Human Development Research Paper 2010/33

Review of HDI Critiques and Potential Improvements

Milorad Kovacevic

United Nations Development Programme
Human Development Reports
Research Paper

February 2011

Human Development Research Paper 2010/33

Review of HDI Critiques and Potential Improvements

Milorad Kovacevic

United Nations Development Programme
Human Development Reports
Research Paper 2010/33
February 2011

Review of HDI Critiques and Potential Improvements

Milorad Kovacevic

Milorad Kovacevic is Head of Statistics at the Human Development Report Office of the United Nations Development Programme. E-mail: milorad.kovacevic@undp.org

Acknowledgment**:** The paper has benefitted from discussions and input from other members of the Statistical and Research Team, most notably, Alison Kennedy and Ricardo Fuentes. The careful reading and subsequent comments by Jeni Klugman and Clara Garcia Aguna have tremendously improved the content and readability.

Comments should be addressed by email to the author(s).

Abstract

This paper reviews critiques of the Human Development Report over the last twenty years. The critiques are mostly related to the choice of indicators, to high correlation of HDI components, functional form of the HDI including normalization of component indicators, aggregation vs. multiplication, and issues related to weighting. A thorough analysis of the critiques leads us to a proposal for modification of the HDI.

Keywords: Composite index, well-being, marginal substitutability, equal weighting, aggregation.

JEL classification: C43, O15

1. INTRODUCTION

The Human Development Index (HDI), first introduced in the 1990 Human Development Report (UNDP: 1990), was in response to the need for a measure that could better represent human achievements in several basic capabilities (what people can do and be) than income based indices of growth and development and could provide a credible alternative to them (Kelly, 1991, Anand and Sen, 1994). The consensus at that time was that the multi-dimensional character of human development was neglected in the typical measures of economic development, mostly GDP and GNP.

At the onset, the HDI followed six basic principles as guidelines (Ul Haq, 1995), that is to (i) measure the basic purpose of human development - to enlarge people's choices; (ii) include a limited number of variables to keep it simple and manageable;(iii) be composite rather than a plethora of separate indices; (iv) cover both social and economic choices; (v) be sufficiently flexible in both coverage and methodology to allow gradual refinements, once better alternatives became available; (vi) not be inhibited by lack of reliable and up-to-date data series.

The HDI was conceived to cover achievements in three basic dimensions - longevity, education and living standard. To capture these three dimensions, the 2009 HDI (as in 2009 HDR)[1] employs four indicators: life expectancy at birth; adult literacy rate; combined gross enrolment for primary, secondary, and tertiary education; and GDP per capita in US$ adjusted by Purchasing Power Parity.

The HDI has never made a claim to be a comprehensive measure of human development or well-being – but rather a summary alternative to economic measures. "The concept of human development is broader than any measure of human development. Thus although the HDI is a constantly evolving measure, it will never perfectly capture human development in its full sense" (UNDP: 1993, page 104). Anand and Sen (1994:12) admit that the Human Development Index "has been concerned only with the enhancement of very basic capabilities of people". They also recognize that for industrialized and developed countries, with high income and high literacy, the only differences are due to small variation in life expectancy, and in order to capture the differences between "high and similar levels of achievements of basic capabilities, it becomes relevant to assess performance using more refined capabilities" (Anand and Sen 1994:13).

The paper is structured as follows. Section 2 provides a brief history of critiques and evolution of the HDI. Section 3 reviews the critique of the choice of indicators. Section 4 examines another set of critiques related to statistical relationship between the component indicators and HDI, section 5 explores the issues surrounding critiques of aggregation used for computation of the HDI including the choice of functional form and weighting. In Section 6 we summarize the critiques and propose a new index. The numerical examples and analysis is based on data presented in 2009 HDR. (UNDP: 2009).

This paper was initially intended to also cover issues of missing dimensions in the current form of the HDI, such as political freedom, sustainability, security and vulnerability. However, we will not discuss these issues in-depth, so to better focus on the current three-dimensional HDI. A review of critiques related to the HDI's silence towards distributional inequality is addressed by Kovacevic (2010) and will not be summarized here. Critical issues of uncertainty, sensitivity and validity of the HDI, which generally receive minimal attention, are the subject of another paper (Garcia Aguna and Kovacevic, 2010).

[1] For changes in the content of the HDI please see Table 1.

2. BRIEF HISTORY OF CRITIQUES AND EVOLUTION OF HDI

Ever since the HDI was first published, it has drawn critiques from many sides. Some critiques claim that it uses the wrong variables, and that it is not reflecting the human development idea accurately (Lind, 1992, Dasgupta and Weale, 1992, Srinivasan , 1994, Sagar and Najam, 1998, Chibber and Laajaj, 2007). Sagar and Najam (1998) claimed that “the HDI presents a distorted picture of the world”. Others argued that the HDI depicts an oversimplified view of human development by relying on only a few indicators often derived from data of low quality (Murray, 1993, Srinivasan, 1994).

In response to critiques of this kind, the UNDP developed additional complementary tools such as the Human Poverty Index, the Gender-Related Development Index, and the Gender Empowerment Measure. However, although these indices complement the HDI’s explanatory power, they have not been widely used.

A possible alternative to having separate indices would be to incorporate more dimensions into the HDI itself. However, others argue that the current composition of the HDI is stable and that adding new dimensions may affect such stability. For example, Streeten, 1994, voices concerns against the addition of political and human rights variables to the HDI on the grounds of differential objectivity with which these dimensions are measured, differential volatility, and the importance of each concept in its own.

One long standing stream of critique has been that a pure economic indicator (GDP, GNI or similar) focused on economic growth alone is sufficient for any serious discussion regarding human development, especially given the intrinsic high correlation of all the HDI components among themselves as well as with the HDI, (McGillivray, 1991, Srinivansan, 1994a, Ravallion, 1997). Ogwang (1994) argues that the HDI doesn’t reveal anything beyond that portrayed by the GDP or by the life expectancy alone. Sugden (1993) questions if there is an alternative to “measurements of real income, and the kind of practical cost-benefit analysis which is grounded in Marshallian consumer theory“ that can be used to assess “the rich array of functionings that Sen takes to be relevant.”

There are also critiques of statistical quality and methodological soundness of the HDI, which highlight the measurement errors and biases inherent in the international data (Srinivansan, 1994b), violation of the evidence-based character of the HDI by increasing reliance on mathematical interpolations, imputations, and modeling (Lind, 1992, Ogwang, 1994), as well as the arbitrariness of the choice of aggregation and weighting strategy (Kelley, 1991, Srinivasan, 1994a, Ravallion, 1997).

Other critics (Diener and Suh, 1997) prefer a multidimensional “dashboard” approach to a scalar composite index as a way to avoid an arbitrary choice of the functional form and an arbitrary weighting scheme, but also not to lose information unnecessarily due to aggregation across indicators. Although, the use of multiple indicators allows the researcher to observe an object of interest from multiple angles, it does not allow a parsimonious understanding of the phenomenon under consideration. Indeed, one can find arguments for both – for composite indices and against them. In Saisana et al. (2005), the arguments in favour of composite indices are listed as follows - Composite indices can be used to summarise complex or multi-dimensional issues, in view of supporting decision-makers; they can provide the big picture; they facilitate the task of ranking countries on complex issues; composite indices can help attracting public interest, and they could help to reduce the size of a list of indicators. As arguments against the composite indices, Saisana et al. state that “composite indices may send misleading, non-robust policy messages if they are poorly constructed or misinterpreted or may invite politicians to draw simplistic policy conclusions." Also, the construction of composite indicators involves stages where judgements are necessary such as the selection of component indicators, choice of functional model, weighting, etc. The perceived arbitrariness in the context or in the process could lead to more frequent disagreement among, say, countries about composite indices than on individual indicators.

"[…] it is hard to imagine that debate on the use of composite indicators will ever be settled […] official statisticians may tend to resent composite indicators, whereby a lot of work in data collection and editing is "wasted" or "hidden" behind a single number of dubious significance. On the other hand, the temptation of stakeholders and practitioners to summarise complex and sometimes elusive processes (e.g. sustainability, single market policy, etc.) into a single figure to benchmark country performance for policy consumption seems likewise irresistible." (Saisana et al., 2005).

Some critics have developed and proposed alternative and novel indices (among others Noorbakhsh, 1998, Sagar and Najim, 1998, Lind, 2002, Chakravarty, 2003, Despotis, 2005, Herrero et al, 2010), which are essentially HDIs modified in one way or another.

Linear aggregation over three dimensions implies perfect substitutability, essentially meaning that a country can compensate for deficiency in one dimension by achievements in another (Desai, 1991, Sagar and Najim, 1998).

A large group of critiques is concerned with the fact that the current HDI presents averages and thus conceals wide disparities in distribution of human development in overall population. They suggest inequality adjustments to the HDI (Hicks, 1997, Foster et al, 2005, Seth, 2009).

In summary, the critiques have centered on two general areas, first, how to define human development and how to observe and measure its components and determinants, and second, how to aggregate the different indicators to obtain a commonly acceptable single index of human development in order to measure its improvement.[2]

Over time, the detailed composition of each index in the HD family has been subject to change as methodological advances have been incorporated. Recognizing and accepting the valid and valuable critiques, the HDI has been modified on different occasions. These changes included (Raworth and Stewart 2002) broadening the scope of the education component by adding another indicator (gross enrolment ratio) to increase variability since literacy data did not allow for differentiation at the top of the distribution. Also, modifications in normalization of component indicators were made by switching from relative maximum and minimum values to fixed goalposts to allow time-series analysis.

---

[2] There is an additional group of critiques, related somewhat to the second area, but interested mostly in information and statistical properties of the HDI and their improvements (Ivanova et al., 1999, Despotis, 2005, Zhou et al., 2007).

Table 1. Evoulution in calculation of the HDI 1990-2009

Year	Number of ranked countries	Health	Education	Standard of Living	Type of Index
1990	130	Life expectancy (LE) in years: maximum and minimum taken from data set at 78.4 and 41.8.	Adult literacy rate (ALR) for 25+: maximum=99%, minimum taken from data set at 12%	Logarithm of Gross Domestic Product (GDP): maximum=3.68 ($4786 per capita in $PPP); minimum (taken from data set)=2.34 ($220)	Deprivation
1991	160	LE: maximum and minimum taken from data set at 78.6 and 42.0.	ALR(%) and mean years of schooling (MYS) weighted 2/3 and 1/3. Minimum and maximum taken fom data.	Adjusted GDP pc obtained with the Atkinson formula with the minimum set at $4829 pc; maximum and minimum of discounted GDP taken from data set at $5070 and $350	Deprivation
1992	160	As of 1991	As of 1991	As of 1991	Deprivation
1993	173	As of 1991	As of 1991	As of 1991	Deprivation
1994	173	LE: maximum and minimum set at 85 and 25.	ALR(%), and MYS weighted 2/3 and 1/3, MYS: maximum and minimum set at 15 and 0	Maximum and minimum of discounted GDP was fixed at $5385 and $200 which was equivalent to fixing the real GDP to $40,000 and $200	Achievement
1995	174	As of 1994	ALR(%) and combined enrolment ratio (primary, secondary, and tertiary, capped at 100, in %), weighted 2/3 and 1/3 respectively	Maximum for formula set at $5448 pc which is equivalent to $40,000, minimum is set at $100	Achievement
1996	174	As of 1994	As of 1995	As of 1995	Achievement
1997	175	As of 1994	As of 1995	As of 1995	Achievement
1998	174	As of 1994	As of 1995	As of 1995	Achievement
1999	174	As of 1994	As of 1995	Logarithm of GDP pc: maximum =$40,000, minimum =$200	Achievement
2000	174	As of 1994	As of 1995; ALR taken from age 15 + combined enrolment ratio uncapped	As of 1999	Achievement
2001-2009	174 to 182	As of 1994	As of 2000	As of 1999	Achievement

The major change was done to the income component by replacing the Atkinson's income discounting form (that heavily discounted any income above the world's average) by the simple logarithmic form of GDP per capita with fixed maximum and minimum values. The *process* of calculating the HDI and data collection also changed as a response to critiques. In particular data selection methods and meta-data compilation were improved, and a rule not to include estimated data in the official HDI tables was established (even if these estimates were used to compute the HDI). The chronology of changes in the HDI is presented in Table 1.

Despite its many challenges and critiques, the HDI has been adopted as an analytical tool to measure development by various groups with very different missions and agendas, ranging from international organizations, government departments, academic researchers, to civil society, non-government organizations and advocacy groups. Many other indices such as the Disaster Risk Index of the United Nations Environment Programme[3] or the Prevalent Vulnerability Index of the Inter American Development Bank[4] make use of the HDI as an indicator by itself and integrate it into their index design.

3. CHOICE OF INDICATORS

Human development concepts are not measurable directly, although their specific manifestations are observable. There is a vast literature about socio-economic indices; however, the consensus is that there is no fully objective way of selecting the relevant indicators for their construction. The choice of indicators is guided by normative principles, in the first place, but also by a set of particular quality criteria aimed at assurance of integrity, methodological soundness, and high accuracy and reliability of the resulting composite index. These goals further dictate the quality dimensions for component indicators and variables, which include: conceptual relevance and intuitive validity, non-ambiguity, availability, reliability of data sources, internal reliability and logical interconnection, non-redundancy, and consistency. There is also a general principle that the information contained in a composite index should be easy to communicate and should be of interest to a broad spectrum of potential users – from policy analysts and decision makers, to media and general public. This alone implies that the information has to be restricted to a manageable quantity of indicators. The HDI has been criticized in the past for both a poor choice of indicators and for their inappropriate combination. We'll address some of these critiques in the next subsections.

Several critiques point out that the distinction between indicators which represent and measure inputs, outputs, outcomes and impacts is important and their mixing should be done carefully so that they measure, in combination, what is supposed to be measured. Booysen (2002) classifies indicators as those focused on ends (i.e., outputs) as opposed to means (i.e., inputs). He notes that the HDI is specifically designed to include both inputs and outputs, and that some variables, like literacy, represent both an end and a mean.

Veenhoven (2005) wrote "There are two ways to assess how well people live. One is to consider to what extent the country provides conditions deemed essential for a good life. In this approach the emphasis is on societal input… The other approach is to assess how well people thrive. In this approach the

[3] http://www.grid.unep.ch/activities/earlywarning/DRI/

[4] http://www.iadb.org/exr/disaster/pvi.cfm?language=EN&parid=4

emphasis is on output." Applying this dichotomy to education alone we can say that, for example, building schools, enrolling children to school programs, hiring and training teachers, organizing the continuing education for adults,…, represent the inputs into the process of developing a knowledgeable society, and thus represents the 'presumptions' of human development. Educated and trained people, generally, knowledgeable members of society, are the outputs, and represent an 'apparent' achievement of human development in one of its core dimensions. The outcomes are also the positive changes in the population expressed through a well-trained and innovative labour force, who can fully participate in civil and political activities, and who can provide and use services effectively. The impacts are reflected by improvements in the well-being of all.

Diener and Suh (1997) and Chibber and Laajaj (2007), among others, suggest that the HDI dimensions and their indicators should have an outcome focus: "The focus point should be the desired outcome rather than the means to reach these outcomes. This is essential for two reasons: first it leaves each country a complete freedom in the way to achieve the desired outcomes, and second, whenever the outcome is not directly observed, figures may not reflect the actual situation in the country and the index may lead to wrong incentives. (Chibber and Laajaj, 2007) "

Ryten (2000) argues that the HDI components need a structure: "the ratio of actual to eligible entrants; the efficiency of transformation of inputs into outputs; and the ratio of actual to potential outputs." He gives an example for education where the three conditions translate themselves into: an enrolment ratio into the first year of each stage of the formal education system; a ratio of graduates to entrants; and a ratio of graduates to the relevant portion of the population (after netting out duplication). Ryten admits that implementing the same structure in the case of health is not easy. He suggests that infant and maternal mortality rates represent the input indicators, and the life expectancy measures the output. For the measure of efficiency of transformation of input into output, Ryten proposes "the healthfulness": something like "the number of painless days" which resembles to the concept of health adjusted life expectancy. For the living standard dimension, Ryten sees as the input variable the "Employed as proportion of total labour force." Transformation from input to output is measured by "labour productivity" expressed as the number of hours worked over the total number of hours. The output is measured by the GDP per capita in purchasing power parity dollars. Ryten, however, doesn't elaborate on how to combine the indicators into an index.

Other critiques are concerned with the fact that the HDI indicators are a mix of stocks and flows (Kelly, 1991, Ryten, 2000). Stocks are the variables measured at one specific time point and represent quantities existing at that time, which may have been accumulated in the past. Flows are the variables measured over an interval of time and are expressed in units per time unit. Stock variables in the 2009 HDI are adult literacy and life expectancy and the flow variables are per capita GDP and gross enrollment ratios. Economic stocks and flows variables have different units and cannot be meaningfully compared, added, or subtracted. However, in the case of education indicators, the literacy rate and the gross enrolment rate, are already given as percentages, thus they are dimensionless, and their mixing may not always be meaningless.

Another problem related to the stock or flow nature of indicators is the responsiveness of the HDI to changes in the respective domains of society. Due to a lag in the impact of policy changes in education and health on literacy and life expectancy, the index has been criticized for measuring the outcomes of past efforts rather than the effects of present or recent policy changes.

In the domain of education and health, the HDI has been criticized because it pays attention exclusively to quantities and ignores the quality of the processes and outcomes.

As mentioned earlier, there is a set of quality criteria for the selection of variables and indicators that is suggested in the literature (McGranahan, 1995, Rose, 1995, Booysen, 2002, Nardo et al. 2008), mainly grouped around conceptual relevance, accuracy and reliability, consistency, and availability. An often asked question (e.g., Anand and Sen, 1994) is whether the current indicators differentiate (discriminate) sufficiently both the developed and developing countries. Some of these properties will be examined in more details in section 4. The statistical quality of data used in the construction of the HDI has often been a target of critique. While the data used in the construction of the HDI come from reputable international sources, there is still a space for improvement.

In the next three subsections we review specific critiques for each HDI dimension separately.

3.1 Education Indicators: Review of Critiques and Ideas

From the HD perspective both initial education at the early stages of a person's life prior to entry into the world of work, as well as continuing education and expansion of knowledge throughout a person's life should be equally important, as human development concerns the expansion of capabilities.
The HDI has relied on two indicators for the dimension of knowledge – adult literacy and gross enrollment ratio for school age children. The exceptions are 1990 when only literacy was used, and from 1991 to 1994, when in addition to literacy, mean years of schooling of adult population was used.

The quantity and quality of education received by the population is very important in today's society. Basic education increases the efficiency and the participation of each individual. Moreover, members of the society who have received little formal education can carry out only simple manual work and find it much more difficult to adapt to more advanced technologies. Lack of basic education can therefore become a constraint on development, with the society finding it difficult to move up through the value chain by producing and consuming more-sophisticated products and services, and thus the entire society lags behind.

Literacy is considered a human capital stock variable, and in some sense it represents the outcome of the education process. The adult literacy rate measures the proportion of people aged 15 years and older who declare they are able to read and write, thus it is based on a self-reported personal assessment and judgment that may be influenced by the social circumstances in which people describe themselves in a way that is different from what objectively their situation is. Also, these rates do not provide information on whether individuals are able to read and write texts of varying degrees of difficulty and content. Moreover, they do not cover neither numeracy skills[5], nor information literacy[6]. That is, the current definition of literacy doesn't even attempt to cover the plurality of literacy which

[5] Numeracy skills enable individuals to perform short mathematical tasks that require computing, estimating, and understanding notions of shape, length, volume, currency and other measures.
[6] The Alexandria Proclamation of 2005 recognizes information literacy as "a basic human right in the digital world" as it empowers individuals "in all walks of life to seek, evaluate, use and create information effectively to achieve their personal, social, occupational and educational goals." "Beacons of the Information Society", High Level Colloquium on Information Literacy and Lifelong Learning, Bibliotheca Alexandria, Egypt, November 2005. UNESCO (2008) Towards Information Literacy Indicators

refers to "the complex interaction between literacy and many areas of social life. The identification of multiple illiteracies suggests how those interactions translate into specific sets of skills. (UIS 2009)".

Lind (2002) argues that while "literacy is essential for development it is unimportant as a measure of development". It cannot be measured reliably (Harkness, 2004). Even, if measured, "can it be compared between writing systems as different as Chinese and English?" Literacy becomes a non-differentiating variable for most countries, e.g., in the HDR 2009, for 102 countries it was ranging from 90% to 100% and the average literacy rate was 83.9%

There are better alternative indicators, although they are available only for some countries. Between 1994 and 2003 several OECD countries participated in the International Adult Literacy Survey (IALS) in order to assess functional literacy level of the adult population[7] which is defined in general terms as "the ability to understand and employ printed information in daily activities, at home, at work and in the community - to achieve one's goals, and to develop one's knowledge and potential." The IALS employed a sophisticated methodology to classify the surveyed adults to one of its five levels of literacy. A person was considered functionally illiterate if classified at level 1 which was broadly specified as "Level 1 indicates persons with very poor skills, where the individual may, for example, be unable to determine the correct amount of medicine to give a child from information printed on the package."(Kirsch, 2001) In total, only 22 countries were covered by the IALS. Analyses of IALS data have shown that differences in levels of literacy matter both economically and socially: "literacy affects, inter alia, labour quality and flexibility, employment, training opportunities, income from work and wider participation in civic society.(OECD, 1998)."

The mean years of schooling of the adult population is another possibility for assessing the education attainment of the adult population. It was a component indicator of the HDI in the period from 1991 to 1994. However, it poses many problems itself, mainly related to the significant variation in duration of each level of education across countries and across time that have to be accounted for in order to achieve comparability. Barro and Lee (2000, 2010) take this variation into account by using information on the typical duration of each level of schooling within countries and an estimation methodology relying on cohort analysis applied to census data and modified by the UNESCO information on school enrollment. However, an issue of comparability between national school systems and programs still remains as an important source of uncertainty. In spite of its limitations and the debates on the relative merits of one form of education over another, mean years of schooling have provided a robust measure for the input of education into the formation of human capital (Chiswick, 1997).

The other indicator used in the HDI to proxy the knowledge dimension is the combined gross enrolment ratio. Harkness (2004) criticizes the use of enrollment data rather than school completion rates which she sees as a more output-oriented indicator. In their critique of the HDI indicators, Chibber and Laajaj (2007) echo Harkness: "The recent emphasis on enrollment rates, for example in the HDI and in the Millennium Development Goals, has often encouraged the creation of large classes, and little follow-up of real learning within countries. When the data will allow it, net rather than gross enrollment ratios would be a step toward the focus outcome, because it excludes grade repetitions, a sign that limited knowledge has been acquired during the year." In the same spirit, the MDG Task Force on Education

[7] The adult population is civilian, non-institutionalized population, aged 16 to 65.

and Gender Equality[8], for example, has criticized the use of enrolment data rather than school completion rates: "Completion of schooling is a significant problem. While enrollment has been increasing, many children drop out before finishing the fifth grade (UNESCO 2004b). In Africa, for example, just 51 percent of children (46 percent of girls) complete primary school. In South Asia 74 percent of children (and just 63 percent for girls) do so."

A possible alternative to the school enrollment rates is School Life Expectancy (SLE) which is conceptually very similar to life expectancy. It is defined as the total number of years of schooling which a child of a certain age can expect to receive in the future, assuming that the probability of his/her being enrolled in school at any particular age is equal to the current enrolment ratio for that age. Thus it makes sense to consider the SLE at the primary school entrance age. This indicator, according to the UNESCO "shows the overall level of development of an educational system in terms of the number of years of education that a child can expect to achieve[9]." In terms of statistical computation it relies on data on school enrolment at all levels of education and on demographic data about population of official school-age. Relatively higher SLE indicates that children can expect to spend more years in education. It must be noted that the expected number of years does not necessarily coincide with the expected number of grades of education completed, because of grade repetition. There is, however, a cautionary note attached to the utility of school life expectancy for inter-country comparison for several reasons: (i) neither the length of the school-year nor the quality of education is necessarily the same in each country; (ii) as this indicator does not directly take into account the effects of repetition, it is not strictly comparable between countries with automatic promotion and those allowing grade repetitions; (iii) depending on the country, the enrolment data may not account for many types of continuing education and training. An idea suggested by the UNESCO is to complement this indicator with the repetition rates.

Some critiques, like Lloyd and Hewitt (2009), note that the content and quality of education matter more than the enrollment and completion rates. They find that primary school completion rates are not necessarily correlated with educational outcomes (as measured by literacy). Also, in their 2006 paper "A Millennium Learning Goal: Measuring Real Progress in Education," Filmer, Hasan and Pritchett (2006) analyzing the MDG 2 recognize that a goal of school completion is very important but that at the same time it doesn't necessarily include learning achievement, and argue that the *outcomes* of learning achievement matters. They demonstrate that "even in countries meeting the MDG of primary completion, the majority of youth are not reaching even minimal competency levels, let alone the competencies demanded in a globalized environment."

Therefore enrollment and completion indicators are not necessarily good or consistent predictors of outcomes. A 1999 study of six African nations reveals a range of relationships (Ellis, 2003). Kenya had the lowest completion rate, at 63 percent, but 65 percent of its sixth grade pupils achieved minimum literacy skills—a better outcome than in any other country. Malawi's completion rate was almost identical to Kenya's, at 64 percent, yet only 22 percent of its sixth grade pupils demonstrated minimum literacy skills.

---

[8] **Task Force on Education and Gender Equality (2005).** Toward universal primary education: investments, incentives, and institutions. *UN MDG. Available at http://www.unmillenniumproject.org/documents/Education-complete.pdf*

[9] http://www.uis.unesco.org/glossary/

Table 2. List of Education Indicators According to Availability[*]

Knowledge, Education							
Indicator	**Source**	**Number of countries**	**Trends (Y/N)**	**Frequency of data availability at country level**	**Earliest year available**	**Latest year available**	**Stock/Flow**
Adult literacy rate (% of pop aged 15+) (ALR)	UIS	≈180[a]	Y	irregular	1975	2015	Stock
Youth literacy rate (% of pop aged 15-24 years)	UIS	≈150	Y	irregular	1975	2008	Stock
Adult population (25+) with at least upper secondary level of education (ISCED 3) (%)	UIS	≈90; If combined with Barro-Lee ≈150	Y	irregular			Stock
Mean years of schooling of adults aged 25+ (years)	Barro/Lee [Centre for International Development, University of Harvard]	≈146	Y	Every 5 years	1955	2010	Stock
Gross enrolment ratio in secondary and tertiary education (%)	UIS	≈170	Y	annual	1999 (or 5-year intervals from 1970)	2008	Flow
Gross enrolment ratio in primary, secondary and tertiary education (%) GER	UIS	≈180	Y	annual	1999 (or 5-year intervals from 1970)	2008	Flow
School life expectancy of a child of school entrance age (years)	UIS	≈150	Y	annual	1999 ?(or 5-year intervals from 1970)?	2008	Stock
Gross intake rate to last year of primary (%)	UIS	≈170	Y	annual	1999	2009	Flow

[a]ALR was not available for about 20 developed countries. The HDRs assume 99% literacy for them.
[*]As of May 15, 2010

This again broadens the discussion and pushes us to consider how multiple measures of a concept, such as education, can play an essential role in our understanding of human capabilities.

Certainly, it would make more sense to use the indicators of the outcomes of the education system: to "assess mastery of the desired competencies of an entire age cohort—both those in school and out of school (Filmer,D., et al., 2006)," something similar to results from PISA, an international survey that covers students who are between 15 and 16 at the time of the assessment, regardless of the grade or type of institution in which they are enrolled. However, the ability to measure the learning outcomes at the national (not to mention the international level) is generally weak. Instead, the focus is typically on the number of children enrolled. Table 2a below shows the current availability of different educational indicators.

3.2. Health Indicators: Review of Critiques and Ideas

Health is central to human development. The most basic personal and societal decisions reflect the salient desire of everyone to live a long and healthy life. However, it is not always clear what are the determinants of longevity and it is even less sure how to achieve such a goal. Some analysts (Pritchett and Summers, 1996) argue that advances in economic growth and technology will overall increase our ability to live longer and healthier lives by automatic reduction of prevalence of fatal diseases. Economic growth and unprecedented technological advances have brought extraordinary declines in mortality rates at all ages so that much of the world's population has experienced significant increase in life expectancy.

Long lives and high incomes are highly correlated in a statistical sense. This has been documented in a variety of research over the last fifty years or so. In his well cited paper, Preston (1975) constructed curves showing dependence of length of life on income, thus indicating an association but cannot infer the direction of causality. Preston had in mind a causality running from income to length of life. However, he also stressed how much average life expectancy has been increasing at any given level of income. According to Preston's data, middle-income countries today, for example, have reached per capita income levels close to that of the United States around 1900 ($892 in 1963 dollars) when life expectancy in the United States was only about 50 years, whereas in many middle-income countries today life expectancy exceeds 75 years.

However, the mechanisms of the relationship between economic growth and health outcomes still remain unclear. Despite massive investments and interventions, preventable diseases continue to kill millions of people each year and high childhood mortality rates persist in many parts of the world. Theories and empirical findings about the determinants of population health that involve socio-economic arguments, as well as epidemiological and medical models and a variety of inter-connecting research hypothesis are probably the most contested of our times (Oeppen and Vaupel 2002). In any case, everyone shares a sense of the importance of improving health outcomes as keys to a prosperous and fulfilling life (Sen, 1998, Sachs et al., 2001).

Sen (1998) has emphasized the importance of mortality data in any assessment of human well-being because "mortality information has (1) intrinsic importance (since a longer life is valued in itself), (2) enabling significance (since being alive is a necessary condition for our capabilities), and (3) associative relevance (since many other valuable achievements relate-negatively-to mortality rates)."

Indicators derived from mortality rates provide a good picture of overall population health. These indicators include infant and child mortality (the probability of dying between birth and 1, and between 1 and 5 years of age, respectively), adult mortality (the probability of dying between 15 and 60 years of age) and overall life expectancy at birth.

Many researchers have pointed out the quality problems of mortality data on which the life expectancy computation is based (Chamie, 1994). Heilig et al.(2008) reveal that only 34 countries, representing about 17% of the world population in 2008, have mortality databases according to international standards. The remaining more than 80% of the world population rely on data from censuses and surveys, and for "the vast majority of these countries, mortality is estimated on the basis of infant and child mortality data provided by various data sources – using model assumptions for the underlying age patterns." Recently, Anthopolos and Becker (2010) have shown dramatic undercounting of infant mortality that might have a significant effect on the estimated life expectancies. This further implies that mortality data are most likely to present a downwards bias in countries where mortality is highest.

Life expectancy, and its variants, has been the metric most commonly used by human development researchers and analysts, although life expectancy does not capture all the aspects of the individual's current health that may affect functionings and limit capabilities. However, indicators of health quality are simply more difficult to obtain. A large number of health indicators have been used, and it seems that they served mostly to underscore the difficulty in measuring the multiple dimensions of the healthy state that influence the human capital potential.

Table 2b: List of health indicators according to availability*

Health dimension							
Indicator	**Source**	**Number of countries**	**Trends (Y/N)**	**Frequency of data availability at country level**	**Earliest year available**	**Latest year available**	**Stock/ Flow**
Life expectancy at birth (LE)	UNPD	≈180	Y	Every 5 yrs	1950	2030	Stock
Life expectancy at age 10	UNPD	≈180	Y	Every 5 yrs	1950	2030	Stock
Life expectancy at age 15	UNPD	≈180	Y	Every 5 yrs	1950	2030	Stock
Health adjusted life expectancy at birth	WHO	≈180	N	Irregular	2003	2003	Stock
Children (under 5) survival rate per 1000 live births	UNICEF MDG	≈180	Y	Every 5 yrs	1990	2007	Flow
Children (under 5) not underweight for their age (%)	UNICEF MDG	≈140	Y	Every 3 yrs	1990	2008	Flow

*As of May 15, 2010

A need for a summary measure of population health which can capture both fatal and non-fatal health outcomes led the World Health Organization[10] to construct the Healthy Adjusted Life Expectancy as an estimate of how many years a person might live in "good" health. "These estimates are based on country life tables, analyses of 135 causes of disability for 17 regions of the world and 69 health surveys in 60 countries. (WHO 2009)." The estimates of healthy life expectancy are even more uncertain than those for life expectancy, because it is difficult to ensure comparable measurements of disability across countries and account for limitations in the data. Moreover the WHO doesn't plan to compute the HALE regularly.

3.3 Income Indicators: Review of Critiques and Ideas

The third HDI component – standard of living, command over resources, or simply the income dimension, has been extensively criticized. Simply, it is very difficult to measure it. As Anand and Sen (2000) wrote "it requires data on access to land, credit, income and other resources. But given the scarce data on many of these variables, we must for the time being make the best use of an income indicator."

The HDI component of living standards has always been measured by the Gross Domestic Product (GDP)[11] per capita in Purchasing Power Parity Dollars (PPP$)[12]. The GDP has been criticized as limited in many ways: it has the limitations inherent in its own construction; it is also limited through the inherent faults of price indices and the purchasing power parity conversions. The way the GDP is constructed, by excluding those goods and services not traded in markets, ignores household production such as taking care of own children, housekeeping and food production for own consumption. Such activities are universally important, especially in developing countries. GDP excludes informal, unreported, and illegal economic activities, which are omnipresent and often very important in developing countries. GDP includes all effects of economic activities, whether they are positive or negative without discrimination, often leading to double accounting for an economic activity for causing of problems, such as pollution, as well as for rectifying those problems. Critiques point out that GDP underreports the true cost of economic activity as many economic activities treat natural resources as free. Some critiques rightfully point that the level of living standard can be increased through increased levels of leisure which is not accounted for in GDP. (Jorgenson, 1997, Islam and Clarke, 2002). Osberg and Sharpe (2005) have suggested using the Index of Economic Wellbeing (IEWB)[13] in place of GDP, as a more complete index of command over resources.

---

[10] WHO: *The world health report 2004 –Changing history*

[11] The gross domestic product (GDP) is the market value of all final goods and services made within the borders of a country in a year, that is:GDP= Private consumption + Gross private investments +Government spending +Export-Import.

[12] Purchasing Power Parity (PPP) is currency conversion rate that both convert to a common currency and equalize the purchasing power of different currencies, so that the differences in price levels between countries are eliminated.

[13] The IEWB has four dimensions - current effective per capita **consumption** flows, net societal **accumulation** of stocks of productive resources, income **distribution**, and economic **security**. This elaborate index draws from tens of indicators, such as - expenditure private and public (government), estimated per capita cost of leisure, number of hours worked per person, hourly wages, total current net stock of fixed capital, etc. While the IEWB has been calculated and used for comparison of the OECD countries (see http://www.csls.ca/iwb/oecd.asp) it is very difficult to calculate it for other countries due to sparse data.

As an increasingly common feature of a globalised economy, the GDP is prone to distortions due to declarations of large multinational companies that operate in small economies, since it doesn't capture correctly the relevant flows of compensation of employees, property income and taxes on production and imports across economies. Several critiques (Sweeney, 1999, Morse, 2003) pointed out that the use of Gross national income (GNI) [14] may be a better choice since it excludes exported goods and services. For example, the profits of a US-owned company operating in India will count towards US GNI and Indian GDP, but will not count towards Indian GNI or US GDP.

The recent Report by the Commission on the Measurement of Economic Performance and Social Progress[15] mentions that "Income flows are an important gauge for the standard of living, but in the end it is consumption and consumption possibilities *over time* that matter". The Report then recommends using income or consumption rather than production; moreover, the Report suggests the use of wealth to take into account the temporal dimension. Finally the Report suggests emphasizing the household perspective. One specific recommendation focuses on the aggregate actual household consumption[16]. This variable adjusts household consumption by adding the individual government-provided services (individual refers to goods and services such as hospital and cultural services; collective to goods and services such as defense and environmental protection). Alternatively, one could use final consumption expenditure plus general government final consumption expenditure, as commonly reported by the World Bank - the difference is that collective goods and services provided by the government would be taken into account. These indicators are closely related to GDP as they are part of it.

Empirical research has shown that the distribution of consumption can be quite different from the distribution of income. Indeed, the most pertinent measures of the distribution of material living standards are probably based on jointly considering the income, consumption and wealth position of households or individuals.

The Commission also recommended that instead of using the averages of income, consumption and wealth it might be more insightful to use the medians as measures that pertain better to the "typical" individual or household than the mean (which may be distorted by extreme values at either end of the distribution). By using the medians one captures, in some general sense, the income distribution aspect as well since half of all individuals are above the median income and half below. If all the increase is allocated in the top 10%, median income may remain unchanged, while mean income increases. However, there is a symmetrical counter argument which says that if there is an increase in poverty at the lower 10% the median will remain unchanged, thus failing to detect an important deterioration of living conditions. Overall, medians and means would give different pictures of what is happening to societal well-being. The Commission also recognizes that in practice, moving from means to medians may be difficult given that medians require the micro data typically derived from household income surveys, whereas macro-economic measures from the national accounts are based on a range of

---

[14] The GNI consists of: Private consumption + Gross private investment + Government spending + Net income from assets abroad + gross exports of goods and services- gross imports of goods and services – indirect business taxes

[15]The Commission chaired by Stiglitz, Sen and Fitoussi wrote up with the Report (2009) www.stiglitz-sen-fitoussi.fr

[16] Actual final consumption of households is the aggregate of goods and services acquired through expenditure by households themselves; those acquired as social transfers in kind from Non Profit Institutions Serving Households, and those acquired as social transfers in kind from general government (on individual consumption)

different macro-economic sources and may not pertain to the same population. Many of the important properties of means, as well as the theories developed around the concepts such as welfare standards may not translate directly to the median-based measures.

Table 2c: List of Income Indicators According to Availability*

Standard of Living							
Indicator	**Source**	**Number of countries**	**Trends (Y/N)**	**Frequency of data availability at country level**	**Earliest year available**	**Latest year available**	**Stock/Flow**
Gross National Income per capita ($PPP) (GNI)	WB-WDI	Over 170	Y	Every year		2008	Flow
Total consumption expenditure per capita in $PPP. (TOTCON)	WB-WDI	166	Y	Every year		2008	Flow
Total household consumption expenditure per capita in $PPP (HHCON)	WB-WDI	165	Y	Every year		2008	Flow
Gross Domestic Product ($PPP) (GDP)	WB-WDI	Over 180	Y	Every Year		2008	Flow

*As of May 15, 2010

4. CORRELATION STRUCTURE OF THE HDI

4.1 Redundancy

Several critiques pointed out the redundancy in the information provided by the HDI and its components due to their high correlation. McGillivray (1991) has questioned the usefulness of the HDI since his analysis revealed that the HDI "is significantly and positively correlated with each of its component variables individually." He concluded that "as a consequence, assessing inter-country development levels on any one of these variables yields similar results to those that the index itself yields, and more profoundly the index largely provides us with little more information regarding inter-country development levels than the more traditional indicator - GNP per capita, alone provides. " These findings lead to the conclusion that the HDI is "yet another redundant composite inter-country development indicator." A good composite index should have components which are themselves insignificantly correlated. McGillivray's concerns were echoed by Srinivasan (1994). Ogwang (1994) used the Principal Component Analysis[17] on the HDI component indices and found that the Life expectancy index is the single best predictor of the HDI score. The significance of his finding is that the simplified index of HD is not the GDP index but one based on Life expectancy. Noorbakhsh (1998a),

[17] Principal Component Analysis (PCA) has been used to transform a larger set of correlated variables into a smaller set of uncorrelated variables, called principal components, that account for most of variation in the original data set.

however, has argued that the component indices of the HDI are not highly correlated with each other, nor with the HDI and need to be kept in the index.

Social indicators are often (highly) correlated with economic indicators of development. Lind (2002) noticed that "compound social indicators are sometimes criticized because their components are correlated. Such critique is unfounded." Booysen (2002) recommends that a weak correlation between an indicator and the index should result in the exclusion of the respective indicator from the process.

Saisana et al. (2005) pointed out the ambivalence in understanding of the impact of a high correlation among indicators. One school of thoughts treats the high correlation as something that needs to be corrected by some statistical intervention such as Principal Component Analysis or Factor Analysis[18]. On the other hand, practitioners of multi-criteria decision analysis tend to consider the existence of correlations as an indication that the correlated indicators measure the same latent construct but reflecting on different aspects of it in a non-compensable manner. Underlying indicators with high internal correlation give a composite index with values and rankings very robust on changes in the selection of weights, the normalisation method and other steps involved in its construction.

Ivanova et al. (1999)[19] have checked on the validity of the HDI construction by looking at correlations of HDI with its components with the idea that an index has to prove its advantages over its component measures to justify its existence. In other words, it is important to find out whether the HDI provides any additional information than its components alone do. To test this, they have regressed each component of the HDI against the other two. They found that the education attainment (EA) measured by the mean years of schooling (used as an indicator from 1991 to 1994) and the GDP explain a large portion of the variation in LE, and concluded that either the LE or the combination of EA and GDP can explain approximately the same proportion of the variation of the HDI. Further to this finding, they concluded that the overall ranking by the HDI would not change significantly if the GDP and EA are excluded from the index and the countries are ranked only by the LE.

A simple correlation analysis of the HDI[20] and its indicators reveals significant linear relationship between the index's components and its score (Table 3) when taken over all countries. Life expectancy (LE) shows the strongest relationship with the HDI (r = 0.92) whereas GDP has the weakest relationship (r = 0.71). We added to this analysis two logarithmized variables, *ln(GDP)* and *ln(LE)*, as well. Clearly, the correlation of the HDI with the *ln(GDP)* is higher than the one with the original GDP because the GDP enters the HDI as logarithmically transformed. In contrast, the *ln(LE)* has weaker association with the HDI than the original LE. All the estimated correlation coefficients are obtained as significant at 0.01 level based on the sample of 182 countries. Note that all computations are done using the values from the HDR 2009. The GDP denotes always the GDP per capita in Purchase Power Parity International dollars ($PPP). The correlation coefficients are given in Table 3 below. We first present correlations for all 182 countries and than for each of quartile groups where countries are grouped according to their level of HDI.

---

[18] The goal of principal components analysis (PCA) is to reveal how different variables are associated. This is achieved by transforming correlated variables into a new set of uncorrelated variables using a covariance matrix or the correlation matrix.

[19] Ivanova, I., Arcelus, F.J., Srinivasan, G. (1999). An Assessment of the Measurement Properties of the Human Development Index. *Social Indicators Research*, 46, 157-179.

[20] The HDI is defined as before 2010: $\left[\mathrm{GDP_x} + \mathrm{LE_x} + \frac{2}{3}\mathrm{ALR_x} + \frac{1}{3}\mathrm{GER_x}\right]/3$ where subscript x means index.

Table 3. Correlation coefficients for the HDI and its indicators

a) All countries (182)

	hdi	le	alr	ger	gdp	lngdp	lnle
hdi	1.0000						
le	0.9154	1.0000					
alr	0.8762	0.7350	1.0000				
ger	0.8863	0.7729	0.7970	1.0000			
gdp	0.7131	0.5827	0.4855	0.5880	1.0000		
lngdp	0.9354	0.7821	0.7294	0.7918	0.8370	1.0000	
lnle	0.9018	0.9969	0.7304	0.7556	0.5508	0.7585	1.0000

b) Most developed countries (46)

	hdi	le	alr	ger	gdp	lngdp	lnle
hdi	1.0000						
le	0.7116	1.0000					
alr	0.2360	0.0184	1.0000				
ger	0.5248	0.3831	0.1932	1.0000			
gdp	0.4421	0.0095	-0.1482	-0.1025	1.0000		
lngdp	0.6005	0.1677	-0.1630	-0.0414	0.9606	1.0000	
lnle	0.7089	0.9999	0.0151	0.3837	0.0088	0.1667	1.0000

c) Upper-middle developed countries (46)

	hdi	le	alr	ger	gdp	lngdp	lnle
hdi	1.0000						
le	0.4384	1.0000					
alr	0.2282	-0.1615	1.0000				
ger	0.3905	-0.0379	0.1653	1.0000			
gdp	0.6306	-0.0410	-0.1617	-0.0542	1.0000		
lngdp	0.6798	-0.0324	-0.1647	0.0209	0.9733	1.0000	
lnle	0.4366	0.9994	-0.1687	-0.0459	-0.0389	-0.0311	1.0000

d) Lower-middle developed countries (45)

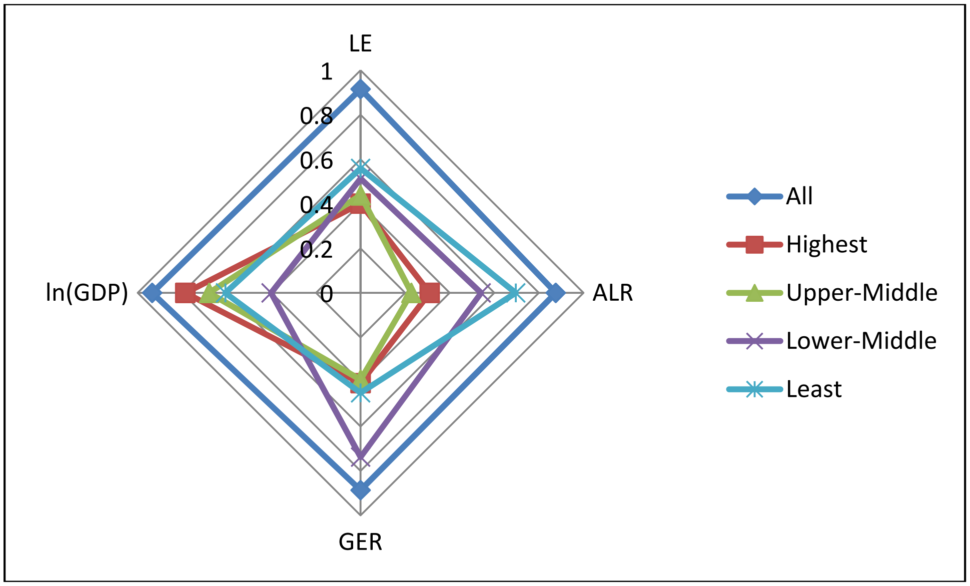
	hdi	le	alr	ger	gdp	lngdp	lnle
hdi	1.0000						
le	0.5132	1.0000					
alr	0.5412	0.0350	1.0000				
ger	0.7377	0.2492	0.5283	1.0000			
gdp	0.2036	-0.5218	-0.0474	0.0388	1.0000		
lngdp	0.4014	-0.3405	-0.1058	0.2110	0.8840	1.0000	
lnle	0.4916	0.9982	0.0358	0.2447	-0.5522	-0.3715	1.0000

e) The least developed countries (45)

	hdi	le	alr	ger	gdp	lngdp	lnle
hdi	1.0000						
le	0.5592	1.0000					
alr	0.6962	0.0538	1.0000				
ger	0.4490	0.0055	0.4413	1.0000			
gdp	0.5269	-0.1023	0.2481	0.1752	1.0000		
lngdp	0.6066	0.0461	0.1813	0.1188	0.8983	1.0000	
lnle	0.5495	0.9985	0.0485	-0.0058	-0.1133	0.0369	1.0000

Apparently, the correlation structure for quartile groups differ from the structure for all countries taken together. The first column shows how the correlation between the HDI and its components varies across the groups. For the Highest and Upper-Middle developed countries, the correlation of the HDI with the adult literacy (Alr) becomes weak. Graph 1 presents the correlations between the HDI and its component indicators for all countries together (All) and each quartile group.

Graph 1. Correlations of the HDI and its component indicators for all countries together and for each quartile group



The correlation structure of the HDI's dependence on its components is very much alike for the highest and the upper-middle developed countries, while the lower-middle and the least developed countries differ between themselves and from the upper two groups. A direct conclusion from this exercise can be that there is no redundancy in construction of the HDI since the correlation varies in different parts of country distribution. The presented results refer to 2009, however due to slow changes in variables the findings can be extended to other years as well.

4.2 Multi-dimensionality and Internal Consistency

Some of the critiques of the HDI high correlation of the HDI with its components also claim that the HDI is only a one-dimensional index, contrary to the conjecture of its founders that it covers three dimensions of human development (Ogwang and Abdou, 2002).

Without going into more details of the Principal Component Analysis (PCA) one can see from Table 4 that the first principal component (Comp1) accounts for approximately 83% of the variation in all four HDI component dimensions. The last 3 principal components explain the remaining 17% of the variation. The rule is not to take into consideration the principal component with the eigenvalue less than one. So, indeed, by applying the PCA to the HDI component indicators one can conclude that all four indicators measure the same dimension. Note also that the weights in the first principle component are almost equal. We will come back to this result in the next section.

Table 4. Results of PCA of the four HDI components

Rotation: (unrotated = principal) Rho = 1.0000

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	3.30449	3.01493	0.8261	0.8261
Comp2	.289561	.0667602	0.0724	0.8985
Comp3	.222801	.0396561	0.0557	0.9542
Comp4	.183145	.	0.0458	1.0000

Principal components (eigenvectors)

Variable	Comp1	Comp2	Comp3	Comp4	Unexplained
le	0.4977	-0.4587	0.7331	-0.0669	0
alr	0.4931	0.7318	0.1635	0.4411	0
ger	0.5092	0.1957	-0.2949	-0.7845	0
lngdp	0.4999	-0.4644	-0.5907	0.4307	0

An alternative way to investigate the degree of association among variables is to use the Cronbach coefficient alpha (C-alpha)[21], which is the most common estimate of internal consistency of items in a model. In the case of the HDI, the C-alpha can be used to check the internal consistency of the HDI indicators. If the C-alpha is high, then there is evidence that the individual indicators are measuring the same underlying (latent) construct. Table 5 below presents values of C-alpha for all 4 indicators and their different combinations. The obtained values can be used to determine the number of dimensions of the HDI.

Table 5. Cronbach alpha for the HDI component indicators

Indicators	C-alpha	Interpretation
All four indicators	0.8073	Measure the same concept since C-alpha>0.75
Without *ln(GDP)*	0.8789	*ln(GDP*) represents a different dimension from other three indicators
Only ALR and GER	0.8849	ALR and GER belong to the same dimension which is different from dimensions to which other indicators belong
Only LE and *ln(GDP)*	0.3334	*LE and ln(GDP) represent different dimensions*

We see that the C-alpha increased the value when *ln(GDP)* was excluded. This means that the *ln(GDP)* belongs to a dimension different from dimension(s) to which other three indicators might belong. This statement is equivalent to saying that *ln(GDP)* and other three indicators are not belonging to the same dimension. When we excluded the LE as well, the C-alpha increased further, indicating that the remaining two indicators (ALR and GER) represent the same dimension which is different from dimension(s) to which *ln(GDP)* and LE might belong. The final check was whether LE and *ln(GDP)* belong to the same dimension and since the value of C-alpha is significantly reduced we concluded that the LE and lnGDP are not representing the same dimension.

The conclusion from this exercise is that the HDI indicators *are* grouped in three dimensions , thus confirming the multidimensionality of the HDI, and that the high correlation between the indicators is not a sign of redundancy but rather a sign of internal consistency.

---

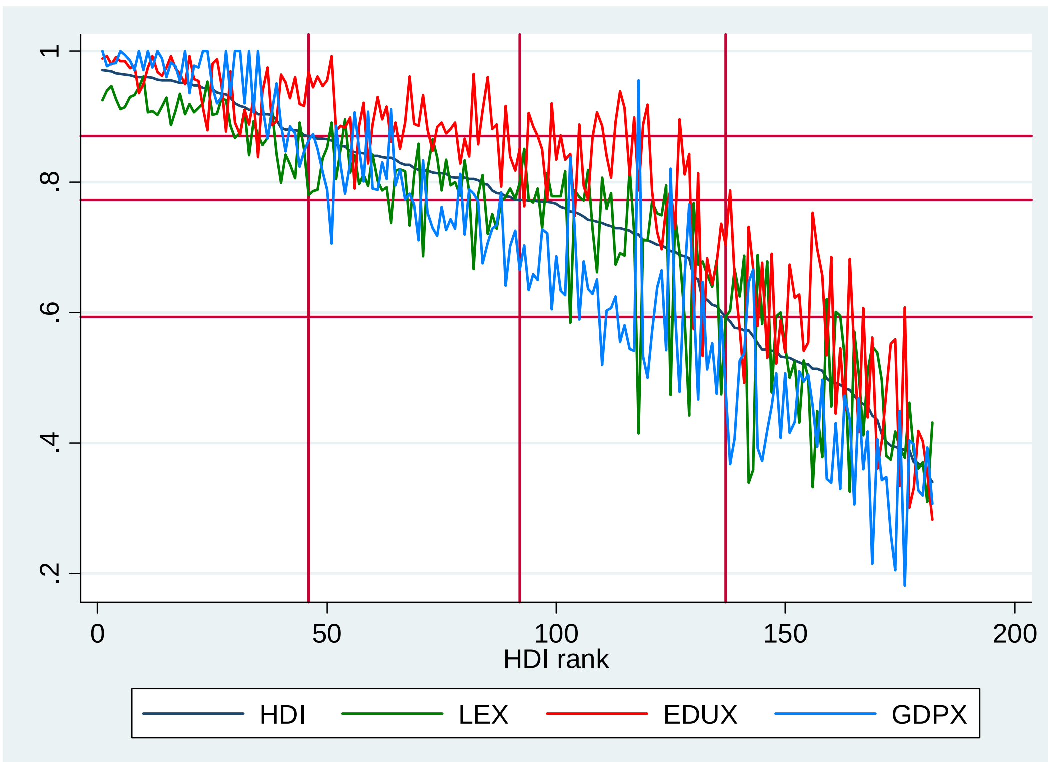
[21] Cronbach coefficient of internal reliability (C-alpha) is defined as $\alpha = \frac{K}{K+1}\left[1 - \frac{\sum_k V(X_k)}{V(\sum_k X_k)}\right]$, where *K* denotes the number of indicators, X_k,*k=1,…K*, and V denotes the variance.

4.3 Differentiation (Discrimination) Power of the HDI

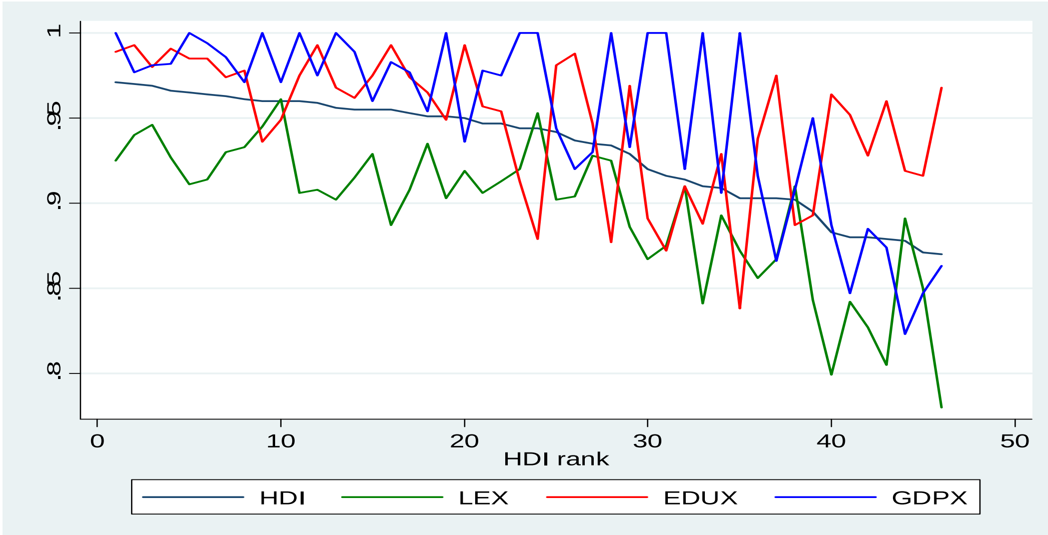
Several critiques (e.g., Anand and Sen, 1994) raised a question about the ability of the chosen indicators to differentiate countries, especially at the top and at the bottom of the distribution. Some critiques (Hoyland, Moene,Willumsen, 2010) find that the difference between the developed countries is only the random measurement noise and that information contained in the HDI can't differentiate developed countries.

Graph 2a presents the distribution of component indices and the HDI in 2009. The vertical lines represent the HDI quartile groups (Highest, Upper-Middle, Lower Middle, Low), while the horizontal lines represent the HDI values observed as demarcation values between the groups.

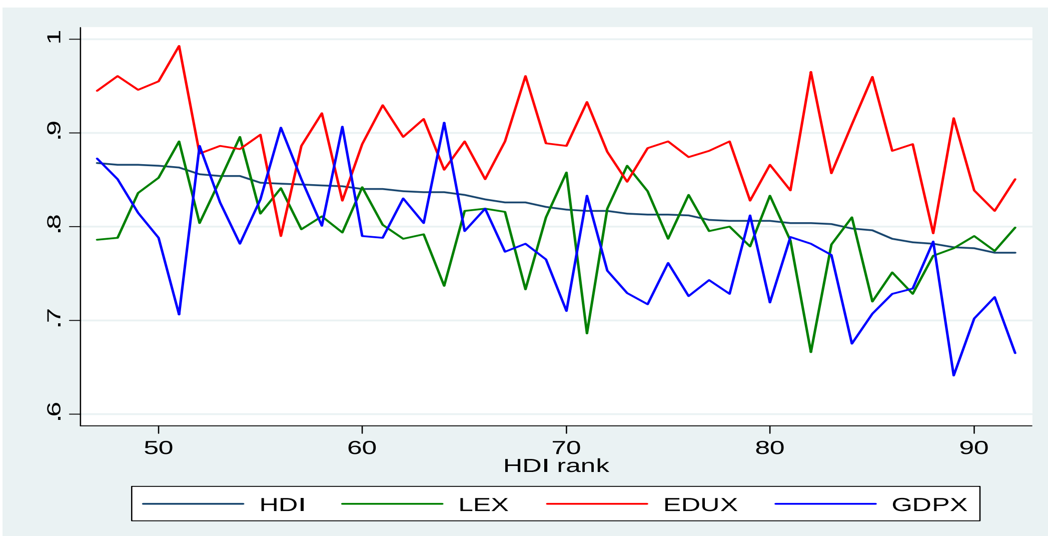
Graph 2a. HDI and the component indices



Graph 2b. HDI and component indices for 46 countries from the highest HDI (first quartile) group



Graph 2c. HDI and component indices for 45 upper-middle HDI (second quartile) countries



A closer look at the distribution in two top groups (Graphs 2b and 2c) indicates very small relative variability in, at least, two out of three component indices. We produce some basic statistics that might indicate some reasons of weak differentiation among countries. In Table 5 below we present the average value, the standard deviation, and the coefficient of variation for each indicator and component index at the level of each quartile group. In addition, for the HDI quartile we calculate the ratio between the range of HDI values and the number of countries in corresponding group, R/n. This ratio indicates an average distance between the countries in the group. A smaller distance means a weaker differentiation.

Table 6. Basic statistics for the HDI, its component indices and component indicators for all countries (182) and the quartile groups

		HDI	LEx	EDUx	GDPx	LE	ALR	GER	Ln(GDP)	GDP	Ln(LE)
All	Mean	0.739	0.725	0.798	0.693	68.5	83.6	72.3	8.77	13481	0.81
Range:	SD	0.174	0.170	0.176	0.216	10.2	19.0	17.3	1.33	15849	0.13
(0.340-0.971)	CV(%)	23.5	23.5	22.1	31.3	14.9	23.9	24.0	15.1	117.0	3.8
n=182	R/n	0.003	0.004	0.004	0.004						
Highest	Mean	0.932	0.896	0.948	0.952	78.5	97.4	89.6	10.39	35486	0.94
Range:	SD	0.031	0.042	0.040	0.051	2.7	2.7	9.7	0.42	15760	0.03
(0.870-0.971)	CV(%)	3.4	4.6	4.2	5.4	3.4	3.9	10.9	4.0	44.4	0.8
n=46	R/n	0.002	0.002	0.002	0.002						
Upper-Middle	Mean	0.822	0.799	0.890	0.779	73.2	94.3	78.6	9.3	11430	0.87
Range:	SD	0.027	0.047	0.046	0.064	3.1	4.9	8.6	0.4	4633	0.03
(0.772-0.868)	CV(%)	3.3	5.8	5.1	8.3	4.2	5.2	10.9	4.2	40.5	1.0
n=46	R/n	0.002	0.002	0.002	0.002						
Lower-Middle	Mean	0.713	0.710	0.807	0.623	67.6	85.6	70.9	8.34	5249	0.809
Range:	SD	0.052	0.104	0.093	0.100	6.2	11.5	7.9	0.60	4814	0.081
(0.593-0.772)	CV(%)	7.2	14.6	11.6	16.0	9.2	13.5	11.2	7.2	91.7	2.3
n=45	R/n	0.004	0.004	0.004	0.004						
Low	Mean	0.481	0.490	0.543	0.408	54.4	56.8	49.4	7.05	1385	0.630
Range:	SD	0.071	0.108	0.128	0.100	6.5	16.3	10.5	0.80	979	0.098
(0.340-0.586)	CV(%)	14.7	22.1	23.5	24.5	11.9	28.8	21.2	8.5	70.7	3.0
n=45	R/n	0.005	0.005	0.005	0.005						

From Table 6 we see that the HDI in the first two quartile groups is the least variable in both the absolute (R/n) and the relative (CV%) sense. The same pattern repeats for all component indices. When compared across component indices, the most variable component is the GDP index (GDPx). High similarity in variation of component indices in first two quartile groups results in very tight HDI scores. The average distance between scores in the third quartile group is twice the average distance in the first two, and in the forth group is 2.5 times larger. In order to make better differentiation between countries it is necessary to have more variability in all indices across countries.

To assess the impact of normalization on the differentiation power of the component indices we compare the variability of the indicators with the variability of the corresponding component indices . Across all countries the normalization applied to the LE increased the CV of LEx 1.6 times (=23.5/14.9), the normalization applied to the education indicators reduced slightly the CV of EDUx, and the normalization of the ln(GDP) (together with capping), on average, increased the CV of GDPx two times (=31.2/15.1). The situation is quite different for the first two quartile groups: the CV of LEx increased only 1.35 times for the first and 1.39 for the second group, the CV of the EDUx is reduced by almost 30%, and the CV of the GDPx for the first group increased 1.4 times and for the second about 2 times. Generally, the normalization increased the differentiation power of all indicators except for the education indicators. However, the upper half of countries benefitted a little from normalization[22].

[22] The last two columns of Table 6 contain the basic statistics of the GDP variable and the ln(LE). The idea was to emphasize that the logarithmic transformation of the GDP, as well as of the LE, significantly reduces the variability of the corresponding indicators for the top groups of countries, and thus reduces the differentiation power of the resulting indices.

A conclusion regarding differentiation of developed countries is that the applied indicators have to be more variable across these countries, and the normalization procedure has to increase the variability of the resulting indices in a more or less uniform fashion across all countries.

Anand and Sen (1994) suggest that for countries at different level of human development it may not be necessary to use the same indicators. The justification is based on the fact that high development countries differ very little in literacy rates, and that the difference in life expectancy at birth is becoming rapidly smaller over time. They suggested the following use of indicators:

Table 7. Anand and Sen (1994) proposal for differential choice of indicators

Level of Development		
Low	Medium	High
Life expectancy	Life expectancy Under 5 mortality rate	Life expectancy Under 5 mortality rate Mothers mortality
Adult literacy	Adult literacy Secondary school enrollment	Adult literacy Secondary school enrollment, Tertiary school enrolment
Log GDP (pc) up to int. poverty line	Log GDP (pc) up to int. poverty line Incidence of poverty	Log GDP (pc) up to int. poverty line Incidence of poverty Gini-corrected mean national income

Such an approach combined with differential weighting of common indicators would most likely solve the problem of differentiation of countries at different level of development but at the expense of loss of universality and simplicity.

5. CRITIQUES RELATED TO THE FUNCTIONAL FORM OF THE HDI

Once a set of indicators of dimensions of human development are decided, the next step is to translate them into an index that measures the success or the failure of a country in achieving a desired level of development and to assess how this improves over time. It has been suggested by critiques that the formula used to calculate the HDI is arbitrary, unjustifiable, and incorrect (Chowdhury, 1991, Hopkins 1991, Kelley 1991, Ogwang 1994, and Sager and Najam 1998.) The HDI's components are combined using a simple, unweighted mean – Sager and Najam (1998) write that "the scheme of arithmetic averaging of the dimensions runs counter to the notion of their being essential and, therefore, non-substitutable."

Composite indexing involves several steps after selecting the indicators – functional transformations, scaling (or normalization), weighting, and aggregation. These steps are not necessarily in a sequence, and often they lead to adjustments of the previous steps, alterations, and corrections.

A composite index is characterized by three sets of decisions (Decanq and Lugo, 2008): transformation functions, the weights for combining the component indices, and the amount of substitutability between dimensions. In the following we will review critiques of the HDI form with respect to these sets of decisions and parameters.

5.1 Functional Transformations

Appropriate transformation functions for composite indices usually serve several purposes. First, if the original distribution is skewed, the transformation functions can be employed to make the shape of distribution more normal-like and avoid giving an excessive relative importance to extreme values. For positively skewed distributions, like income, a concave monotone increasing functions such as the logarithm can be very useful. The second purpose is to rescale component variables measured in different measurement units to a common basis before they can be sensibly aggregated. The rescaling will be addressed in subsection 5.2. In the context of social evaluation, one may request a transformation so that the return (gain) of an indicator diminishes at higher levels of attainment.

5.1.1 Capping

Capping is a functional transformation of an indicator defined as: $f(x) = \begin{cases} x, & x \leq c \\ c, & x > c \end{cases}$, and it is aimed at reducing the influence of extreme values. The income variable has been capped at \$40,000 since 1994. At the time when the capping of income was introduced it was not binding. However, in the HDR 2009, even 13 countries had the GDP per capita (\$PPP) above the limit of \$40,000. The HDI has also caps on life expectancy at 85 which is for now non-binding, since there is no country with LE higher than 85 years. The cap for literacy is set at 99% and for gross enrolment is 100%. Both caps for education indicators are binding.

5.1.2 Logarithmic Transformation of Income Variable

In addition to capping, the income enters the HDI as a logarithmically transformed variable. The idea was to consider diminishing marginal utility of income using the logarithmic transformation, which means that the logarithmic transformation will bring closer the notion that an increase of GDP per capita by \$100 in a country where the average income is only \$500 has a much greater impact on the standard of living than the same \$100 increase in a country where the average income is \$5,000 or \$50,000. The logarithmic transformation was used in the first HDR. The motivation for logarithms was stated as "A further consideration is that the indicator should reflect the diminishing returns to transforming income into human capabilities. In other words, people do not need excessive financial resources to ensure a decent living. This aspect was taken into account by using the logarithm of real GDP per capita for the income indicator. (HDR 1990)." The logarithms were replaced by a modified Atkinson step-wise concave transformation of GDP which was used until 1999, when it was replaced back by the logarithmic transformation. This function was criticized by Trablod-Nubler (1991) for its inadequate diminishing of marginal returns. Luchters and Menkhoff (1996) pointed out that the Atkinson step-wise function is not differentiable. The change back to the logarithmic transformation was also motivated by the fact that the Atkinson formula was too severe on middle-income countries.[23] The replacement of the Atkinson formula was explained by Sager and Najam (1998) in the following way "The overall application of the GDP adjustment artificially depresses the relative affluence for wealthy nations so that the gap between the rich and poor countries seems much narrower than it actually is. The result is that the standard-of-living index presents a falsely equitable picture of a world which in fact is more inequitable than ever… As long as it is below that threshold the focus is on ensuring survival and not on adding to human development."

HDR 1999 (p159) lists three advantages of the logarithmic transformation over the Atkinson formula: "the discounting is less severe; all levels of income are discounted uniformly; and middle-income

[23] An excellent review of the treatment of income dimension and implications is given in Anand and Sen (2000).

countries receive recognition for increases in income that, under the Atkinson formula, would have been very heavily discounted."

Using the basic statistics from Table 8 we check on validity of the above statement. We quantify the severity of discounting as a simple ratio of HDI quartile group means to the overall mean for the GDP, ln(GDP) and GDPx (index), we compare this ratio across the quartile groups to assess the uniformity of discounting.

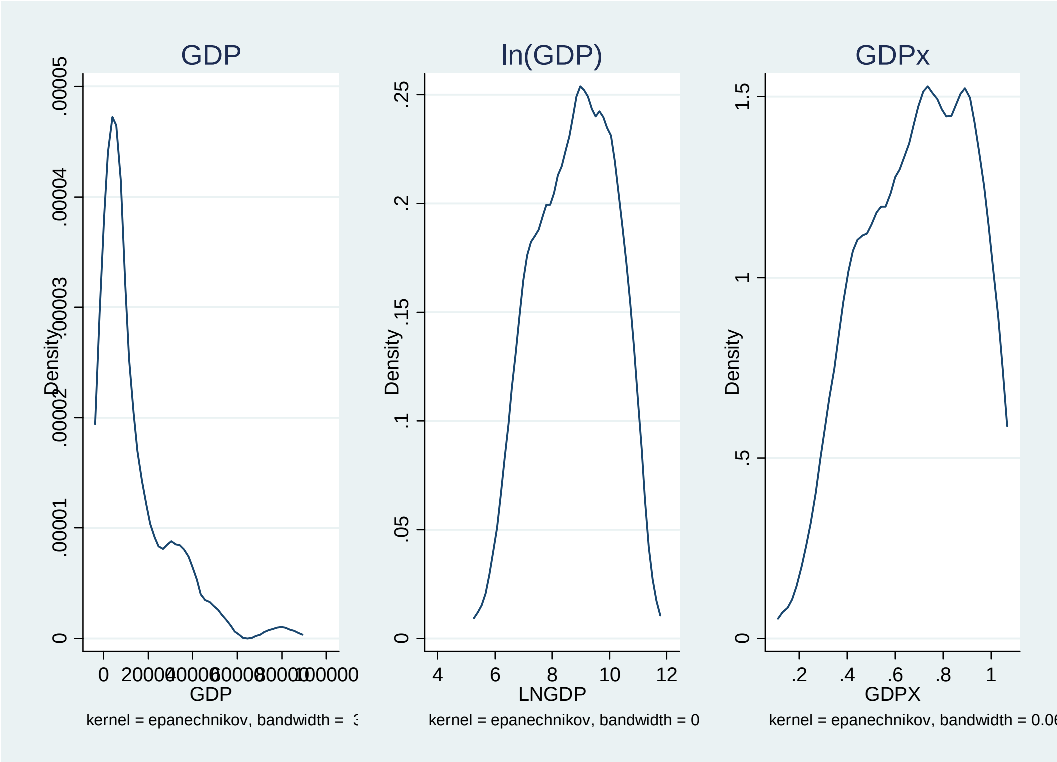
Table 8. Effects of logarithmic transformation of GDP

		GDP	Ln(GDP)	GDPx
All countries (182)	Overall mean	13481	8.77	0.693
	% of overall mean	100	100	100
Highest HDI (46)	Mean	35486	10.39	0.952
	% of overall mean	263.2	118.4	137.4
Upper-Middle (46)	Mean	11430	9.27	0.779
	% of overall mean	84.8	105.7	112.4
Lower-Middle (45)	Mean	5249	8.34	0.623
	% of overall mean	38.9	95.0	89.9
Low (45)	Mean	1385	7.05	0.408
	% of overall mean	10.3	80.4	58.9

From Table 8 we see that the mean GDP for Lower-Middle developed countries is about 38.9% of the overall mean. However, after the logarithmic transformation it becomes 95% of the overall mean of logarithms, and finally after normalization, the mean GDP index is 89.9% of the overall mean GDP index. Apparently, the middle-lower and low developed countries get a considerable boost from logarithmic transformation, which is then slightly reduced by the applied rescaling (normalization) procedure. The directions of change in the first quartile group is opposite: The relative advantage of high GDP countries (2.63 times higher GDP on average) is considerably reduced by logarithmic transformation and capping at $40,000 (only 1.18 times higher ln(GDP) on average) and then the reduction is corrected by the applied rescaling. The real beneficiaries of the combination of logarithmic transformation and rescaling are the countries in the low, lower-middle and upper-middle development groups.

The logarithmic transformation and the normalization move both ends of GDP distribution towards the overall mean. Graph 3 presents the densities of the GDP, ln(GDP) and GDP index. It is evident that logarithm "corrects" most of the skewness of the original distribution. The truncation effect, evident for the density of GDPx, is due to capping.

Graph 3. Densities of distributions of the GDP per capita: original, log-transformed and after rescaling (with capping at $40,000)

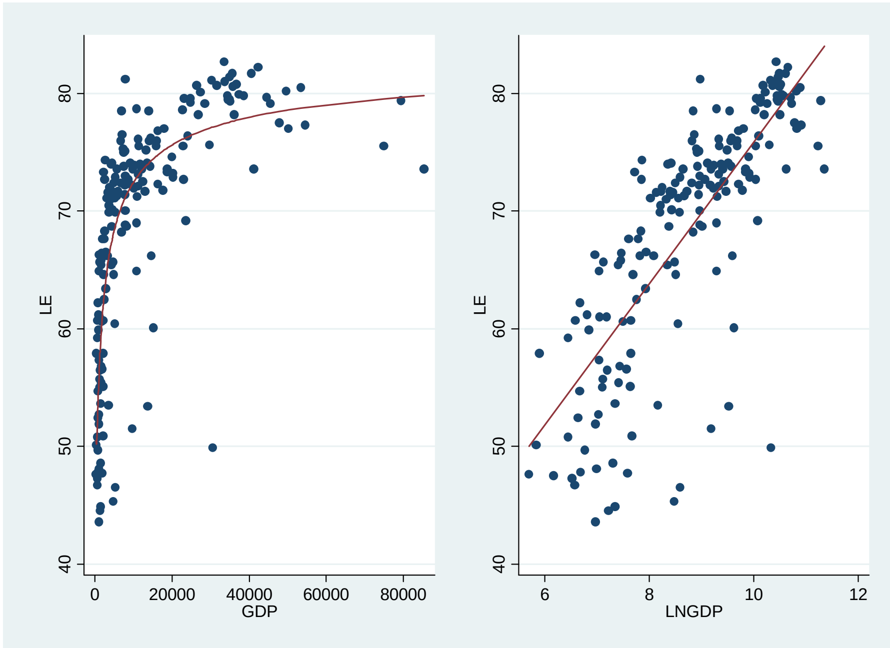


The use of the logarithmic transformation of the per capita GDP is methodologically consistent with the idea that per capita income has sharply decreasing returns to scale and is consistent way of representing this normative view. The use of the log functional form in the HDI is fairly transparent, and is the only functional transformation used on the variables.

5.1.3 Logarithmic Transformation of Other Component Indicators

Income is treated differently from the other variables because of the accepted assumption in economics that increases in income have a diminishing marginal effect on human well-being. Some critics of the HDI have raised the question of why life expectancy and literacy are not transformed to take their diminishing returns into account (Kelley 1991, Acharya and Wall 1994, Srinivasan 1994, Paul 1996, and Noorbakhsh 1998a.) Stenton (2007) quotes an earlier work of Hicks and Streeten (1979) who addressed these concerns noting that for other social indicators, "skewness at the upper end is more limited than it is for income per head… There is practically no limit to how much income a man can receive, but the maximum life span is limited." According to Stanton "they also point out that some social indicators, like life expectancy, capture the costs of both national affluence – for example, heart disease – and destitution." At the country level, life expectancy is also concave in income (Preston, 1957). Graph 4 illustrates this. Similar findings concerning individual health and income are reported in the literature (Marmot and Smith 1991; Marmot and Bosma 1997). The concavity of LE in income indicates a diminishing effect of income on life expectancy. The concavity of LE disappears when the logarithmically transformed income is used.

Graph 4. Concavity of LE in GDP and its linearity in ln(GDP)



As we pointed out in the introduction of this section, one of the main reasons for the functional transformation of indicators is to make their distributions more normal to avoid an excessive impact of the extreme values. The question is whether the logarithmic transformation of life expectancy makes the left-skewed distribution more normal and if not what would be the appropriate transformation. Apparently, as demonstrated by Graph 5, neither the logarithmic transformation, nor the rescaling aftermath makes the distribution of LE across countries more normally distributed.

Although there is no evident a statistical benefit of transforming the LE logarithmically, if such transformation is justified on some other (non-statistical) grounds it may be acceptable given that the transformation doesn't change the shape of the original distribution.

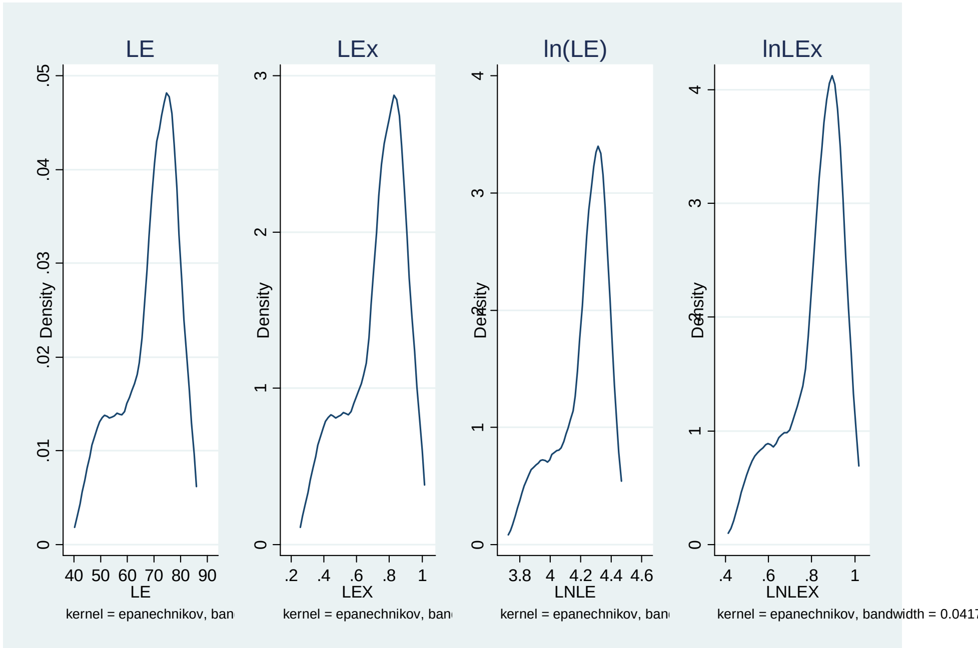
For left skewed distributions, the statistical literature recommends a power transformation, with the power parameter being greater than 1, for example the Cox-Box[24] power transformation could be applicable. However, an acceptable interpretation of any power transformation within the context of HDI poses a problem.

In the case of education indicators, the adult literacy rate and gross enrolment rate have quite different distributions: the adult literacy rate (ALR) is left skewed as it is the case with the distribution of LE, while the gross enrolment ratio (GER) is more normally distributed (see Graph 6). The logarithmic transformation of the ALR wouldn't make the ALR distribution more symmetric, and it would make a GER distribution less symmetric. So, again, there is no statistical reason to apply the logarithmic transformation to education indicators.

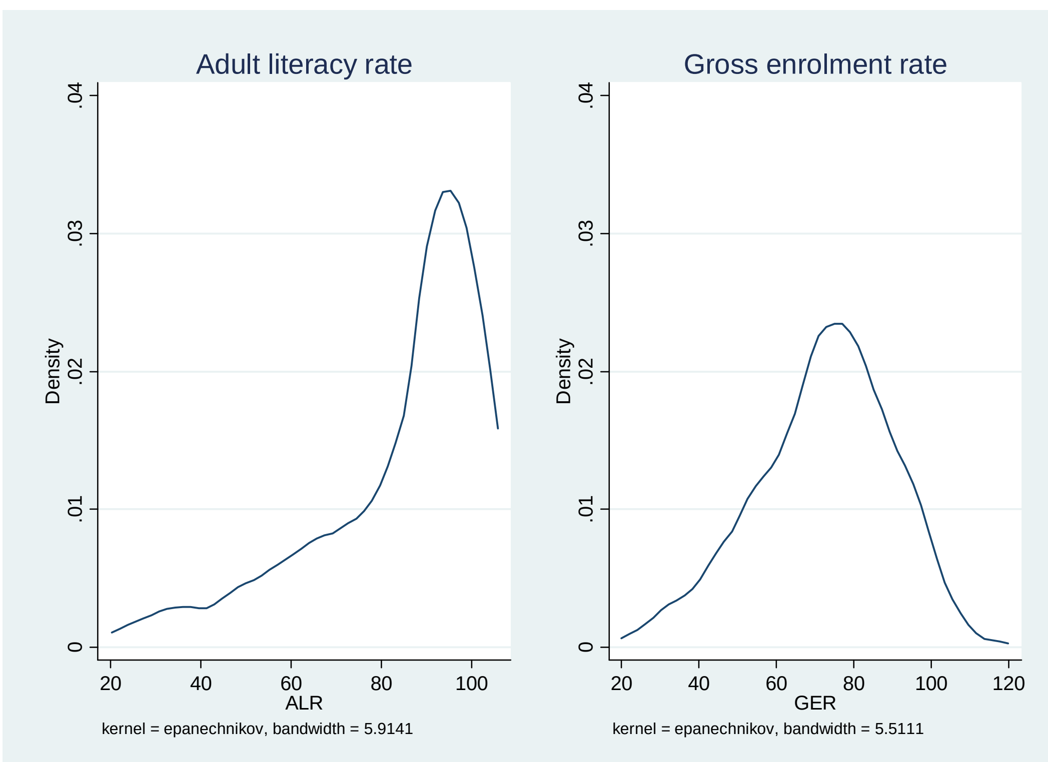
---

[24] Box-Cox (1964) transformation $Y_i^{(\lambda,\alpha)} = \frac{(y_i+\alpha)^{\lambda}-1}{\lambda[gm(y_1 \ldots y_n)]^{\lambda-1}}$, where λ≠0, α is the shift parameter and gm(.) denotes a geometric mean.

Graph 5. Densities of distributions of the LE: original, after rescaling (with Max=85 and Min=25), log-transformed and after rescaling the ln(LE).



Graph 6. Shape of distribution of education parameters



5.2 Scaling (Normalization) of Indicators

Scaling of indicators is necessary for comparability on the same scale. Each indicator is presented in its own units - income in dollars, life expectancy in years, literacy in % of the population, enrollment rate in % of enrolled school-age children. Only after rescaling they can be combined into a single scalar – a composite index. Scaling of indicators is a linear transformation of indicators and is usually applied as the last transformation before combination with other components into a composite index:

$$I(x_c, m_x, M_x) = \frac{x - m_x}{M_x - m_x} = x\frac{1}{M_x - m_x} - \frac{m_x}{M_x - m_x}$$

A key critique of the HDI specification is about the "moving goalposts." (McGillivray and White 1993, Aturupane et al. 1994, Doessel and Gounder 1994, UNDP 1994, Paul 1996, Noorbakhsh 1998a, Bagolin and Comim, 2008). From 1990 to 1993 (see Table 1) the HDI had minimum and maximum values as the rescaling goalposts for all three components based on temporal-variable criteria - the observed minimum and maximum in the current year. Calculating the component indices using minimum and maximum values that change each year makes it difficult to compare between years. Kelley (1991) also noted that such an approach "assumes that little or no progress in human development can be made by the developed countries." Since 1994, all indicators of the HDI were standardized with respect to the fixed goalposts through min-max rescaling. This procedure allows indicators to be expressed in terms of their distance to these points and hence their distance to an "ideal" status. The ideal state of HD as expressed by the current four indicators is: LE at 85 years, literacy at 100%, gross enrolment ratio at 100%, and the GDP per capita in ($PPP) at 40,000. A point of "no HD achieved" is defined as: LE at 25 years, Literacy and Gross enrolment ratio at 0%, and the GDP per capita in ($PPP) at $100.

Fixing the goalposts appeared to solve the problem of comparison of achievements of one country or between countries over time (Anand and Sen, 2000). However, due to frequent data revisions between subsequent years, the Human Development Reports always advise that the yearly published HDI scores are not strictly comparable. In order to account for this problem, HDRs give revised estimates of the HDI values for purposes of comparison studies. The estimates of the HDI are further revised and given with 5-year intervals which are comparable across countries and time. Thus comparable estimates of the indices are not available on a year-to-year basis, however, revised comparable estimates for selected years 1985, 1990, 1995 etc., are given in succeeding HDRs.

A viable alternative to rescaling via fixed goalposts is to use the maximum and minimum values observed over the period for which the time-series of the HDI and its components are presented. In such a case each indicator for a country *a* is rescaled as:

$$I(x(a), m_{x,T}, M_{x,T}) = \frac{x(a) - m_{x,T}}{M_{x,T} - m_{x,T}}, \qquad (1)$$

where $m_{x,T}$ and $M_{x,T}$ denote minimum and maximum of indicator *x* across all countries and over the considered time interval, say T. To maintain the comparability between existing and new data, the HDI has to be always recalculated for existing data.

As a linear transformation, the rescaling of an original indicator has certain important properties: (i) normalization: $I(m, m, M) = 0$ and $I(M, m, M) = 1$; (ii) monotonicity: given *m* and *M*, an increase in *x* increases the index *I*; (iii) translation invariance: $I(x+c, m+c, M+c) = I(x, m, M)$; (iv) homogeneity: for c>0, $I(cx, cm, cM) = I(x, m, M)$.

The idea of rescaling of life expectancy, education, and the natural log of income for conversion into indices is also to establish a range of variation from 0 to 1. When using the predetermined fixed posts, the rescaled indices may range within different subintervals of (0,1). For example, in 2009, the range of the life expectancy index was between 0.306 and 0.961, for education index - between 0.282 and 0.993, and for income - between 0.182 and 1.000.

Thinking in terms of power of differentiation between countries (see Section 4.3) using the component indices, we can see that the average distance between all the countries with respect to the LE index is 0.0036 (given that the range for the 182 countries is 0.306-0.961), with respect to education index it is 0.0039, and with respect to income index it is 0.0044. This implies that differentiation by income index is the most significant driver of differences in the HDI – which, somehow, defies the underlying motivation of the HDI and the claim that all three dimensions are of equal importance.

The HDI scores are extremely sensitive to the choice of goalposts. Panigrahi and Sivramkrishna (2002) note that it affects scores and rankings. "The sensitivity to limits is not only that countries move from one level of development to another with their new HDI values, but the possibility of an *interchanging* or *re-ordering* of HDI-based country ranks. In other words, for given indicator values, with one set of limits we may have a situation where country 1 is ranked better than country 2; but with a different set of limits, country 2 is ranked better than country 1." (See also Herrero et al (2010, page 9).

5.3 Effects of transformation and rescaling

In the following we check several options that may improve the performance of the HDI. We are using the same indicators that were used in the HDR 2009: the GDP per capita (in$PPP), the adult literacy and gross enrolment rate, and life expectancy. The GDP enters rescaling as logarithmically transformed, while other indicators are tried both, untransformed and logarithmically transformed. There is no capping of any indicator. The rescaling is done in three different ways – using only the observed minima and maxima, reducing the minima for 5% to avoid 0 score on any component index, and symmetrically increasing the maxima for 5% to give a leeway to leaders to continue excelling in HD. We compare performance of the resulting HDIs at the level of quartile groups by calculating some descriptive statistics and studying the distributional shapes of HDI and its components.

Table 9 contains the nine combinations of functional transformations and rescaling we examined. We applied them to indicators that were used in the HDR 2009. So, for example HDI5 has ln(LE), EDU, and ln(GDP) and all are rescaled by using the corresponding minimum reduced by five percent and the maximum. The rescaling is defined by (1) with a modification for minimum. We repeat this analysis in the next section with a choice of indicators proposed for 2010.

Table 9. Combinations of logarithmic transformation and rescaling

	min, max	0.95 min, max	0.95 min, 1.05 max
LE, EDU, ln(GDP)	HDI1	HDI2	HDI3
ln(LE), EDU, ln(GDP)	HDI4	HDI5	HDI6
ln(LE), ln(EDU), ln(GDP)	HDI7	HDI8	HDI9

The next table (Table 10) contains the basic summary statistics for nine HDIs displayed for all countries and for quartile groups.

i) With respect to the mean value, HDI8 is the largest over all countries as well as over the quartile groups, while the HDI3 is the smallest. The maximum of the HDI8 range of values is the highest, while the maximum for the range of HDI3 values is the lowest. This finding can be extended to all three HDIs (HDI7, HDI8, HDI9) being with the highest mean and maximum values, and HDI1, HDI2 and HDI3 being with the lowest means and maxima. Evidently, using the logs in all three dimensions increases the values of indicators.

ii) Using the modified minima and maxima (multiplied by 0.95 and 1.05 respectively) as rescaling parameters, the values of HDIs are reduced. If only minima are modified, the HDI increases. Keeping the rescaling parameters equal to minimum and maximum, the HDI results in values in the middle.

iii) The range of variation is an interval defined by minimum and maximum of the HDI. The larger range implies a better differentiation between country scores. At the level of the entire distribution of the HDI across countries, the average distance between countries, *R/n,* remains pretty steady at 0.004 regardless of the functional transformation and rescaling parameters. The same uniformity of *R/n* holds for the quartile groups with the high developed and the lower-middle developed countries. The upper-middle and the low developed quartile groups have a variation of values of R/n across different HDIs: for the upper-middle group HDI1, HDI2 and HDI3 outperformed other HDIs, and for low developed countries the HDI7 had the best performance.

iv) Similarly to the range of variation, two measures of variability - standard deviation and the coefficient of variation are good indicators of differentiation power of the HDI. Here too, the higher variation indicates the higher power of differentiation. The HDI1 has uniformly the highest SD and CV while the HDI9 has the lowest. We can also say that the HDI1, HDI2 and HDI3 outperform other choices of HDIs with respect to the differentiation power.

v) It seems that the HDI1, HDI2, and HDI3, using only the GDP logarithmically transformed, have better differentiation properties from others.

Table 10. Comparison of 9 HDIs based on different combinations of transformations and rescaling

		HDI1	HDI2	HDI3	HDI4	HDI5	HDI6	HDI7	HDI8	HDI9
All (182)	Mean	0.611	0.621	0.586	0.628	0.639	0.608	0.662	0.674	0.649
	SD	0.220	0.214	0.202	0.216	0.209	0.200	0.207	0.199	0.192
	CV	35.971	34.453	34.530	34.470	32.711	32.821	31.333	29.526	29.581
	Min	0.114	0.136	0.128	0.125	0.156	0.151	0.142	0.174	0.167
	Max	0.928	0.931	0.882	0.935	0.938	0.896	0.952	0.955	0.921
	R/n	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
Highest HD (46)	Mean	0.854	0.857	0.810	0.862	0.866	0.826	0.883	0.887	0.855
	SD	0.045	0.044	0.041	0.041	0.040	0.038	0.038	0.037	0.036
	CV	5.244	5.087	5.116	4.780	4.613	4.653	4.307	4.140	4.166
	Min	0.776	0.780	0.736	0.791	0.798	0.761	0.812	0.818	0.787
	Max	0.928	0.931	0.882	0.935	0.938	0.896	0.952	0.955	0.921
	R/n	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Upper-Middle (46)	Mean	0.713	0.719	0.679	0.730	0.737	0.702	0.761	0.768	0.739
	SD	0.036	0.036	0.034	0.034	0.033	0.032	0.032	0.031	0.030
	CV	5.117	4.943	4.997	4.722	4.543	4.600	4.202	4.008	4.034
	Min	0.653	0.660	0.621	0.672	0.680	0.646	0.709	0.718	0.690
	Max	0.774	0.779	0.736	0.789	0.794	0.757	0.810	0.816	0.787
	R/n	0.003	0.003	0.002	0.003	0.002	0.002	0.002	0.002	0.002
Lower-Middle (45)	Mean	0.568	0.579	0.545	0.590	0.602	0.572	0.630	0.643	0.618
	SD	0.073	0.071	0.067	0.071	0.069	0.065	0.066	0.063	0.060
	CV	12.899	12.268	12.198	12.114	11.402	11.372	10.507	9.807	9.770
	Min	0.411	0.425	0.400	0.437	0.453	0.431	0.485	0.501	0.480
	Max	0.652	0.660	0.621	0.671	0.680	0.646	0.709	0.718	0.690
	R/n	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
Low (45)	Mean	0.280	0.299	0.282	0.299	0.322	0.306	0.346	0.371	0.357
	SD	0.085	0.083	0.078	0.090	0.086	0.081	0.096	0.092	0.088
	CV	30.485	27.734	27.639	29.911	26.662	26.606	27.821	24.729	24.653
	Min	0.114	0.136	0.128	0.125	0.156	0.151	0.142	0.174	0.167
	Max	0.410	0.424	0.396	0.436	0.450	0.426	0.484	0.500	0.479
	R/n	0.007	0.006	0.006	0.007	0.007	0.006	0.008	0.007	0.007

5.4 Aggregation

The aggregation of the HDI component indices at the level of a country is done by the most widespread linear aggregation:

$$HDI(x,y,z) = \frac{1}{3}\left[I(x,m_x,M_x) + I\left(y,m_y,M_y\right) + I(z,m_z,M_z)\right]. \quad (2)$$

A more general form of a well-being index expressed as a general mean can be presented as

$$M(x_1,\dots,x_K) = \begin{cases} \left\{\sum_k^K w_k[f(I(x_k))]^\beta\right\}^{\frac{1}{\beta}}, \text{for } \beta \neq 0 \\ \prod_k^K [f(I(x_k))]^{w_k}, \ \text{for } \beta = 0. \end{cases} \quad (3)$$

For sake of simplicity the component indices are expressed without an explicit writing of minimum and maximum used for rescaling, i.e., $I(x,m_x,M_x) = I(x)$. The weights w_k sum to one. Real-valued function *f (.)* brings an additional flexibility to the general mean towards satisfying important properties some of which will be stated below. Parameter β is related to the elasticity of substitution between the component indices, $\sigma = \frac{1}{1-\beta}$. (See Arrow et al. , 1961.)

When β=1, the case of the HDI, the elasticity of substitution is infinite thus making the dimensional indices to be perfect substitutes Expression (2) suggests that the degree of substitutability is the same for all pairs of dimensions which may not be sensible assumption to make. The marginal substitutability is discussed in more details in section 5.6. In the case of the HDI, function *f*(.) is an identity function.

Some authors explored different axiomatic characterizations of the HDI as an index number, mostly based on Sen (1976) axiomatic ordinal conceptualization of a welfare function. The HDI satisfies a list of axioms termed "basic requirements for a good index": (i) normalization - if achievements in all dimensions are equal to *d* then HDI = *d;* (ii) linear homogeneity- if all achievements are doubled or tripled, HDI is doubled or tripled, respectively; (iii) symmetry in persons - personal identity does not matter; (iv) symmetry in dimensions - all dimensions are equally important; (v) monotonicity- if an achievement in one dimension increases, while others remain same, HDI increases; (vi) population replication invariance - if the same population is replicated more than once, the HDI remains the same; (vii) subgroup consistency - if the human development of one subgroup rises and the other is unaltered, then overall human development must rise; (viii) continuity – HDI does not change abruptly due to a change in any of its dimension.

Some authors add more axioms to gain better properties for the social evaluation index. For example, Chakravarty (2003) requests that - (ix) for any indicator, "the achievement difference is greater at lower attainment levels, given that the values of other attributes remain fixed." He then suggests that for β =1, function *f(.)* in (3) is a power function, i.e., $f(I(x_k))$ = $[I(x_k)]^r$, 0<*r*<1. When *r*=1, the index is equal to HDI (assuming the equal weighting), but the requested property is not satisfied.[25] Herrero et al (2008) introduce another set of axioms to characterize a family of social evaluation indices – (x) minimal lower boundedness, that requires the index to take on its minimal value whenever a characteristic attains its minimum level across all members of the society; (xi) separability, that establishes that the ranking of two social states in which one of the characteristics takes on identical values, does not change with a change of this common value by another common vector, and (xii) component-wise quasi-concavity, that makes the evaluation index non-increasing with the dispersion of the different variables. They then prove that the HDI based on the arithmetic means does not satisfy these axioms, and that the only index that satisfies them is the geometric mean.

5.5 Critiques of Equal Weighting

"Since any choice of weights should be open to questioning and debating in public discussions, it is crucial that the judgments that are implicit in such weighting be made as clear and comprehensible as possible and thus be open to public scrutiny. "(Anand and Sen, 1997)

One of major critiques about the HDI's aggregation concerns the equal weights assigned to the three components. HDR 1991 (UNDP 1991: 88) justifies the equal weighting procedure by explaining that the

[25] For the limiting case, when *r*→0, Kakwani (1993) was suggesting the logarithmic transformation of the following form $[I(x_k)] = \log\left(\frac{x_k}{m_x}\right) / \log\left(\frac{M_x}{m_x}\right)$.

three indices are equally important, so that they deserve equal weights. Streeten (1994), echoes this line and defends use of a simple average stating that it is a good tool for focusing on decreasing gaps between countries, and that there is a political appeal to a simple method. Kelly (1991) recognizes that it is difficult to justify any set of weights but that testing the sensitivity of the HDI and ranking to alternative weights would have been useful.

Several critics have found the explanation of this kind unpersuasive. Chowdhury (1991), for example, notes a paradox: "If a composite index is sensitive to weights, then one must be able to offer a solid defense of one's chosen weights if the index is to be taken seriously. On the other hand, if the index is relatively robust, this would imply that the components are correlated, so that aggregation is pointless – any component would carry pretty much the same information."

Decanq and Lugo (2009) distinguish three approaches to set the component weights in a composite index: normative, data-driven, and hybrid. The approach taken by the HDI is normative, i.e., it depends on the value judgments about the trade-offs and are not based on the actual distribution of the achievements in the society under analysis. Data-driven weights are a function of the distribution of the achievements in the society and are not based, at least explicitly, on any value judgment about how the trade-offs between the dimensions should be. Hybrid approaches depend both on the distribution of the achievements and on some form of valuation of these achievements.

In the same venue, some critiques are concerned with the fact that the HDI component indices have different variability which is completely ignored by the HDI being a simple average (Biswas and Caliendo, 2001). Panigrahi and Sivramkrishna (2002) suggest standardizing each indicator to have the same variance before combining them. Noorkbakhsh (1998b) compares several different methods of deriving a composite index, including the PCA and Borda method, and concludes that given that the ranks based on all compared methods are very similar, the equal weighting of the HDI is acceptable. HDR 1993 also reports the results of PCA studies and concludes that these support equal weighting. In the following subsection we are addressing issues related to unequal variability of the component indices and its impact on differential weighting.

As noted in subsection 5.2, the range of variation of a component index affects its implicit weight in the composite index. The implicit weight is actually masked by the explicit equal weight. A higher variation of a component index implies its higher implicit weight. If the idea is to keep all components equally weighted, a correction for the differences in implicit weights is needed. In Table 11 we present the range of variation and the standard deviation of each component index (LEx, GDPx and EDUx). We also present the range of variation for the HDIs.

Table 11. Range of variation and standard deviation for the HDI and its components

	Fixed goalposts*	
	Range	SD
HDI(LEx,GDPx,EDUx)	(0.340, 0.971)	0.174
LEx	(0.310, 0.961)	0.170
GDPx	(0.182, 1)	**0.216**
EDUx(ALRx,GERx)	(0.282, 0.993)	0.176
ALRx	(0.262, 1)	**0.190**
GERx	(0.255, 1)	0.173

* For LE: [25,85]; for ALR and GER: [0,100]; for GDP(pc) [100, 40,000]

It is clear that the GDP index varies in the largest interval and has the largest standard deviation, and consequently its implicit weight is higher. If the objective is to have the final weights equal, the overall GDP weight would need a correction by having a lower index weight.

5.6 Marginal Substitutability

It is well-known that different achievements in same dimensions may lead to the same level of Human development, because the small changes in the achievements in different dimensions can compensate each other (Ravallion, 1997, Lind 2004).

In the case of an equally weighted arithmetic mean, such as HDI, the marginal rate of substitution between two dimensions, say *x* and *y*, denoted as $R_{x,y}$, is the amount of dimension x a country is "willing" to trade-off for a unit of dimension *y* while maintaining the same level of development, i.e.,

$$R_{x,y} = \frac{\partial HDI(x,y,z)}{\partial y} / \frac{\partial HDI(x,y,z)}{\partial x}$$

If *x* is income, then

$$\frac{\partial HDI(x,y,z)}{\partial x} = \frac{1}{3} \frac{1}{\ln(M_x) - \ln(m_x)} \cdot \frac{1}{x},$$

where M_x and m_x are the maximum and the minimum income across all the countries, respectively. If *y* is life expectancy, then

$$\frac{\partial HDI(x,y,z)}{\partial y} = \frac{1}{3} \frac{1}{M_y - m_y}$$

Thus the HDI allows the following rate of compensation in GDP pc ($PPP) for one year of life :

$$R_{x,y} = \frac{\ln(M_x) - \ln(m_x)}{M_y - m_y} \cdot x$$

From this expression it is clear that the rate of compensation (or substitution) between dimensions which are expressed in their original units, due to the log transformation, linearly depends on the level of income achieved by the country. If we substitute the values of goalposts with the fixed values used over the last 10 years, we compute the rate of substitution as being about (6/60=)10% of the country income, so a country can compensate a difference of one year in life expectancy with 10% of its GDP.

Similarly, we can derive the marginal rate of compensation between other pairs of dimensions:
GDP(x) for one percent of Literacy (z):

$$R_{x,z} = \left(w_z / \frac{1}{3}\right) \frac{\ln(M_x) - \ln(m_x)}{M_z - m_z} \cdot x$$

Here w_z represents the explicit weight of the literacy in the HDI. In the HDI until 2010, literacy is weighted by 2/9, with the maximum being set at 100 and minimum at 0, so the overall rate of compensation is (4/100)=4%, meaning that a country can use an increase of about 4% of its GDP pc($PPP) to compensate for an 1% of literacy of its population.

The compensation rate between non-income (non-log transformed) dimensions is constant in absolute terms. For example, the rate of trade-off between life expectancy (y) and literacy (z) is

$$R_{y,z} = 2/3\,\frac{M_y - m_y}{M_z - m_z}$$

Using the values from the goalposts used in recent years the compensation rate is 0.4. Thus for a 1% of literacy a country has to trade 0.4 years of LE.

An important finding from this analysis is that the rate of trade-off between any two HDI indicators is independent of the other two indicators, thus allowing the assessment of the marginal contribution of each variable separately.

By computing the marginal rates of substitution of health and education with income, it is possible to obtain the implicit prices of the components. While this approach has a tradition in estimating the monetary equivalents of health and education in the economic literature, it is not the case with the well-being literature. Such a price-based analysis of substitution between HDI components was first proposed by Srinivasan (1994a) and echoed by Ravallion (1997). However, the price-based reasoning is in contradiction with the overall idea of Human Development and shouldn't be a part of any HD strategy to achieve higher HDI as it was stated in the HDR 1993 (page 110). The HDI is not a production or a utility function to be maximized, it is rather a capability index and the trade-offs should be interpreted differently. Foster and Sen (1997) argued that even though the implicit prices can be obtained from the form of the HDI they are inappropriate for well-being comparisons because the well-being indices (e.g., HDI) are not price-based indices at the first place.

5.7 Array instead of Composite Index

There are critics that are opposed to the idea of aggregating indicators into a single composite index. Their key objection to aggregation is what they see as the arbitrary nature of the weighting process by which the variables are combined, but also a reduction in information contained in the composite index. Saisana and Tarantola (2002) listed several objections against composite indices (CI), such as: (i) CI may send misleading, non-robust policy messages if they are poorly constructed or misinterpreted. (ii) The simple "big picture" results which CI shows may invite politicians to draw simplistic policy conclusions. (iii) The construction of CI involves stages where judgment has to be made: the selection of sub-indicators, choice of model, weighting indicators and treatment of missing values etc. (iv) There could be more scope for disagreement among member states about CI than on individual indicators, because the selection of sub-indicators and weights could be the target of political challenge. (v) The CI increase the quantity of data needed because data are required for all the sub-indicators and for a statistically significant analysis.

Thus in order to make the HDI useful for policy evaluation and targeting, sensitivity analyses are needed so that only robust variants of the HDI are used. Also, the HDI should be always presented and used in combination with its component indices, especially in attempts to draw sophisticated policy conclusions. Whenever a judgment is used in the process of designing the composite index, it should be made transparent and based on sound statistical principles, as much as possible.

6. SUMMARY AND A NEW PROPOSAL

We reviewed a number of discussions about the composite well-being and performance indices, the Human Development Index in particular, published over the last twenty years. They ranged from those related to the choice of indicators to those relevant to the structure of the index itself. The existing frameworks differ in which important concepts they focus to, and it is hard to imagine a framework that includes all important concepts and relationships.

The paper also highlights that human development is complex, heterogeneous and qualitative evolution that typically resists simplification. Aggregated indices are supposed to aid an understanding and presentation of such evolution, but may also mislead and need a careful interpretation. The purpose of the HDI is to help researchers, policy analysts and, generally, governments and citizens to think systematically about development and related public policies. So the communicability of the index remains the key feature.

A careful review of the HDI critiques, lead us to a new proposal of the HDI, one that is considered a better measure that quantifies the progress in three essential dimensions of human development. It recognizes also availability of better data today than twenty years ago. A modified version of the HDI should better capture the current and potential future capabilities in education of modern societies and the income that is available to the country. Also we would like the new measure to reduce the variation in the implicit weights and to allow less substitutability between dimensions. Finally, the new index should award more the even development across dimensions, that is if the two countries have the same average of the dimensional indices, a country with less heterogeneity across dimensions should be ranked better.

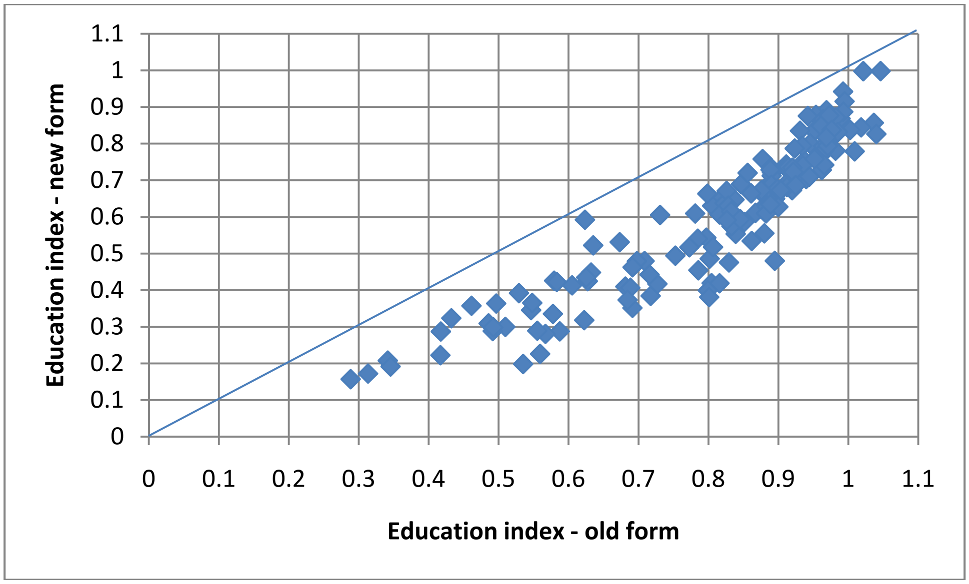
The following changes are proposed: In the knowledge dimension mean years of schooling replaces literacy, and gross enrolment is recast as expected years of schooling—the years of schooling that a child can expect to receive given current enrolment rates. Mean years of schooling is estimated more frequently for more countries and can discriminate better among countries than the literacy rate (which has become less useful precisely because of the progress observed), while expected years of schooling is consistent with the reframing of this dimension in terms of years. Ideally, measures of the knowledge dimension would go beyond estimating quantity to assessing quality, as several National and Regional Human Development Reports (HDRs) have done. For example, the 2003 Arab States HDR constructed a measure that captures both the quantity and quality of education, adjusting mean years of schooling with average test scores and including indicators related to media, communication and scientists trained. But good measures of education quality do not exist for enough countries—cross-national assessments of science, mathematics and reading levels of young people are valuable but scarce in coverage and irregular in frequency. Graph 8 presents the relationship between the old education index[26] computed from adult literacy and the total gross enrolment and the new one[27] which uses mean years of schooling and school life expectancy.

---

[26] The old education index is computed as arithmetic mean of adult literacy and total gross enrolment ratio taken with weights 2/3 and 1/3 respectively.

[27] The new education index is a geometric mean of the mean years of schooling normalized by the maximum (13.2 years and

Graph 8. New education index vs. old education index

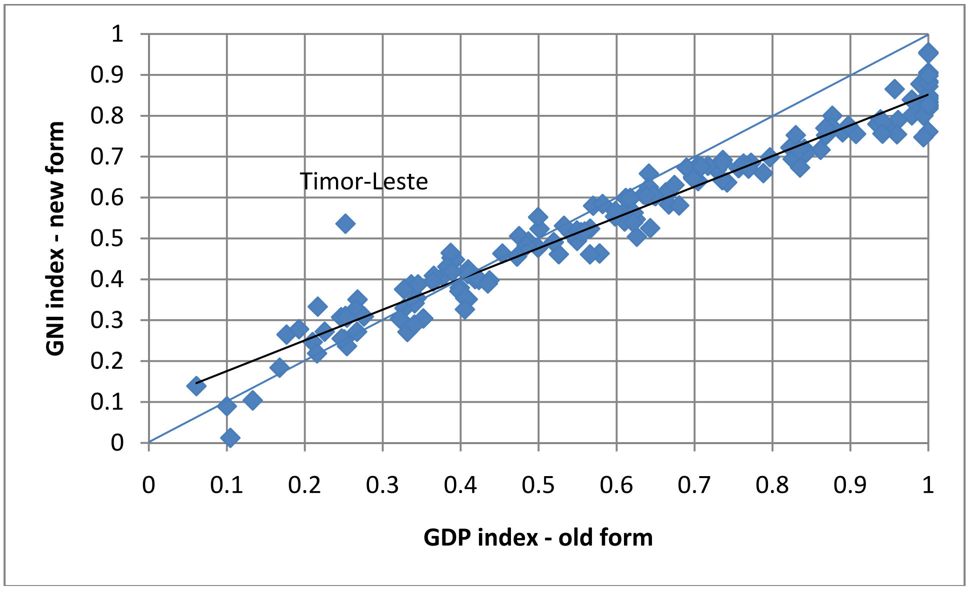


Alternative measures of the ability to enjoy a healthy life were also investigated, but we found no viable and better alternative to life expectancy at birth.

To measure the standard of living, gross national income (GNI) per capita replaces gross domestic product (GDP) per capita. As discussed in section 3.2, in a globalized world differences are often large between the income of a country's residents and its domestic production.

Graph 9 illustrates the relationship between income index based on the GDP and the old normalization using the capping at $40,000 and the minimum of $100 and the new index based on the GNI and using minimum and maximum observed over the last 20 years. Apparently there is no country with the index equal to 1 now, while before there were even 13 countries due to the binding capping. Also, for the high income countries the new income index has lower values than before, while for most of low income countries the new income index is higher. Timor-Leste seems to be an outlier due to much higher GNI than the GDP.

Graph 9. Relationship between the new income index based on the per capita GNI and the old one based on per capita GDP. Both GDP and GNI are given in PPP terms



The practice of using the log of income is maintained: income is instrumental to human development but higher incomes have a declining contribution to human development. This pertains to the diminishing marginal utility of income discussion in section 5.1.1. We have also shifted the maximum values in each dimension to the observed maximum, rather than a predefined cut-off beyond which achievements are ignored.

We also reconsidered how to aggregate the three dimensions, following the critiques described in section 5. A key change was to shift to a geometric mean (which measures the typical value of a set of numbers) - thus the proposal (accepted for 2010 HDI) is the geometric mean of the three dimension indices instead of the arithmetic mean. Poor performance in any dimension is now directly reflected in the HDI, and there is no longer perfect substitutability across dimensions. This method captures how well rounded a country's performance is across the three dimensions. As a basis for comparisons of achievement, this method is also more respectful of the intrinsic differences in the dimensions than a simple average is. It recognizes that health, education and income are all important, but also that it is hard to compare these different dimensions of well-being and that we should not let changes in any of them go unnoticed.

The computation of dimension indices remains based on the normalization by the range of variation of the indicators. Minimum and maximum values (or goalposts) are set in order to transform the indicators into indices between 0 and 1. Because the geometric mean is suggested for aggregation, the maximum value does not affect the relative comparison (in percentage terms) between any two countries or periods of time. The maximum values are set to the actual observed maximum values of the indicators from the countries in the time series, that is 1980–2010. The minimum values will affect comparisons, so values that can be appropriately conceived of as subsistence values or "natural" zeros should be used. Progress is thus measured against minimum levels that a society needs to survive over time. The minimum values are set at 20 years for life expectancy, at 0 years for both education variables and at $163 for per capita GNI. The life expectancy minimum is based on long-run historical evidence from

Maddison (2010) and Riley (2005).[28] Societies can subsist without formal education, justifying the education minimum. A basic level of income is necessary to ensure survival: $163 is the lowest value attained by any country in recorded history (in Zimbabwe in 2008) and corresponds to 45 cents a day, just over a third of the World Bank's $1.25/day poverty line.

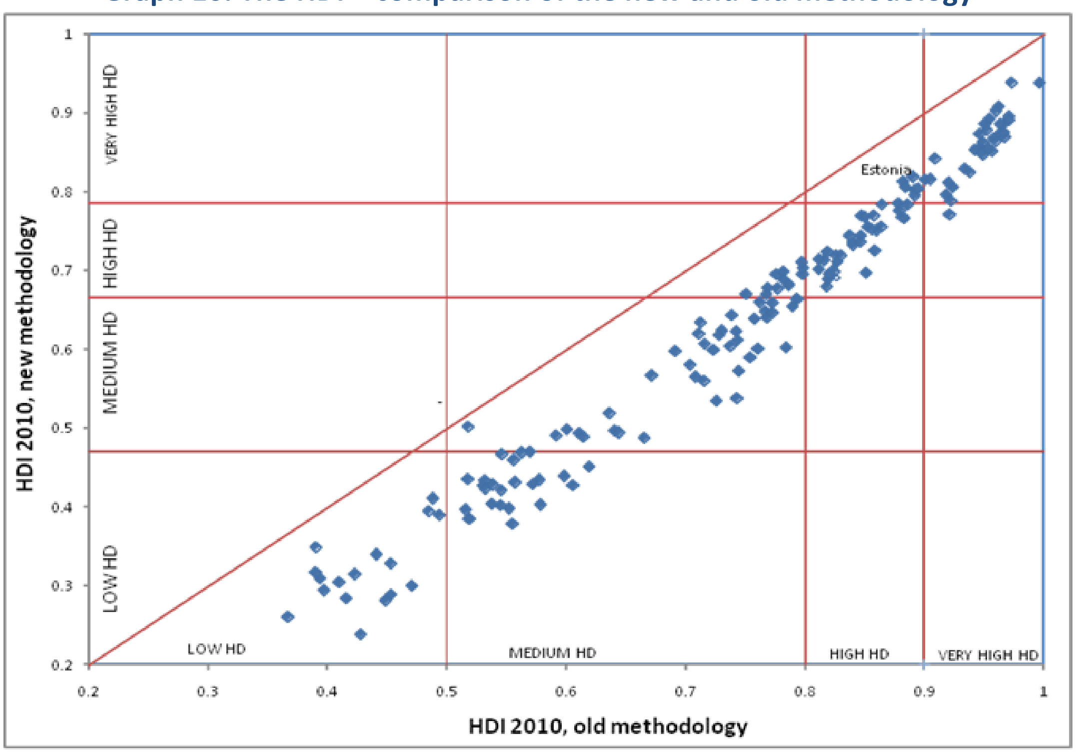
Having defined the minimum and maximum values, the dimension indices are calculated using expression (1). For education, expression (1) is applied to each of the two subcomponents, creating their geometric mean and then applying expression (1) again. This is equivalent to applying equation (1) directly to the geometric mean of the two subcomponents.

The HDI is the geometric mean of the three dimension indices: $(I_{LE} \cdot I_{EDU} \cdot I_{INC})^{1/3}$

As we already pointed out, the geometric mean embodies imperfect substitutability across all HDI dimensions. It thus addresses one of the most serious criticisms of the linear aggregation formula, which allowed for perfect substitution across dimensions. Some substitutability is inherent in the definition of any index that increases with the values of its components.

The methodological improvements in the HDI, using new indicators and the new functional form, result in substantial changes (see Graph 10). Adopting the geometric mean produces lower index values, with the largest changes occurring in countries with uneven development across dimensions. The geometric mean has only a moderate impact on HDI ranks. Setting the upper bounds at actual maximum values has less impact on overall index values and has little further impact on ranks.

Graph 10. The HDI – comparison of the new and old methodology



[28] Lower values have occurred during some crisis situations (such as the Rwandan genocide) but were obviously not sustainable.

Bibliography

Acharya, A., Wall, H.J (1994). An Evaluation of the United Nations' Human Development Index. *Journal of Economic and Social Measurement* 20: 51-65.

Aguna-Garcia, C. and M. Kovacevic (2010). Sensitivity and Reliability Analisys of HDI. Manuscript.

Anand, P. and Sen, A.K. (1994). Human Development Index: Methodology and Measurement. HDR Occasional Papers.

Anand, S., and A. K. Sen (1997): Concepts of human development and poverty: A multidimensional perspective. *Human development papers, Human Development Report Office (UNDP), New York*

Anand,S. and Sen, A.K. (2000).The Income Component of the Human Development Index*. Journal of Human Development*, Vol. 1, No. 1, 2000, 83-106.

Anand, S. and Sen, A.K. (2000). Human Development and Economic Sustainability. *World Development* Vol. 28, No. 12, pp. 2029-2049

Anthopolos, R. and Becker, C.H.(2010). Global Infant Mortality: Correcting for Undercounting**.** *World Development***,** 38, 467-481 Aturupane, H,. Glewwe,P., Isenman, P. (1994) "Poverty, Human Development, and Growth: An Emerging Consensus?" *American Economic Review* 84(2): 244-249.

Arrow, K.J., Chenery, H.B., Minhas, B.S., and Solow, R.M. (1961). Capital-labor substitution and economic efficiency. *Review of Economics and Statistics,* 43(3), p. 225-250.

Aturupane, H., Glewwe, P., and Isenman, P. (1994). Poverty, human development, and growth: An emerging consensus? *American Economic Review, Papers and Proceedings*, 84, 244-249.

Bagolin, I., Comim, F. (2008). Human Development Index (HDI) and its family of indexes: an evolving critical review. *Revista de Economia, v. 34, n. 2 (ano 32), p. 7-28, maio/ago. 2008. Editora UFPR.* Also available at http://ojs.c3sl.ufpr.br/ojs2/index.php/economia/article/view/12293/8511

Barro, R.J. and Lee,J-W. (2000). International Data on Educational Attainment: Updates and Implications. *Working Paper 42.* Center for International Development: Harvard University.

Barro, Robert and Jong-Wha Lee, April 2010, "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.

Biswas, B., Caliendo,F. (2001) "A Multivariate Analysis of the Human Development Index." *Indian Economic Journal* 49(4): 96-100

Booysen, F. (2002). An overview and evaluation of composite Indices of development. *Social Indicators Research* **59:** 115–151.

Chakravarty, S.R. (2003). A Generalized Human Development Index. *Review of Development Economics*, 7, 99-114.

Chhibber, A. and Laajaj, R. (2007). A Multi-Dimensional Development Index: Extending the Human Development Index with Environmental Sustainability and Security. Mimeo. UNDP.

Chiswick, B.R.(1997). Interpreting the coefficient of schooling in the human capital earnings function. *Working Paper of the Human Development Department*, World Bank.

Chowdhury, O.H. (1991). Human Development Index: A Critique. *The Bangladesh Development Studies* 19(3): 125-7.

Dasgupta, P. and Weale, M. (1992). On Measuring the Quality of Life. *World Development,* Vol. 20, No. 1, pp. 119-131.

Decancq, K. and Lugo, M.A. (2010) Weights in Multidimensional Indices of Well-Being: An Overview, *Econometric Reviews* (accepted for publication)

Despotis, D. K. (2005). A reassessment of the Human Development Index via data envelopment analysis. *Journal of the Operational Research Society*, 56(8), 969.980.

Diener, E. and Suh, E. (1997). Measuring Quality of Life: Economic, Social, and Subjective Indicators. *Social Indicators Research*, Vol. 40, No. 1/2, pp. 189-216

Diener, E. and Suh, E. (1997). Sustainability and Human Well Being: Exploring the Connections. *Social Indicators Research*, Vol. 40, No. ½, pp. 189-216

Doessel, D.P., Gounder, R. (1994) "Theory and Measurement of Living Levels: Some Empircal Results for the Human Development Index." *Journal of International Development* 6(4): 415-435.

Filmer,D., Hasan, A. and Pritchett, L. (2006) "A Millennium Learning Goal: Measuring Real Progress in Education," *Working paper # 97, Center for Global Development. World Bank.*

Foster, J.E. and Sen, A.K. (1997).*On economic Inequality.* Oxford University Press. Oxford.

Foster, J. E., Lopez-Calva,L. and M.Szekely (2005). Measuring the Distribution of Human Development: methodology and an application to Mexico. *Journal of Human Development* 6(1): 5-29

Harkness, S. (2004). Social and Political Indicators of Human Well-being. United Nations University: World Institute for Development Economics Research (WIDER). Research paper 33.

Herrero, C., Martinez, R., and Villar, A. (2010). Multidimensional social evaluation: an application to the measurement of human development. *Review of Income and Wealth*, 56, 3, 483-497.

Hicks, D. A. (1997). The Inequality-Adjusted Human Development Index: A Constructive Proposal. *World Development*. Vol. 25. No. 8 pp. 1283-1298.

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

Hopkins,M.(1991). Human development revisited: A new UNDP report. *World Development* 19, pp. 1469–1473.

Hoyland,B., Moene, K., and Willumsen, F. (2009). The Tyranny of International Index Rankings. University of Oslo, Technical paper, http://www.esop.uio.no/research/working-papers/Tyranny.xml

Islam, S. and Clarke, M. (2002). The Relationship between Economic Development and Social Welfare: A New Adjusted GDP Measure of Welfare. *Social Indicators Research*, Vol. 57, No. 2, pp. 201-228

Ivanova, I., Arcelus, F.J. and G. Srinivasan. (1999). An assessment of the measurement properties of the human development index. *Social indicators research* **46:** 157–179

Kelley, A. (1991). The Human Development Index: "Handle with Care". *Population and Development Review*, Vol. 17, No. 2, pp. 315-324.

Kirsch, I. (2001) The International Adult Literacy Survey (IALS): Understanding What Was Measured. *Research Report*. *Educational Testing Service*. Princeton, NJ. http://www.ets.org/Media/Research/pdf/RR-01-25-Kirsch.pdf

Kovacevic, M. (2010). Measurement of inequality in Human Development – A review. *Human Development Research Paper 35. UNDP-HDRO.*

Lind, N.C. (1992). Some thoughts on the human development index. *Social lndicators Research* 27: 89--101, 1992.

Lind, N.C. (2004). Values Reflected in the Human Development Index. *Social Indicators Research*, Vol. 66, No. 3, pp. 283-293.

Luchters, G. and Menskhoff, L. (1996) Human Development as Statistical Artifact. *World Development.* Vol. 24, No. 8. pp. 1385-1392.

Lloyd, C.B, Hewett, P.(2009). Educational inequalities in the midst of persistent poverty: Diversity across Africa in educational outcomes. *Journal of International Development 21.*

Marmot, Michael.G., and H. Bosma (1997) "Contribution of Job Control and Other Risk Factors to Social Variations in Coronary Heart Disease Incidence." *Lancet* 350(9073).

Marmot, Michael G., and G.D. Smith (1991) "Health Inequalities Among British Civil Servants: The Whitehall II Study." *Lancet* 337(8754).

McGillivray, M. (1991). The Human Development Index: Yet Another Redundant Composite Development Indicator? *World Development,* Vol. 19, No. 10, pp. 1461-1468.

McGillivray, M. and White,H. (1993). Measuring development? The Undp's human development index. *Journal of International Development: Vol. 5, No. 2, 183-192.*

McGranahan, D. (1995). Measurement of development: Research at the UN's Research Institute for social development. *International Social Science Journal,* 143, pp. 39–57.

Morse, S. (2003). For better or for worse, till the human development index do us part? *Ecological Economics,* 45, 281-296.

Murray, J. L. (1993). Development data constraints and the Human Development Index. In D. G. Westerndorff and D. Ghai (eds) Monitoring Social Progress in the 1990s, pp. 40–64.

Nardo, M., M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, and E. Giovannini (2008). Handbook on constructing composite indicators: Methodology and user guide. OECD publishing.

Noorbakhsh, F. (1998a). A Modified Human Development Index. *World Development,* Vol. *26, No. 3,* pp. *517-528*

Noorbakhsh, F. (1998b). The human development index: some technical issues and alternative indices. *Journal of International Development*, 10, 589-605.

OECD (1998). Literacy in the information age. Final report of te International Adult Literacy Survey. Statistics Canada.

Oeppen, J., Vaupel, J.W.. (2002). Broken Limits to Life Expectancy. *Science* 296, 1029–1031.

Ogwang, T. (1994) The choice of principle variables for computing the Human Development Index. *World Development*, 22(12), pp. 2011–2014.

Ogwang, T., Abdou, A. (2003) The choice of principal variables for computing some measures of human well-being. *Social Indicators Research,* 64, 139-152.

Osberg, L. and Sharpe, A. (2005). How should we measure the "economic" aspects of well-being? *Review of Income and Wealth*. Series 51, Number 2, 311-336.

Paul, S. (1996) "A Modified Human Development Index and International Comparison." *Applied Economics Letters* 3(10): 677-682.

Preston, S.H. (1975).The Changing Relation between Mortality and Level of Economic Development. *Population Studies*, Vol. 29, No. 2, pp. 231-248

Pritchett, L., Summers, L.H. (1996). Wealthier is Healthier. *Journal of Human Resources*, 841–868.

Ravallion, M. (1997). Good and Bad Growth: The Human Development Reports. *World Development.* Vol. 25. No. *5.* pp. 63 1-638.

Raworth, K. and Stewart, D. (2002). Critiques of the Human Development Index: A review. In Fukuda-Parr and Kumar (eds.) *Readings in Human Development*. Oxford University Press, New Delhi, 2003.

Rose, R. (1995). Making progress and catching up: Comparative analysis for social policy-making. *International Social Science Journal,* 143, pp. 113–125.

Ryten, J. (2000). The Human Development Index and Beyond: Which Are the Prerequisites for a Consistent Design of Development Indicators – Should There Be a Human Development Index? International Association of Official Statistics Conference "*Statistics, Development and Human Rights*" Montreux, May 4-8, 2000.

Sachs, J.D. and the WHO Commission on Macroeconomics and Health. (2001). *Macroeconomics and health: investing in health for economic development.* Geneva: *Report of the Commission on Macroeconomics and Health*, World Health Organization.

Sagar, A. and Najam, A. (1998). The human development index: a critical review. *Ecological Economics* 25, 249–264

Saisana, M., Tarantoorbakhshola, S., and Saltelli, A. (2005). Uncertainty and Sensitivity Techniques as Tools for the Analysis and Validation of Composite Indicators. *Journal of the Royal Statistical Society A*, 168(2), 1-17,

Sen, A.K. (1998). Mortality as an Indicator of Economic Success and Failure. *Economic Journal*, Vol. 108, No. 446, pp. 1-25

Seth, S. (2009). Inequality, Interactions, and Human Development. *Journal of Human Development and Capabilities, 10* , 375-396.

Srinivasan, T.N. (1994a). Human Development: A New Paradigm or Reinvention of the Wheel? *The American Economic Review*, Vol. 84, No. 2, pp. 238-243

Srinivasan, T.N. (1994b). Data Base for Development Analysis: An overview. *Journal of Development Economics* 44 (1994) 3-27.

Stanton, E. (2006). Accounting for Inequality: A Proposed Revision of the Human Development Index. *Working Paper 119*. Political and Economic Research Institute, University of Massachusetts, Amherst.

Stiglitz, J. E., A. Sen, and J.-P. Fitoussi (2009). Report by the Commission on the measurement of economic performance and social progress, mimeo

Streeten, Paul (1994) "Human Development: Means and Ends." *American Economic Review,* 84(2): 232-237.

Sugden, R (1993). Welfare, Resources, and Capabilities: A Review of Inequality Reexamined by Amartya Sen. *Journal of Economic Literature*, 31(4): 1947-62.

Sweeney, P.(1999). *The Celtic Tiger. Ireland's continuing economic miracle*. Oak Tree Press, Dublin

Ul Haq, M. (1995). *Reflections on Human development*. Chapter 4. Oxford University Press.

UIS (UNESCO Institute for Statistics) (2009). "The Next Generation of Literacy Statistics. Implementing the Literacy Assessment and Monitoring Programme (LAMP). Technical Paper 1.

Veenhoven, R. (2005). Apparent Quality-of-Life in Nations: How Long and Happy People Live. *Social Indicators Research*, Vol. 71, No. 1/3, pp. 61-86.

WHO: *The world health report 2004 –Changing history*

Zhou, P., Ang, B.W., and Poh, K.L (2007). A mathematical programming approach to constructing composite indicators. Ecological economics, 62, 291–297