs); the third (***C3***) includes 32 countries (two LICs, 20 LMICs and 10 UMICs); the forth (***C4***) has 15 countries (five LMICs and 10 UMICs); and the fifth (***C5***) includes 14 countries (one LMICs and 13 UMICs).[17] **Map 1** provides a simple representation of the development taxonomy derived from this cluster results. As it can clearly be seen in the map, the development cluster are scattered across the geographical regions, with the two least development groups (*C1* and *C2*) mainly located in sub-Saharan Africa and south-east Asia.

Therefore, ***C1*** includes the poorest countries (according to income per capita), followed by ***C2*** and ***C3***; whereas ***C4*** and ***C5*** include the countries with the highest incomes. However, our development taxonomy differs notably from the usual income classification. Thus the rank analysis between the variables *GNI per capita* and the cluster membership shows that both classifications have a limited level of coincidence, with a statistically significant Spearman coefficient equal to 0.49.

Notably many countries commonly labelled "emerging economies" are not in the emerging economies clusters because they retain characteristics of poorer countries. We find that there is no simple "linear" representation of development levels (from low to high development countries). We find that each development cluster has its own and characteristic development issues. There is no group of countries with the best (or worst) indicators in all development dimensions. It thus would be more appropriate to build "complex" development taxonomies on a five-year basis than ranking and grouping countries in terms of per capita incomes, as this will offer a more nuanced image of the diversity of challenges of the developing world and policy responses appropriate to different kinds of countries.

A more precise interpretation of the five clusters obtained in the analysis involves examining the cluster centroids (i.e. the clustering variables' average values of all countries in a certain cluster). This comparative procedure enables us to analyze the data on the basis of the grouping variable's values. According to **Table 6** the five development clusters can be described as follows:

- *Cluster 1*: ***High poverty rate countries with largely traditional economies***. These countries have the highest poverty and malnutrition headcounts; however, the income inequalities are less acute than in *C3* and *C5*. On average, the agricultural sector contributes one third to the GDP, although their exports of primary products are low. Moreover, they are the lowest productivity and innovation of the economies in the dataset. They have the second poorest governance indicators and the lowest CO2 per capita emissions. Many of these economies are highly dependent on external flows (mainly ODA).

- *Cluster 2*: ***Natural resource dependent countries with little political freedom****.* Those countries with severely constrained political freedoms, high dependency on natural resources and moderate inequality (relative to the average for all developing countries). These countries rank second (after *C1*) in terms of poverty, malnutrition, non-agricultural GDP, productivity, innovation capacities, and CO2 per capita emissions. However, the income inequalities are less acute than in *C1*, *C3* and *C5*.

[17] LIC, LMIC and UMIC World Bank country classifications as of 2011.

- *Cluster 3*: ***External flow dependent countries with high inequality****.* Those countries with high dependency on external flows and high inequality, and moderate poverty incidence (relative to the average for all developing countries). These countries rank third in terms of poverty, malnutrition, non-agricultural GDP, productivity, innovation capacities, and CO2 per capita emissions. However, they are the economies with the second highest Gini index (after *C5*), the lowest ratio of primary exports, the second highest external finance, the second best score in the governance indicators (although still below the world average) and the second best democracy indicator.

- *Cluster 4*: ***Economically egalitarian emerging economies with serious challenges of environmental sustainability and limited political freedoms.*** Those countries with most equal societies, with moderate poverty and malnutrition but serious challenges of environmental sustainability and political freedoms. These countries rank forth in terms of poverty, malnutrition, non-agricultural GDP, productivity, innovation capacities, and external finance. However, they have the second highest participation of primary products, the second worst governance indicators, the worst democracy indicator and they are the most polluting countries of the sample relative to population.

- *Cluster 5*: ***Unequal emerging economies with low dependence on external finance****.* Those countries with the highest inequality and the lowest dependency on external finance. These countries have the lowest poverty and malnutrition headcounts, and the highest non-agricultural GDP, labour productivity, innovation capacities, and governance and democracy indicators. They have the lowest dependency on external finance.

It is important to note, as in any development classification, there are countries that do not perfectly fit their assigned development groups. The most notable case in the above taxonomy is India, which is the biggest and the second 'richest' (in terms of per capita GNI) country of cluster *C1*[18]. In general terms, India is above the group's average in most of the development proxies: its Gini coefficient is considerably lower (34), GDP in non-agricultural sectors is 16 percentage points higher, exports of primary products are almost three times lower, scientific articles production is five times larger, external finance is almost four times lower, and governance and democracy indicators are better. However, India has 'poorer' indicators in terms of malnutrition (with a rate of 43.5%) and CO2 per capita emissions (which are five times greater than *C1*'s average). In short, *C1* is the 'most similar' group in relation to the 'atypical' development values of India.
Furthermore, it is worth noting that there are also important "development gaps" across the clusters, in terms of the 11 development indicators. A simple way to explore the magnitude of these gaps is comparing the deviations of each group of countries from the overall group of developing countries (i.e. the ratio between each cluster's centroids and the developing countries' average). Hence **Graph 3** shows that both *C1* and *C2* (the two groups that include those countries with the worst development indicators) have more acute problems of poverty and malnutrition than the average developing country –although they are close to the average in terms of inequality–. On the other hand, their labour productivity, innovation capacities, governance indicators and CO2 per capita emissions are well below the average. The most important differences between *C1* and *C2* are in terms of primary exports (much higher in *C2*¸ 2.2 times over the overall average), quality of democracy (higher in *C1*¸although still below the overall average) and dependency on external finance (higher in *C1*¸ 1.5 times over the overall average). *C3*, *C4* and *C5* are all below the developing countries' average in terms of poverty and malnutrition headcounts, and non-agricultural GDP. However, there are striking differences across these three clusters: *C5*'s innovation capacity exceeds 3.5 times the overall average; *C4*'s political freedoms are more limited than the average; *C4* and *C5* economies are more productive and polluting; primary exports are lower in *C3* and *C5*; and *C4* and *C5* are well below the average in terms of external finance.

[18] Therefore, India is one of the last countries in joining *C1* according to the agglomeration schedule (see Appendix 3).

Map 1. Taxonomy of the developing world by clusters

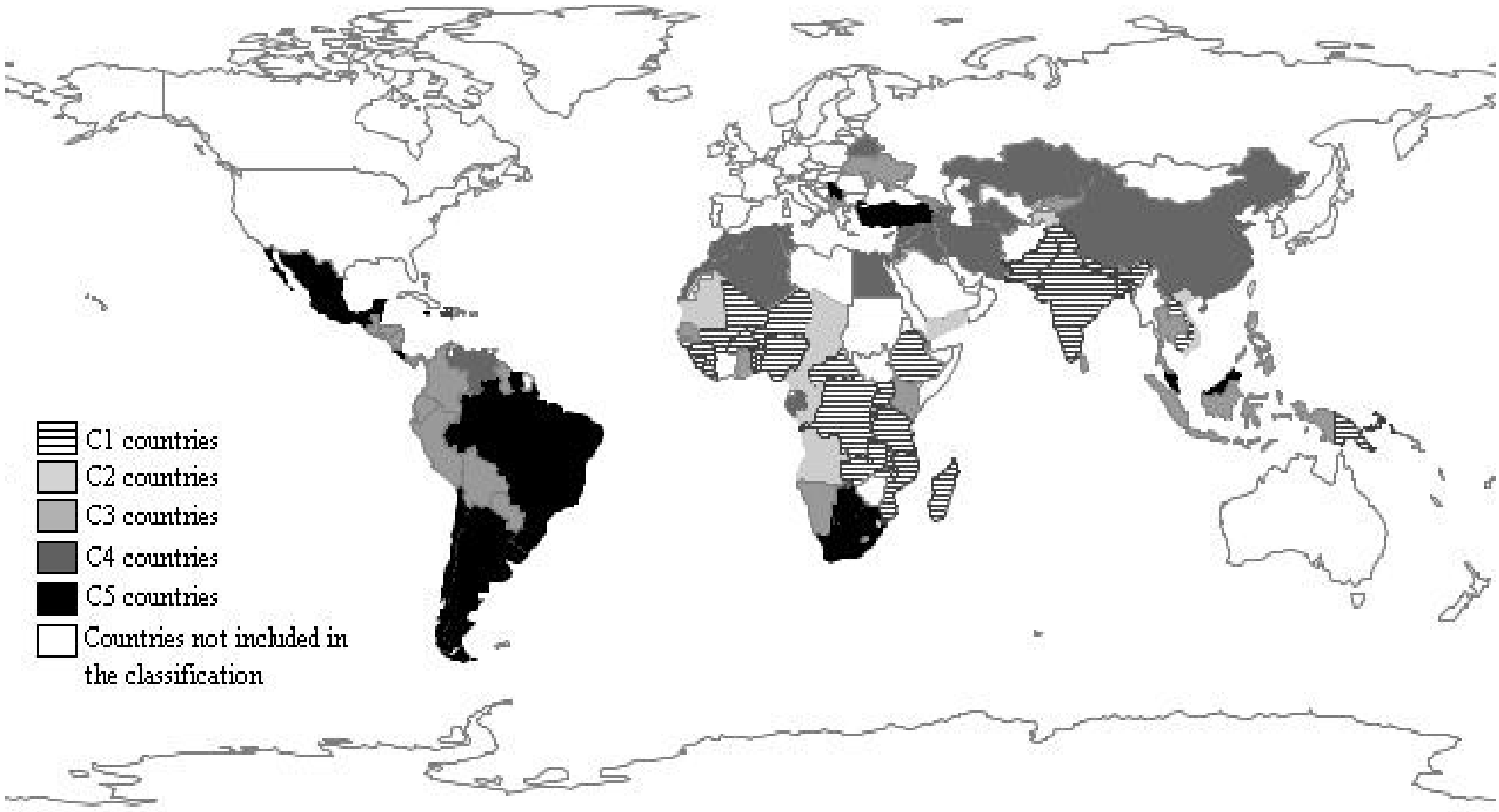
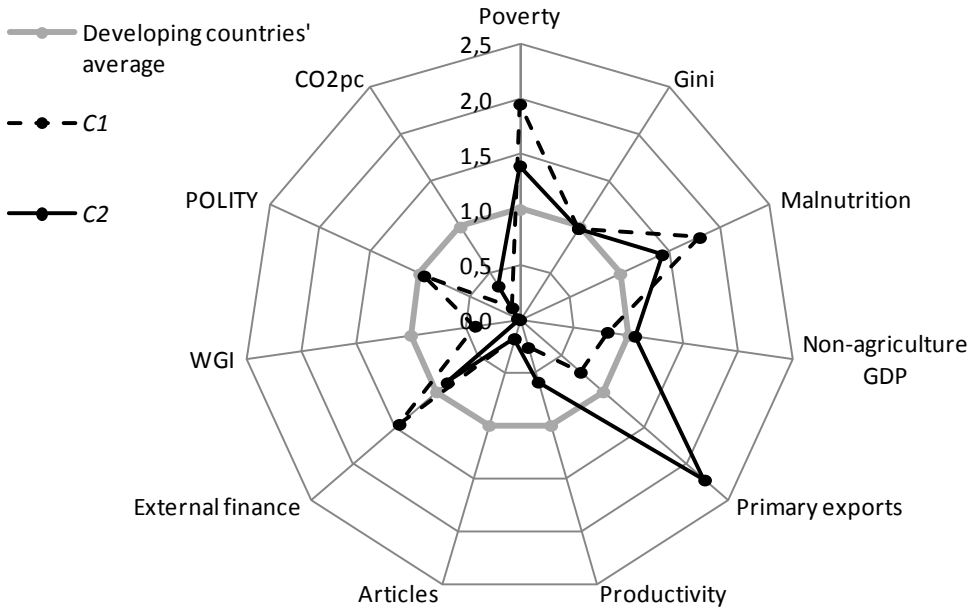


Table 6. Cluster centroids

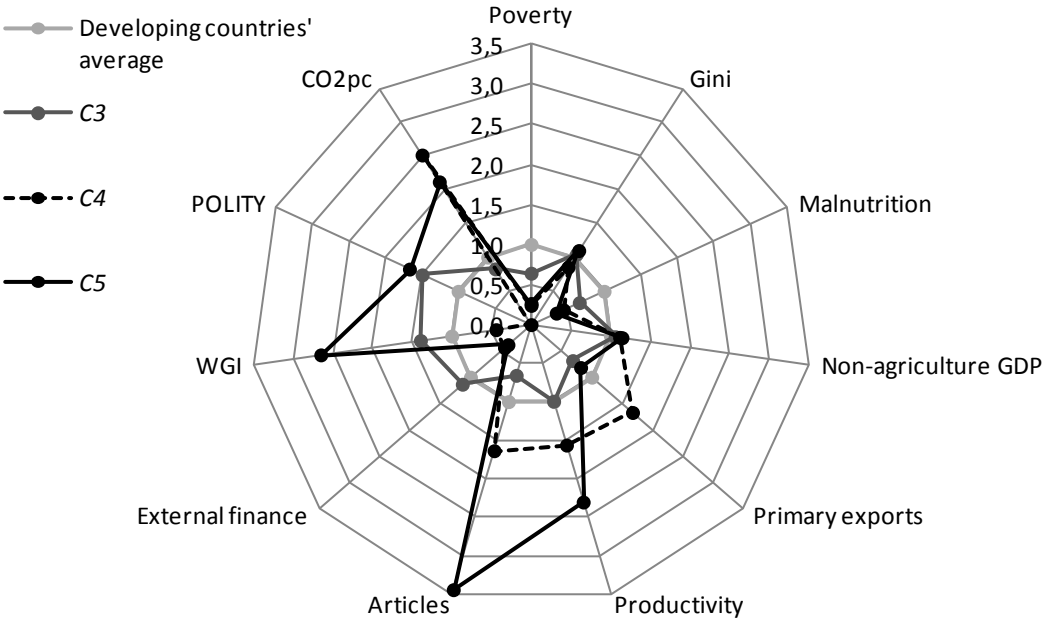
Development clusters		Poverty	Gini	Malnutrition	Non-agriculture GDP	Primary exports	Productivity	Articles	External finance	WGI	POLITY	CO2pc	*For reference*: GNIpc
C1	Mean	74.97	41.55	25.77	65.17	12.52	2,515.25	2.83	22.88	-0.77	3.06	0.25	614.19
	N	31	31	31	31	31	31	31	31	31	31	31	31
	Std. Desv.	11.60	6.97	8.54	11.20	12.36	1,537.94	3.27	22.29	0.37	4.19	0.28	314.96
C2	Mean	53.57	41.49	20.36	85.71	38.16	5,646.59	2.89	13.78	-0.95	-3.89	0.71	1,675.56
	N	9	9	9	9	9	9	9	9	9	9	9	9
	Std. Desv.	17.38	7.32	11.34	7.19	19.32	2,397.29	2.61	15.55	0.30	2.52	0.46	1,128.92
C3	Mean	24.58	44.20	9.48	85.98	11.76	9,512.29	10.49	17.91	-0.34	7.06	1.61	2,984.06
	N	32	32	32	32	32	32	32	32	32	32	32	32
	Std. Desv.	17.19	8.11	7.20	7.01	8.18	4,620.72	13.41	13.16	0.32	1.92	1.50	1,653.35
C4	Mean	9.19	35.96	6.36	90.50	28.74	14,978.55	26.09	6.93	-0.76	-4.07	4.91	4,934.00
	N	15	15	15	15	15	15	15	15	15	15	15	15
	Std. Desv.	8.17	6.41	3.35	4.55	20.57	4,597.42	24.32	8.63	0.45	3.92	3.21	2,466.06
C5	Mean	10.10	46.36	4.94	92.92	14.03	22,059.14	54.84	6.02	0.20	8.36	4.13	7,487.14
	N	14	14	14	14	14	14	14	14	14	14	14	14
	Std. Desv.	10.44	6.99	3.83	3.37	9.81	4,332.84	40.54	6.70	0.44	1.15	2.20	2,087.71
Developing countries' average	Mean	38.34	42.22	14.36	81.20	17.18	9,571.20	15.93	15.79	-0.51	3.39	1.95	3,053.86
	N	101	101	101	101	101	101	101	101	101	101	101	101
	Std. Desv.	30.27	7.82	11.18	13.50	15.66	7,498.51	25.72	16.76	0.51	5.49	2.42	2,774.27

Graph 3. Difference respect the average of C1 and C2



Note: Ratios between each cluster's centroids and the developing countries' average. *Governance* and *Polity* are previously re-scaled to avoid negative values.

Graph 4. Difference respect the average of C3, C4 and C5



Note: Ratios between each cluster's centroids and the developing countries' average. *Governance* and *Polity* are previously re-scaled to avoid negative values.

4.4. The global distribution of poverty

How is world poverty distributed by the five clusters? The distribution of global poverty by LICs and MICs is as follows: The proportion of the world's $1.25 and $2 poor accounted for by MICs is, respectively, 74% and 79% and the distribution of global poverty is thus (Sumner, 2012a; 2012b):

- Half of the world's poor live in India and China (mainly in India).
- A quarter of the world's poor live in other MICs (primarily populous LMICs, such as Pakistan, Nigeria and Indonesia).
- A quarter (or less) of the world's poor live in the remaining LICs.

How does the distribution by clusters compare? The clusters classification has important implications in terms of the developing world's population distribution (**Table 7**): Almost 41% of the developing countries' population is concentrated in *C1*, which includes some of the most populated countries of the world (India, Pakistan, Nigeria and Bangladesh); 35% is concentrated in *C4* (due to China), and the remaining 27% is scattered across *C3*, *C5* and – to a more limited extent– *C2*.

Table 7. Estimates of the distribution of global poverty, and poverty incidence, $1.25 and $2 (2008)

	$1.25 poverty line			$2 poverty line		
	Accumulated poor (millions)	**Participation in global poverty (%)**	**Poverty incidence (%)***	**Accumulated poor (millions)**	**Participation in global poverty (%)**	**Poverty incidence (%)***
East Asia and Pacific	265.4	21.5	14.3	614.3	26.1	33.2
Europe and Central Asia	2.1	0.2	0.5	9.9	0.4	2.4
Latin American and the Caribbean	35.3	2.9	6.9	67.4	2.9	13.1
Middle East and North Africa	8.5	0.7	2.7	43.8	1.9	13.9
South Asia	546.5	44.3	36.0	1,074.7	45.6	70.9
Sub-Saharan Africa	376.0	30.5	47.5	547.5	23.2	69.2
LICs	316.7	25.7	48.5	486.3	20.6	74.4
LMICs	711.6	57.7	30.2	1,394.5	59.2	59.1
LMICs minus India	285.6	23.1	23.4	569.4	24.2	46.7
UMICs	205.5	16.7	8.7	476.6	20.2	20.3
UMICs minus China	32.5	2.6	3.2	82.3	3.5	8.0
45 fragile states (OECD 2011)	412.3	33.4	40.3	684.0	29.0	66.9
Least developed countries	324.0	26.3	46.4	505.0	21.4	72.2
Quartile 1 (poorest GDP PPP pc)	454.6	36.8	45.6	680.8	28.9	68.3
C1	889.5	72.1	41.7	1,556.0	66.0	73.2
C2	35.7	2.9	22.8	80.6	3.4	47.4
C3	64.6	5.2	15.9	149.8	6.4	35.2
C4	41.7	3.4	10.7	84.8	3.6	25.7
C5	201.0	16.3	4.1	482.4	20.5	9.3
Developing countries, total	1,233.8	100.0	22.8	2,357.5	100.0	43.6

Source: Data processed from PovcalNet (2012) and World Bank (2012). The cutler classification (*C1* to *C5*) includes 72.7% of the developing countries, and 95.3% of the population of the developing world; * Population weighted averages.

In contrast, the distribution in terms of poverty is even more skewed: almost two thirds of the world's poor live in *C1* (the 'high poverty' countries) (but one should remember a third of world poverty is accounted for by India), 18% live in *C4* (the group with overall good development indicators but bad governance), 10.6% live in *C3*, and the remaining 5.5% live in *C2* and *C5*. All in all, the participations of *C1* and *C2* in poverty are larger than their participations in population, due to their higher incidence of poverty.

The above eludes to the fact that world's poor are heavily concentrated. 80% of world poverty is in 10 countries and 90% of world poverty is in 20 countries (Sumner, 2012a; 2012b).

Table 8 shows the position in our taxonomy of the 20 countries that account for 90% of world poverty. In C1 are India, Nigeria, Bangladesh, DRC, Pakistan and Tanzania, Malawi, Nepal, Uganda, Madagascar, Mozambique and Ethiopia. In C2 are Vietnam and Angola. In C3 are Indonesia, Kenya and the Philippines. In C4 is China. In C5 is Brazil. In terms of global poverty this suggest much of the issues of C1 should be considered. However, substantial numbers of the world's poor are in quite different countries.

Table 8. Top 20 poor countries (by number of $1.25/day poor people), and country classifications by GNI per capita and by our taxonomy

	% World $1.25 Poor (2008)	% World $2 Poor (2008)	Country classification (based on data for calendar year) (2009)	Our taxonomy
1. India	34.5	35.0	LMIC	C1
2. China	14.0	16.7	UMIC	C4
3. Nigeria	8.1	5.4	LMIC	C1
4. Bangladesh	6.0	5.3	LIC	C1
5. Congo, Dem. Rep.	4.5	2.6	LIC	C1
6. Indonesia	4.2	5.2	LMIC	C3
7. Pakistan*	2.3	5.2	LMIC	C1
8. Tanzania	1.4	1.6	LIC	C1
9. Philippines	1.3	1.6	LMIC	C3
10. Kenya	1.2	1.1	LIC	C3
11. Vietnam	1.1	1.6	LMIC	C2
12. Uganda	1.1	0.9	LIC	C1
13. Madagascar	1.1	0.7	LIC	C1
14. Mozambique	1.0	0.8	LIC	C1
15. Ethiopia*	0.9	1.8	LIC	C1
16. Brazil	0.8	0.9	UMIC	C5
17. Angola	0.8	0.5	LMIC	C2
18. Malawi	0.8	0.6	LIC	C1
19. Nepal	0.8	0.8	LIC	C1
20. Sudan*	0.7	0.8	LMIC	Not included
Top 10	79.2	79.5		
Top 20	86.6	89.1		

Source: Data processed from PovcalNet (2012) and WDI (2011). Note: * = The poverty data listed in PovcalNet (2012) for these countries in 2008 appears lower than one might expect suggesting caution (see also discussion in Sumner, 2012b, and for rates by national poverty lines see Gentilini and Sumner (2012).

5. CONCLUSIONS

This paper has used a cluster analysis to identify five types of developing country using a set of indicators covering definitions of development based on the history of thinking about 'development' over the last 50 years across: 'development as structural transformation'; 'development as human development'; 'development as democratic participation and good governance'; and 'development as sustainability'. We find that there are five types of developing country as follows:

Cluster 1: ***High poverty rate countries with largely traditional economies.*** Those countries with the highest poverty and malnutrition headcounts, the lowest indicators for productivity and innovation and mainly agricultural economies, with severely constrained political freedoms and high dependence on external flows (primarily ODA).

Cluster 2: ***Natural resource dependent countries with little political freedom****.* Those countries with severely constrained political freedoms, high dependency on natural resources and moderate inequality (relative to the average for all developing countries).

Cluster 3: ***External flow dependent countries with high inequality****.* Those countries with high dependency on external flows and high inequality, and moderate poverty incidence (relative to the average for all developing countries).

Cluster 4: ***Economically egalitarian emerging economies with serious challenges of environmental sustainability and limited political freedoms.*** Those countries with most equal societies, with moderate poverty and malnutrition but serious challenges of environmental sustainability and political freedoms.

Cluster 5: ***Unequal emerging economies with low dependence on external finance.*** Those countries with the highest inequality and the lowest dependency on external finance.

Furthermore, it is worth noting that this development taxonomy differs notably from the usual income classification. The rank analysis between the variables *GNI per capita* and the cluster membership shows that both classifications have a limited level of commonality (less than 50%). This supports the case for considering the multidimensionality of development when building an international taxonomy.

Such "development taxonomies" are useful because they help us to identify relatively homogeneous groups of countries that share similar development characteristics and are useful for guiding international development assistance. However, building a development classification is not a simple task: once we overcome the over-simplistic income-based classification of the developing world, we find that there is no simple "linear" representation of development levels (from low to high development countries). We find that each development cluster has its own and characteristic development issues. There is no group of countries with the best (or worst) indicators in all development dimensions. It thus maybe more appropriate to build "complex" development taxonomies on a five-year basis than ranking and grouping countries in terms of per capita incomes, as this will offer a more nuanced image of the diversity of challenges of the developing world.

Given its multidimensional nature, the analysis carried out in this paper seeks to provide input in to thinking about post-2015 debates on approaching thinking about goal setting for different types of country. In this sense, the identification of relatively homogeneous groups of countries in terms of development issues can encourage 'dynamics of peer-progress' between countries of the same group, allowing them to collectively identify specific

development strategies for the group, and therefore going beyond the ‘one-size-fits-all’’ approach that the current MDG agenda has been perceived to be (Tezanos, 2011).

As Seers (1972: 32) noted: “The most important use of development indicators is to provide the targets for planning”. Therefore, if it is possible to identify reasonable homogeneous development groups, it will be easier to provide appropriate development goals for each group and thus design a more tailored post-2015 agenda.

References

Alkire, S.; Roche, J.; Santos, E. and Seth, S. 2011. Multidimensional Poverty Index 2011, Oxford: OPHI

Alkire, S. 2002. Valuing Freedoms. Oxford: OUP.

Baster, N. 1979. 'Models and Indicators', in Cole, S. and Lucas, H. (eds.), Models, Planning and Basic Needs. Pergamon: Oxford.

Beckerman W. 1994. Sustainable Development': Is it a Useful Concept? Environmental Values. **3**:191-209

Beckerman W. 1995. How would you like your 'sustainability', sir? Weak or strong? A reply to my critics. Environmental Values. **4:** 169-179.

Chandy, L. and Gertz, G. (2011) Poverty in Numbers: The Changing State of Global Poverty from 2005 to 2015, Policy Brief 2011-01, Washington, DC: Global Economy and Development at Brookings, The Brookings Institution

Common M. 1996. Beckerman and his critics on strong and weak sustainability: confusing concepts and conditions. Environmental Values. **5:** 83-88.

Calinski, T. and Harabasz, J. (1974) "A dendrite method for cluster analysis", Communications in Statistics - Theory and Methods, 3(1): 1–27.

Daly H. 1995. On Wilfred Beckerman's critique of sustainable development. Environmental Values. **4**: 49-55.

Daly H. 1996. Beyond Growth. Beacon Press: Boston MA.

Daly H. 2002. Sustainable Development: Definitions, Principles and Policies. Invited address, April 30, World Bank: Washington DC.

Deaton, A. (2011) 'Measuring Development: Different Data, Different Conclusions', paper presented at the 8th AFD-EUDN Conference, Paris

Deaton, A. (2010) 'Price Indexes, Inequality, and the Measurement of World Poverty', American Economic Review 100.1, 5–34

Deaton, A. and Heston, A. (2010) 'Understanding PPPs and PPP-based National Accounts', American Economic Journal 2.4, 1–35

Development Assistance Committee (DAC) (2011) DAC List of ODA Recipients, available at http://www.oecd.org/dac/stats/daclist

Development Assistance Committee (DAC) (2012) Development database on aid form DAC members: DAC Online, OECD.Stat, available at http://www.oecd.org/dataoecd

Dower N. 1994. Worth Sustaining? Environmental Values **3**:159-160.

Easterly W, Rebelo S. (1993). Fiscal policy and economic growth: An empirical investigation.

El Serafy S. 1996. In defence of weak sustainability: a response to Beckerman. Environmental Values. **5:** 75-81.

Everitt, B.S., Landau, S., Leese, M. and Stahl, D. (2011) Cluster analysis, John Wiley & Sons, Chichester.

Fischer, A.M. (2010) 'Towards Genuine Universalism within Contemporary Development Policy', IDS Bulletin 41, 36–44

Glassman, A. Duran, D. and Sumner, A. (2011) Global Health and the New Bottom Billion: What Do Shifts in Global Poverty and the Global Disease Burden Mean for GAVI and the Global Fund?, CGD Working Paper, Washington, DC: Center for Global Development

Gore, C. 2000. The rise and fall of the Washington consensus as a paradigm for. developing countries. World Development, 28 (5), 789-804.

Graham, J., Amos, B. and Plumptre, T. (2003) Principles of Good Governance in the 21st Century, Ottawa: Institute on Governance.

Guillaumont, P. (2009) Caught in a Trap: Identifying the Least Developed Countries. Economica: Paris.

Hartwick J. (1977). Intergenerational equity and the investing of rents of exhaustible resources. American Economic Review. 67(5): 972-974

Hands, S. and Everitt, B.S. (1987) "A Monte Carlo study of the recovery of cluster structure in binary data by hierarchical clustering techniques", Multivariate Behavioral Research, 22: 235–243.

Heston, A., Summers, R. and Aten, B. (2011) Penn World Table, Version 7.0, Center for International Comparisons of Production, Income and Prices, University of Pennsylvania, available at http://pwt.econ.upenn.edu/php_site/pwt_index.php

Hicks J. 1939. Value and Capital. Clarendon Press: Oxford.

Hicks N, Streeten P. 1979. Indicators of Development: The Search for a Basic Needs Yardstick. World Development. **7**: 567–580.

Hyden, G., Court, Julius and Mease, K. 2004. Making Sense Of Governance. ODI: London.

ILO (International Labour Organization). 1976. Employment, Growth and Basic Needs: A One-World Problem. Geneva: ILO.

ILO 1977. Meeting Basic Needs: Strategies for Eradicating Mass Poverty and Unemployment. Geneva: ILO.

International Monetary Fund (IMF) (2011) World Economic Outlook Database, available at http://www.imf.org/external/pubs/ft/weo/2011/02/weodata/index.aspx

Kanbur, R. and Sumner, A. (2011a) Poor Countries or Poor People? Development Assistance and the New Geography of Global Poverty, Working Paper 2011-08, Ithaca, NY: Charles H. Dyson School of Applied Economics and Management, Cornell University

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2011) The Worldwide Governance Indicators (WGI) project, available at http://info.worldbank.org/governance/wgi/index.asp

Lewis A. 1954. Economic development with unlimited supplies of labour. The Manchester School. **22**: 3-42.

Lewis A. 1955. The Theory of Economic Growth. Irwin: Homewood IL.

Marshall, M.G. and Jaggers, K. (2011) Polity IV Project: political regime characteristics and transitions, Center for Global Policy, School of Public Policy, George Manson University, available at http://www.systemicpeace.org/polity/polity4.htm

McGranahan D, Pizarro E, Richard C. 1985. Measurement and Analysis of Socio-Economic Development: An Enquiry into International Indicators of Development and Quantitative Interrelations of Social and Economic Components of Development. United Nations Research Institute for Social Development (UNRISD): Geneva.

Milligan G.W. and Cooper, M. (1985) "An examination of procedures for determining the number of clusters in a data set", Psychometrika, 50(2): 159–179.

Mooi, E. and Sarstedt, M. (2011) A concise guide to market research, chapter 9: "Cluster analysis", Springer-Verlag, Berlin.

Neumayer E. (2003). Sustainability and well-being indicators. Paper presented at United Nations University/WIDER Project meeting on 'Social development indicators'. Helsinki. 29 May.

Nielsen, L. (2011) "Classifications of Countries Based on Their Level of Development: How it is Done and How it Could be Done", IMF Working Paper, WP11/31. International Monetary Fund.

Nussbaum, M. 2000. Women and Human Development: the Capabilities Approach. Cambridge University Press: Cambridge.

Pearce D, Markandya A, Barbier E. (1989). Blueprint for a Green Economy. Earthscan: London.

Peña, D. (2002) Análisis de datos multivariantes, McGraw-Hill, Madrid.

Seers, D. 1969. The Meaning of Development. International Development Review, 11: 2-6.

Seers, D. 1963. The Limitations of the Special Case. Bulletin of the Oxford Institute of Economics and Statistics. 25 (2):77–98; reprinted in Martin, K. and Knapp, J. (eds.). 1967. The Teaching of Development Economics: Its Position in the Present State of Knowledge. Frank Cass: London: 1-27.

Seers, D. 1972. What are we Trying to Measure? Journal of Development Studies 8 (3):21-36; reprinted in Baster, N. (ed). 1972. Measuring Development: The Role and Adequacy of Development Indicators. Frank Cass: London:21-36.

Sen, A. 1999. Development as Freedom. Oxford University Press: Oxford.

Sneath, P.H.A. and Sokal, R.R. (1973) Numerical Taxonomy. W.H. Freeman, San Francisco.

Solow R. 1974 Intergenerational equity and exhaustible resources. Review of Economic Studies. 29-46.

Solow R. 1986. On the intertemporal allocation of natural resources. Scandinavian Journal of Economics. 8: 141-149.

Solt, F. (2009) "Standardizing the World Income Inequality Database", Social Science Quarterly, 90(2): 231-242. SWIID Version 3.1, December 2011, available at http://www.siuc.edu/~fsolt/swiid/swiid.html

Streeten, P. 1984. Basic Needs: Some Unsettled Questions. World Development, 12 (9):973-80.

Sumner, A. 2004. Economic Well-being and Non-economic Well-being: A Review of the Meaning and Measurement of Poverty. WIDER Research Paper. UNU WIDER: Helsinki.

Sumner, A. (2012a) 'Where do the Poor Live?', World Development 40.5, 865-877

Sumner, A. (2012b) 'From Deprivation to Distribution: Is Global Poverty Becoming A Matter of National Inequality?, IDS Working Paper, Brighton: IDS

Sumner, A. and Tribe, M. (2008) International Development Studies. Sage, London.

Tezanos, S. and Quiñones (2012): "¿Países de renta media? Una taxonomía alternativa del desarrollo de América Latina y el Caribe", Revista Iberoamericana de Estudios de Desarrollo, 2 (forthcoming on September 2012).

Tezanos, S. (2011): "Más allá de 2015: Objetivos de Desarrollo del Milenio y desafíos para la nueva agenda internacional de desarrollo", SISTEMA, Revista de Ciencias Sociales, 220: 33-58.

United Nations Conference on Trade and Development (UNCTAD) (2012) UNCTADstat, available at http://unctadstat.unctad.org/ReportFolders/reportFolders.aspx

UNDP (United Nations Development Programme). 1990-2011. Human Development Report. Oxford University Press for the UNDP: New York.

United Nations Development Programme (UNDP) (2011) Human Development Report 2011. Sustainability and Equity: A Better Future for All, Palgrave Macmillan, New York.

UNRISD (United Nations Research Institute on Social Development) 1970. Contents and Measurement of Socioeconomic Development. UNRISD: Geneva.

Ward, J.H. (1963) "Hierarchical groupings to optimize an objective function", Journal of the American Statistical Association, 58: 236–244.

World Bank (2012a) How we classify countries: a short history, available at http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:204870

70~menuPK:64133156~pagePK:64133150~piPK:64133175~theSitePK:239419~isCURL:Y,00.html

World Bank (2012b) World dataBank, available at http://data.worldbank.org

World Bank (2012c) PovcalNet, available at http://iresearch.worldbank.org/PovcalNet/index.htm

APPENDIX 1. Descriptive statistics of the dataset

	N	Minimum	Maximum	Mean	Std. deviation
Poverty	111	0.13	94.64	37.33	29.87
Gini	112	18.62	66.64	41.92	7.81
Malnutrition	124	0.55	45.30	14.60	11.51
GDPpc	132	291.57	20,554.87	5,042.88	4,339.42
Non-agriculture	131	40.39	98.28	82.16	13.11
Primary	133	0.13	75.95	15.88	15.73
Productivity	123	418.97	49,436.06	10,019.24	8,495.45
Articles	134	0.04	126.00	15.63	24.82
External	135	-7.53	445.12	21.06	40.48
WGI	136	-2.35	1.16	-0.49	0.63
POLITY	117	-9.00	10.00	2.92	5.85
CO2pc	136	0.02	13.51	2.10	2.50
***For reference*:**					
GNIpc	133	170.00	13,280.00	3,387.14	3,065.83
Population	138	0.01	1,338.30	40.51	156.00
Valid *N*	101				

APPENDIX 2. Cluster method

Given the type of data used in this cluster analysis (11 continuous variables), 3 possible clustering algorithms are the nearest neighbour method, the farthest neighbour method and the Ward's method (Everitt *et al*., 2011, Peña, 2002; Mooi and Sarstedt, 2011). Since there is no objective criterion for selecting the most appropriate method, the selection depends largely on the interpretability of the final results (Sneath and Sokal, 1973, Everitt *et al*., 2011, Peña, 2002; Mooi and Sarstedt, 2011).

In our analysis we use the method proposed by Ward (1963), in which the fusion of two clusters is based on the size of an error sum-of-squares criterion. The objective at each stage is to minimize the increase in the total within-cluster error sum of squares. Specifically, the Ward's method begins by calculating, for each cluster, the means for all variables. Then, for each country, the squared Euclidean distance to the cluster means is calculated. These distances are summed for all of the cases. At each step, the two clusters that merge are those that result in the smallest increase in the overall sum of the squared within-cluster distances. The coefficient in the agglomeration schedule is the within-cluster sum of squares at that step, not the distance at which clusters are joined. In practical terms, the Ward's method has been proven to be especially suitable for building clusters with similar sizes, when no outliers are present (Hands and Everitt, 1987; Everitt *et al*., 2011, Peña, 2002; Mooi and Sarstedt, 2011).

Finally, it should be mentioned that the cluster solutions of our analysis are reasonably "robust". As recommended by Mooi and Sarstedt (2011) we verify the robustness of the cluster analysis by means of the following 3-step check: firstly, we evaluate the stability of the results by using different clustering procedures, distance measures and standardization methods on the same data and we test whether these yield similar development taxonomies. However, it should be bear in mind that –as noted, among many others, by Everitt et al. (2011), Peña (2002) and Mooi y Sarstedt (2011)–, it is common for results to change even when the cluster solution is adequate, so some degree of variation is expected when changing the cluster procedure. Secondly, we change the order of the countries in our dataset and re-run the analysis to check the results' stability. Obviously, the results should not depend on the order of the dataset; otherwise, there may be outliers that influence the results of the change in order. And thirdly, we replace one of the variables by a new variable that we did not use in our benchmark analysis.

The first check shows moderate variations in the results. In particular: *i*) changing the clustering procedure, from Ward to the single linkage (nearest neighbour), only affects 17 out of the 101 countries (all of them are changes to the nearest cluster in terms of development); *ii*) changing the distance measure, from square Euclidian distance to Chebychev distance, only renders 14 differently classified countries; and *iii*) changing the standardization method, from range -1 to 1 to the simple *z* standardization, only renders 14 differently classified countries.

The second check shows no variation in the results: for example, changing the order of the countries in the dataset (from alphabetical order, to an increasing order of GNI per capita) does not affect the classification.

Finally, we check the implications of dropping the variable *GDPpc* instead of *productivity* (as explained in section 4.2.). We re-run the analysis using the per capita income and the results only differs in 5 countries.

APPENDIX 3. Agglomeration schedule

Stage	Cluster combined		Coefficients	Stage cluster first		Next stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	59	67	0.01	0	0	31
2	24	77	0.032	0	0	17
3	30	33	0.055	0	0	13
4	10	61	0.085	0	0	31
5	89	95	0.116	0	0	16
6	44	70	0.147	0	0	27
7	57	65	0.181	0	0	37
8	15	58	0.215	0	0	65
9	12	76	0.249	0	0	27
10	46	80	0.287	0	0	12
11	69	71	0.325	0	0	50
12	46	78	0.364	10	0	54
13	30	74	0.404	3	0	17
14	50	92	0.445	0	0	78
15	41	82	0.493	0	0	68
16	79	89	0.543	0	5	62
17	24	30	0.598	2	13	38
18	32	66	0.653	0	0	32
19	28	49	0.713	0	0	30
20	7	73	0.773	0	0	45
21	5	37	0.835	0	0	64
22	62	99	0.897	0	0	47
23	40	72	0.96	0	0	48
24	20	91	1.027	0	0	60
25	4	14	1.095	0	0	41
26	29	84	1.165	0	0	76
27	12	44	1.236	9	6	63
28	11	42	1.308	0	0	63
29	38	52	1.382	0	0	54
30	9	28	1.458	0	19	44
31	10	59	1.536	4	1	80
32	32	87	1.614	18	0	56
33	17	34	1.694	0	0	66
34	53	64	1.778	0	0	42
35	31	90	1.863	0	0	67
36	3	27	1.95	0	0	89
37	57	96	2.038	7	0	79
38	24	39	2.131	17	0	67
39	18	88	2.225	0	0	47
40	2	35	2.321	0	0	59
41	4	97	2.418	25	0	73
42	1	53	2.519	0	34	64
43	19	55	2.628	0	0	71
44	9	13	2.737	30	0	58
45	7	45	2.848	20	0	74
46	8	47	2.96	0	0	53
47	18	62	3.075	39	22	72
48	26	40	3.19	0	23	69
49	60	85	3.308	0	0	70
50	16	69	3.426	0	11	65
51	36	43	3.547	0	0	62
52	81	93	3.671	0	0	83
53	8	98	3.803	46	0	82
54	38	46	3.936	29	12	76
55	75	101	4.072	0	0	69
56	23	32	4.211	0	32	78
57	21	100	4.35	0	0	87

58	9	63	4.491	44	0	70
59	2	48	4.635	40	0	91
60	20	54	4.783	24	0	66
61	51	94	4.942	0	0	75
62	36	79	5.102	51	16	84
63	11	12	5.272	28	27	85
64	1	5	5.452	42	21	79
65	15	16	5.632	8	50	74
66	17	20	5.816	33	60	84
67	24	31	6.001	38	35	85
68	25	41	6.203	0	15	86
69	26	75	6.454	48	55	80
70	9	60	6.733	58	49	92
71	19	68	7.02	43	0	81
72	18	86	7.32	47	0	87
73	4	22	7.621	41	0	77
74	7	15	7.935	45	65	93
75	6	51	8.251	0	61	82
76	29	38	8.569	26	54	81
77	4	83	8.92	73	0	83
78	23	50	9.282	56	14	95
79	1	57	9.676	64	37	88
80	10	26	10.074	31	69	86
81	19	29	10.489	71	76	96
82	6	8	10.916	75	53	91
83	4	81	11.357	77	52	92
84	17	36	11.831	66	62	93
85	11	24	12.333	63	67	88
86	10	25	12.907	80	68	90
87	18	21	13.509	72	57	89
88	1	11	14.309	79	85	96
89	3	18	15.11	36	87	97
90	10	56	15.996	86	0	94
91	2	6	16.967	59	82	95
92	4	9	18.044	83	70	98
93	7	17	19.152	74	84	94
94	7	10	20.438	93	90	100
95	2	23	21.918	91	78	97
96	1	19	23.551	88	81	98
97	2	3	26.969	95	89	99
98	1	4	31.336	96	92	99
99	1	2	39.982	98	97	100
100	1	7	57.909	99	94	0