Human Development Research Paper 2010/35

Measurement of Inequality in Human Development – A Review

Milorad Kovacevic

United Nations Development Programme
Human Development Reports
Research Paper

November 2010

Human Development
Research Paper
2010/35

Measurement of Inequality in Human Development – A Review

Milorad Kovacevic

United Nations Development Programme
Human Development Reports
Research Paper 2010/35
November 2010

Measurement of Inequality in Human Development – A Review

Milorad Kovacevic

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

Comments should be addressed by email to the author.

Abstract

It is widely accepted that country-averages of income, literacy, life expectancy and other indicators conceal widespread human deprivation and inequality. The measures of human development based on these indicators are also averages, and therefore mask disparities in the overall population. While the Human Development Index (HDI) itself is well accepted as a summary measure of HD capabilities and achievements, there is no a consensus about how to measure inequality in the HD distribution within a country. The conceptual difficulties, as well as the lack of appropriate disaggregated data, are customarily given as major obstacles for not adjusting the HDI for inequality. The objective of this paper is to first review some recent developments in measuring inequality in the distribution of multidimensional indices such as the HDI, and second - to present a practical implementation of the Alkire and Foster (2010) adaptation of the Foster, Lopez-Calva, Szekely (2005) method. The paper will first attempt to summarize the normative issues around the importance of accounting for inequality in opportunities for and outcomes of human development. Then it will review different approaches to accounting for inequality when quantifying HD. A special emphasis is placed on data requirements for each of the approaches. Consequently, data availability for the disaggregated analysis in the international context is thoroughly examined. Ease of interpretation is an important consideration. Finally, the paper describes the inequality-adjusted HDI and provides its limited sensitivity analysis.

Keywords: Human development index, multidimensional inequality, micro data, aversion to inequality

JEL classification: D63, O15, I0, I3, C8

1 INTRODUCTION

The conceptual framework of human development is based on Armartya Sen's capability approach. Human development is broadly defined as "a process of enlarging people's choices and enhancing their capabilities" (2007 Human Development Report (HDR)). The notions of poverty and inequality can also be defined within the human development framework: poverty reflects failure to enhance basic capabilities while inequality depicts disparities in the capabilities enjoyed by individuals that enable them to do or be what they value in their lives. "Capability may be limited by a lack of personal resources, but also by the context in which the individual is operating – the economic, social, political, cultural, and environmental conditions." (Burchardt, 2006).

The pursuit of equality has three general modes – equality of process (meaning that all individuals are subject to the same treatment), equality of opportunity and equality of outcomes. This paper will mostly dwell on the methods for measuring equality of outcomes.

The idea of equality of opportunities refers to starting points, and includes such things as access to nutrition and basic services, education, and jobs. Access to opportunities is frequently determined by characteristics beyond individual control, such as gender, ethnicity, race, socio-economic status of parents, etc. In the 1995 HDR, equality of opportunity was defined as one of three essential components of Human Development: "At the heart of this concept are three essential components: Equality of opportunity for all people in society; Sustainability of such opportunities from one generation to the next; Empowerment of people so that they participate in-and benefit from-development processes." (2005 HDR, page 1).

In terms of equality of outcomes (or results), we refer to differences in such things as the incomes people earn, the health they enjoy, their acquired knowledge and experience, the security they possess, and so on.

Annual Human Development Reports (HDR) have highlighted many achievements of human development and well-being across the world. Some achievements are studied at the level of individual dimensions, while others are discussed in terms of composite indices. The most well known composite index is the Human Development Index (HDI) which is based on three dimensions of human achievements – health, knowledge and standard of living. The HDI is well accepted as a summary measure of HD achievements. It is transparent, and simple to calculate and interpret. However, it has often been criticized for ignoring inequality in the distribution of human development across populations; particularly because uneven development is a major concern in the evaluation of human development and "there is evidence that many, if not all, people put some intrinsic value on equality as an end in itself" (Sen, 1992).

Inequalities in health, education, and income – key components of human development –deeply impact progress towards increasing human development. "The world's richest 1% of people receive as much income as the poorest 57%... The income of the world's richest 5% is 114 times that of the poorest 5%." (2001 HDR, page 19). "In developing countries, almost 60 percent of all births take place with no health professional in attendance. In one-third of all countries, 20 percent of the population or more lacks even the most basic literacy." (2005 HDR).

It is clear that both inequalities in country-average HDIs and within-country HD inequality matter for the evaluation and analysis of HD achievements. To this end, measuring inequality is an important first step towards addressing inequality in HD. Conceptual difficulties, as well as a lack of appropriate disaggregated data, are customarily given as major obstacles for documenting inequality. In addition, there hasn't been a broad consensus about how to measure inequality in HDI distribution within a country (see Foster, Lopez-Calva and Szekely, 2005, for a review), even if credible data are available. This paper will address some of these obstacles with an emphasis on data issues.

Generally, methods for measuring inequality have been developed in relation to the unequal distribution of income and wealth across a population. Inequality in the distribution of other characteristics and resources is often recognized but rarely measured: "Our physical and social characteristics make us immensely diverse creatures. We differ in age, sex, physical and mental health, bodily prowess, intellectual abilities, climatic circumstances, epidemiological vulnerability, social surroundings, and many other respects… Such diversities, however, can be hard to accommodate adequately in the usual evaluative framework of inequality assessment. As a consequence, this basic issue is often left substantially unaddressed in the evaluative literature." (Sen, 1992, page 28). Accordingly, inequality measures reported in the *HDR* have been mainly about income distribution. The first *HDR* (1990) stated that the average measurements of all three dimensions of human development "conceal wide disparities in the overall population," but that compared to income inequality, the "inequality possible in respect to life expectancy and literacy is much more limited: a person can be literate only once, and human life is finite." (1990 HDR, page 12) While it is true that health and education inequalities are quantitatively more limited than income inequality, such inequalities do exist and it should be equally important to incorporate their unequal distribution within countries. Hicks (1997), for example, notes that "there is significant life-span inequality, ranging from infants who die at birth or before age one, to persons who die at ages over 100 years."

There have been a very few l attempts to reformulate the HDI so that it accounts for both the average achievement in HD relevant dimensions in a country, and for inequality in the distribution of HD achievements (Hicks, 1997, Foster et al, 2005, Stenton, 2008). These adjustments to the overall HDI have been particularly useful for international comparisons of disparities among countries. For examining HDI inequalities within countries, a more useful approach is to calculate separate HDIs for different groups. National Human Development Reports have carried out such analysis – disaggregating HDI by regions, by ethnic groups, by racial groups, and most recently (2006 HDR) by income quintiles. Such disaggregations lead to a better understanding of human development.

Everyone should have the opportunity to be educated, to be nourished, to have access to clean water and other basic services, and to live a long and healthy life. Equally-developed human capabilities and equally-distributed opportunities can ensure that HD progress is not lopsided and that its benefits are equitably shared. HDRs that disaggregate analyses of HD across ethnicity, race, gender, and region, have focused on the link between observed inequalities and predetermined circumstances. The next step in inequality analysis would be to separate inequality into two parts: one that can be attributed to circumstances beyond the control of persons and one that is related to personal choices, efforts and talents.

The remainder of this paper is organized as follows. In section 2 we briefly review the role, methods, and data requirements of disaggregated analyses published in past Human Development Reports. Section 3 summarizes the recent literature on measurements of equality of opportunities for human development. Section 4 introduces the Income-Adjusted HDI, an index published in the HDR between 1991 and 1993. Section 5 reviews three distribution-sensitive modifications of the HDI: Hicks' (1997) Inequality-adjusted HDI; an index based on a general mean of general means proposed by Foster, Lopes-Calva and Szekely (2005); and an association-sensitive HDI proposed by Seth (2009). Section 6 describes a variant of the Foster, Lopes-Calva and Szekely (FLS) method that has been devised by Alkire and Foster (2010) and applied to the HDI 2010. It also provides details of measuring inequalities in each dimension of the HDI and contains a limited sensitivity analysis of the IHDI. We conclude in section 7.

2 DISAGGREGATED ANALYSIS OF HD

National human development data disaggregated by geographical or administrative units; by social groups according to gender, ethnicity or rural/urban divide; by economic delineation on rich and poor; or by some sort of wealth quintile, may reveal significant disparities in HD within countries as expressed by human development indices. These are disparities which have been labeled as categorical inequality, group inequality, or between-group inequality (Tilly, 1999, Stewart, 2002). Within the context of Human Development and its analysis, there have been few attempts to express categorical inequality with a single index, the exception being indices related to gender disparities[1]. Just recently, as part of the burgeoning literature on equality of opportunities, several indices of inequality of opportunities have emerged. We will review a few of these indices in the next section.

Different aspects of inequality can be distinguished through different types of disaggregation. With so many inequalities in multiethnic and otherwise divided societies, a disaggregated HDI profile is essential to eventually understand the underlying sources of tension and potential causes of future conflict. Frances Stewart (2002, page 2) notes that most analyses of poverty and

[1] For a good review of gender inequality indices see the paper by Permanyer (2009).

inequality focus on the individual: these analyses are, "concerned with the numbers of individuals in poverty in the world as a whole, not with who they are, or where they live." In a discussion of the origins of violent conflict, Stewart (2002: 3) goes on to distinguish between "vertical" and "horizontal" dimensions of inequality: "It is my hypothesis that an important factor that differentiates the violent from the peaceful [countries] is the existence of severe inequalities between culturally defined groups, which I shall define as horizontal inequalities to differentiate them from the normal definition of inequality which lines individuals or households up vertically and measures inequality over the range of individuals – I define the latter type of inequality as vertical inequality."

The HDI and other indices from the HD family that have been calculated for specific regions, urban-rural sub-populations, and racial/ethnic groups within countries, have depicted horizontal inequalities in almost all cases. For example, a table in the *2007/2008 HDR,* shows Kenya's HDI disaggregated to regions affected by draught to illustrate that, in some cases, vulnerability is directly linked to climate shocks. Disaggregated HDIs in this example show a close fit between food emergencies linked to drought and districts where human development is low (2007/2008 *HDR*, Table 2.1, page 80).

Similarly, disaggregating a country's Human Poverty Index (HPI) by region has identified concentrations of impoverishment. For example "In the Islamic Republic of Iran in 1996 the disaggregated HPI showed that human deprivation in Tehran was only a quarter that in Sistan and Baluchestan. The HPI for urban Honduras in 1999 was less than half that for rural areas. For English speakers in Namibia in 1998 the HPI was less than one-ninth that for San speakers." (2001 HDR, page 15). Disaggregation of the Gender Empowerment Measure (GEM) in national human development reports show that gender differences within a country can also be large. For example, "the GEM for the Puttalam district in Sri Lanka in 1994 was less than 8% of that for Nuwara Eliya" (2001 *HDR*). Rural-urban differences interact with regional disparities. For example "In China, urban Shanghai would rank 24 in the global HDI league, just above Greece, while rural Guizhou Province would rank alongside Botswana."(2006 *HDR*, page 271).

The 2003 *HDR* (page 47) gives a brief summary of a variety of disaggregated analyses reported in National Human Development Reports since 1992. More recently, Gaye and Jha (2009) reviewed conceptual and measurement innovations in the National and Regional HDRs between 1998-2009. They pointed out many interesting disaggregated analyses of HD done across different groups. The NHDR of El Salvador (2008), for example, divided the working population according to their "decent work" status into four groups: unemployed, underemployed, fully employed but without fair remuneration or social protection, and fully employed with social protection. The HDI was then calculated separately for each group. As expected, the unemployed had the lowest HDI score (0.664) and the fully employed with social protection had the highest HDI score (0.855).

Grimm, Harttgen, Klasen, and Misselhorn (2006, 2008) generated a separate HDI for each income quintile. This kind of analysis allows one to track the progress in HD for the income-poor, middle class, and income-rich. The results showed that across all countries, inequality in HD by income quintile was very high, and was particularly high in developing countries, especially in Africa. In this study, inequality was measured by the ratio between the HDI for the richest quintile and the poorest quintile (80/20 ratio)[2].

It is clear that a comparative disaggregated analysis over many countries requires a common denominator – the same population groups, the same number of groups, etc. Such universally defined groups can be urban-rural, gender, age-sex groups, grouping by income and consumption quintiles, or education levels. However, it is not a simple task to obtain credible disaggregated data especially for developing nations. In the following section, we give an example of disaggregated analysis of HD across the income distribution quintile groups conducted by Grimm et al. (2006) using data from 13 developing and two developed countries. They extended

[2] Quantile ratios are straightforward indicators of inequality that are easy to interpret. For example, if the 80/20 ratio is equal to 2, then the average person of the richest 20 percent of the population achieves human development two times as high as the average person of the poorest 20 percent. The 80/20 ratio can be decomposed; it is equal to the product of 80/50 ratio and the 50/20 ratio. This decomposition tells to what extent the 80/20 ratio is driven by inequality in the upper part of the distribution versus inequality at the bottom part. However, the quantile ratios are insensitive to outliers either in the very top or very bottom tail of the distribution, and also they do not reflect what happens in the middle of the distribution.

their methodology in 2008 to 21 developing and 11 OECD countries, (Grimm et al., 2009). In reviewing their analysis the emphasis will be on data issues.

Data Issues in Disaggregated Analysis: An Example

Disaggregated analysis generally requires data at the individual level or household level. For their research, Grimm et al. combined data from different household surveys. For developing countries, household income surveys (HIS) were used to calculate education and gross domestic product (GDP) indices for each income quintile group, and Demographic and Health Surveys (DHS) were used to calculate the life expectancy index. These surveys were conducted on different samples so the data sets didn't refer to same households. The data sets were then merged by a statistical matching method using variables that were available in both surveys. These variables included household structure, education and age of the household head, area of residence, housing characteristics and the like. The correlation between household income per capita and a set of household variables used for statistical matching was estimated and used to generate a proxy for the income of households in the DHS. For the two developed countries in the study, Finland and the United States, GDP and education data were from the Luxembourg Income Study, and life expectancy data were taken from published empirical work.

The authors used two alternative statistical matching techniques. The first technique estimates the correlation between income and a set of household characteristics which are available in the HIS and the DHS, and then uses this correlation pattern to predict income for the households covered by the DHS. The quality of such a matching process depends heavily on the available data with all characteristics observed so that the correlation pattern can be properly estimated. Also the data quality and consistency of both surveys is important for predicting and imputing the 'missing' variables. The second technique uses the ''asset index'' for parsing both data sets, assuming that the asset index is a good proxy for income (see Filmer and Scott, 2008). This measure is often used to get an idea of the living standard of households interviewed in the DHS.

The quality of subsequent analyses depends on the quality of the data set produced by statistical matching and imputation. For example, violation of the conditional independence assumption[3] may result in a biased imputation, which will further bias the analysis. So the matching/imputation has to be done with the utmost care utilizing as much relevant auxiliary information as possible (Rassler, 2002).

The next step is to decide which variables are most appropriate for calculating indices for the three dimensions of the HDI by income quintile. The health component was based on infant mortality data from the Demographic and Health Surveys. The observed data of infant mortality were then combined with the Ledermann (1969) model life tables to estimate life expectancy for each quintile income group. The education index was based on adult literacy and school enrolment data available directly from the household income surveys for each income quintile. The GDP index was calculated using the income variable from the household income survey expressed in US dollars in purchasing power parity (PPP) terms using conversion factors based on price data from the International Comparison Program surveys (World Bank). This income per capita is rescaled using the ratio between GDP per capita and estimated household income per capita. The average adjusted income per capita for each quintile is then transformed and standardized into the GDP index at the level of quintile.

In summary, disaggregated analysis of the HDI requires distributional data which are often not available in the ready-to-use form. They have to be created by combining different sources and by using statistical modeling. The quality of data and the validity of statistical models will determine the quality of the subsequent analysis of the HDI.

[3] It must be remarked that matching/imputation methods are explicitly or implicitly based on a statistical model of relationships among variables. Thus a wrong specification of the model leads to seriously biased final results. In fact, since the typical situation of statistical matching consists of lack of simultaneous information on all three variables ($\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$) the only model we are able to reasonably estimate is the one based on the hypothesis that, roughly speaking, information on the variable $\boldsymbol{X}$ is sufficient to determine pair of $\boldsymbol{Y}$ and $\boldsymbol{Z}$. More formally it means that $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are statistically independent conditionally on $\boldsymbol{X}$, i.e. $P(\boldsymbol{Y},\boldsymbol{Z}|\boldsymbol{X})=P(\boldsymbol{Y}|\boldsymbol{X})P(\boldsymbol{Z}|\boldsymbol{X})$. This hypothesis is known as *Conditional Independence Assumption* (CIA). (Rassler, 2002). In the example of matching of DHS and HIS, the statistical matching algorithm assumes that income and health variables are conditionally independent given the set of common variables.

3 MEASURING INEQUALITY OF OPPORTUNITIES

Equal opportunity is generally understood as the notion that success in life should reflect a person's choices, efforts and talents, not his background defined by a set of predetermined circumstances at birth, such as, gender, race, place of birth, family origins, etc. Circumstances are exogenous to the individual, by definition, and differences in circumstances are argued to be morally irrelevant to outcomes, while person's choices and efforts can lead to morally justifiable differences in achievements (Roemer, 1998). The equal opportunity principle is conceptually simple – circumstances at birth should not matter for a person's chances in life. The main idea of theories of equality of opportunity lies in the distinction between circumstances that constrain a person's opportunities and the person's choices that may also affect a particular outcome. In statistical terms, in an equal opportunity society there is no statistically significant association between circumstances and important life outcomes.

In socio-economic literature, a common way to study equality of opportunities is through intergenerational income correlation (Corak, 2006). This means investigating whether the parent's income is highly correlated with the income of the adult child. A high correlation would suggest that equality of opportunity is unlikely to be present. "In the United States almost one half of children born to low income parents become low income adults. This is an extreme case, but the fraction is also high in the United Kingdom at four in ten, and Canada where about one-third of low income children do not escape low income in adulthood. In the Nordic countries, where overall child poverty rates are noticeably lower, it is also the case that a disproportionate fraction of low income children become low income adults. Generational cycles of low income may be common in the rich countries, but so are cycles of high income. Rich children tend to become rich adults. Four in ten children born to high income parents will grow up to be high income adults in the United States and the United Kingdom, and as many as one third will do so in Canada. (Corak, 2006)"

Following Roemer's (1993, 1998) development of an algorithm for a practical separation of effects of opportunities from the effects of efforts, there have been several attempts to decompose inequality in income distribution, most notably Roemer et *al* (2003), World

Development Report (2006), Lefranc et *al*. (2008), Bourguignon, Ferreira and Menéndez (2007), Ferreira and Gignoux (2008), and Barros et al. (2009). The book by Barros et al. (2009) also examines inequality in opportunities for educational achievements in Latin America and the Caribbean. There have been no published attempts to measure inequality of opportunities in human development.

A practical way to measure inequality of opportunity is to decompose inequality of outcomes into a portion resulting from circumstances that lie beyond the individual's control, and a residual component that contains differences due to the individual's effort, choices, talent and luck. The first component accounts for inequality in opportunity, while the second reflects inequality due to individual heterogeneity. Individual heterogeneity includes everything that is not contained in the circumstance variables, primarily including personal effort, choices, talent, or luck.

The simplest procedure for measuring inequality of opportunity consists of: (i) identifying the most important variables that describe individuals' exogenous circumstances and which affect the measured outcome; (ii) partitioning these variables into categories; (iii) classifying individuals into groups defined by these categorical variables; and (iv) decomposing the total inequality of outcomes to the inequality between groups, *IB(y,c)*, that can be attributed to inequality of opportunity, and the differences in the outcomes existing after controlling for these circumstances, i.e., the inequality within groups *IW(y,c)*, that is attributed to individual heterogeneity. Therefore, the absolute level of inequality of opportunities can be expressed as inequality between groups, *IB(y,c)*. When divided by overall inequality in the population, a relative measure of inequality of opportunity is obtained. Obviously, the choice of inequality measures is limited to those that can be decomposed by population subgroups.[4]

A true measure of inequality of opportunity would require that all relevant circumstance variables are included, which is impossible to accomplish in reality. That is why the assessment

[4] The Gini index, the most widely used inequality index, is not additive decomposable. The two indices that satisfy the additive decomposability are the two Theil indices - the Theil mean log deviation index, E(0), and the Theil entropy index, E(1). They, however, differ in their sensitivity to inequality in different parts of the distribution. The entropy measure, E(1), is most sensitive to inequality in the top range in the distribution, while the mean log deviation measure, E(0), is most sensitive to inequality in the bottom range of the distribution. (Shorrocks, 1980)

of inequality of opportunity is always conditional on observed circumstances, which are often limited to quite basic aspects at best. However, inclusion of as many circumstance variables as possible increases the number of groups while decreasing the sizes of groups, leading to insufficient data for estimation of the corresponding means and inequality measures. A good balance between including many circumstance variables and maintaining large group sizes is needed.

3.1 Measuring inequality of opportunity using the Theil Index[5]

For a reasonably large sample of individuals, n, and a relatively small number of circumstance groups, m, it is feasible to obtain the measure of inequality of opportunity as a ratio

$$IO_0(y,m) = E_0(y,m)/E_0(y),$$

where $E_0(y)$ and $E_0(y,m)$ are respectively defined as the Theil mean log deviation index and its component due to between-group inequality:

$$E_0(y) = \sum_1^n \ln\left(\frac{\mu}{y_i}\right)/n, \text{ and } E_0(y,m) = \sum_1^m \mathrm{w_j} \ln\left(w_j \frac{\mu}{\mu_i}\right). \quad (1)$$

Here, y_i denotes the characteristic of interest, say HDI, for the i-th individual; μ and μ_i are the overall mean and the mean for the j-th group, respectively, and w_j is the population share of the j-th group. Similarly one can use the entropy measure and define a measure of inequality of opportunity that is more sensitive to inequalities in the lower part of the distribution.

[5] Ferreira and Gignoux (2008) noticed that the estimates of between-group inequality can differ because of different paths of the decomposition: (i) estimating the between-group inequality directly by comparing the group-specific means, or (ii) estimating the within-groups inequality by rescaling the individual outcomes with the group-means so that all between-group inequality is suppressed. In the second case, the between-group inequality is estimated as a residual inequality after subtracting the within-groups inequality ratio from one. The only inequality measure that is path-independent is the Theil mean log deviation (Foster and Shneyerov, 2000).

Bourguignon et al. (2007) and Ferreira and Gignoux (2008) develop a family of indices of the same ratio type but use a parametrically standardized model of distribution of y which accounts for both the circumstance variables, C , and the "efforts" variables, E:

$$\ln y = C\alpha + E\beta + u$$
$$E = HC + v,$$

α and β are the unknown parameters, and H is a matrix of unknown parameters linking the circumstance variables to the "efforts" variables. This matrix explicitly allows for the fact that some of the "effort" variables are affected by circumstances. A typical example is that education is affected by family background. Finally, u and v are the error terms assumed to be centered at 0 with constant variance. The reduced form of the equations above allows relatively straightforward estimation of the parameters using either OLS or ML leading to a parametrically estimated distribution of y, by $\tilde{y}_i = \exp\,(\bar{C}_i\hat{\gamma} + \hat{\varepsilon}_i)$, and also by $\tilde{\tilde{y}} = \exp\,(C_i\hat{\gamma})$. These smoothed values are then used in the Theil index (1) for calculation of either the between-group inequality, $E_0(\tilde{\tilde{y}}, m)$, or the within-group inequality $\bar{E}_0(\tilde{y}) = \sum_j^m E_0(\tilde{y}_i, j)$. In the later case, the inequality of opportunity is obtained as a residual $IO(y) = 1 - \bar{E}_0(\tilde{y})/E_0(y)$.

Bourguignon et al. (2007) apply their approach to income distribution data in Brazil. They use five circumstance variables which lie beyond the control of the individual—father's and mother's education, father's occupation, race, and region of birth. They also decomposed the effect of opportunities into a direct effect on earnings and an indirect component, which works through the "effort" variables. They find that parental education is the most important circumstance affecting earnings, but the occupation of the father and race also play a role.

3.2 Human Opportunity Index for children: The D-Index

Barros et al. (2009) analyze inequality of opportunity in Latin America and Caribbean with a special focus on opportunities for children. The circumstance variables they used were gender, area of residence, the educational attainment of the family head, per capita family income, single-parent or two-parent household, and the number of siblings under 16 years. These six circumstance variables were used for analysis of basic opportunities. They identified as basic opportunities the services that are critical for children's development and are exogenous for the child: adequate housing as measured by access to electricity, water and sanitation, and an opportunity for basic education measured by the completion of sixth grade on time and by school attendance.

The D-index of inequality of opportunity constructed by Barros, Molinas, and Saavedra (2008) is a version of a dissimilarity index which quantifies the deviation from the mean due to circumstances. Namely, if p_j is the percentage of children with access to a given opportunity (e.g., to complete 6th grade by age of 14) who live in circumstance (group) j, which represents w_j share of population of interest, then the D-index of inequality of opportunity is given by

$\widehat{D} = \sum_1^m w_j \left|p_j - \bar{p}\right|/(2\bar{p})$, where $\bar{p} = \sum_1^m w_j\, p_j$.

Alternatively, p_j can be estimated from a logistic regression as the probability of having access to a particular opportunity, conditional on a person's circumstances. For example, inequality in opportunity to finish the sixth grade on time across the circumstance groups in Guatemala in 2005 was 0.27, while in Costa Rica it was three times lower, only 0.09.

The authors also proposed the Human Opportunity Index (HOI) for Children, as the opportunity coverage rate discounted for inequality of opportunity, that is $O = \bar{p}(1 - D)$. For example, at the national level, in Paraguay only ($\bar{p}$ =57% children under age 16 had access to clean water in 2005. However, inequality of access to water across the circumstance groups in Paraguay was (D=)20% which gave the HOI for children for access to water in Paraguay of only 45.6%.

3.3 The Gini Opportunity Index

Lefranc, Pistolesi and Trannoy (2008) analyzed the relationship between income inequality and inequality of opportunities for income acquisition in nine developed countries during the 1990s. They defined equality of opportunity as the situation where income distributions conditional on parental education and occupation cannot be ranked according to stochastic dominance criteria.

Stochastic dominance is assessed using nonparametric statistical tests[6]. They found disparities in the degree of equality of opportunity across countries and a strong evidence of correlation between inequality of outcomes and inequality of opportunity. The U.S. and Italy showed the highest inequality in both outcomes and opportunities. They found that inequalities in opportunities in Sweden and Norway were not statistically significant, i.e., the income distributions conditional on social origin are almost the same as unconditional distributions. They extended the Gini index to a scalar "Gini" index of opportunities (*GO(y)*):

$$GO(y) = \sum_j^m \sum_{k>j} p_j p_k \mu_j \left[\left(1 - G_j\right) - \mu_k (1 - G_k)\right] / \mu ,$$

where p_j, μ_j and G_j refer to the proportion of people with a particular opportunity in the jth circumstance group, the group mean income and the within-group Gini index, respectively, and μ is the overall mean income. The properties of *GO(y)* are yet to be studied[7].

3.4 Inequality in human development opportunity: Discussion

Inequalities in the distribution of outcomes such as standard of living, educational attainment, and health status can be attributed to differences in circumstances, at least partially. Therefore, some portion of inequality of HD outcomes as measured by the HDI might be explained by inequality in opportunities.

Formally disentangling circumstances from personal efforts and choices may require complex structural equation modeling with a necessary identification of mediation factors, i.e., intervening variables that may modify the impact of chosen circumstances differentially for different people. These mediation factors are usually the omitted circumstances. Upon defining the circumstance groups for human development, and assuming that the HDI is a one-

[6] Lefranc et al. (2008) apply the stochastic dominance criteria by performing two tests independently: (1) they test the null hypothesis of equality of two conditional distributions obtained under different circumstances (the first order stochastic dominance test); (2) they test the second order dominance by comparing the Generalized Lorenz curves of two conditional distributions. They set a decision making rules to decide about equality of opportunity based on results of these tests.

[7] It can be shown that GO=0 although the opportunities are unequal. For example, for m=2, $\mu_1 = 3 > \mu_2 = 2$, $G_1 = 2/3 > G_2 = 1/2$, and $p_1 \neq p_2$, GO=0. Similarly, it can be shown that GO can be negative!

dimensional measure of HD, a direct application of the decomposed Theil Index or the D-index can provide a straightforward assessment of inequality of opportunity in HD achievements.

However, the impact of circumstance groups on components of the HDI may be differential: mother's education may impact the son's education more than his income, and his health may be even less dependent on mother's education. Accounting for these differences when assessing the inequality of HD opportunities is a formidable task. For policy targeting purposes it may make sense to assess the inequality of opportunity for each component separately, and then to combine these inequalities into a measure of inequality of opportunity for the HDI.

The methodology used for measuring the inequality in opportunities can also be used for disentangling between-group and within-group inequality for any grouping of interest of the population. For example, a question can be whether inequality in HD is larger between provinces or within. The question can also be phrased as what percentage of total inequality in HD can be attributed to disparities between provinces.

4 PARTIALLY ADJUSTED HDI: INCOME DISTRIBUTION-ADJUSTED HDI

The first attempt to modify the HDI to account for inequality in the distribution of one of its dimensions – income, was published in the 1991 HDR, as the ***Income Distribution-Adjusted HDI.*** While it doesn't adjust for inequality in the other two dimensions (health and knowledge) of the HDI, it is an important and inspiring attempt to sensitize the HDI to inequality in the most unequally distributed dimension. Also, conceptualizing inequality in a single dimension brings simplicity in reasoning about the plausible implications of such inequality.

The argument for creating the Income Distribution-Adjusted HDI was that inequality in the dimension of income was the most significant and the least bounded, and therefore, accounting for inequality in this dimension provides a good sense of overall inequality.

Moreover, the correlation between income inequality expressed by the Gini index and the other HDI components reveals a certain dependency. Table 4.1 shows the correlations between the income Gini index and the HDI, as well as its components, calculated from 2007 data[8]. The last column presents the intervals of Gini variation and its median values.

Table 4.1. Correlations of the income Gini index with HDI and HDI components

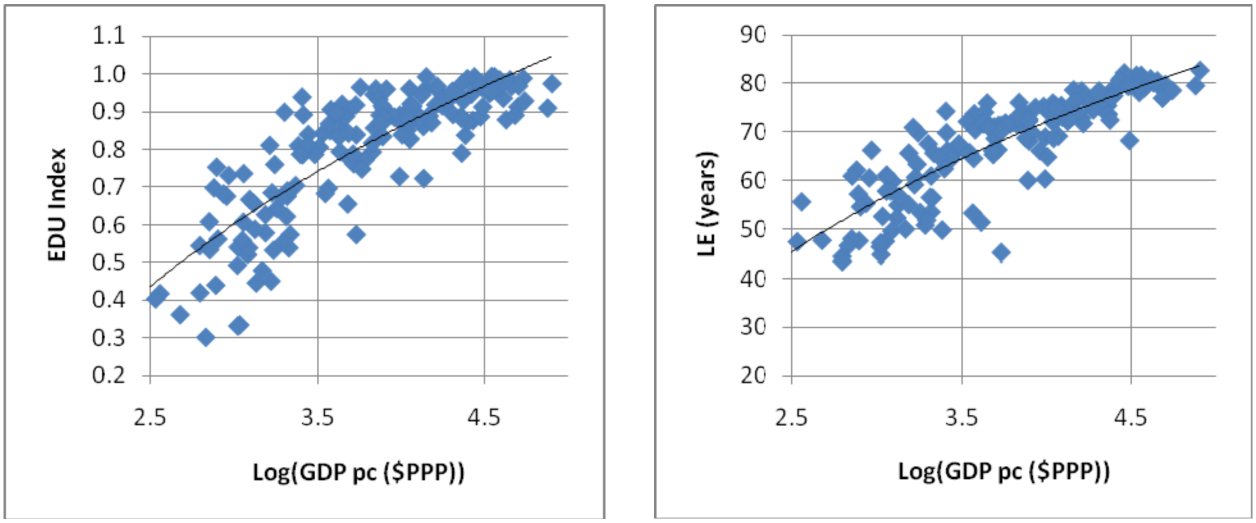
	HDI	LE	EDU	GDP	Range and median value of income Gini index
All Countries (141)	-0.335	-0.361	-0.255	-0.324	(24.7, 74.3), 39.5
High HD Countries (58)	-0.515	-0.240	-0.593	-0.536	(24.7, 58.5), 35.5
Medium HD Countries (61)	-0.122	-0.280	-0.053	0.083	(28.2, 74.3), 42.5
Low HD Countries (22)	0.094	-0.171	0.175	0.099	(29.8, 52.6), 41.2

Source: 2009 HDR. Author's own computations

It is evident that the income Gini index is negatively correlated with the HDI and all of its components at the global level, essentially confirming that the higher inequality in income is associated with lower values of HD as measured by the HDI. In other words, increased income inequality may be seen as an obstacle to human development. When countries are stratified according to their HD level[9], the correlation between the income Gini index and the HDI remains negative except for 22 low HD countries where the correlation changes the sign but is reduced to almost a zero. It is also evident that the negative correlation is strongest for 58 high HD countries. For these countries the education component is the most negatively associated with income inequality: the higher the level of income inequality the lower the education index. For medium HD countries the negative correlation with inequality in the income component is strongest for the health component measured by the life expectancy index. Figure 4.1 shows that life expectancy (in years) and the education index are concave in income (logarithm of GDP per capita in $PPP) at the global level, so then a redistribution of income towards more equality would be associated with an increase in the average level of health and the average level of education index. Based on the observed ranges of variation, the Gini itself is the most unequal across the medium HD countries, and the least across the low HD countries.

[8] HDR 2009 reports the Gini index only for 141 countries.
[9] We use the three-class classification: High (HDI≥0.8, Medium (0.5≤HDI<0.8) and Low Human Development (HDI<0.5).

Figure 4.1. Global Distribution of Life Expectancy and Education Index as Function of Income



Source: HDR 2009 (Author's own computations and graphs)

The adjustment of the income component of the HDI for a particular country in the 1991-1993 HDR was done by multiplying income by (1-$\boldsymbol{G^I}$) where $\boldsymbol{G^I}$ is the Gini index for the income distribution. The income component itself was already adjusted for diminishing returns, so that the distributional adjustment modified the income further by the degree of inequality in its distribution. This adjustment is also known in the literature as the Sen welfare standard[10]. The Gini index is then interpreted as the loss in welfare due to inequality, and is expressed as a percentage of the maximum achievable welfare. The inequality adjusted GDP was then used to compute the HDI. [11]

The Income Distribution-Adjusted HDI was calculated for a small number of countries (53) in the 1991 HDR, and for subsequent years until it was discontinued in 1994. In 1991 only 25 countries had the Gini coefficient reported, among them 17 also had the estimated ratio of the

[10] Sen (1976) also calls this function "the real national income."

[11] The income component in HDR 1991-1999 was based on Atkinson's formulation of the welfare function: $W(y) = y^{1-\epsilon}/(1-\epsilon)$. Parameter ε was set to 0 for incomes below poverty line so that in such a case there were no diminishing returns from income, and $W(y) = y$. For incomes between multiples of the poverty line, $ay^* \leq y < (a+1)y^*, a \geq 1$, parameter ϵ was given as $\epsilon = a/(a+1)$, and the adjusted income was calculated as $W(y) = \sum_{a=1}^{A} a(y^*)^{1/a} + (A+1)(y - Ay^*)^{1/(A+1)}$, with $A = [y/y^*]$ (an integer part) . When calculating the Income Distribution-Adjusted HDI for a country, the life expectancy and education indices remained as they were in the HDI, while the GDP index was modified as: $GDP_{ind}^* = \frac{(1-G^I)W(y)-367}{5075-367}$, where 367 and 5075 were the goalposts at that time. Thus, the Income Distribution-Adjusted HDI had the form: $IDAHDI = \frac{1}{3}(LE_{ind} + EDU_{ind} + GDP^*{}_{ind})$.

income share of the richest quintile to the poorest quintile (80/20 ratio). It has been shown that the logarithm of the 80/20 ratio and the Gini coefficient are strongly linearly related. This relationship[12] was used to estimate the Gini for the remaining 28 countries. However, the 80/20 ratio itself has often been dismissed as an inequality measure (Sen, 1973) since it only looks at the differences in tails of the distribution and it is not sensitive to income transfers between two points.

Since 1994 the HDR hasn't reported the Income Distribution-Adjusted HDI, although each year the information necessary for its computation has been available in the HDR's statistical tables. For example, the 80/20 ratio was presented in tables on human distress for industrial countries in 1995 and 1996, and then from 1998 to 2000 the real GDP for the poorest and richest quintiles were given in the table on human poverty profiles for developing countries. The Gini coefficient was also reported and discussed for developed countries, as well as for the CIS and Eastern European, countries. Since 2001, a new table on inequality in income or consumption (expenditure) has reported the 90/10 ratio (10% richest to 10% poorest), 80/20 ratio, as well as the Gini coefficient for over 100 countries (e.g., in 2009, the Gini was reported for 141 countries). In the 2009 HDR, only the 90/10 ratio was reported. Using the 1991 method it is possible to calculate the Income Distribution-Adjusted HDI for most years.[13]

Note that if the Gini adjusted GDP index follows the same original pattern of income adjustment in IAHDI, that is, first to adjust for diminishing returns, then for inequality, and then to be normalized, it would take the form

$$GDP^{*}_{ind} = \frac{(1 - G^{I})\log(GDP) - \log(100)}{\log(40{,}000) - \log(100)}$$

[12] The World Bank's World Development Indicators publication uses this relationship when estimating the Gini index for some countries (World Development Indicators, 2009).

[13] A word of caution is in order when dealing with the distributional data at the international level. Data on distribution of income or consumption, from which the Gini index and the income quantiles are estimated, are collected in national household surveys and compiled by the WB Development Research Group. These data refer to different years, often cover different variables, based on different definitions, and different sample sizes. Similarly, data for high-income countries come from the Luxembourg Income Study database (comprising data from over 50 surveys from 35 countries) and exhibit similar variability.

This adjustment, however, produces negative values of GDP^*_{ind} whenever $DP < 100^{1/(1-G^I)}$. For example, if the Gini index is equal to 0.5, the GDP has to be greater than $10,000 to have a positive value of GDP^*_{ind}. Hence, other patterns of adjustment must be considered. We discuss one possibility in Section 5.1.

5 ADJUSTMENTS FOR INEQUALITIES IN ALL THREE DIMENSIONS: A REVIEW

In this section we examine three different proposals for adjusting the HDI for inequality. All three require disaggregated data for the HDI dimensions. The starting point can be the relationship between the univariate welfare standard $S(x)$ and inequality measure $G(x)$, where the inequality measure is conveniently chosen to be the Gini index: $S(x) = \mu(x)[1 - G(x)]$. The maximum of the welfare standard is reached when there is no inequality in distribution and everyone receives the same amount of x. Conversely, the welfare standard $S(x)$ is interpreted as the mean of x discounted by the level of inequality in x. The Sen welfare standard $S(x)$ satisfies the basic properties of a well defined welfare function (see Foster et al., 2005) except subgroup consistency.[14] Foster and Schorrocks (1991) showed that the Gini index (and consequently the Sen welfare standard) were not subgroup consistent.

5.1 Hicks' Inequality-Adjusted HDI (IAHDI)

Hicks (1997) notes that there is a significant life-span inequality and that literacy and school enrollment are not distributed equally. He proposed an ***Inequality-Adjusted HDI*** (***IAHDI***) which was adjusted for inequality of distribution in each of three dimensions.

While the Gini index is the most often applied to analyzing inequality in distribution of income, consumption, wealth, assets, or land holdings, it can also be applied to any other distribution function. In the space of education, Hicks considers the total number of years of schooling of

[14] The subgroup consistency means that if $S(x)$ increases (declines) in one subgroup and remains un-changed in the rest of population, then the overall $S(x)$ has to increase (decline).

population as being a stock, so that each person holds some share of that total. Analogously, Hicks defines life-span attainment as the age at death, and considers age-at-death statistics to be the best measure available for determining life-span attainment and its distribution. He computes the Gini index of inequality for each dimension and then discounts the normalized component indices by multiplying them by (1-G) where G is the corresponding Gini index. Hicks also contemplates the possibility of differential weighting of the inequality adjustment factors by allowing a separate weight λ^x for each dimension *x*, under a general condition that
$\lambda^x(1-G^x) < 1$, and $\sum_x \lambda^x = 1$

$$IAHDI = \frac{1}{3}[\lambda^{LE}(1-G^{LE})LE_{ind} + \lambda^{L}(1-G^{E})E_{ind} + \lambda^{I}(1-G^{I})GDP_x].$$

For illustration of the proposed index, Hicks set all three λ^x equal to 1. An apparent difference of the Hicks' IAHDI from the Income Distribution-Adjusted HDI, besides the adjustment of all three component indices, is that the adjustments were done to the component indices ***after*** normal-ization, and not to the indicators.

The aggregation is first done within each dimension separately and then over dimensions. Foster et al. (2005) use a general expression for Hicks' IAHDI,

$$IAHDI = \mu(S(x), S(y), S(z))$$

to emphasize different stages in aggregation. [15] Here, $\mu(.)$ denotes the arithmetic mean, while *S*(.) denotes the Sen welfare for the particular dimension, $S(x) = \mu(x)[1 - G(x)]$. [16]

---

[15] Using the same notation the original HDI is simply a mean of the means, $HDI = \mu(\mu(x), \mu(y), \mu(z))$, and as such it is independent of the order of aggregation.

[16] A change of the order of aggregation in the expression above leads to a completely different index that captures the inequality between the means of dimensions and not between the individuals:

$$IAHDI^* = S(\mu(x), \mu(y), \mu(z)) = HDI(1 - G(\mu(x), \mu(y), \mu(z)))$$

Hicks combined aggregate data at the national level with distributional (grouped) data for the income, health and education dimensions. He used aggregate data on literacy, life expectancy and GDP (per capita in $PPP) for computation of the component indices in the same way as they were used for computation of the HDI. The distributional data on age-at-death, years of schooling and income were used to calculate the Gini indices which were used to rescale the original component indices. He used data for 20 developing countries.

Health distributional data were obtained from mortality statistics in the *U.N. Demographic Yearbook 1992, Special Topic: Fertility and Mortality Statistics*. Age-at-death data were grouped by age into ten classes and the class midpoint number of years was assigned to each person in that class. The Gini index for life-span attainment was calculated using trapezoidal rule of numerical integration. It ranged between 0.15 and 0.63 for the 20 countries used in the illustration.

Education distributional data were taken from Ahuja and Filmer (1995), who used the Barro-Lee data set on the highest degree attained, mostly compiled from different UNESCO publications. The data were classified into 6 categories (“no education”, “some primary”, “completed primary”, “some secondary”, “completed secondary”, and “some higher education”). Hicks, following Ahuja and Filmer, assigned to these categories 0, 3, 6, 9, 12, and 15 years of schooling respectively, and from these data he approximated the Gini index for the education component for each country considered using the same trapezoidal rule as for health. The Gini index ranged between 0.32 and 0.65.

The distributional data on income were obtained from the World Development Report (WDR) 1995 in quintile shares plus the top decile. From these six data points the Gini index was obtained by the same trapezoidal rule. The Gini index for the income component ranged between 0.28 and 0.60. Apparently, there were countries in the Hicks’ illustration that exhibited higher inequality in dimensions other than income.

One disadvantage of distributionally grouped data is the suppression of within-group inequality such that the between-group Gini underestimates the real extent of inequality. A practical advantage of the Hicks index is that while it requires the distributional grouped data for all three dimensions for each country, they are not necessarily across the same groups - the distribution of health dimension was across the age-groups, the distribution of education was over education categories, and the income distribution was over the income quintiles.

Hicks presented inequality in the Human Development Index by calculating the percent-age loss due to inequality:

$$\%LOSS = \frac{HDI - IAHDI}{HDI} \times 100$$

The loss ranged between 29.6 and 56.6 percent.

The LOSS is actually a weighted mean of Ginis for three dimensions with the weights being equal to the shares of corresponding dimension indices in the HDI.[17] These shares vary from country to country. The percentage contribution of inequality in a dimension, say E , to total HDI inequality is found to be:

$$\%P^{LE} = S^{LE} \times \frac{G^{LE}}{G^{HDI}} \times 100$$

One may also wish to know the marginal contribution of each dimension to inequality. In other words, by what percentage would inequality increase if a dimension index increased by x percent? Lerman and Yitzhaki (1984) show that the elasticity of the Gini coefficient with respect to the component share, say *S*, is given by:

$$\epsilon^{LE} = S^{LE} \times \frac{G^{LE} - G^{HDI}}{G^{HDI}}.$$

[17] The loss due to inequality
$LOSS = \frac{HDI - IAHDI}{HDI} = (S^{LE}G^{LE} + S^{E}G^{E} + S^{I}G^{I})/3,$

where $S^{LE} = \frac{LE_{ind}}{HDI}$, $S^{E} = \frac{E_{ind}}{HDI}$, and $S^{I} = \frac{GDP_{ind}}{HDI}$

Hence, whenever the Gini of a HDI dimension is greater than the overall Gini index, an increase in that dimension (holding everything else constant) will increase inequality. In particular, if the share of LE increases by 1 percent, overall inequality will increase by ϵ^{LE} percents.

As noted before, the Hicks method can be applied to distributional grouped data. If distributional data are available at the level of individuals or households the Hicks method can also be applied in this case. The weakness of the method remains the lack of subgroup consistency and the equal weight the Gini index places throughout the distribution instead of emphasizing the lower part of the distribution.

5.2 Inequality accounted for by the General Mean of General Means

Using the Atkinson's measure of inequality[18] instead of the Gini coefficient in the expression of the Sen welfare standard, Foster et al. (2005) arrived at Atkinson's class of welfare functions $W_{\varepsilon}(x) = \mu(x)[1 - I_{1-\varepsilon}(x)$ This class of welfare functions satisfies all the basic properties of a well defined welfare standard, plus it satisfies the subgroup consistency which the Sen welfare standard does not satisfy as pointed in the previous subsection.

Indeed, from the expression of the Atkinson's measure of inequality one can see that $\mu_{1-\varepsilon} = \mu(x)[1 - I_{1-\varepsilon}(x)$. It can be interpreted as the (arithmetic) mean of x discounted for the inequality in distribution measured by the Atkinson's measure with parameter ε. In this way the

[18] Atkinson (1970) studies inequality in one-dimensional distribution by using the class of general means of order ε defined as

$\mu_{1-\varepsilon}(x,a) = (a_1 x_1^{1-\varepsilon} + \cdots + a_n x_n^{1-\varepsilon})^{\frac{1}{1-\varepsilon}}$, for ε≠1,

where a is a vector of weights which are all positive and sum to one. These means are distribution sensitive: for ε>0 the mean puts more weight on the lower part of the distribution, for ε=0 it is neutral, and for ε<0 it is upper-sensitive. The higher the ε the more emphasis is on the lower part of distribution so that $\mu_{1-\varepsilon}(x)$ is smaller than the neutral arithmetic mean. Therefore, the order **ε** is interpreted as an aversion towards inequality across persons. The larger inequality in distribution leads to smaller value of the ratio $\frac{\mu_{1-\varepsilon}(x)}{\mu_1(x)}$ Atkinson defines the family of inequality measures as

$I_{\varepsilon} = 1 - \mu_{1-\varepsilon}/\mu_1$.

adjustment of the mean is done for a single dimension. In order to provide a combined adjustment to all three dimensions, some kind of averaging over dimensions is needed. The simple arithmetic mean over the dimensions, as in the case of the Hicks' index, $\mu(\mu_{1-\varepsilon}(x), \mu_{1-\varepsilon}(y), \mu_{1-\varepsilon}(z))$, results in an index that does not satisfy the subgroup consistency (Foster et al., 2005).[19]

However, the general mean of the general means of the same order defines a family of composite indices and maintains all the properties of well defined welfare function including the subgroup consistency (Foster et al., 2005):

$$H_\varepsilon = \mu_{1-\varepsilon}(\mu_{1-\varepsilon}(x), \mu_{1-\varepsilon}(y), \mu_{1-\varepsilon}(z)) \text{ , for } \varepsilon \neq 0..$$

Also, the order of aggregation when using the general mean of the general means of the same order can be altered with the value of the index remaining unchanged. This property of a composite index is known as path independence.

For example, H_2 is a harmonic mean of harmonic means ($\varepsilon = 2$):

$$H_2 = \frac{3}{\frac{n}{\frac{a_1}{x_1} + \cdots + \frac{a_n}{x_n}} + \frac{n}{\frac{a_1}{y_1} + \cdots + \frac{a_n}{y_n}} + \frac{n}{\frac{a_1}{z_1} + \cdots + \frac{a_n}{z_n}}}$$
$$= \frac{3n}{a_1\left(\frac{1}{x_1} + \frac{1}{y_1} + \frac{1}{z_1}\right) + \cdots + a_n\left(\frac{1}{x_n} + \frac{1}{y_n} + \frac{1}{z_n}\right)} = \mu_{-1}(D)$$

where D denotes all the data irrespectively of dimension.

[19] Note that the HDR's Gender Related Development Index (GDI) is of the form $\mu(\mu_{1-\varepsilon}(x), \mu_{1-\varepsilon}(y), \mu_{1-\varepsilon}(z))$ with $\varepsilon = 2$. However, because there are only two units (female and male group) across which the HDI is calculated the subgroup consistency the argument of Foster et al. (2005) about subgroup inconsistency doesn't apply.

Foster et al. (2005) extend the family of Atkinson inequality measures, I_ε, to a new multidimensional measure of inequality

$$I_\varepsilon(D) = 1 - \frac{\mu_{1-\varepsilon}(D)}{\mu_1(D)}. \qquad (2)$$

Using the link between welfare and inequality they write the new index H_ε as:

$$H_\varepsilon(D) = H_0(D)[1 - I_\varepsilon(D)] = HDI[1 - I_\varepsilon(D)] \qquad (3)$$

i.e., the level of human development according to the new class of indices is the original HDI discounted by the level of inequality among all the entries measured by the multidimensional Atkinson measure. We refer to (2) and (3) in this paper as the FLS inequality measure and FLS index, respectively.

To illustrate the FLS measure, Foster et al. used data from Mexico. The income and education component came from a sample from the 2000 Population Census. The data were available at the individual level (over 10 million records). The data on infant survival rates were available at the municipality level. For each individual in the sample, the income, education, and health indices were constructed. The household income per capita was the ratio-adjusted to the state GDP per capita. The FLS index, the new HDI index, H_ε, was computed for each state by aggregating first within each dimension (across persons), and then by aggregating across dimensions. Two values of ε were used, $\varepsilon = 0$ and $\varepsilon = 3$. For $\varepsilon = 0$ the corresponding index is neutral to inequality and is equivalent to the original HDI, while $\varepsilon = 3$ results in the inequality sensitive index. For $\varepsilon = 3$, the value of corresponding national Mexican HDI was reported as H_ε=0.4912 which was 20% lower than the original HDI. The reduction in HDI due to inequality presents itself as an important development focus.

As in the case of Hicks' IAHDI, it is possible to apply the FLS measure even when data are only available as group aggregates and the groups are not necessarily the same across dimensions. However, as we pointed out earlier, by using the group aggregates, within-group inequality will not be accounted for.

Seth (2009) argues that the FLS index is not strictly association-sensitive but is weakly association-sensitive due to its path independence.[20] A strictly association-sensitive index must aggregate first across dimensions and then across persons, and it must not be path independent.

5.3 Association-sensitive inequality index

The role of composite indices is essentially to synthesize a multidimensional distribution into a one-dimensional distribution preserving to some extent the relevant information about the original distribution. It seems desirable that any measure of inequality in distribution of a composite index should be sensitive to both inequality in the distribution of each component dimension across population, and the correlation between the component dimensions.

The first attempts for formal measurement of multidimensional inequality were made by Kolm (1977) who provided the multivariate generalizations of the Pigou-Dalton principle of transfer. Atkinson and Bourguignon (1982) developed dominance principles for ranking multivariate distributions according to the magnitude of inequality. These criteria account for correlations between the components of the multivariate distribution. In terms of social preference, the weaker associations between dimensions are valued more, so that the wealthiest person is not the healthiest and the most educated.[21] This notion is formalized by Tsui (1995) into a criterion named "correlation increasing majorization." A good review is given in Lugo (2005). Seth (2009) reviews the criteria and principles of multidimensional dominance. Decancq, Decoster and Schokkaert (2009) applied several variants of inequality measures based on general means of general means in their study of the evolution of world inequality in well-being.

---

[20] The strict association-sensitivity is related to the axiom of strictly decreasing under increasing association or to the alternative axiom of strictly increasing under increasing association. The meaning of the first axiom is that under the association-increasing transfer the composite index is decreasing. The second axiom says that under the association-increasing transfer the composite index is increasing too. The weak variant of either axiom allows no change in index under the association-increasing transfer.

[21] This notion is echoed in recent paper by Thomas Pogge (2009): "a credible index of development must be sensitive to whether an increase in literacy goes to landowners or the landless, an improvement in medical care goes to children or to the aged, an increase in enrollment to privileged university students or to children in slums, an increase in life expectancy to the elite or to the marginalized, enhanced physical security to males or to females."

To proceed further with the idea of multivariate inequality, let us first define more formally the two-stage aggregation mentioned in the previous section. The first stage aggregation is over K dimensions for each individual, and the second stage aggregation is across individuals, i=1,…,n. The first stage aggregation produces a value of the composite index at the individual level:

$$\mu_\beta(w,i) = (w_1 x_{1,i}^\beta + \cdots + w_K x_{K,i}^\beta)^{1/\beta}, \text{ for } \beta \neq 0. \qquad (4)$$

The weights w_k reflect the extent to which each dimension contributes to individual well-being and they sum to 1. Parameter β is related to the degree of substitutability between dimensions[22]. The dimensions can be transformed by real-valued functions if needed. In such a case the equation (4) takes a more general form

$$\mu_\beta(w,i) = \left\{w_1\left[f_1\left(x_{1,i}\right)\right]^\beta + \cdots + w_K\left[f_K\left(x_{K,i}\right)\right]^\beta\right\}^{1/\beta}.$$

Different choices in terms of weights, transformations, and values of β will lead to different composite indices.[23] [24]

The first stage aggregation reduces the dimensions to a one-dimensional composite index validated at the individual level. Our goal is to measure inequality in the distribution of $\mu_\beta(w,i)$ across persons. In this situation we can use a one-dimensional Atkinson-Kolm-Sen inequality index along the lines introduced in the previous section.

The second stage aggregation is across individuals using a general mean of order ε defined as

---

[22] The elasticity of substitution between each pair of dimensions is given by $1/(1\text{-}\beta)$. Evidently, for $\beta \to 1$, the elasticity $1/(1\text{-}\beta)$ tends to ∞, and the substitution of dimensions is perfect, with the infinite elasticity . When $\beta \to -\infty$, the elasticity $1/(1\text{-}\beta) \to 0$ meaning that the dimensions are the perfect complements, i.e., with zero-elasticity. The common restriction is that $\beta \leq 1$.

[23] Note that the HDI uses certain transformations of the original indicators – linear transformations for education and life expectancy dimensions, and a combination of logarithmic and linear transformation for the income component. It uses the equal weights, plus parameter β is equal to 1, making the dimensions perfectly substitutable.

[24] For β =0, equation (2) takes the form of a geometric mean $\mu_0(w,i) = \prod_k x_{k,i}^{w_k}$

$$\mu_{1-\varepsilon}(a,w,\beta) = \{a_1[\mu_\beta(w,1)]^{1-\varepsilon} + \cdots + a_n[\mu_\beta(w,n)]^{1-\varepsilon}\}^{\frac{1}{1-\varepsilon}}, \text{ for } \varepsilon \neq 1, \quad (5)$$

where a is a vector of individual weights which sum to one. As before, ε is interpreted as "aversion" towards inequality across population. Larger inequalities in distribution lead to smaller values of the ratio $\mu_{1-\varepsilon}(a,w,\beta)/\mu_1(a,w,\beta)$ The Atkinson-Kolm-Sen family of inequality measures is

$$I_{\varepsilon,\beta} = 1 - \mu_{1-\varepsilon}(a,w,\beta)/\mu_1(a,w,\beta) \qquad (6)$$

The criterion of strict correlation-increasing majorization is satisfied if $\varepsilon + \beta > 1$(Bourguignon and Chakravarty, 2003).

Note that if $\varepsilon + \beta = 1$, equation (6) reduces to equation (2), and equation (5) becomes a general mean of general means of the same order, i.e., FLS index.

Seth (2009) used the same data as Foster et al. (2005) to calculate the HDI indices and their inequalities for 33 Mexican states. Seth applied three sets of parameters ε and β: (i) ε=0, $\beta = 1$, resulting in the original HDI; (ii) ε =3, β =-2, this combination creates general means in both stages that are of the same order, as in the FLS index; (iii) ε =4, β =-1, this combination satisfies the necessary condition for the criterion of strict correlation-increasing maximization. Thus the inequality measurement stemming from (iii) is sensitive to the association between the dimensions. The elasticity of substitution is placed at a moderate level of ½ leading to aggregation across the dimensions in the form of a harmonic mean. Also this value of β (in combination with $\varepsilon = 4$) "penalizes" for increasing the inter-dimensional association. In this case, the aversion towards inequality across populations $\varepsilon = 4$ is considerable. Seth illustrated this feature with the example of a Mexican state whose HDI score dropped from 0.254 to 0.244 due to a higher association between dimensions.

Based on the index that is sensitive only to distribution across individuals, Seth (2009) argues, that assuming that the dimensions are perfect substitutes, governments should target the poorest person in terms of overall achievement. If the dimensions are perfect complements then policy makers should target the person with the least achievement in any dimension and assist that person to improve the most deprived dimension. Using the association sensitive index with carefully selected parameters, governments should target the person with the smallest value of this index since it is already adjusted for the socially acceptable degree of substitution between dimensions.

5 INEQUALITY-ADJUSTED HDI: AN APPLICATION OF THE FLS METHODOLOGY

In the 2010 Human development report background research paper "Designing the Inequality-Adjusted Human Development Index (IHDI)", Alkire and Foster proposed a modification to the methodology used by Foster, Lopez-Calva, and Szekely (2005)to adjust the HDI for inequality in the distribution of each dimension across populations. Their proposed measure is referred to as the Inequality-adjusted HDI (IHDI). It is computed as a geometric mean of geometric means, calculated across the population for each dimension separately. The IHDI accounts for inequalities in HDI dimensions by 'discounting' each dimension's average value according to its level of inequality. The IHDI is equal to the HDI when there is no inequality across the population, but falls below the HDI as inequality rises. In this sense, the IHDI reflects actual levels of human development (by accounting for inequality), while the HDI can be viewed as an index of 'potential' human development (or the maximum level of IHDI) that could be achieved if there was no inequality. The percentage 'loss' in potential human development due to inequality can be measured as the difference between the HDI and the IHDI.

In the following subsections we review Alkire and Foster's proposal in light of the revised HDI that is included in the 2010 HDR. We address and justify some approximations that were necessary in the process of computing the IHDI. We describe the treatment of extreme values in distributions, provide a limited sensitivity analysis, and give a thorough description of data sources used for computing the 2010 IHDI.

6.1 The new HDI

The concept and the basic approach to the creation of the composite index for measuring human development in the 2010 HDR remains the same as for the original HDI. However, the indicators used to measure knowledge and a decent standard of living have changed, as have the education dimension indices. The method of aggregating dimension indices has also changed. The new HDI measures achievements in the three dimensions by combining a country's life expectancy at birth, expected years of schooling for children, mean years of schooling for adults aged 25 years and older, and gross national income per capita.

A key change in the functional form of the index was to shift from calculating the arithmetic mean to calculating the geometric mean of three dimension indices. Also, the dimension indicators are now rescaled based on the observed maxima over the period for which the HDI trends are presented (1980-2010) and by the minima set to subsistence levels or natural zeros.

The dimension indicators are transformed using the maximum levels for all sub-components observed over the period for which the HDI trends (from 1980) are presented. The 2010 HDI is thus based on maximum life expectancy of 83. 2 years, mean years of schooling of 13.2 years, expected years of schooling of 20.6 years, and GNI per capita (PPP$) of 108,211. Minimum levels for the dimension indicators are set as follows: life expectancy at 20 years; both education variables at 0, and GNI per capita at PPP US$ 163, which is the observed minimum. The choice of the minimum values is motivated by the principle of natural zeros below which there is no possibility for human development. This way of normalizing has the effect of making the component sub-indices of these dimensions vary along a similar range. For more information about the changes in the HDI see Technical note 1 in 2010 HDR.

Having defined the minimum and maximum values, the dimension indices are calculated as follows:

$$\text{Dimension index} = \frac{\text{actual value} - \text{minimum value}}{\text{maximum value} - \text{minimum value}} \qquad (7)$$

For education, equation (7) is applied to each of the two education subcomponents (mean years of schooling and expected years of schooling), then the geometric mean of the resulting indices is calculated. Finally, equation (7) is re-applied to the geometric mean of the indices using 0 as the minimum and the highest geometric mean of the resulting indices for the time period under consideration as the maximum (for the period from 1980 to 2010 the maximum mean was 0.951).

The income component is logarithmically transformed to account for the diminishing returns of income. Accordingly, the logarithm of the actual minimum and maximum values is taken to normalize the income indicator.

The aggregating of the component indices is done using equal weights. Note that equal weights are used to aggregate the education sub-indices too:

$$HDI = (LE_x)^{1/3}(EDU_x)^{1/3}(INC_x)^{1/3} \qquad (8)$$

Here the subscript *x* denotes the index.

6.2 Approximation of underlying distributions

Ideally, the HDI should be based on indicators that represent national averages across a population. For example, one can think about mean years of schooling as the average years of schooling across adult members of the population. When thinking about inequality in the education dimension of the HDI, one can easily picture inequality in the distribution of mean years of schooling. In other dimensions assessing the distribution is less straight forward. The HDI relies on country-level aggregates obtained from the system of national accounts for income and the life tables for life expectancy. Hence, the IHDI must draw on alternative sources of distributional data on living standards, education, and health. Given that the IHDI is based on approximations of the real HDI indicators, the calculated inequalities can serve ***only*** *as approximations* of inequalities in the HDI distribution.

Expected years of schooling is derived from the gross enrolment ratio which is based on a binary variable making the estimation of inequality in its distribution difficult. Inequality in the education dimension is thus based only on inequality in the distribution of mean years of schooling. Distributions of years of schooling for adult populations are found in a variety of nationally representative household surveys, which are harmonized and stored in international databases including: the Luxembourg Income Study, Eurostat's European Union Survey of Income and Living Conditions, the World Bank's International Income Distribution Database, the United Nations Children's Fund's Multiple Indicators Cluster Survey Database, the U.S. Agency for International Development's Demographic and Health Survey, the World Health Organization's World Health Survey, and the United Nations University's World Income Inequality Database. Appendix 3 contains a full list of countries and the source of data used for 2010 IHDI estimations.

Inequality in the standard of living dimension is assessed from the distribution of disposable household income per capita, or household consumption per capita both of which are obtained from nationally representative household surveys and are available in international databases. For countries where neither income nor consumption were available but the assets index was calculated for surveyed households, the income was imputed based on an asset index matching methodology (Harttgen and Klasen, 2010). By their very nature, income and consumption yield different levels of inequalities, with income inequality being higher than inequality in consumption. Income seems to correspond more naturally to the notion of "command over resources", but income data pose technical challenges because of the greater presence of zero and negative values. Consumption data are arguably more accurate in developing countries, less skewed by high values, and directly reflect the conversion of resources. In an ideal world, one would be consistent in the use of either income or consumption data to estimate inequality. However, to obtain sufficient country coverage, it was necessary to use both. Final IHDI estimates are slightly influenced by whether data are income or consumption based.

Inequality in the health dimension is estimated using the abridged life tables provided by UNDESA (2009d). Because it is based on a special approach we describe it in detail below.

6.3 Inequality in life length estimated from the UN Model Life Tables

There is an important conceptual issue in relation to 'inequality' in the domain of public health where the term usually denotes the differences in health between social groups. Sometimes this type of inequality is called health inequity (Marmot, 2007). Within the context of measuring inequality in the human development distribution, there is a strong voice for including a health dimension because the development of society, rich or poor, can be judged by how fairly health is distributed across the social spectrum. This type of reasoning reflects on socio-economic inequalities in health, while also incorporating a notion of fairness and justice. Quantifying health inequity has proven to be difficult because of the lack of data, and also because of the conceptual complexity.

In our approach to quantifying inequality in the distribution of the expected length of life, we are not explicitly taking into account the socio-economic stratification of the population. However, the approach implicitly conflates both disparities in distributions across groups and differences in expected lengths of life based on observed or modeled age-specific mortality rates. Improvements in the socio-economic and environmental conditions of the population affect the length of life of individuals. Our inequality measure doesn’t capture inequality in individual lengths of life but rather in expected lengths of life. First, there is a “mortality experience” of the population captured by age-specific mortality rates. The “inequality in health dimension” of the IHDI is essentially the inequality in expected life lengths obtained from the mortality rates which thus reflects the uncertainty of dying in the period for which the HDI pertains. The life expectancy is the weighted average of expected lengths of life which is obtained from the same distribution (expressed by life tables) whose inequality we are assessing.

The life tables have been used to estimate inequality in the distribution of age at death or equivalently the length of life. Illsey and Le Grand (1987) computed a Gini index from distributions of death by age in real populations. Several other authors have computed a Gini index and other measures of inequality in the distribution of age at death using the life tables (Hanada, 1983, Anand and Nanthikesan, 2001, Shkolnikov et al. 2003, Smits and Monden, 2009. In the context of inequality assessment for the HDI distribution, Hicks (1997) used the life tables to compute a Gini for length of life distribution, and similarly, Stanton (2007) used the abridged life tables.

Generally, a **model life table** follows a hypothetical cohort of 100,000 people born at the same time as they progress through successive ages, with the cohort shrinking from one age to the next according to a set of death rates by age until all people eventually die. Such a table is used for computation of **life expectancy at birth** broadly defined as the average number of years a group of people born in the same year can be expected to live under the constant-mortality assumption, i.e., mortality is maintained constant at the level estimated for the reference year or period.

The most recent UN life tables[25] were produced for the 2008 Revision of the United Nations *World Population Prospects*. The tables were produced under two mortality assumptions – the normal mortality assumption where mortality is projected on the basis of models of change in life expectancy produced by the United Nations Population Division. The selection of a model for each country is based on recent trends in life expectancy by sex. The HIV/AIDS mortality assumption was made for countries where HIV prevalence among persons aged 15 to 49 was equal to or greater than one per cent at any point between1980 and 2007.Their mortality rates are projected by modelling the course of the epidemic and projecting the yearly incidence of HIV infections. In total, 58 countries were considered to be affected by the HIV/AIDS epidemic in the UNDESA (2009).

[25] UN life tables are based on " the most consistent international collection of life tables that are included in the Human Mortality Database (HMD) created and maintained by the Department of Demography at the University of California, Berkeley (USA) and the Max Planck Institute for Demographic Research in Rostock (Germany) … However, for the vast majority of 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. The data may include direct and indirect measures from censuses and surveys. The UN Population Division is collaborating since 2004 with WHO and The World Bank with an inter-agency process led by UNICEF that produces common child mortality estimates…"[Heilig et al. (2008)]

A typical life table consists of several columns/variables. Here we mention those that are relevant for our computation:

Age - the initial age of the age interval $(x, x+n)$, where *x* is the age and *n* is the length of the interval.[26]

$l(x)$ - number of survivors at age *x* (of the hypothetical starting cohort of 100,000)[27]

${}_n a_x$- average number of years lived in the age interval $(x, x+n)$ by those dying during the age interval.

e_x – life expectancy at age *x*. Life expectancy at birth is given by e_0

We first compute the *proportion of the hypothetical cohort* which dies in the age interval $(x, x+n)$:

$$w_n(x) = \frac{l(x) - l(x+n)}{100{,}000}$$

so that $\sum_{x=0}^{85} w_n(x) = 1$ Then, we approximate the expected age of dying for those who die in the age interval $(x, x+n)$ by

$$A_n(x) = x + {}_n a_x$$

Therefore, the distribution of the expected age at death is defined by $\{A_n(x), w_n(x)\}$. The average expected age at death (or the life expectancy at birth) is obtained as

$$M_0 = \sum_{x=0}^{85} w_n(x) \cdot A_n(x).$$

The general mean of order ε for the distribution of expected age at death is given as

$$M_\varepsilon = \{\sum_{x=0}^{85} [w_n(x) \cdot A_n(x)]^{1-\varepsilon}\}^{\frac{1}{1-\varepsilon}} \text{ for } \varepsilon \neq 1,$$

and for ε=1,

[26] For the abridged LT the interval *n* equals five years with the exception of the first interval (one year), second interval (four years) and the last interval (open ended).

[27] Such a hypothetical population is denoted as *radix* in demography.

$$M_1 = \prod_{x=0}^{85} [A_n(x)]^{w_n(x)}$$

The Atkinson measure of inequality with the aversion parameter ε=1 for the distribution of expected length of life is computed as

$$I_1 = 1 - M_1/M_0.$$

As an illustration we present several countries with similar life expectancies, but very different distributions of expected age at death (implying that estimated inequalities in distribution are different). Within a group of countries with similar life expectancies – for example Mongolia, Pakistan and the Russian Federation, levels of inequality can be very different (Figure 6.1 and Table 6.1).

Figure 6.1. Cumulative distribution of age-at-death for seven countries

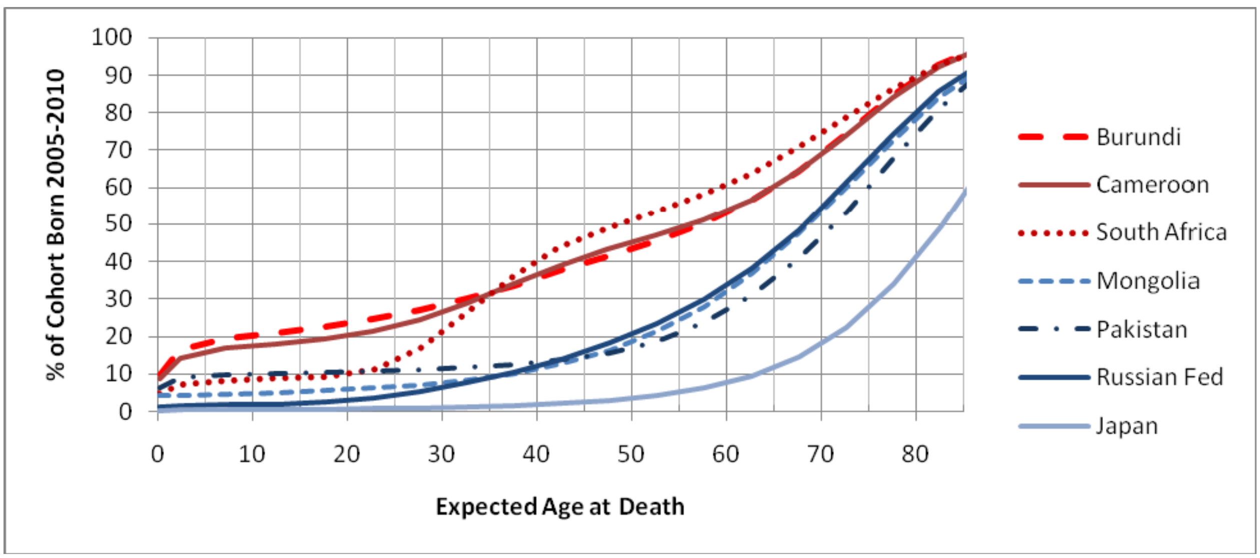


Table 6.1. Life expectancy and the inequality measure for seven countries

	Life expectancy	Atkinson measure (ε=1)
Japan	83.2	0.039
Burundi	51.4	0.478
Cameroon	51.7	0.444
South Africa	52.0	0.302
Mongolia	67.3	0.226
Pakistan	67.2	0.329
Russian Federation	67.2	0.115

There is a relationship between life expectancy and inequalities in distribution -the longer life expectancy is, the more likely it is that there are very few short lives and the vast majority of lives are long and of a similar length. Hence, there is a reason to study inequality in the distribution of length of life (from Life tables) conditionally on life expectancy. There is also a strong negative correlation between under 5 mortality and life expectancy (-0.90 in 2010). The Atkinson measure, which is sensitive to values at the lower end of the distribution, will be higher for countries with higher under 5 mortality (correlation is equal to 0.96) and thus inequality will be higher for countries with shorter life expectancies (correlation is -0.95). It is possible to control for life expectancy when assessing inequality by looking at inequality in the distribution of the length of life after age 5. Smits and Monden (2009) use the distribution after age 15.

Our approach is to consider inequality in the entire distribution of the expected length of life without exclusion of any part of the distribution. Excluding newborns or children under 5 would mean capturing only partial inequality in the distribution. It should not be ignored that people die at a younger age in Pakistan than in Russia at such a rate that inequality is tripled while life expectancy stays at the same level (see the example above). Finally, our measure of inequality in the distribution of age at death can be used to correct life expectancy for largely preventable premature deaths, and help differentiate between countries even with the same average life expectancy.

6.4 Adjusting for extreme values in distribution of mean years of schooling and income

As stated earlier, general means are not defined for zero and negative values for $\varepsilon \neq 0$. Therefore, we have to decide how to handle cases with zero years of schooling, as well as reported zero and negative incomes. The Atkinson measure is sensitive to values at the lower end of distribution, so any intervention at the lower end will have a disproportionate impact on the assessed inequality.

6.4.1 Adjusting the years of schooling

We examined several options for adjusting the distribution of years of schooling so that the Atkinson measure could be computed. However, all of them alter the distribution to some extent and thus interfere with inequality.

The first option was to replace the zero value with a small positive value that would be kept constant across countries. In such a case the ratio between the geometric and arithmetic mean takes the form:

$$\alpha^{(\sum w_i)} \prod (x_i)^{w_i} / \left[\alpha \sum w_i + \sum w_i x_i\right]$$

i goes over original zeros and j over non-zeros. An alternative to this option is to apply the country - specific replacement to each national data set. The question is how to determine the optimum value of α. It seems that the effect of α depends on the number of zero values, or more precisely on the share of the adult population with zero years of education. The smaller the share is, the smaller the impact of the choice of α on the value of the Atkinson measure. The table below represents three countries with very different distributions of mean years of schooling estimated from UNICEF's Multiple Indicators Cluster Survey:

Table 6.2. Distribution of years of schooling in three developing countries

	Year of MICS	Mean	Median	Share of adult population with zero years of schooling (in %)
Burkina Faso	2006	1.3	0	85.8
Suriname	2006	7.2	7	12.9
Montenegro	2005	10.6	12	3.3

Figure 6.2. Approach 1: Change of Atkinson measure depending on the value of replacement for zero observations

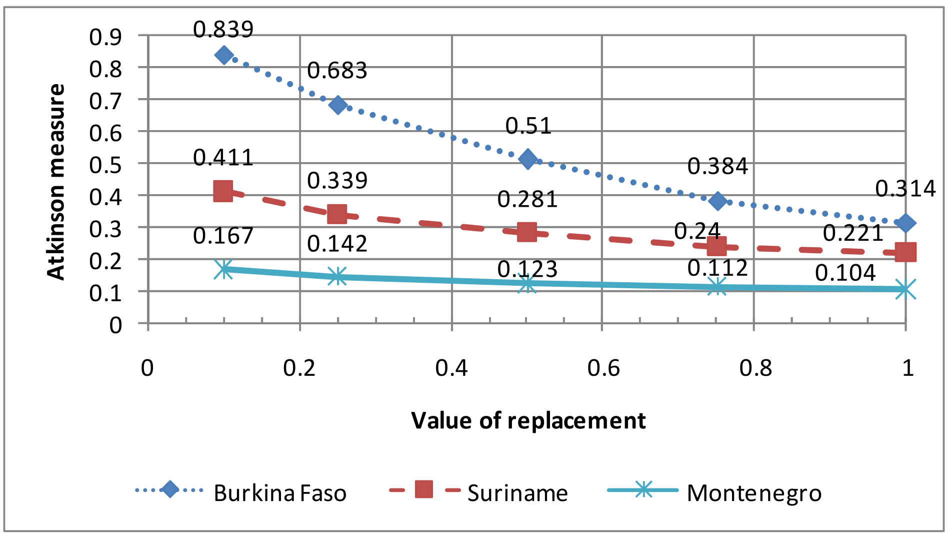


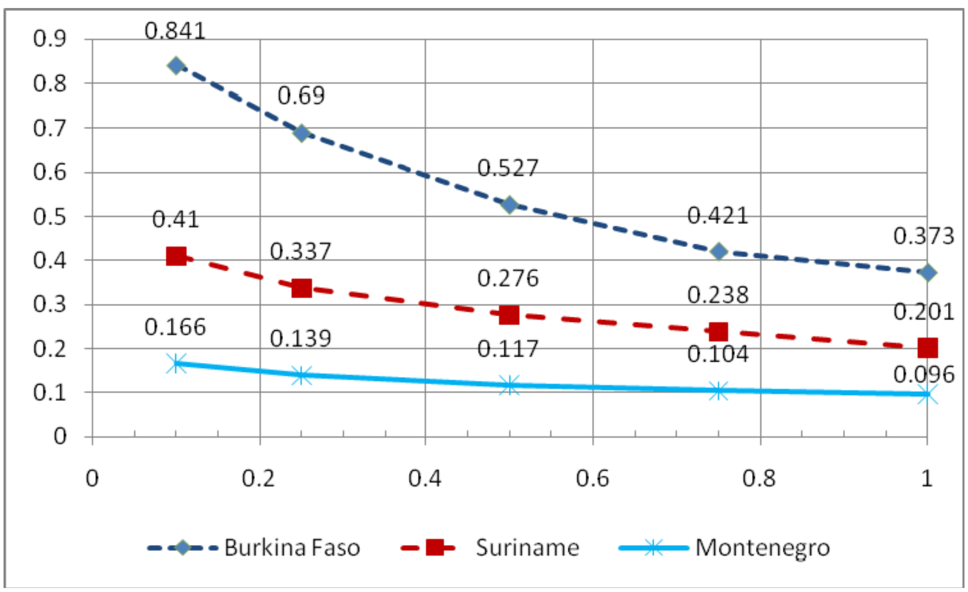
Figure 6.2 shows changes in the value of the Atkinson measure depending on the replacement value α. We used five different values (α=0.1, 0.25, 0.5, 0.75, 1). Obviously, using a larger value of α reduces the inequality, and the reduction varies across countries. For example by increasing α from 0.25 to 0.75, the Atkinson measure for Burkina Faso is reduced from 0.683 to 0.384 which represents a reduction of 44%, while the Atkinson measure for Montenegro, for the same change in α, reduces the measure by only 21%.

The other approach was to translate the entire distribution away from the zero value by adding the same constant value to all the observed values and not only to zero values. It is known that the Atkinson measure doesn't satisfy the translation invariance principle[28], but our reasoning was that it would be a fair and just approach to treat the differences in education "additively" so that the difference between having one year of schooling and not having any -remains one year. In this case the ratio between the geometric and arithmetic mean takes the form:

$$\alpha^{(\sum w_i)} \prod (x_i + \alpha)^{w_i} / \left[\alpha + \sum w_i x_i\right]$$

In the example of the three countries shown above for the same five values of α, the estimated inequality using this technique is presented in Figure 6.3.

Figure 6.3. Approach 2: Change of Atkinson measure depending on the value of replacement for zero observations



The estimated values of the Atkinson measure obtained using the two approaches differ very little. The decision was made to use the second approach with α =1. The reasoning was the following – while the two approaches are slightly different; the second approach leaves an

[28] Translation invariance principle says that the measured inequality does not change by changing all values of distribution by the same amount as long as the changes do not lead to a negative or zero value. This is regarded as a (politically) leftist axiom for inequality view.

impression of fairness, and by using α =1 we are penalizing countries the least, which, we felt, was important for the introduction of the new IHDI index.

Another option that was considered was to compare the Atkinson measure obtained for different α and the Gini index which is not sensitive to zero values, and choose α which is slightly above the Gini index in order to emphasize the lower part of the distribution. This reasoning would lead to selecting different values of α for different countries and make the countries comparison impossible. Also, it doesn't take into account that the Gini index is generally situated between the Atkinson measures obtained for ε=1 and ε=2.

6.4.2 Treatment of negative, zero and extreme incomes/consumption

In our approach to treating the outliers in income data, both extremely high values as well as negative and zero values, we didn't differentiate between income and consumption based data. We considered four strategies in dealing with the outliers. They are combinations of the following approaches to treating negative and zero values on the one hand and extremely high values on the other hand.

For the treatment of negative and zero values we considered two options: a) replacement with the minimum non-zero value; and b) replacement of negative values, zero values and the first 0.5 percentile of non-zero values by the minimum of the second 0.5 percentile.

For the treatment of extreme high values we considered two options: c) keep them as is; and d) trim the highest 0.5% of values.

Strategy 1 combines options (a) and (c); strategy 2 combines options (a) and (d); strategy 3 combines options (b) and (c); and strategy 4 combines options (b) and (d).

We applied these four strategies to data from the Luxembourg Income Study and found a systematic ordering of strategies 1 and 4 across 13 countries. Strategy 1 always results in the largest Atkinson measure, whilst strategy 4 in the smallest. Strategies 2 and 3 are not ordered systematically – depending on the country these values alter. We chose strategy 4 for this year's exercise – because it produces the smallest value of inequality – one can refer to this as the lower boundary of inequality in the distribution of the living standard dimension. A table with the full results of the analysis is given in the Appendix 2.

6.5 Computing the Inequality-adjusted HDI

There are three steps to computing the IHDI.

Step 1. Measuring inequality in the underlying distributions

The Inequality-adjusted HDI (IHDI) draws on the Atkinson (1970) family of inequality measures and sets the aversion parameter ε equal to one.[29] In this case the inequality measure is $A = 1 - g/\mu$, where g is the geometric mean and μ is the arithmetic mean of the distribution. This can be written:

$$A_x = 1 - \frac{\sqrt[n]{X_1 \ldots X_n}}{\bar{X}} \qquad (9)$$

where $\{X_1, \ldots, X_n\}$ denotes the underlying distribution in the dimensions of interest. A_x is obtained for each variable (life expectancy, years of schooling and disposable income or consumption per capita) using household survey data and the life tables.[30]

Step 2. Adjusting the dimension indices for inequality

The mean achievement in a dimension, $\bar{X}$, is adjusted for inequality as follows:

$$\bar{X}^* = \bar{X}(1 - A_x) = \sqrt[n]{X_1 \ldots X_n} .$$

Thus $\bar{X}^*$, the geometric mean of the distribution, reduces the mean according to the inequality in distribution, emphasizing the lower end of the distribution.

---

[29] The inequality aversion parameter guides the degree to which lower achievements are emphasized and higher achievements are de-emphasized

[30] A_x is estimated from survey data using the survey weights, $\hat{A}_x = 1 - \frac{X_1^{w_1} \ldots X_n^{w_n}}{\sum_1^n w_i X_i}$, where $\sum_{\mathbf{1}}^{n} w_i = \mathbf{1}$.

However, for simplicity and without loss of generality, equation 1 is referred as the Atkinson measure.

The inequality-adjusted dimension indices, I_{I_X}, are obtained from the HDI dimension indices, I_X, by multiplying them by $(1 - A_x)$, where A_x is the corresponding Atkinson measure:

$$I_{I_X} = (1 - A_x) \cdot I_x .$$

The inequality-adjusted income index, $I^*_{I_{Income}}$, is based on the unlogged GNI index, I^*_{Income}. This enables the Inequality-adjusted HDI to account for the full effect of income inequality.

Step 3. Computing the Inequality-adjusted HDI

The IHDI is the geometric mean of the three dimension indices adjusted for inequality. First, the Inequality-adjusted HDI that includes the unlogged income index (*IHDI**) is calculated:

$$IHDI^* = \sqrt[3]{I_{I_{Life}} \cdot I_{I_{Education}} \cdot I^*_{I_{Income}}} = \sqrt[3]{(1 - A_{Life}) \cdot I_{Life} \cdot (1 - A_{Education}) \cdot I_{Education} \cdot (1 - A_{Income}) \cdot I^*_{Income}}$$

The HDI based on unlogged income index (*HDI**) is then calculated. This is the value that *IHDI** would take if all achievements were distributed equally:

$$HDI^* = \sqrt[3]{I_{Life} \cdot I_{Education} \cdot I^*_{Income}} .$$

The percentage loss to the *HDI** due to inequalities in each dimension is calculated as:

$$Loss = 1 - \frac{IHDI^*}{HDI^*} = 1 - \sqrt[3]{(1 - A_{Life}) \cdot (1 - A_{Education}) \cdot (1 - A_{Income})} .$$

Assuming that the percentage loss due to inequality in the income distribution is the same for both average income and its logarithm, the *IHDI* is then calculated as:

$$IHDI = \left(\frac{IHDI^*}{HDI^*} \right) \cdot HDI$$

which is equivalent to

$$IHDI = \sqrt[3]{(1 - A_{Life}) \cdot (1 - A_{Education}) \cdot (1 - A_{Income})} \cdot HDI .$$

6.6 Aversion to inequality

The proposed method for calculating IHDI is implicitly based on the Atkinson measure of inequality and thus depends on the assumed level of societal aversion to inequality. Moreover the method assumes that this aversion is constant over all three HDI dimensions – income, education and health. Thus, it excludes the possibility that a society can be more tolerant of income inequality, and less so of inequality in the distribution of education and health. Also, different countries may have different levels of aversion to inequality in different dimensions. By selecting only one value for the aversion parameter, we impose that value on all societies as a norm. Foster et al. (2005, page 22) wrote “Perhaps one limitation of the distribution-sensitive HDI we present is that it requires the definition of a parameter, which determines the social aversion to inequality. The intuitive interpretation of this parameter is not straightforward”. An important question is raised about how to empirically determine the level of aversion expressed by parameter ε across dimensions and across populations.

The meaning and the interpretation of parameter ε is relatively well studied in the literature on income inequality, especially as it relates to the idea of redistribution of income by a “leaky bucket” approach (Okun, 1975). Okun’s hypothetical experiment was about a transfer of $4000 per head from the top 5 percent giving, in principle, $1000 to each of the bottom 20 percent of the income distribution. The idea is to reduce inequality in the income distribution in accordance with Rawls’ difference principle, which insists that “All social values-liberty and opportunity, income and wealth, and the basis of self respect – are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone’s advantage.(Rawls,1971, page 62)” But some of the money “leaks” away in the process of redistribution. Okun asked how much leakage would society accept before abandoning the proposed redistribution. Okun also points to a possible interpretation of “leakiness” as a manifestation of inefficiency of society.

Atkinson (1983) conjectures that the answer to Okun’s question can be used to “back up” a value of aversion parameter ε. He expresses the answer to Okun’s question in the form of

$$\frac{1}{x} = g \cdot \varepsilon \qquad (10)$$

where x represents the proportion of the transfer received by the recipient group and g represents the ratio of the average income of the group from which the transfer is made to the recipient group. For example, in a country where a hypothetical transfer is from the upper 10 percent to the lowest 20 percent of income holders and the ratio of income levels is 4, g=4; and if society will not tolerate less than half of the transfer going to the poor (x=0.5) then ε=0.5. But if society is ready to tolerate even higher inefficiencies in the transfer in order to make the distribution less unequal, for example it would tolerate a leakage of 75%, then x=0.25, implying ε=1. If aversion to inequality is such that society will accept only x=0.125, then ε=2, and so on. Note, however, that the ratio of average incomes of the upper 10 percent to the lowest 20 percent varies from country to country in a rather wide range, showing that the same tolerance to inefficiency in order to reduce inequality may lead to different values of parameter e. See the table below:

Table 6.3. Value of parameter ε for different values of g and x in Equation 10.

G	x=0.5	x=0.25	x=0.125
2	1	2	4
4	0.5	1	2
6	0.33	0.67	1.33
8	0.25	0.5	1
10	0.2	0.4	0.8

While the “leaky bucket” analogy can help our understanding of the meaning of societal aversion to inequality in income distribution, it doesn’t provide a practical way of assessing the aversion, nor does it provide justification for a normative choice of one value over another.

Amiel et al (1999) measured the individuals' attitudes to inequality aversion using survey data, based on the leaky-bucket experiment, for several groups of students in Australia and Israel. They found that these estimates are substantially lower than the values typically used by those measuring inequality. Further-more, a welfare function based on the Gini inequality measure is generally found to give a better fit than forms based on constant relative or constant absolute inequality aversion. In a recent paper about the estimation of inequality aversion, Pirtilla and Uusitalo (2010) wrote that existing evidence of inequality aversion relies on data from classroom experiments where subjects face hypothetical questions. They went a step further and estimated the magnitude of inequality aversion using representative survey data, with questions related to the real-economy situations the respondents face. Their results reveal that income and wage inequality aversion can be measured in a meaningful way using survey data, but the magnitudes of the estimates depend dramatically on how inequality aversion is measured.

What is the meaning of aversion to inequality in the distribution of other HDI dimensions, expected life-length and years of schooling? Can the leaky bucket analogy be used for these dimensions? While reduction of inequality in income distribution is possible through a transfer of income from rich to poor, reduction of inequality in life-length distribution is not – we cannot reduce the longevity of some to gain in length of life for others. Although the aversion to inequality of length of life exists, it cannot be assessed easily, nor can it be set normatively. Similar arguments can be repeated for education where it is difficult to imagine reducing the number of years of schooling for some to increase years for others. An approach different from the "leaky bucket" is needed to understand, identify and interpret aversions to inequality in non-income HDI dimensions and their levels.

While the HDI can be viewed as an index of the "potential" human development that could be obtained if achievements were distributed equally among the population, the IHDI is the actual level of human development (accounting for inequality in the distribution of achievements across people in a society). The IHDI will be equal to the HDI when there is no inequality in the distribution of achievement or the society does not have aversion to inequality, i.e., $\varepsilon=0$.

6 CONCLUSION

In this paper we reviewed several approaches to assessing the disparities in HD and to quantifying inequality in the distribution of HD achievements. We treated HD as a multidimensional concept, but focused on three key dimensions – income, health and education. We recognize that the complexity of the concept of HD inequality goes beyond inequality in income or wealth. The data needed for assessing inequalities in the distribution of HD achievements, or opportunities for HD, should be at the level of the household or person and should be rich in content beyond the core HD dimensions. Surveys should include data on the marginalized and deprived because they represent the lower end of distributions, and are often missed by official statistics because they are non-respondents, or hard to reach subpopulations. For modeling purposes, as well as for independent validation of consistency of primary sources, data collected by alternative sources are also needed. Sometimes a rush for analysis of a hot, policy relevant, or politically sensitive topic, without a clear conceptual or measurement framework, may prompt the use of data sets and modeling techniques that are not appropriate, leading to biased analysis and results. Studies of inequality in the distribution of HD achievements are generally not of that kind, although they require high quality data sets, imputation and prediction of certain variables (e.g., household life expectancy, household income adjusted for public spending, etc.) that are not observed at the level of the household. A special effort is needed to interpret the estimated inequalities and their causes.

We recognize the complexity of integrating inequality into the HDI. High inequality in any dimension should lower the index value for that dimension, and hence its contribution to the HDI. Although the idea of the Gini-corrected HDI is rather intuitive it has not been widely used due to the difficulties of calculating Gini indices from the grouped data. Foster et al. (2005) chose an axiomatic approach to derive a distribution-sensitive HDI. They suggested a two-step procedure assuming that data are available at the household level. Later, Seth (2009) proposed a generalization which is sensitive to associations between the dimensions.

Table 7.1, below, summarizes the methods for measuring inequality that were reviewed in this paper along four criteria: the lowest disaggregation required, coverage (approximate number of countries for which inequality can currently be measured), intuitive appeal and ease of interpretation.

Table 7.1. Some properties of the reviewed inequality measures

	The lowest disaggregation required	**Arbitrary parameters**	**Approximate number of countries (based on 2009** HDR**)**	**Intuitive appeal and ease of interpretation**
IDAHDI	Only Gini for income component	No	140	Only partial adjustment, easy to interpret
IAHDI (Hicks)	Gini for each component; an approximation can be obtained from distributions of grouped data; grouping can be different for each dimension	No	90 Distribution of education across school levels attained, distribution of health from mortality tables	Intuitively appealing, easy to interpret, lacks the sub-group consistency, and is insensitive to degree of association between dimensions
FLS index	Ideally household/ individual data. An approximation can be based on distributions of grouped data for each dimension; grouping can be different for each dimension	One aversion to inequality parameter	90 Distribution of income across quintile groups, distribution of education across the school levels attained, distribution of health from mortality tables 50 Alternatively using the household/ individual data from DHS (Klasen group's project)	Intuitively appealing, relatively difficult to interpret the notion of aversion parameter, insensitive to association between dimensions
Seth	Household/ individual data	Two aversion parameters	50 Using the household/ individual data from DHS (Klasen group's project)	Difficult to interpret (two aversion parameters), not intuitively appealing, good statistical properties, association-sensitive

Generally, the problem with the FLS-based approach is, however, that the generalized mean may not seem very intuitive for many users of the HDI. It also raises the question of how to determine inequality aversion. The attraction of the FLS index is that it allows comparison with the unadjusted HDI, and computation of the loss due to inequality across populations and across dimensions. This loss function can be used as a measure of inequality in the distribution of HDI. The approach resulting in the Inequality-adjusted HDI proposed by Alkire and Foster (2010) is essentially a modification of the FLS method implemented with a combination of household survey data and the life tables, as reviewed in section 6.

Studies of disparities, unequal progress, and unequal access and deprivation, should always be done by disaggregation over relevant groups. The visualization of such information seems to be equally important as its quantification. Many of the disparities revealed by the disaggregated analysis might already be known, but the HDI can reveal them even more starkly. "Disaggregation by social group or region can also enable local community groups to press for more resources, making the HDI a tool for participatory development." (2004 HDR)

Although the concept of equality of opportunity is very relevant to the ideas of HD, measuring it with an index seems to be very difficult due to difficulties in identification and disentangling the circumstances from the individual heterogeneity. However, the methodologies developed around the idea of measuring inequalities in opportunity can be readily used for quantifying between and within group inequalities.

Acknowledgment

This paper greatly benefited from inputs from Jeni Klugman, discussions with Sabina Alkire and James Foster, computation by Zachary Gidwitz, and reading and editing by Astra Bonini.

Bibliography

Ahuja, V., Filmer, D. (1995). Educational Attainment in Developing Countries New Estimates and Projections Disaggregated by Gender. Background paper for *World Development Report 1995.*

Alkire, S. and Foster, J. (2010). Designing the Inequality-Adjusted Human Development Index (IHDI). *2010 HDR Background Paper No. 10.*

Amiel, Y., Y.J. Creedy and S. Hurn, (1999), Measuring attitudes towards inequality. *Scandinavian Journal of Economics*, 101(1), 83-96.

Anand, S., Nanthikesan, (2001). A Compilation of Length-of-Life Distribution Measures For Complete Life Tables. *Working Papers Series*, 10(7). Harvard Center for population and development Studies, Cambridge,MA, USA

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244-263.

Atkinson, A.B. and Bourguignon, F. (1982). The comparison of multi-dimensioned distributions of economic status. *Review of Economic Studies,* 49, pp. 183–201.

Barros, R., Ferreira, F., Vega, J., & Chanduvi, S. (2009). *Measuring Inequality of Opportunities in Latin America and Caribbean.* The World Bank: Conference Edition.

Bourguignon, F., and S. R. Chakravarty. (2003). The measurement of multidimensional poverty. *Journal of Economic Inequality*, 1 (1): 25-49.

Bourguignon, F, Ferreira, F., Menéndez, M. (2007). Inequality of Opportunity in Brazil. *Review of Income Wealth,* **53** (6): 585-618.

Burchardt, T. (2006). Foundations for measuring equality: A discussion paper for the Equalities Review. *Technical Paper 111, CASE, London School of Economics* .

Corak, M. (2006). Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility. *IZA Dp N31993* .

Davidson, R. ,Duclos, J.-Y. (2000). Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. *Econometrica, 68(6)* , 1435-1464.

Decancq, K. D., Decoster, A., Schokkaert, E. (2009). The Evolution of World Inequality in Well-being. *World Development Vol. 37* , 11-25.

Ferreira. F.H.G. and Gignoux, J. (2008).*The Measurement of Inequality ofOpportunity: Theory and an application to Latin America*. Policy Research Working Paper 4659.The World Bank, Development Research Group, Poverty Team.

Filmer,D. and Scott,K. (2008). Assessing Asset Indices. *World bank Policy Research Working Paper No. 4605, World Bank.*

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

Foster, J. E. (2000). Path Independent Inequality Measures. *Journal of Economic Theory,91* , 199-222.

Foster, J.E. and Sen, A. (1997). *On Economic Inequality,* Oxford University Press, New Delhi.

Foster, J.E. and Shorrocks, A.F. (1991). Subgroup consistent poverty indices, *Econometrica,* 59, pp. 687–709.

Foster,J.E., Shneyerov,A.A. (2000). Path Independent Inequality Measures. *Journal of Economic Theory 91*, 199-222

Gakidou,E., King, G. (2002). Measuring Total Health Inequality: Adding Individual Variation to Group-Level Differences. *International Journal for Equity in Health* 2002, (Available at: http://www.equityhealthj.com/content/1/1/3

Gakidou,E.E., Murray,C.J.L, Frenk,J. (2000). Defining And Measuring Health Inequality: An Approach Based On The Distribution Of Health Expectancy. *Bulletin of the World Health Organization, 2000, 78 (1), 42-54*

Gaye, A., and Jha, S. (2009). A review of Conceptual and Measurement Innovation in National and regional HDR, 1998-2009. *HDRO Background Paper.*

Gerhard K. Heilig, G.K., Gerland,P., Li, N., Kantorova, V., Gjaltema, T., and Lattes, P. (2008). The 2008 Revision of the United Nation World Population Prospects: Challenges in Estimating and Projecting the World's Population." United Nations population Division Document presented at IUSSP Population Conference.

Grimm, M., Harttgen, K., Klasen, S., & Misselhorn, M.(2006). A Human Development Index by income groups. HDR 2006. Occasional Paper.

Grimm, M., Harttgen, K., Klasen, S., & Misselhorn, M.(2008). A Human Development Index by income groups. *World Development 36, 2527-2546.*

Grimm, M., Harttgen, K., Klasen, S., & Misselhorn, M., Munzi , T., Smeeding, T. (2009). Inequality in Human Development: An Empirical Assessment of 32 Countries. *Social Indicators Research*, Published on line.

Hanada, K. (1983). A Formula Of Gini's Concentration Ratio And Its Application To Life Tables. *Journal of Japan Statistical Society*, 13, 95-98.

Harttgen, K. and Klasen, S. (2010). A household-based Human development Index. *2010 HDR Background Paper.*

Hicks, D. (1997). The inequality adjusted Human Development Index: a constructive proposal. *World Development, 25* , 1283–1298.

Illsey, R., Le Grand, J. (1987). The Measurement Of Inequality In Health. In: *Health and Economics*. Ed. By A. Williams. London: Macmillan, 13-36.

Kolm, S.-C. (1977). Multidimensional Egalitarianism. *The Quarterly Journal of Economics*, 91 (1): 1-13.

Lerman,R.I.,Yitzhaki, S. (1984) A Note on the Calculation and Interpretation of the Gini Index. *Economics Letters, vol. 15, 363-368.*

Lefranc, A, Pistolesi. N, Trannoy, A. (2008). Inequality of Opportunities vs. Inequality of Outcomes: Are Western Societies All Alike? *Review of Income and Wealth,* 54, 513-546.

Ledermann, S. (1969). Nouvelles tables-types de mortalite. Travaux et documents, Cahier no. 53. Paris: INED and PUF.

Lugo, M.A. (2005). Comparing Multidimensional Indices of Inequality: Methods and Application. *Society for Study of Economic Inequality*. *Working Paper* 2005-14.

Marmot, M. (2007) Achieving health equity: from root causes to fair outcomes. *Lancet* 370(9593):1153-1163.

Okun, A.M. (1975). *Equality and Efficiency. The Big Tradeoff*. Brookings Institution Press

Permanyer, I. (2009). The Measurement of Multidimensional Gender Inequality: Continuing the Debate. *Social Indicators Research,* DOI 10.1007/s11205-009-9463-4

Pirttila,J. and Uusitalo, R. (2010). "A Leaky Bucket" in the real world: Estimating Inequality aversion using survey data. Economica 77(305):60-76.

Pogge, T. (2009). Developing Morally Plausible Indices of Poverty and Gender Equity: A Research Program. (Manuscript, personal communication.)

Rawls,J. (1971). *A Theory of Justice*.The Belknap Press of Harvard University Press.
Rassler, S. (2002). *Statistical Matching: a Frequentist Theory, Practical Applications and Alternative Bayesian Approaches*, Lecture Notes in Statistics, New York: Springer Verlag.
Roemer, J. (1993). A pragmatic theory of responsibility for the egalitarian planners. *Philos. Public Aff., 22,* 146-166.
Roemer, J. E. (1998). *Equality of Opportunity. Cambridge.* Cambridge, MA: Harvard University Press.
Roemer, J. E. (2006). Review Essay, 'The 2006 world development report: Equity and Development'. *Journal of Economic Inequality,4* , 233-244.
Sen, A. K. (1992). *Inequality reexamined.* Cambridge, MA: Harvard University Press.
Sen, A. K. (1973). *On Economic Inequality*, New York, Norton.
Seth, S. (2009). Inequality, Interactions, and Human Development. *Journal of Human Development and Capabilities, 10* , 375-396.
Shorrocks, A. (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica, Vol.48* , 613-625.
Shorrocks, A. (1982). Inequality Decomposition by Factor Components . *Econometrica, Vol.50* , 193-211.
Shkolnikov, V.M., Andreev, E.E., Begun, A.Z. (2003). Gini Coefficient As A Life Table Function: Computation From Discrete Data, Decomposition Of Differences And Empirical Examples. *Demographic Research, Volume 8, Article 11 (*available at www.demographic-research.org)
Smits, J., Monden, C. (2009) Length of Life Inequality Around the Globe. *Social Science and Medicine*, 68, 1114-1123
Stanton, E. (2007). Accounting for Inequality: A proposed Revision of the Human Development Index. Working Paper Series 119. PERI. University of Massachusetts, Amherst, MA, USA
Stewart, F. (2002). Horizontal Inequality: a Neglected Dimension of Development. *Annual Development Lecture.* Helsinki: WIDER
Tilly, C. (1988). *Durable Inequality*. Berkeley: University of California Press.
Tsui, K.-Y. (1995). Multidimensional generalisations of the relative and absolute inequality indices:The Atkinson-Kolm-Sen approach. *Journal of Economic Theory,67* , 251-265.
UNESCO: *Global Education Digest 2009.* UNESCO.
UNDESA (2009). *World Population Prospects: The 2008 Revision*. New York
UNDP (1990-2009): Human Development Report, New York and Oxford: United Nations Development Programme.
UNDP (2008). El Salvador's National HDR
World Bank (2009): *World Development Indicators.* Washington, DC: World Bank.
World Bank (2006): *World Development Report 2006: Equity and Development*. Washington, DC: The World Bank and Oxford University Press.

Appendix 1. List of surveys and other data sources for assessment of inequality by country

Country Name	**Survey Year**	**Survey Name**			**Income (y), consumption (c), imputed income (y*)**
		Income or consumption	**Education**		
Afghanistan	..	..	National Risk and Vulnerability Assessment (NRVA), 2003	a	..
Albania	2002	Living Standard Measurement Survey (LSMS)		a	c
Angola	1999	Inquerito aos Agregados Familiares Sobre Despesas e Receitas		a	c
Argentina	2006	Encuesta Permanente de Hogares (EPH)		a	y
Armenia	2003	Integrated Survey of Living Standards		a	c
Australia	2003	SIHC		b	y
Austria	2007	EU/SILC		c	y
Azerbaijan	2002	Azerbaijan Survey of Living Conditions		a	c
Bahamas	2001	Bahamas Survey of Living Conditions		a	c
Bangladesh	2005	Household Income Expenditure Survey (HIES)		a	c
Barbados	1996	Continuous Labour Force Sample Survey (CLFSS)		a	c
Belarus	2002	Household Income and Expenditure Survey		a	c
Belgium	2007	EU/SILC		c	y
Belize	1995	Survey of Living Conditions		a	c
Benin	2003	Questionnaire des Indicateurs de Base du Bienêtre (QUIBB)		a	c
Bhutan	2003	Living Standards Survey		a	c
Bolivia	2002	Encuesta Continua de Hogares - Condicionas de Vida (ECH)		a	y
Bosnia and Herzegovina	2001	Living Standards Measurements Survey (LSMS)		a	c
Brazil	2005	Pesquisa Nacional per Amostra de Domicilios (PNAD)		a	y
Bulgaria	2001	Integrated Household Survey		a	c

Country Name	Survey Year	Survey Name			Income (y), consumption (c), imputed income (y*)
		Income or consumption	Education		
Burkina Faso	2003	Questionnaire des Indicateurs de Base du Bienêtre (QUIBB)		a	c
Burundi	1998	Etude Nationale Sur les Conditions de Vie des Populations		a	c
Cambodia	2004	Household Socio-Economic Survey		a	c
Cameroon	2007	Enquête Camerounaise Auprès de Ménages (ECAM)		a	c
Canada	2004	SLID		b	y
Central African Republic	2003	Enquete Sur Les Conditions De Vie Des Menages		a	c
Chad	2002	Enquete sur la Consommation et le Secteur Informel au Tchad (ECOSIT)		a	c
Chile	2006	Encuesta de Caracterización Socio-económica Nacional (CASEN)		a	y
China	2004	WIDER	WHS, 2003	f	..
Colombia	2004	ECH		a	y
Comoros	1996	DHS	Enquête Intégrale Auprès des Ménages (EIM), 1996	d	y*
Congo	2005	Enquete Congolaise Aupres Des Menages Pur L'evaluation De La Pauvrete		a	c
Congo (Democratic Republic of the)	2005	Enquete Nationales Aupres Des Menages		a	c
Costa Rica	2006	Encuesta de Hogares de Propósitos Múltiples		a	y
Côte d'Ivoire	2002	Enquête Niveau de Vie des Ménages		a	c
Croatia	2004	Labour Force Survey		a	c
Cyprus	2007	EU/SILC		c	y
Czech Republic	2007	EU/SILC		c	y
Denmark	2007	EU/SILC		c	y
Djibouti	2002	Echantillon Maitre et Enquête Préliminaire sur la Pauvreté		a	c
Dominican Republic	1997	Encuesta Nacional de Fuerza de Trabajo		a	y

Country Name	Survey Year	Survey Name			Income (y), consumption (c), imputed income (y*)
		Income or consumption	Education		
Ecuador	1995	Encuesta Periódica de Empleo y Desempleo (EPED)		a	y
Egypt	2005	Labor Market Survey		a	c
El Salvador	2005	Encuesta de Hogares de Propósitos Múltiples (EHPM)		a	y
Equatorial Guinea	..	..	MICS, 2000	e	..
Estonia	2007	EU/SILC		c	y
Ethiopia	2000	Welfare Monitoring Survey		a	c
Fiji	..	..	Population Census, 2006	a	..
Finland	2007	EU/SILC		a	y
France	2007	EU/SILC		a	y
Gabon	2005	Enquete Gabonaise Pur L'evaluation Et Le Suivi De La Pauvrete		a	c
Gambia	1998	Household Poverty Survey		a	c
Georgia	2002	Monitoring of Households		a	y*
Germany	2007	EU/SILC		c	y
Ghana	2005	Ghana Living Standards Surveys Round Four (GLSS4)		a	c
Greece	2007	EU/SILC		c	y
Guatemala	2006	NSLC		b	y
Guinea	1994	Enquête Intégrale Sur les Conditions des Vie de Ménages		a	c
Guinea-Bissau	2006	Multiple Indicator Cluster Survey (MICS)		a	y*
Guyana	1992	Multiple Indicator Cluster Survey (MICS)		a	y
Haiti	2005	DHS		a	y*
Honduras	2003	Encuesta de Permanente de Hogares de Propósitos Múltiples (EPHPM)		a	y
Hungary	2007	EU/SILC		c	y
Iceland	2007	EU/SILC		c	y
India	2004	Socio-economic Survey		a	c
Indonesia	2002	Survei Sosial Ekonomi Nasional		a	c
Iraq	..	..	Household Socio Economic Survey		..

Country Name	Survey Year	Survey Name			Income (y), consumption (c), imputed income (y*)
		Income or consumption	Education		
			(IHSES), 2006		
Ireland	2007	EU/SILC		c	y
Israel	2005	HES		b	y
Italy	2007	EU/SILC		c	y
Jamaica	2002	Jamaica Survey of Living Conditions and Labor Force Survey		a	c
Jordan	2002	Household Income Expenditure Survey		a	c
Kazakhstan	2006	MICS		e	y*
Kenya	2005	Welfare Monitoring Survey III		a	c
Korea (Republic of)	2006	HIES/FHIE		b	y
Kyrgyzstan	2002	Living Standards Measurement Survey		a	c
Lao People's Democratic Republic	2002	LAO Expenditure and Consumption Survey (LECS)		a	c
Latvia	2007	EU/SILC		c	y
Lesotho	2004	DHS		d	y*
Liberia	2007	Core Welfare Indicators		a	c
Lithuania	2007	EU/SILC		c	y
Luxembourg	2007	EU/SILC		c	y
Madagascar	2004	Enquête Aupres des Ménages		a	c
Malawi	2005	Integrated Household Survey (IHS)		a	c
Malaysia	1998	Household Expenditure Survey		a	y
Maldives	2004	Vulnerability and Poverty Survey		a	c
Mali	1994	Enquête Légère Intégrée Auprès des Ménages		a	c
Marshall Islands	1999	Census of Population and Housing		a	y
Mauritania	2000	Enquête Permanente sur les Conditions de Vie des Ménages (EPCV)		a	c
Mexico	2006	Encuesta Nacional de Ingreso-Gasto de los Hogares (ENIGH)		a	y
Micronesia (Federated States of)	2000	Census of Population and Housing		a	y

Country Name	Survey Year	Survey Name			Income (y), consumption (c), imputed income (y*)
		Income or consumption	Education		
Moldova (Republic of)	2005	Household Budget Survey		a	c
Mongolia	2002	Living Standards Measurement Survey (LSMS)		a	c
Montenegro	2005	MICS		e	y*
Morocco	1998	Enquête Nationale sur les Niveaux de Vie des Ménages		a	c
Mozambique	1996	Inquerito Nacional aos Agregados Familiares Sobre as Condicoes de Vida		a	c
Namibia	2007	DHS		d	y*
Nepal	2003	Living Standars Survey II		a	c
Netherlands	2007	EU SILC		c	y
Nicaragua	2001	Encuesta Nacional de Hogares sobre Medición de Niveles de Vida (EMNV)		a	c
Niger	2006	DHS		d	y*
Nigeria	2003	Living Standards Surveys		a	c
Norway	2007	EU/SILC		c	y
Pakistan	2001	Pakistan Integrated Household Survey (PIHS)		a	c
Palau	2000	Census of Republic of Palau		a	y
Panama	1995	Encuesta de Hogares (EH)		a	y
Paraguay	2006	Encuesta Permanente de Hogares (EPH)		a	y
Philippines	2002	Annual Poverty Indicators Survey		a	c
Poland	2007	EU/SILC		c	y
Portugal	2007	EU/SILC		c	y
Romania	2002	Family Budget Survey		a	c
Russian Federation	2003	Survey of Household Welfare and Participation in Social Programs		a	c
Rwanda	2005	Enquête Intégrale Sur les Conditions des Vie de Ménages		a	c
Sao Tome and Principe	..		Enquête Sur les Conditions des Vie de Ménages, 2000	a	..
Senegal	2001	Deuxieme Enquête Senegalese Aupres de Menages		a	c
Serbia	2005	MICS		e	y*
Sierra Leone	2003	Integrated Household Survey		a	c

Country Name	Survey Year	Survey Name			Income (y), consumption (c), imputed income (y*)
		Income or consumption	Education		
Slovakia	2007	EU/SILC		c	y
Slovenia	2007	EU/SILC		c	y
South Africa	1998	DHS	General Household Survey, 2005	d	y*
Spain	2007	EU/SILC		c	y
Sri Lanka	2002	Household Income and Expenditure Survey		a	c
Suriname	2001	Expenditure Household Survey (EHS)		a	c
Swaziland	2000	Household Income and Expenditure Survey		a	c
Sweden	2007	EU/SILC		c	y
Switzerland	2004	ERS		b	y
Syrian Arab Republic	2004	Household Income And Expenditure Survey		a	c
Tajikistan	2003	Living Standard Survey		a	c
Tanzania (United Republic of)	2001	Household Budget Survey		a	c
Thailand	2006	Socio-economic Survey		a	c
The former Yugoslav Republic of Macedonia	2005	MICS		e	y*
Timor-Leste	2001	Household Survey 2001		a	c
Togo	2006	Indicateurs De Base Du Bien Etre		a	c
Tonga	..	..	Population Census, 2006	a	..
Trinidad and Tobago	2006	MICS		e	y*
Tunisia	WIDER2000	Enquête National Sur la Population et l'Emploi		a	..
Turkey	2003	DHS		d	y*
Turkmenistan	1998	Living Standard Measurement Survey		a	c
Uganda	2002	Socio-economic Survey		a	c
Ukraine	2005	Household Living Conditions Survey		a	y
United Kingdom	2007	EU/SILC		c	y
United States	2004	CPS-LIS		b	y
Uruguay	2006	ECH LIS		b	y
Uzbekistan	2002	DHS		d	y*

Country Name	Survey Year	Survey Name			Income (y), consumption (c), imputed income (y*)
		Income or consumption	Education		
Venezuela (Bolivarian Republic of)	1995	Encuesta de Hogares por Muestreo (EHM)		a	y
Viet Nam	2002	Household Living Standard Survey		a	c
Yemen	2005	Household Budget Survey		a	c
Zambia	2003	Living Conditions Monitoring Surve		a	c
Zimbabwe	2005	DHS		d	y*

Note:
a-Survey is found in the World Bank's International Income Distribution Database
b-Survey is stored in the Luxembourg Income Study Database
c-Survey is part of the EU Survey of Income and Living Conditions
d-DHS database is used. Income is imputed using the assets index method (Harttgen and Klasen, 2010).
e-UNICEF's Multiple Indicators Cluster Survey used to estimate the inequality in MYS.
f-China and Tunisia's inequalities are estimated from the distribution percentiles reported in WIDER data base.

Appendix 2. Estimated measures of inequality in income distribution for four different estimation strategies.

(Means and medians are given in local currency units)

Country	Year	Strategy	P90/P10	P90/P50	P10/P50	P75/P25	GE(0)	GE(1)	GE(2)	Gini	A(0.5)	A(1)	A(2)	Mean	Median
Australia	2003	#1	4.472	2.313	0.517	2.316	0.210	0.202	0.271	0.341	0.096	0.190	0.540	19588	15535
Australia	2003	#2	4.425	2.308	0.522	2.31	0.194	0.197	0.267	0.338	0.092	0.176	0.343	19687	15600
Australia	2003	#3	4.426	2.293	0.518	2.3	0.198	0.180	0.208	0.330	0.089	0.179	0.533	19183	15479
Australia	2003	#4	4.371	2.286	0.523	2.294	0.181	0.176	0.205	0.327	0.085	0.166	0.332	19281	15535
Brazil	2006	#1	11.748	3.387	0.288	3.404	0.522	0.563	1.253	0.533	0.236	0.407	0.655	6798	4162
Brazil	2006	#2	11.642	3.372	0.29	3.375	0.511	0.559	1.244	0.531	0.234	0.400	0.626	6834	4200
Brazil	2006	#3	11.375	3.303	0.29	3.362	0.484	0.494	0.837	0.514	0.216	0.384	0.638	6453	4133
Brazil	2006	#4	11.285	3.283	0.291	3.339	0.473	0.489	0.830	0.511	0.214	0.377	0.608	6487	4166
Canada	2004	#1	4.85	2.104	0.434	2.216	0.218	0.212	0.302	0.343	0.100	0.196	0.481	22432	18685
Canada	2004	#2	4.741	2.101	0.443	2.204	0.203	0.207	0.298	0.340	0.096	0.183	0.354	22544	18725
Canada	2004	#3	4.784	2.079	0.435	2.2	0.203	0.186	0.224	0.330	0.091	0.184	0.470	21903	18625
Canada	2004	#4	4.685	2.075	0.443	2.19	0.188	0.181	0.221	0.327	0.088	0.171	0.341	22014	18685
Switzerland	2004	#1	4.067	1.956	0.481	2.263	0.175	0.174	0.230	0.316	0.082	0.160	0.419	38128	32922
Switzerland	2004	#2	4.029	1.95	0.484	2.257	0.166	0.170	0.227	0.313	0.080	0.153	0.286	38309	33041
Switzerland	2004	#3	4.037	1.95	0.483	2.248	0.163	0.154	0.177	0.305	0.076	0.150	0.410	37400	32718
Switzerland	2004	#4	3.998	1.938	0.485	2.23	0.154	0.151	0.174	0.302	0.073	0.143	0.275	37578	32922
Colombia	2004	#1	15.022	4.044	0.269	3.832	0.633	0.688	1.731	0.580	0.280	0.469	0.717	2756567	1423500
Colombia	2004	#2	14.481	4.045	0.279	3.755	0.620	0.682	1.719	0.577	0.277	0.462	0.694	2772928	1432000
Colombia	2004	#3	14.446	3.922	0.272	3.796	0.585	0.597	1.114	0.558	0.256	0.443	0.700	2588938	1405500
Colombia	2004	#4	13.802	3.878	0.281	3.721	0.572	0.592	1.105	0.556	0.253	0.436	0.676	2604351	1423500
Guatemala	2006	#1	16.45	3.898	0.237	4.256	0.618	0.642	1.545	0.567	0.268	0.461	0.732	11185	6187
Guatemala	2006	#2	16.102	3.873	0.241	4.204	0.606	0.637	1.534	0.565	0.266	0.454	0.694	11244	6236
Guatemala	2006	#3	16.151	3.84	0.238	4.213	0.566	0.541	0.902	0.542	0.241	0.432	0.715	10458	6150
Guatemala	2006	#4	15.762	3.813	0.242	4.158	0.554	0.536	0.895	0.540	0.239	0.425	0.674	10513	6200
Israel	2005	#1	7.317	2.543	0.348	2.788	0.306	0.304	0.492	0.410	0.139	0.263	0.603	43743	33417
Israel	2005	#2	7.099	2.535	0.357	2.769	0.293	0.300	0.488	0.407	0.137	0.254	0.449	43954	33561
Israel	2005	#3	7.155	2.498	0.349	2.777	0.284	0.261	0.319	0.395	0.126	0.247	0.591	42354	33220
Israel	2005	#4	6.966	2.49	0.357	2.757	0.271	0.257	0.316	0.392	0.124	0.237	0.433	42559	33422

Country	Year	Strategy	P90/P10	P90/P50	P10/P50	P75/P25	GE(0)	GE(1)	GE(2)	Gini	A(0.5)	A(1)	A(2)	Mean	Median
Korea	2006	#1	5.107	2.177	0.426	2.298	0.227	0.217	0.289	0.352	0.104	0.203	0.497	11300000	9392600
Korea	2006	#2	5.022	2.176	0.433	2.286	0.213	0.213	0.286	0.349	0.100	0.192	0.365	11400000	9416052
Korea	2006	#3	5.031	2.151	0.427	2.286	0.218	0.203	0.247	0.344	0.098	0.196	0.490	11100000	9358800
Korea	2006	#4	4.946	2.145	0.434	2.274	0.204	0.198	0.243	0.341	0.095	0.184	0.357	11200000	9392600
Mexico	2004	#1	10.423	3.39	0.325	3.216	0.512	0.610	2.707	0.529	0.238	0.401	0.652	27316	15990
Mexico	2004	#2	10.174	3.412	0.335	3.194	0.502	0.605	2.691	0.527	0.236	0.394	0.616	27459	16000
Mexico	2004	#3	10.185	3.324	0.326	3.187	0.452	0.464	0.793	0.498	0.204	0.364	0.627	25364	15888
Mexico	2004	#4	9.975	3.325	0.333	3.156	0.442	0.460	0.787	0.496	0.201	0.357	0.589	25497	16000
Peru	2004	#1	17.218	3.661	0.213	4.371	0.612	0.618	1.384	0.559	0.262	0.458	0.729	3752	2175
Peru	2004	#2	16.697	3.653	0.219	4.326	0.601	0.614	1.375	0.557	0.260	0.451	0.707	3770	2188
Peru	2004	#3	16.769	3.571	0.213	4.36	0.569	0.540	0.953	0.538	0.241	0.434	0.715	3544	2165
Peru	2004	#4	16.234	3.561	0.219	4.305	0.558	0.536	0.946	0.536	0.238	0.427	0.692	3561	2178
United States	2004	#1	6.457	2.401	0.372	2.602	0.311	0.301	0.479	0.403	0.138	0.268	0.963	22181	17213
United States	2004	#2	6.325	2.394	0.379	2.588	0.289	0.296	0.474	0.400	0.134	0.251	0.472	22291	17296
United States	2004	#3	6.341	2.365	0.373	2.589	0.290	0.264	0.365	0.388	0.134	0.251	0.472	21485	17127
United States	2004	#4	6.217	2.358	0.379	2.575	0.268	0.260	0.360	0.385	0.122	0.235	0.457	21592	17213
Uruguay	2004	#1	8.765	2.898	0.331	3.084	0.379	0.396	0.694	0.462	0.175	0.315	0.767	61895	42743
Uruguay	2004	#2	8.565	2.884	0.337	3.066	0.371	0.393	0.689	0.460	0.173	0.310	0.510	62173	42963
Uruguay	2004	#3	8.598	2.857	0.332	3.077	0.354	0.352	0.502	0.447	0.162	0.298	0.760	59725	42503
Uruguay	2004	#4	8.42	2.844	0.338	3.05	0.347	0.349	0.498	0.445	0.160	0.293	0.495	59994	42768