Courant Research Centre
‘Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis’

Georg-August-Universität Göttingen
(founded in 1737)

Discussion Papers

No. 75

A Human Development Index at the Household Level

Kenneth Harttgen, Stephan Klasen

March 2011

Wilhelm-Weber-Str. 2 · 37073 Goettingen · Germany
Phone: +49-(0)551-3914066 · Fax: +49-(0)551-3914059

Email: crc-peg@uni-goettingen.de Web: http://www.uni-goettingen.de/crc-peg

A Human Development Index at the Household Level[1]

February, 2011

Kenneth Harttgen, University of Göttingen.
Address: University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany (Phone: +49-551-398170; Fax: +49-551-397302; Email: k.harttgen@wiwi.uni-goettingen.de)

Stephan Klasen (corresponding author), University of Göttingen.
Address: University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany (Phone: +49-551-397303; Fax: +49-551-397302; Email: sklasen@uni-goettingen.de)

[1]We would like to thank Sabina Alkire, Joachim von Braun, Francois Bourguignon, James Foster, Jeni Klugman, Milorad Kovacevic and Francisco Rodriguez as well as participants at workshops and conferences in Oxford, Hannover, New York, and Gttingen for helpful comments on earlier versions of the paper. Funding from UNDP in support of this work is gratefully acknowledged.

A Household-Level Human Development Index

Abstract

One of the most serious weaknesses of the human development index (HDI) is does not take into account the distribution of human development within a country. All previous attempts to capture inequality in the HDI have also used aggregate information and there exists no HDI at the household level. We provide a method for proxying the HDI at the household level. This allows the analysis of the HDI by any kind of population subgroups and by household socioeconomic characteristics as well as to apply any kind of inequality measure across population subgroups and over time. We illustrate our approach for 15 developing countries. Inequality in the HDI is stunningly large for some countries, driven mostly by very high inequality in the education and income components of the HDI. The inequality in human development is larger than previously reported which is largely due to the new procedures for calculating the HDI used in the 2010 Human Development Report. Inequality in the HDI is largest in poorer countries, particularly in Sub-Saharan Africa.

Key words: Human Development Index, Income Inequality, Differential Mortality, Inequality in Education.

A Human Development Index at the Household Level

1 Introduction

The HDI is a composite index that measures the average achievement in a country in three basic dimensions of human development: health, education, and standard of living. Since last year's 20th anniversary Human Development Report, the indicators used to calculate the index are the following: life expectancy at birth for the health component, mean years of schooling for adults and expected years of mean schooling for children for the education component, and the log of the PPP-adjusted Gross National Income per capita for the standard of living component (UNDP 2010).

In its 20 year history, it has become one of the most widely used indicators for comparisons of welfare, which is linked to its grounding in multidimensional well-being measurement as in its transparency and simplicity. At the same time, there have been many criticisms, ranging from the neglect of some relevant dimensions (see e.g. Anand and Sen (1992), Ranis, Stewart and Samman (2006)), implied substitution possibilities and trade-offs linked to allegedly arbitrary weighting schemes (see e.g. Kelley (1991), Srinivasan (1994) and Ravallion (1997))[1] as well as the neglect of distribution of achievements within a country (see e.g., Sagar and Najam (1998), Grimm et al. (2008)). It is this last issue that we address in this paper by proposing and applying a method that allows the calculation of the HDI at the household level and thus allows a direct assessment of inequality in the HDI, and decompositions of inequality by subgroups.

There are some papers that address the inequality in the HDI. Anand and Sen (1992) and Hicks (1997) suggested to discount each dimension index by one minus the Gini coefficient for that dimension before the arithmetic mean over all three is taken. Therefore, high inequality in one dimension lowers the index value for that dimension and, hence its contribution to the HDI. Although the idea of such a discount factor is rather intuitive, the Gini-corrected HDI has not been widely used, largely due to data constraints.

Related to this approach and based on further work by Alkire and Foster (2010), the 2010 Report UNDP did provide, for the first time, an inequality-adjusted HDI for a large number of countries. The IHDI 'discounts' the average achievements in a dimension by inequality in the dimension. If, for example, income is distributed highly unequally in a country, the inequality-adjusted achievement will be lower than in another country with the same average income but lower inequality.[2] What this IHDI tries to measure, therefore, is the extent to which inequality reduces well-being in a dimension and thus human development overall, which is an interesting and important aim. But it is important to point that it remains an aggregate index that simply adjusts human development by inequality in its dimensions. It does not, therefore, allow an assessment of human development at the household level, the aim of this paper; nor does it allow an assessment of inequality levels in the HDI and its components, or decompositions by subgroups. Similarly, it cannot consider the joint distribution of different human development dimensions. For example, as long as inequality in education and incomes are the same in two countries it does not matter whether all education poor have high incomes in one country, and in the other one the education poor are also income poor.[3]

There is a second difference of the IHDI to what we propose here. The IHDI considers the inequality in actual life lengths associated with a particular life expectancy as the relevant inequality in health. For example, if average life expectancy is 60, there will be some who die at age one, some at age five, some at age 25, and many at ages 60 and above. The inequality in achieved life lengths (only some of which is related to socioeconomic inequalities) is considered the relevant inequality here.[4] In contrast, we will try to measure socioeconomic differentials in life prospects and use them as the relevant indicator in our measure.

There are more closely related precursors to our work. Grimm et al. (2008, 2009) produce an HDI for different income quintiles. Based on a method and computations described in detail in Grimm et al. (2006), the HDR 2006 presented a HDI for all five income quintiles for a sample of 11 OECD countries and 21 developing countries. Harttgen and Klasen (2011) calculate the HDI separately for internal migrants versus non-migrants.[5] Both of these papers calculate the HDI separately for the specified groups. One cannot use these results to calculate unconditional inequality measures such as the Gini coefficient for the HDI or its components, or calculate it for other groups.

In short, all previous attempts to capture inequality in the HDI either use aggregate inequality information or use aggregate group information. To date, there exists no HDI at the household level based only on information coming from the household level. This paper provides a method and an illustration for calculating the HDI at the household level. This will allow a large range of previously unavailable analysis to yield new insights with respect to levels and changes of human development. It immediately allows comparisons across population subgroups (e.g. urban, rural), by income and other population groups like the mentioned papers. Furthermore, it provides a completely new opportunity to analyze differences in the HDI between household specific characteristics.[6] In addition, having calculated an HDI at the household level, one could calculate any kind of inequality measure of the HDI, compare it across space and time, and decompose it within and between groups. Also, one could apply the methods used by the IHDI to generate components indices and an HDI that penalizes for inequality between people within each dimension, and between dimensions.

When constructing such a household-level HDI, data availability is the key constraint one needs to address. Today household income surveys are widely undertaken and provide data on income distribution. However, it is much more difficult to get data on inequality in life expectancy and educational achievements and hardly any household surveys exist that contain high-quality data on health, education, and incomes in one survey.

In addition, there are some conceptual issues to face when moving calculating an HDI at the household level. First, life expectancy is an aggregate indicator summarizing current mortality conditions that cannot be estimated directly at the household level. At the same time, mortality information at the household level can be used in an imputation or simulation techniques to generate life expectancies at the household level. Second, no information on educational enrolment exists for households without children.

These data and conceptional constraints allows us only to calculate a proxy HDI at the household level. But such an HDI proxy at the household level is important because it allows to take explicitly study the inequality in human development between population subgroups and by socioeconomic status. It also immediately considers the joint distribution of disadvantages across dimensions at the household level, which cannot be assessed using the IHDI or previous approaches to study distribution-sensitive human development achievements.

The objective of this paper is to illustratively demonstrate the feasibility of such an approach. We illustrate our approach for 15 developing countries using DHS survey data. Because of the comprehensive availability of the DHS data, the household-based HDI can be calculated for than 50 countries for several time periods. The reminder of this paper is organized as follows. Section two presents our methodology. Section three presents the sample of countries and presents the results. Section four concludes.

2 Methodology

2.1 Calculating the GNI index

For our analysis we rely on DHS data where information on education and mortality is available. We start with the calculation of the GNI component of the HDI. Since we do not have information on income or expenditure in the DHS data sets that can be used for our analysis, we consider an alternative approach to determine the socioeconomic status of a household, which we use as a proxy for the income component of the HDI. In particular, we combine an asset index approach in defining well-being proposed by Filmer and Pritchett (2001) and Sahn and Stifel (2000) with an income simulation approach. We thereby simulate income levels for each household in the DHS data sets to overcome the problem that the DHS do not contain information on income or expenditure.

We proceed in five steps. In a first step, we calculate an asset index fou households included in the DHS. The main idea of this approach is to construct an aggregated unidimensional index over the range of different dichotomous variables of household assets capturing housing durables and information on the housing quality that indicate the material status (welfare) of the household.[7] For the estimation of the weights and for the aggregation of the index, we use a principal component analysis proposed by Filmer and Pritchett (2001), relying on the first principal component of our asset index.[8] In particular, as components for the asset index we include dichotomous variables whether the following assets in a household exist or not: radio, TV, refrigerator, bike, motorized transport, capturing household durables and type of floor material, type of wall material, type of toilet, and type drinking water capturing the housing quality and we calculate the asset indices separately for each country and period.[9]

The use of the asset index as a proxy for material welfare (i.e. income or expenditures) requires some further discussion. A large body of literature exists using an asset index to explain inequalities in educational outcomes (e.g. Ainsworth and Filmer 2006; Bicego et al 2003), health outcomes (e.g. Bollen et al. 2002; Schellenberg et al 2003), child malnutrition (e.g. Sahn and Stifel 2003; Tarozzi and Mahajan 2005), child mortality (e.g. Sastry 2004) when data on income or expenditure is missing. In addition, asset indices are used to analyze changes and determinants of poverty (Harttgen and Misselhorn 2006; Sahn and Stifel 2000; Stifel and Christiaensen 2007; World Bank 2006).

The literature has also examined whether it can serve as an appropriate proxy for income or expenditure. For example, Stewart and Simelane (2005) validate the use of the asset index as a proxy for income to predict child mortality in South Africa. They find a very close relationship between income and the asset index. The recent paper by Filmer and Scott (2008) provides an excellent validation of the use of various asset index methods by comparing how asset index outcomes match to results using per capita expenditures with respect to the ranking of households and with respect to inequality analysis outcomes in education, health care use, fertility, and child mortality. They show that inferences about inequalities in education and health are robust to the use of the asset index. The gradient of the outcomes of the asset index closely follows the outcome using per capita expenditures.[10] Using Mexico's national income and expenditure survey, McKenzie (2005) also shows that assets provide a reliable proxy for inequality in education, after controlling for household income and demographic characteristics. However, Howe et al. (2009) provided a systematic review of studies that used the asset index and found that the asset index is not performing well in proxying consumption expenditures. Filmer and Scott (2008) argue that targeting of social program to the lowest population quintile on the basis of the asset index would therefore only partly reach the same households. They found that the assets index identifies especially the more rural and smaller households as deprived, compared to per capita expenditure. However, they conclude that because the gradient of the economic status is similar to that of per capita expenditure using the asset index would not lead to a misleading targeting and that using the asset index even allows to identify the most deprived households in terms education, health, and labor force participation.

Hence, the asset index should not be seen as a perfect proxy for income or expenditure. But, in the absence of information on household income or expenditure, the ownership of assets can be used to measure inequalities in living standards based on an asset index. Thus, the welfare ranking based on an asset index do not lead to the exact welfare ranking based on per capita expenditure but the gradient of both measures is similar.[11]

In fact, one may also argue that the asset index even allows to identify the most deprived households better than incomes since assets may be a better proxy for long-term income than annual income as the asset index is less vulnerable to fluctuations over time than income or expenditure.[12] Therefore, using the asset index provides a good indicator of longer-term living standards, which is in line with the basic idea of the HDI.

While this literature has demonstrated the use of the asset index to measure poverty or study the determinants of health or education, the use of the asset index to derive a welfare distribution has to address some additional challenges. First, the asset index might not correctly reveal standard of living differences between urban and rural areas. The asset index can bias estimates of living standards upwards in urban areas as many assets are cheaper there and preferences for such assets are higher. As a result, urban households typically own more (and other) assets than rural households.

Second, by definition, the asset index is a discrete function. This discreteness causes difficulties estimating a continuous distribution. The ability to use the asset index as a proxy for a continues variable such as income or expenditure depends, therefore, on the number of assets incorporated in the asset index (Howe et al 2009; McKenzie 2005). This problem can be reduced by increasing the numbers of indicators incorporated in the index. In our case, similar to McKenzie (2005), we incorporate least 25 assets into the index.[13]

Third, the distribution of assets is invariably more equal than the distribution of income it is meant to proxy. This is an issue we will explicitly deal with below.

In a second step, we derive a log normal distribution (LN) based on the respective country specific mean income per capita and the respective Gini coefficient reported in the PovcalNet database of the World Bank (and based on a household income or expenditure survey). Formally, the log-normal distribution $LN(\mu,\sigma)$ is defined as the distribution of the random variable $Y = \exp(X)$, where $X \sim N(\mu,\sigma)$ has a normal distribution with mean μ and standard deviation σ. It can be shown that the density of $LN(\mu,\sigma)$ is

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot e^{-(log(x)-\mu_i)^2/2\sigma^2}, \qquad x > 0, \tag{1}$$

and its mean and variance are given respectively by

$$E(Y) = e^{\mu+\sigma^2/2}, \qquad Var(Y) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}.^{14} \tag{2}$$

The parameters μ and σ of $LN(\mu,\sigma)$ can be determined from the average income $E(Y)$ and the Gini coefficient G as follows.

$$\sigma = \sqrt{2}\,\Phi^{-1}\left(\frac{G+1}{2}\right), \qquad \mu = \log(E(Y)) - \sigma^2/2.$$

Hence, with the two parameters, the mean income per capita and the Gini, we are able to estimate μ and σ of the density function of the log normal distribution for each country. While likely to provide only an approximation, the assumption of the log normal distribution of the income distribution is often used in the empirical literature to estimate income distributions (see, e.g. Chotikapanich et al. 1997; Bourguignon 2003; Schultz 1998; and Milanovic 2002, 2006). Clearly, this does not imply that we accept the log-normal as the true distribution for income data, it rather means that the data at hand do not contain enough information to fit a more sophisticated model.[15]

In a fourth step, we then simulate household income per capita based on the asset index distribution.[16] In particular, we attach to each quantum of the asset index distribution the respective income value from the income distribution and derive for each asset index value the respective simulated income value. This means that, for example, percentile 1 in the asset index distribution becomes the value of income of percentile 1 in the income distribution.[17]

Then, in a fifth step, we calculate the household specific GNI component of the HDI. To eliminate differences in price levels across countries we express household income per capita y_h calculated from the HIS, in USD PPP using the conversion factors based on price data from the latest International Comparison Program surveys provided by the World Bank (2005).[18] Then, we rescale y_h^{PPP} using the ratio between the mean income per capita and the GNI per capita expressed in USD PPP (taken from the general HDI).[19]

Once these adjustments are done, it is straightforward to calculate the household specific GNI index, using the new minimum and maximum values of the HDI.[20]

2.2 Calculating the education index

In the next step, we calculate the education index of the HDI at the household level. Since the 2010 HDR the education component consists of mean years of education of adults aged 25 and older and of expected years years of education of children at each level of education of the official school age.

The expected years of education measures the 'school life expectancy' and is defined as the total number of years of schooling which a child at a certain age can expect to achieve, assuming that her/his enrollment probability does not change in the future.[21] In principle, the expected years of education a simply another expression of the former enrolment component of the HDI.[22] Hence, we need to calculate the household mean years of education of adult household members aged 25+ and the age specific enrolment rates for children at the household level to calculate the expected years of education.

The data constraint we face when calculating the education index is that enrolment information is only available for households that have school-age children. The main challenge that arises here is the question of how to compare the value for the education component of households where we just have information on the adult's years years of education with those where we have information also on enrolment. We provide two possible solutions.

First, one could simply drop the observations for which we do not have information on enrolment, i.e. the households without children in that particular age range. Simply deleting the missing values might lead to biased results if the remaining cases are not representative for the entire sample (e.g. Schaefer and Graham 2002). The second approach is to use an imputation-based approach to fill the missing values of the age specific enrolment information.[23]

There is a broad literature on the statistics of imputation (e.g., Graham and Hofer 2000; Rubin 1987; Stern and Russel 2001; Schaefer 1997; Graham et al 2003; Schaefer and Graham 2002; Allison 2007). The major shortcoming of a simple mean substitution or regression-based single imputation (besides depending on the quality of the regression) is that the variance is still underestimated and thus standard errors and significance test can still be biased (e.g. Graham et al (2003); Landerman et al. (1997)). The solution to deal with this issue is to rely on multiple imputation (Rubin 1977, 1987; Schaefer 1997). The idea is to repeat the imputation process, producing multiple 'complete' data sets. The values are drawn from the Bayesian posterior distribution of the parameters. Because of the random term, the estimates of the parameters will slightly differ and this variability can then be used to adjust the standard errors upwards (Allison 2007). These analysis results are then combined to one overall analysis resulting in the prediction of the missing values (Wayman 2003). It has been shown that multiple imputation performs favorably (see, e.g. Schaefer and Graham 2002; Schaefer 1997; Wayman 2003). We use multiple imputation to impute the missing values of enrolment rates; the regression-model used is country-specific and includes a large range of socioeconomic determinants of enrolment rates for which information is available in the survey.[24] Table A.1 provides an example of such a regression model which shows that it provides quite a good fit and reasonable parameter estimates. Please note that we do not use the asset index in the regression model as this is used already to generate the income component. This way we want to avoid in-built correlation between the components of the HDI.[25]

After imputing the age specific enrolment rate, we calculate the expected years of education of children and the mean years of education for adult household member for each household in the data set. Then, we derive the household specific education index of the HDI. Following the new approach of the 2010 HDR, we calculate the household specific mean years of education index and the expected years of education index using a minimum value of zero and the respective maximum value, which is set to the actual maximum observed value of the indicators from the Barro and Lee (2010) data set on years of education between 1890 and 2010.[26] In contrast to the former education component, the new education component is calculated as the geometric mean of the two educational subindexes.[27] We report the results for the education component using only the small sample without imputation of missing values on enrolment as well es the the results for the imputation of the enrolment.

2.3 Calculating the health index

To calculate the health index, we combine information on child mortality with model life tables and use again an imputation-based approach to calculate mortality rates at the household level. The reason for this imputation is twofold. First, we need to overcome the problem that for households without children this procedure cannot be implemented. Second, we need to obtain an estimate of child mortality that has a more continuous character, because otherwise we would have only limited variation in the data since in most household either none, one or two children died resulting in a household specific mortality rates clustered around zero (for which no life expectancy is computable), and values such as 0.25, 0.33, or 0.5. We proceed in three steps to calculate the health component.

First, we regress child mortality on a set of basic household and community socioeconomic characteristics using a using a discrete time proportional hazard model with a piece-wise constant baseline hazard function to control for censored data. Table A2 shows the regression results exemplarily for Burkina Faso. All covariates show the expected sign.[28] Then, we use the prediction of child mortality for all households (and not only on those without children).

We should emphasize that the imputation for the health component (as well as the enrolment rates for households without children) implies that we are no longer calculating a household-based HDI for the particular individual households included in the DHS. But instead we are calculating a household-based HDI for household types described by the socioeconomic characteristics in our regressions. For example, a household where no child has died but, based on the socioeconomic characteristics, faces a substantial risk of child mortality will get a health component that reflects this risk (rather than a perfect score since no child has died so far).[29] We believe that this is what is really of interest to policy-makers who want to know disparities in the HDI by household types (and groups) rather than the particular individuals included in the survey.

Please note also that the imputation approach could generate some in-built correlation between the education and health component, since the regressions are partly based on the on the same characteristics; as we show in Table A.3 below, however, the correlations are rather low and not very sensitive to the exclusion of particular covariates that appear in both regressions. We explicitly did not include the asset index into the regression to avoid an extra spurious correlation between the health component and the income component of the HDI, since the income component is calculated based on the asset index values.[30]

Second, after having estimated the household specific mortality rate, we apply the recently provided modified logit life table systems by Murray et al (2003) to estimate the household specific life expectancy at birth.[31] The advantages of the modified logit life tables by Murray et al (2003) compared, for example to Princeton Model Life Tables (Coale and Demeny 1983) or the older Ledermann model life tables (Ledermann 1969) which were used in Grimm et al (2008, 2009), are that they are very flexible and rely on more than 1800 recently available life tables; they also reflect more recent mortality experiences across the world, including the impact of the HIV/AIDS epidemic on mortality.

Third, after having estimated life expectancy for each household in the DHS data, we can then calculate the household specific health index of the HDI using again the new upper and lower limits.[32]

2.4 Calculating the household-based HDI

Once the three dimension indices are calculated, we simply calculate the household specific HDI, following the HDR 2010, by taking the geometric mean of the three dimension indices to calculate the HDI. Using the geometric mean has an in-built 'inequality aversion' across components which implies that individuals whose achievements differs greatly by components will receive a lower score compared to those with more 'balanced' achievements across components. We agree with UNDP that this implication justified the switch from the arithmetic mean used previously to the geometric mean used in the 2010 HDR. We use g to denote the geometric mean and $\mu(y)$ to denote the arithmetic mean of a given distribution y, i.e. household income per capita, and apply this definition also to the education (e) and health (h) component of the HDI. All three dimensions of the HDI can then simply be represented in a three x k matrix D, where the first row is the vector y, followed by e and h. The household based human development index H (where k refers to the number of households in the data) can then be defined as a function $F : D \to R$ from the set of D matrices to the real numbers R and formally expressed as the mean of the means:[33]

$$H(D)_{household} = g[\mu(y), g(e), \mu(h)], \tag{3}$$

which corresponds to the mean achievement in each dimension of the HDI which is than is averaged across dimensions.

To get person-based values, this value is assigned to each household member and the descriptive analysis below is based on this person-level analysis. To be sure, assigning the same HDI to all household members assumes that there is no intra-household inequality in human development which is unlikely to be the case. But with the exception of education which we could measure directly at the individual level, we have no way to study intra-household inequality in health or incomes with the data at hand so that this assumption is the only one we can make. In this sense, it is an underestimate of inequality in human development among individuals.[34]

3 Results

3.1 Overall results

In this section, we present the results of the household-based HDI for our 15 countries. Table 1 shows the mean household-based HDI and its sub-components by country.[35] HDI refers to the mean household-level HDI based on imputed enrolment and child mortality rates. The next column shows the associated ranking of countries using the HDR 2010 ranking list. HDI 'scaled' refers to the mean value of the HDI scaled to the official value of 2010; the difference to the first column shows the impact of the scaling.[36] HDI 'old' refers to the old methodology (which used adult literacy and enrolments for education, used an arithmetic mean, and used different goalposts for the component indices) to calculate the HDI using our data, and the ranking places those countries in the ranking of the 2009 HDR. HDI 'small' refers to the HDI based on the sample of households with children in school age to calculate the expected years of education.

As shown in the table and also discussed in the HDR 2010, there are some level differences in the HDI between the old and new methodology. By using the geometric mean instead of the arithmetic mean, higher values for maximum achievements and a different education component, the new methodology shows considerable lower levels of the HDI than the old methodology. This holds especially for the education component, which is already a geometric mean of the achievements in years of education and expected years of education. The other reason why the education component is considerably lower compared to the old methodology is due to the new indicators. Whereas the formerly used literacy rates, which where already quite high for many developing countries putting many close to the maximum achievable, the years of education achieved of adults and the expected years of education to be achieved based on the current enrollment status of school-aged children is considerably lower (relative to the maximum achievable value) in most countries of our sample. In many households, especially in the Sub-Saharan countries, most adults still have very low years of schooling, leading to a low education component. While the levels of the HDI change considerably, the relative rankings of our countries remain quite similar. Ethiopia and Burkina Faso are at the bottom of the list using both methods and Armenia and Peru are the best performers in our sample using both approaches.

When comparing the HDI using the imputations for enrolment with the one dropping households without children (HDI (small)), Table 1 shows that the one based on the small sample has a generally lower value. The difference is particularly large for Egypt and Zambia; dropping households without children seems to leave a biased sample with worse HDI outcomes. Besides these two countries, however, the ranking of the HDI is quite similar between the two samples. Since we are interested in analysis using the entire populations, we believe that out imputation approach is advantageous. Also, this allows us analyze differences in the HDI by population subgroups in much more detail comparing to the much smaller sample sizes when relying on households with children.[37]

We now turn to the analysis of inequality (unconditional and by groups), the main advantage of our new approach.

3.2 Inequality in the HDI and differentials by population subgroups

We first present and discuss indicators of inequality in the HDI. This we call unconditional inequality; we then present the HDI by sub-groups including incomes conditioned by different population subgroups and household characteristics.

To get a first impression of the results, we present the HDI by HDI deciles and quintiles in Table 2. The differences in the HDI by HDI deciles are staggering but differ greatly by country. In Burkina Faso, India, Ethiopia, Nicaragua and Vietnam, the ratio of top to bottom decile is close to 10 or even above. In the other countries it ranges from four to eight. These differentials are as large or larger than corresponding decile ratios in incomes in developing countries, suggesting that inequality in human development is as large as income inequality.

But this finding requires some more careful interpretation. Closer inspection reveals that it is driven to a considerably extent by very poor human development performance in the lowest decile. This, in turn, is affected heavily by the education component where in the poorest decile many households contain adults with no education (particularly in Africa) and some contain school-age children who are all out of school (in quite a few countries across all regions). The geometric mean of achievements within the education components will be zero if either sub-component is 0, and the overall HDI for the households (which is the geometric mean of all components) will then also be zero. To see the importance of this issue, the next two columns examine the bottom quintile, the top quintile and the associated ratio. In several countries, including Vietnam, Peru and Armenia, the Q5/Q1 ratio is much lower than the D10/D1 ratio suggesting that the problem of 0 education or zero enrolments is confined only to part of the bottom decile. Conversely, in a few countries the Q5/Q1 ratio is nearly as high as the D10/D1 ratio (e.g. in Burkina Faso, India, Pakistan, Nigeria, Nicaragua, Senegal and Zambia) indicating the problem of very little education is also present in the second decile (and actually far beyond in most of these countries).

This discussion suggests that the use of the geometric mean, introduced in 2010 in the calculation of the HDI, leads to much greater measured inequalities in the HDI than using the arithmetic mean within the education component and across components. How are we to interpret this? While the arithmetic mean implies complete substitutability between components (a higher education component can fully compensate for a lower health component), the geometric mean implies imperfect substitution possibilities; that seems in principle a rather reasonable assumption.[38] At the extreme, when the component indicator is zero, the geometric mean implies no substitution whatsoever. While one might find this to be an attractive feature (implying e.g. 'without income, all health and education is worth nothing', or 'without education, all health and income is worthless'), it appears exaggerated to conclude that there is zero human development if the adults in the household have zero education. One way to remedy the problem is to simply give the households a very small number (e.g. 0.001) for a component or sub-component instead of a zero which would still enable some (but highly imperfect) substitution possibilities. This we implemented when calculating inequality measures to which we now turn.

In Table 2, we also calculate the Gini coefficient for the HDI and its subcomponents. In addition, we also provide the Atkinson inequality index (Atkinson 1970) for the HDI and by components, and the Theil inequality measure, which we will later use for decompositions.

Although it is hard to interpret the absolute value of the Gini (see also below), we can compare the outcome across countries and by components. As the discussion above would suggest, the Ginis for the HDI are quite sizable and differ greatly by country. While Peru, Armenia, and Vietnam have a low HDI Gini of only around 0.2, at the other extreme, the Gini of the HDI in Senegal, Burkina Faso, and Ethiopia is close to 0.6 or even above. When examining the Gini by components, we see the overwhelming importance of educational inequality for the differential in Ginis across countries. This is for the reasons described above (even though we replace a zero with a 0.001 in each education sub-component). In contrast, the Ginis for health and income are rather low which we discuss now.

The low values of the Gini coefficient for the income component nicely illustrate how the log transformation of the GNI component reduces inequality. Table 2 provides also the Gini coefficient for the actual income (without logs). We can see that the Gini coefficients for the simulated household per capita income show the expected values in the typical range of income Ginis that correspond closely to the Ginis of the countries taken from PovcalNet.[39] The log transformation thus suggests that the inequality in HDI-relevant standard of living is rather modest and much lower than income inequality. In other words, since higher incomes have falling marginal human development returns (an implication of the log transformation), inequality in income-related human development is indeed smaller than in incomes which, in the logic of the HDI, seems to be a perfectly consistent finding.[40]

Second, we find much lower inequality in life expectancy than in income and education, which is due to the combination of an upward bound, limiting the potential inequality, as well as the more stochastic nature of mortality (compared to incomes). Other studies have found similar results.[41]

The findings for the Atkinson inequality index are very similar to the Gini. Since the Atkinson index is particularly sensitive to the bottom of the distribution, while the Gini is particularly sensitive to the middle of the distribution, some differences emerge. Comparing Armenia and Burkina Faso nicely illustrates the point. In Armenia, the Gini is rather low as the middle of the distribution does well in terms of education, health, and incomes; but due to rather poor achievements in the lowest deciles, the Atkinson index is considerably higher. The converse is the case in Burkina Faso. There the Gini is much higher as also the middle of the distribution has very poor human development achievements which is penalized particularly in the Gini, but less so in the Atkinson index. We also show results for the Theil index which we will later on decompose into within and between group components.

Table 2 also shows that higher values of inequality are generally found for those countries with low levels of human development. For example, Burkina Faso is the country with the highest Gini in the HDI (0.646) and at the same time is shows the lowest value of the HDI in our sample (see Table 1). On the other hand, Peru and Vietnam have the lowest value of the Gini coefficient for the HDI (0.216) while at the same time Armenia shows the highest value of the HDI. These discrepancies are driven primarily by the dramatic inequalities in years of schooling and enrolments in low HDI countries; conversely, in high HDI countries, even (most of the) worst off still have some educational opportunities. Higher inequalities in the health and income component of the HDI further contribute to this relationship between levels and inequality in the HDI. Apart from this rather strong regularity, of particular interest are the outliers to this relationship. For example, Nigeria and Senegal have about the same HDI (See scaled HDI in Table 1), but inequality in the HDI is much higher in Senegal than Nigeria, driven particularly by higher education inequalities there. Also, India's inequality in its HDI is high in the sample (depending on the indicator chosen), much worse than Zambia which has much lower overall human development. The massive inequality in education in India is mostly responsible.[42] Studying these outliers in more detail and linking them to particular development policies and strategies is one of the further analyses one can undertake with these results.

After having discussed levels and drivers of unconditional inequality in the HDI, we now turn to the HDI by population subgroups. In Table 3, we first discuss the HDI by income deciles. There are also great differentials in the HDI by income decile and the differences are similarly dramatic between countries. The differentials are particularly pronounced in Burkina Faso, Nigeria, Ethiopia, Bolivia, and Zambia where the HDI of the richest decile is more than three times higher than the HDI of the poorest decile. At the other end of the spectrum, the differentials are more moderate in Armenia, Indonesia, Vietnam, Egypt and the Kyrgyz Republic.

To compare to earlier results in Grimm et al (2008, 2009) we also report HDIs by income quintile. The differentials in the HDI by income quintile is larger than those reported in Grimm et al (2008, 2009). Similar to the discussion above, the reason for the larger reported inequalities by income decile here is related to the different years of schooling indicator (instead of the previously used adult literacy indicator) and the use of the geometric mean which amplifies inequalities if achievement in one (sub-)component is low.

As with unconditional inequality in the HDI, the inequality by income decile suggests that inequality tend to be higher in settings where the level of human development is relatively low.[43]

Of course, the results by income decile are not unexpected as the income component is part of the HDI. But this clear distributional pattern is also observed when the health index and especially the education index are analyzed by income decile (Table 3). In particular, we find the largest inequality between the poorest and the richest income decile in the education component.

A strong distributional pattern is found for the HDI by education of the household head (Table 4). Households, where the head has no education are considerably worse off in terms of the HDI than better educated households. For example, Ethiopia shows a HDI that is more than 3 times as high for households where the household head has achieved secondary education compared to households where the head has no educational attainment at all (0.141 compared to 0.448). Again, the differentials are particularly large in Africa.

A much weaker pattern is found when looking at the outcomes in the HDI by the age of the household head (Table 4). It appears that households where the head is older experience, on average, a higher HDI than households with younger household heads.[44]

Table 5 shows the HDI by urban and rural areas. Also here, we find a clear trend. As expected, rural areas are worse off than urban areas with respect to human development.

The differences are not as large compared to income deciles but they are always sizable. For example, in Nicaragua, the ratio between rural and urban areas in the HDI is 0.630. The differences tend to be larger in poorer countries, particularly in Africa and are smallest in Egypt, again driven to an important extent by low differentials in education and health there. The large differential in Africa by location is again, particularly influenced by the massive inequalities in the education component. In urban Burkina, for example, the education component is about 10 times higher than in rural areas.[45]

To summarize the foregoing results, we found massive (unconditional) inequality in the HDI which is particularly severe in countries with low overall human development. The very high inequality is largely driven by the inequality in the education component and the implied penalty for inequality across components that is associated with the geometric mean. There are also (in declining order), large inequalities by income decile, education of the household, and location, and relatively small inequalities by age of the household head. These intergroup inequalities will be further examined in the next section.

3.3 Inequality Decompositions

In this section, we decompose the Theil index to see whether overall inequality is largely driven by inequality between the groups or within them. We only do this for the three groups (income decile, location, and education of the head) where inter-group differences were sizable. Even so, as Table 6 shows, the vast majority of total inequality is accounted for by inequality *within* these groups (e.g. *within* income deciles or *within* urban areas); inequality *between* groups accounts typically for less than 20% of total inequality for most countries. Thus for most countries, heterogeneity within groups are larger than between groups, a finding often made for income inequality decompositions as well.

There are three exceptions, however. In the three Sub Saharan African countries that are the worst performers in the overall HDI and also have the highest unconditional inequality in the HDI, the between component reaches up to 40% of total inequality. It thus appears that location, educational background and income decile are in these cases quite powerful predictors of overall inequality in the worst HDI and inequality performers.[46]

4 Conclusion

This paper provides a method and illustration for estimating an HDI at the household level. A household-based HDI provides us with a large range of previously unavailable types of analysis. On the one hand, it immediately allows the analysis of the HDI by any kind of population subgroups and by household socioeconomic characteristics. On the other hand, it allows to calculate any kind of inequality measure, compare it over space and time, and decompose it by groups. The use of DHS survey data allows to apply the method for many countries and time periods so that one could easily move beyond the 15 countries included here. While our results make use of a range of regression and simulation techniques, the results are remarkably robust to variations in those techniques; but clearly, there is scope for further analysis of our approach to estimate HDIs at the household level.

The results of our empirical illustration for 15 developing countries provide new insights with respect to levels and international comparisons of inequality in the HDI as well as inequality in the HDI by subgroups. It also sheds light on some key implications of the new indicators and methods to calculate the HDI in the 2010 report. The key findings are that inequality in the HDI, when using the new procedures, are massive. Particularly in some countries with low human development, inequalities in the HDI are about as large or possibly even larger than income inequalities. This is heavily influenced by massive inequalities in the two education indicators used and the use of the geometric mean to average across components. In other components and other countries, inequality in the HDI is much more moderate. Inequality by subgroups is also substantial, particularly by income groups, location, and education of the head. In low human development countries, inter-group differences between these groups account for a sizable share of total inequality in the HDI, while in other countries, within-group inequality is dominating total inequality.

It is beyond the scope of the paper to analyze the policy-drivers of inequality in HDI in different countries. But our approach is able to identify those countries where inequality in the HDI is a massive problem. Among the particular issues that deserve further attention are, for example, the very high inequality in human development (driven largely by education) in India, the very poor performance in the lowest HDI decile in countries as diverse as Armenia or Vietnam, and the massive urban-rural disparities in the HDI in Ethiopia and Burkina Faso.

References

Ainsworth, M. and D. Filmer (2006). Children's Schooling: AIDS, Orphanhood, Poverty, and Gender. *World Development* 34 (6):1099-1128.

Allison, P. D. (2007). Missing Data. Quantitative application in social sciences No. 07-136, Sage university press.

Alkire, S., and J. Foster (2009). Counting and Multidimensional Poverty Measurement. OPHI Working Paper 7, Oxford Poverty and Human Development Initiative, Oxford, UK.

Alkire, S., and M. Santos (2010). Acute Multidimensional Poverty: A New index for Developing countries. Human Development Research paper No. 11, UNDP, New York.

Anand S. and A. Sen (1992). Human Development Index: Methodology and Measurement. Human Development Report Office Occasional Paper 12, UNDP, New York.

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

Barro, R.J. and J.W. Lee (2010). A New Data Set of Educational Attainment in the World, 1950−2010. NBER Working Paper 15902. Cambridge, MA: National Bureau of Economic Research.

Bicego, G., S. Rutstein, and K. Johnson (2003). Dimensions of the Emerging Orphan Crisis in Sub-Saharan Africa. *Social Science and Medicine* 56 (6): 1235-1247.

Bollen, K. A., J. L. Glanville, and G. Stecklov (2002). Economic Status Proxies in Studies of Fertility In Developing Countries: Does the Measure matter?. *Population Studies*

56 (1): 81-96.

Bourguignon, F. (2003). The Growth Elasticity of Poverty Reduction: Explaining Heterogeneity across Countries and Time Periods. In T. Eichler and S. Turnovsky (eds.). *Growth and Inequality*. Cambridge. MIT Press.

Chant, S. (2008). Dangerous equations? How female-headed households became the poorest of the poor: causes, consequences and cautions. In: Momsen and Janet (eds.). *Gender and development: critical concepts in development studies. Critical concepts in development, 1 - Theory and classics*. Routledge, London, UK.

Chotikapanich, D., R. Valenzuela, and D.S.P Rao (1997). Global and Regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data. *Empirical Economics*, 22: 533-564.

Dreze, J. and and A. Sen (2002). *India - Development and Participation*. Oxford University Press.

Coale A.J. and P. Demeny (1983). *Regional model life tables and stable populations*. 2. ed. New York/ London: Academic Press.

Filmer, D. and L.H. Pritchett (2001). Estimating Wealth Effects without Expenditure Data - or Tears: An Application to Educational Enrollments in States of India. *Demography*, 38 (1): 115-132.

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

Foster J.E., L.F. López-Calva und M. Székely, (2005). Measuring the Distribution of Human Development: methodology and an application to Mexico. *Journal of Human Development*, 6 (1): 5-25.

Graham, J. W., Cumsille, P. E., and Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka and W. F. Velicer (Eds.). Research Methods in Psychology (pp. 87-114). Volume 2 of Handbook of Psychology (I. B. Weiner, Editor-in-Chief). New York: John Wiley and Sons.

Graham, J. W., and Hofer, S. M. (2000). Multiple imputation in multivariate research, in T. D. Little, K. U. Schnabel, and J. Baumert, (Eds.), *Modeling longitudinal and multiple-group data: Practical issues, applied approaches, and specific examples*, (201-218). Hillsdale, NJ: Erlbaum.

Grimm M., K. Harttgen, M. Misselhorn and S. Klasen (2006). A Human Development Index by Income Groups. Ibero-American Institute for Economic Research Discussion Paper No. 155, Ibero-American Institute for Economic Research, University of Gttingen.

Grimm, M., K. Harttgen, M. Misselhorn, and S. Klasen (2008). A Human Development Index by Income Groups. *World Development*, 36 (12): 2527-2546.

Grimm M., K. Harttgen, M. Misselhorn, T. Munzi, S. Klasen, and T. Smeeding (2009). Inequality in Human Development: An empirical assessment of thirty-two countries. *Social Indicator Research* DOI 10.1007/s11205-009-9497-7.

Haddad, L. and R. Kanbur (1990). How Serious Is the Neglect of Intra-Household Inequality?. *Economic Journal*, 100 (402): 866-881.

Harttgen, K. and S. Vollmer (2010). Validate the use of an asset index as a proxy for income. University of Goettingen, mimeo.

Harttgen, K. and M. Misselhorn (2006). A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa, Ibero-America

Institute Discussion Paper No. 152.

Harttgen, K. and S. Klasen (2011). A Human Development Index by Internal Migrational Status. *Journal of Human Development and Capabilities, forthcoming.*

Hicks D.A. (1997). The inequality-adjusted human development index: a constructive proposal. *World Development* 28 (8): 1283-1298.

Holzmann, H., S. Vollmer, and J. Weisbrod (2007). Income Distribution Dynamics and Pro-Poor Growth in the World from 1970 to 2003, Ibero America Institute for Econ. Research (IAI) Discussion Papers 161. Ibero-America Institute for Economic Research, Goettingen.

Houweling, T.A.J., A. E. Kunst, and J. P. Mackenbach (2003). Measuring Health Inequality among Children in Developing Countries: Does the Choice of the Indicator of Economic Status Matter?. *International Journal for Equity in Health* 2 (8).

Howe, L.D., J.R. Hargreaves, S. Gabrysch, and S.R.A. Huttly (2009). Is the wealth index a proxy for consumption? A systemic review. *J Epidemiol Community Health*, 63: 871-880.

Kelley A.C.(1991). The Human Development Index: 'Handle with Care'. *Population and Development Review*, 17 (2): 315-324.

Klasen S. (2006). UNDPs Gender-Related Measures: Some Conceptual Problems and Possible Solutions. *Journal of Human Development.* 7 (2): 243-274.

Klasen, S. (1996). Nutrition, Health, and Mortality in Sub-Saharan Africa: Is there a Gender Bias?. *Journal of Development Studies*, 32: 913-933.

Klasen, S., T. Lechtenfeld, and F. Povel (2010). What about the Women? Female Headship, Poverty and Vulnerability in Thailand and Vietnam. Proceedings of the German Development Economics Conference, Hannover 2010 43, Verein fr Socialpolitik, Research Committee Development Economics.

Landerman, Lawrence R., Kenneth C. Land and Carl F. Pieper (1997). An Empirical Evaluation of the Predictive Mean Matching Method for Imputing Missing Values. *Sociological Methods and Research* 26: 3-33.

Ledermann S. (1969). *Nouvelles tables-types de mortalité.* Travaux et documents, Cahier n. 53, Paris: INED and PUF.

Marcoux, A. (1999). The feminization of poverty: claims, facts, and data needs. *Population and Development Review*, 24 (1): 131-139.

McKenzie, D.J. (2005). Measuring inequality with asset indicators, *Journal of Population Economics*. 18: 229-260.

Milanovic, B. (2002). True world income distribution, 1988 and 1993: First calculation based on household surveys alone. *Economic Journal*, 112: 51-92.

Milanovic, B. (2006). Global Income Inequality: What it is and Why it Matters. *World Economics* 7: 131-157.

Murray C.J.L., B.D. Ferguson, A.D. Lopez, M. Guillot, J.A. Salomon, and O. Ahmad (2003). Modified logit life table system: principles, empirical validation, and application. *Population Studies*, 57 (2): 165-182.

Nguefack-Tsague, G., S.Klasen, and W. Zucchini (2010). On weighting the components of the Human Development Index: A statistical justification. Courant Research

Centre: Poverty, Equity and Growth - Discussion Papers 37, Courant Research Centre PEG, University of Gttingen.

Pan American Health Organization (PAHO) (2001). Measuring Health Inequalities: Gini Coefficient and Concentration Index. *Epidemiological Bulletin / PAHO*, 22 (1).

Ranis G., F. Stewart and E. Samman (2006). Human Development: beyond the HDI. QEH Working Paper Series - QEHWPS135.

Ravallion M. (1997). Good and bad growth: The Human Development Reports. *World Development*, 25 (5): 631-638.

Rubin, D. B. (1977). The design of a general and flexible system for handling non-response in sample surveys. manuscript prepared for the U.S. Social Security Administration.

Rubin, D. B. (1987). *Multiple Imputation for Non-Response in Surveys.* John Wiley and Sons, Inc..

Sagar A.D. and A. Najam (1998). The Human Development Index: A Critical Review. *Ecological Economics*, 25: 249-264.

Sahn, D. E. and D. Stifel (2003). Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. *Review of Income and Wealth* 49 (4): 463-489.

Sahn, D. E. and D. Stifel (2000). Poverty Comparisons over Time and Across Countries in Africa. *World Development* 28 (12): 2123-2155.

Sastry, N. (2004). Trends in Socioeconomic Inequalities in Mortality in Developing Countries: The case of Child Survival in Sao Paulo, Brazil. *Demography* 41 (3): 443- 464.

Schafer. J. L. and Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7 (2), 147-177.

Schaefer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman and Hall.

Schellenberg, J. A., C. G. Victora, A. Mushi, D. de Savigny, D.Schellenberg, H. Mshinda, and J. Bryce (2003). Inequities among the Very Poor: Health Care for Children in Rural Southern Tanzania. *The Lancet*, 361 (9357): 561- 566.

Schultz, T. P. (1998). Inequality in the distribution of personal income in the world: How it is changing and why. *Journal of Population Economics*, 11 (3): 307-344.

Stewart, T. and S. Simelane (2005). Are Assets a Valid Proxy for Income? An Analysis of Socioeconomic Status and Child Mortality in South Africa. Indiana University, mimeo.

Srinivasan T.N. (1994), Human Development: A New Paradigm or Reinvention of the Wheel?. *American Economic Review*, 84 (2): 238-243.

Stifel, D. and L. Christiaensen (2007). Tracking Poverty Over Time in the Absence of Comparable Consumption Data. *The World Bank Economic Review* 21 (2): 317-341.

Tarozzi, A. and A. Mahajan (2005). Child Nutrition in India in the Nineties: A Story of Increased Gender Inequality?. Unpublished manuscript. Duke University and Stanford University.

UNDP (2010), *Human Development Report 2010. The Real Wealth of Nations. Pathways to Human Development* UNDP, New York.

UNESCO (2010), Global Education Digest 2008, New York.

Wayman, J. C. (2003). Mulitple Imputation for Missing Data: What is it and How Can We Use it?. paper presented at the 2003 Annual Meeting of the American Educational Research Association, Chicago.

World Bank (2006). *World Development Report 2007: Development and the Next Generation.* The World Bank and Oxford University Press. Washington, DC.

World Bank (2005). World Development Indicators, CD-ROM. World Bank, Washington D.C..

Notes

[1]See also Nguefack et al. 2011 for a statistical justification for the currently used equal weighting scheme.

[2]For details, refer to UNDP (2010) and Alkire and Foster (2010). This 'discounting' is achieved by taking the geometric mean of achievements in a dimension which is lower than the arithmetic mean if there is inequality (and the difference grows with the amount of inequality). If there is no inequality, the two are the same.

[3]Of course, the inequality in the dimension indices must be based on some household-level data but these are calculated separately for each dimension so that the poor performers in each dimension are not necessarily referring to the same households and are based on separate sources of micro data. See also related work by Foster, López-Calva and Székely (2005) to calculate a distribution sensitive HDI for Mexico.

[4]This is largely due to data availability issues.

[5]See also related work by Foster, López-Calva and Székely (2005) to calculate a distribution sensitive HDI for Mexico.

[6]Although the HDI will be calculated at the household level, we could extend this analysis to the person-level by imputing the HDI of a household to each member. Of course this would ignore intra-household inequality in the HDI which is quite hard to tackle give our approach.

[7]The formular for estimating the asset index is: $A_i = \hat{\gamma}_1 a_{i1} + ... + \hat{\gamma}_n a_{in}$, where A_i is the asset index, the a_{in}'s refer to the respective asset of the household i recorded as dichotomous variables in the DHS data sets and the $\hat{\gamma}$ are the respective weights for each asset that are to be estimated.

[8]An alternative way to estimate the weights for the assets to derive the aggregated index is a factor analysis employed, for example, by Sahn and Stifel (2000). However, the two estimation methods show very similar results.

[9]The asset index is calculated for each individual, weighted by the household size. By also using DHS data, Houweling et al. (2003) analyze how the choice of indicators to be included in the asset index makes a difference in the ranking of households. The authors find significant but very small differences in the ranking of households depending on different sets of indicators.

[10]However, although they do find an overlap, they also show some differences in the ranking of households between the asset index and per capita expenditures in the lowest population quintile. The reason for the differences is that asset indices are less suitable to capture transitory shocks, because assets are a measure of stocks whereas income or expenditure are flow measures. In addition, assets indices are typically derived from public goods at the household level, while expenditures prominently captures the consumption of food.

[11]Similar results were also found by Harttgen and Volmer (2010) who validate the use of the asst index as a proxy for per capita income using LSMS data for several developing countries and compare the household ranking and outcomes of social indicators of human development. They also find some differences in the ranking of households while also here the gradient of the asset index and household per capita income are similar.

[12]In addition, assets are less vulnerable to measurement error based on underreporting (McKenzie, 2005).

[13]This results in a more continues character of the asset density distribution shown in Figure A.1 for each country.

[14]We should briefly discuss the interpretation of the parameters μ and σ, which is different from that of the normal distribution. In fact, from (2) one sees that e^{μ} is proportional to the expectation and $(e^{\mu})^2$ is proportional to the variance, and in fact, e^{μ} is the scale parameter of the log-normal distribution, whereas σ is a shape parameter. Since the Gini coefficient is invariant under changes of scale (it does not matter whether income is measured in Euro or in Dollar), it should be independent of μ and only depend on σ. This is indeed the case: The Gini coefficient G of $LN(\mu, \sigma)$ is given by $G = 2\Phi(\sigma/\sqrt{2}) - 1$, where Φ is the distribution function of the standard normal distribution.

[15]For example, Holzmann et al. (2007) estimate the global income distribution based on the assumption of the log normal distribution using also the Gini and the mean income as two parameters. When testing for log-normality from the quintiles or even from the deciles, they could rejected the hypothesis of log-normality for only less than 0.5% of all countries, and never for one of the population heavy weights China, India, the U.S., Indonesia, and Brazil.

[16]The estimation of the distribution is based on a maximum likelihood estimation technique.

[17]To illustrate this approach, Figure A.1 shows the asset index distribution (left) and also the obtained income distribution following our approach (right) for several countries in our sample.

[18]$y_h^{PPP} = y_h \times PPP$.

[19]$ry_h^{PPP} = y_h^{PPP} \times \left[\frac{GDPPC^{PPP}}{\bar{y}^{PPP}}\right]$.

[20] $Y^h = \frac{\log \bar{ry}^{h,PPP} - \log(163)}{\log(108,211) - \log(163)} \quad \forall\, h = 1, 2, \ldots, K$, where $\bar{ry}^{h,PPP}$ is the household specific arithmetic mean of the rescaled household income per capita.

[21] In particular, for each child of age a, the expected years of education are calculated as the sum of the age specific enrolment ratios to the reference age range a to n (UNESCO 2010) The expected years of education is calculated by $EYS_a^t = \sum_{i=1}^{n} \frac{E_i^t}{P_i^t}$, where $EYS_a^t$is the expected years of education at age a in year t, E_i^t is the enrolment of children at age i $(i = a, a+1, ..., n)$ in school year t, n denotes the upper age limit of the respective school system, and P_i^t denotes the population of age i in school year t.

[22] It shares the same problem as gross enrolment rates in that high rates of repetition will erroneously boost expected years of education.

[23] Please note that in the recently published Multidimensional Poverty Measure, also included in the HDR 2010, parents without children are given perfect score in the enrolment component which, by construction, makes them less susceptible to being multidimensionally poor.

[24] To keep the calculation time manageable, we set the number of imputations to derive the missing values to 20. Please also note that we perform the imputations only for households without children in school-going age. For the others, we use the actual information. Imputing enrolment rates for all households does not change any of the results we present below. Results are available on request.

[25] Remarkably, including the asset index only has a minor impact on the fit of the regressions and the resulting imputations.

[26] $MY^h = \frac{my^h - 0}{my^{max} - 0} \quad \forall\, h = 1, 2, \ldots, K$, and $EY^h = \frac{ey^h - 0}{ey^{max} - 0} \quad \forall\, h = 1, 2, \ldots, K.$, where my^h refers to the household specific mean years of education and ey^h to the household specific expected years of education.

[27] $E^h = \frac{\sqrt{MY^h * EY^h} - 0}{0.951 - 0} \quad \forall\, h = 1, 2, \ldots, K.$

[28] We also tried various other specification and included other covariates, but the results of the predicted outcomes did not change if we add further variables.

[29] In this approach, we also differ from the recently published Multidimensional Poverty Measure (Alkire and Santos 2010) which uses only actual information and thus gives perfect enrolment score to households without children and perfect health score to households where no child has died.

[30] In a robustness analysis we also the asset index into the regression for child mortality. Inclusion of the asset index has an extremely minor impact on the correlation between the income and the health component, only visible in the third decimal place. But since the asset index also has rather little predictive power in the regression we do not include it.

[31] This model is based on a Brass logit approach: $Logit(l_x^h) = \alpha_h + \beta_h * Logit(l_x^s) + \gamma_x \left[1 - \left(\frac{Logit(l_5^h)}{Logit(l_5^s)}\right)\right] + \theta_x \left[1 - \left(\frac{Logit(l_{60}^h)}{Logit(l_{60}^s)}\right)\right] \quad \forall\, h = 1, 2, \ldots, K.$, where x is the age, γ_x and θ_x are parameters of the age specific Standard Life Table, α_h and β_h are country specific parameters, and γ the survival probability from zero to x, 5, and 60. To any value of l_5, the corresponding value for the life expectancy at birth e_0 can be estimated through in iterative procedure.

[32] $H^h = \frac{\hat{e}_0^h - 20}{83{,}2 - 20} \quad \forall\, h = 1, 2, \ldots, K.,$

[33] See Foster et al. (2005) and Alkire and Foster (2010).

[34] For a further discussion of these issues, see Klasen (2006) and Haddad and Kanbur (1990).

[35] For all the results presented in this section, we do not, following also the procedure of UNDP, provide any confidence intervals or significance tests between differences in the outcomes because of space limitations. Standard errors, confidence intervals and significance test can be provided on request. See Table A.4 for the descriptive statistics that are the basis for the calculation of the household-based HDI.

[36] Since UNDP's HDI are based on aggregate data and our HDI is based on micro data, the two sources need not generate the same results due to differences in definitions and measurement error with the aggregate or the micro data. As shown in the Table, there are indeed significant differences between the two. We believe that scaling is appropriate as we are mainly interested in the HDI differentials within a country, rather than in presenting an alternative mean HDI which would be hard to interpret.

[37] See Table A.4.

[38] Mathematically, the geometric mean is identical to the arithmetic mean of the log of the individual components (except for the case of an index value of zero); that is to say, imperfect substitution possibilities is mathematically equivalent to assuming declining marginal returns to each component, which, in general, appears to be a reasonable assumption.

[39] The reason for these small differences is that the asset index distribution is less continuous than the income distribution. This means, for the imputation of the household per capita income we do not take the whole income distribution, but rather draw from the distribution for the values of the asset index distribution.

[40] See Grimm et al (2008, 2009) for a further discussion of this issue.

[41]Unfortunately, only very limited comparable Gini coefficients on health exits in the empirical literature. One exception is a study by the Pan American Health Organization (PAHO) (2001) which calculates Gini coefficients for infant mortality for five countries from Latin America. Also here, the results is quite similar to our results in a sense that Gini coefficients for the health indicators are lower than for income indicators. See also Grimm et al (2008, 2009) for a related discussion.

[42]See, for example, Dreze and Sen, 2002 on poor educational performance in India.

[43]This is plausible to the extent that differences in nutritional status, essential access to housing and clothing are smaller in high HDI countries than in low HDI countries. See also Grimm et al (2008) for further discussion.

[44]However, these results should be treated with caution, because they are also be driven by differences in the shares of households of the respective age ranges and thus the calculation are based in very different numbers of observation, For example, there are many more households with a household head aged between 20 and 29 than aged 60 years or older. See also Table A.6 for the results for the components of the HDI by age of the household head.

[45]We also investigate whether there is inequality in the HDI by sex of the household head. Here we find generally very small differentials sometimes suggesting that female-headed households have slightly higher human development and sometimes finding the reverse. Since the group of female-headed households differs greatly by size across countries and is, in any case, a rather heterogeneous group, these findings are hard to interpret but provide further confirmation to the literature claiming that female-headship per se is not a clear marker of deprivation in many developing countries. See, e.g. Chant (2008), Marcoux (1999) and Klasen, Lechtenfeld, and Povel (2010). Results are available on request.

[46]To some degree, of course, the results are influenced by the fact that the income is part of the HDI and education of the household head if part of overall education. All the more is it surprising how little the between component explains in all the other countries.

Tables and figures

Table 1: Overall HDI and sub-components by country

Country	HDI	Ranking (HDR) (2010)	HDI (scaled)	HDI (old)	Ranking (HDR) (2009)	HDI (small)	GNI index	GNI index (scaled)	Education index	Education index (scaled)	Education index (small)	Education index (old)	Health index	Health index (scaled)	Health index (small)
Armenia (2005)	0.628	99	0.671	0.731	108	0.540	0.476	0.519	0.572	0.693	0.571	0.760	0.909	0.840	0.971
Peru (2005)	0.593	111	0.727	0.727	130	0.536	0.511	0.601	0.526	0.744	0.526	0.723	0.776	0.860	0.883
Indonesia (2003)	0.540	117	0.535	0.670	132	0.463	0.436	0.444	0.435	0.444	0.435	0.687	0.828	0.774	0.921
Vietnam (2002)	0.535	118	0.528	0.664	132	0.474	0.359	0.376	0.481	0.462	0.481	0.693	0.886	0.845	0.954
Egypt (2007)	0.513	119	0.626	0.811	136	0.287	0.507	0.550	0.317	0.554	0.288	0.492	0.843	0.802	0.911
Kyrgyz R. (1997)	0.507	120	0.557	0.670	139	0.442	0.301	0.353	0.559	0.666	0.559	0.806	0.775	0.736	0.810
Bolivia (2003)	0.473	127	0.630	0.303	154	0.333	0.414	0.491	0.352	0.721	0.352	0.622	0.726	0.706	0.829
Nicaragua (2000)	0.466	131	0.560	0.625	149	0.380	0.368	0.436	0.350	0.500	0.350	0.536	0.784	0.806	0.901
India (2005)	0.428	140	0.447	0.552	149	0.301	0.341	0.384	0.264	0.348	0.264	0.500	0.870	0.671	0.983
Pakistan (2007)	0.420	144	0.481	0.603	159	0.301	0.407	0.424	0.266	0.358	0.266	0.435	0.686	0.733	0.807
Nigeria (2003)	0.386	152	0.390	0.404	169	0.277	0.309	0.344	0.332	0.412	0.331	0.534	0.560	0.418	0.711
Zambia (2002)	0.383	154	0.340	0.348	169	0.310	0.259	0.293	0.362	0.394	0.362	0.602	0.598	0.341	0.721
Senegal (2005)	0.317	158	0.404	0.480	172	0.187	0.340	0.374	0.147	0.317	0.147	0.339	0.635	0.557	0.763
Ethiopia (2005)	0.220	168	0.252	0.330	183	0.102	0.186	0.208	0.102	0.153	0.102	0.281	0.560	0.503	0.684
Burkina Faso (2003)	0.220	168	0.278	0.564	182	0.077	0.256	0.281	0.066	0.156	0.066	0.203	0.631	0.489	0.712
Average	0.442	135	0.495	0.565	151	0.334	0.365	0.405	0.342	0.462	0.340	0.548	0.738	0.672	0.837

Note: 'HDI' is based on multiple imputation of school enrolment to calculate the expected years of education and on estimated under five mortality rates at the household level. 'HDI old' is the arithmetic mean of the three subcomponents. 'HDI small' is based on the sample with school-aged children. The ranking refers to the ranks the country would have had in the HDR 2010 for the new calculation methodology and in the HDR 2009 for the old methodology. 'Education old' is based on the old methodology of calculating the education index based on adult literacy and gross school enrolment. 'scaled' refers to the value of the HDI scaled to the official HDI value of the respective survey year.
Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table 2: Inequality in the HDI by country

	Unconditional HDI						Gini index						Atkinson index				Theil index
	Decile		Ratio	Quintile		Ratio						Gini					
Country	1	10	10:1	1	5	5:1	HDI	Educ. index	Health index	GNI index	Simulated income	(PovcalNet) *(%)*	HDI	Educ. index	Health index	GNI index	HDI
Armenia (2005)	0.114	0.875	7.653	0.229	0.843	3.677	0.229	0.299	0.074	0.121	0.363	*33.80*	0.290	0.316	0.026	0.024	0.343
Burkina F. (2003)	0.058	0.696	12.062	0.060	0.565	9.424	0.646	0.881	0.226	0.235	0.384	*39.60*	0.596	0.519	0.127	0.099	0.907
Bolivia (2003)	0.162	1.044	6.434	0.171	0.977	5.707	0.456	0.542	0.169	0.249	0.590	*60.24*	0.545	0.516	0.095	0.132	0.787
Egypt (2007)	0.242	0.985	4.070	0.294	0.939	3.191	0.446	0.605	0.099	0.092	0.294	*32.14*	0.444	0.566	0.036	0.014	0.587
Ethiopia (2005)	0.049	0.620	12.655	0.060	0.511	8.517	0.571	0.786	0.241	0.188	0.232	*29.76*	0.508	0.582	0.133	0.058	0.709
India (2005)	0.063	0.728	11.557	0.067	0.695	10.411	0.418	0.553	0.104	0.150	0.334	*33.32*	0.436	0.493	0.048	0.036	0.572
Indonesia (2003)	0.112	0.730	6.519	0.143	0.704	4.922	0.243	0.359	0.124	0.107	0.297	*30.23*	0.249	0.313	0.050	0.019	0.286
Kyrgyz R. (1997)	0.117	0.764	6.531	0.184	0.722	3.934	0.238	0.285	0.143	0.168	0.319	*35.98*	0.048	0.293	0.056	0.054	0.049
Nicaragua (2000)	0.084	0.836	9.955	0.092	0.792	8.609	0.334	0.454	0.134	0.225	0.499	*50.30*	0.339	0.395	0.062	0.092	0.414
Nigeria (2003)	0.096	0.671	7.010	0.110	0.634	5.762	0.446	0.542	0.264	0.223	0.423	*42.93*	0.467	0.519	0.172	0.118	0.630
Pakistan (2007)	0.110	0.756	6.852	0.111	0.718	6.471	0.403	0.544	0.179	0.123	0.312	*31.18*	0.433	0.501	0.094	0.025	0.568
Peru (2005)	0.207	0.986	4.755	0.374	0.947	2.530	0.216	0.289	0.143	0.171	0.504	*53.01*	0.218	0.226	0.064	0.053	0.245
Senegal (2005)	0.091	0.744	8.213	0.096	0.686	7.131	0.547	0.691	0.198	0.194	0.404	*39.19*	0.562	0.565	0.112	0.070	0.824
Vietnam (2002)	0.072	0.701	9.772	0.211	0.674	3.192	0.216	0.313	0.087	0.155	0.346	*37.55*	0.203	0.276	0.029	0.043	0.227
Zambia (2002)	0.104	0.583	5.618	0.105	0.546	5.218	0.359	0.421	0.217	0.269	0.415	*42.08*	0.337	0.399	0.109	0.155	0.410
Average	0.112	0.781	7.977	0.154	0.730	5.913	0.385	0.504	0.160	0.178	0.381	39.231	0.378	0.432	0.081	0.066	0.504

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table 3: HDI by income (conditional) deciles

							Conditional								
	HDI		Ratio	HDI (by quintile)		Ratio	GNI index		Ratio	Education index		Ratio	Health index		Ratio
Country	1	10	10:1	1	5	10:1	1	10	10:1	1	10	10:1	1	10	10:1
Armenia (2005)	0.558	0.796	1.426	0.574	0.757	1.317	0.334	0.697	2.087	0.619	0.843	1.362	0.842	0.858	1.020
Burkina Faso (2003)	0.090	0.633	7.074	0.108	0.531	4.915	0.089	0.513	5.735	0.018	0.839	47.918	0.459	0.591	1.288
Bolivia (2003)	0.259	0.979	3.779	0.321	0.911	2.836	0.109	0.852	7.778	0.294	1.368	4.648	0.539	0.805	1.493
Egypt (2007)	0.494	0.858	1.738	0.492	0.722	1.467	0.372	0.702	1.888	0.438	1.034	2.359	0.738	0.870	1.179
Ethiopia (2005)	0.154	0.588	3.825	0.153	0.480	3.133	0.115	0.375	3.254	0.064	0.820	12.811	0.493	0.662	1.342
India (2005)	0.305	0.636	2.088	0.326	0.617	1.893	0.237	0.595	2.512	0.186	0.620	3.336	0.643	0.698	1.087
Indonesia (2003)	0.394	0.666	1.689	0.417	0.645	1.549	0.290	0.592	2.042	0.292	0.609	2.088	0.726	0.820	1.129
Kyrgyz Republic (1997)	0.394	0.709	1.801	0.437	0.677	1.550	0.158	0.534	3.371	0.566	0.800	1.414	0.682	0.835	1.224
Nicaragua (2000)	0.260	0.776	2.979	0.305	0.737	2.421	0.120	0.697	5.815	0.213	0.791	3.716	0.692	0.847	1.224
Nigeria (2003)	0.171	0.636	3.727	0.200	0.585	2.929	0.102	0.575	5.643	0.142	0.873	6.146	0.344	0.513	1.493
Pakistan (2007)	0.274	0.684	2.496	0.295	0.646	2.188	0.268	0.582	2.171	0.118	0.682	5.765	0.650	0.807	1.243
Peru (2005)	0.424	0.927	2.184	0.473	0.894	1.890	0.246	0.899	3.652	0.450	0.947	2.107	0.691	0.934	1.353
Senegal (2005)	0.215	0.653	3.041	0.210	0.602	2.865	0.148	0.587	3.974	0.143	0.737	5.170	0.470	0.644	1.369
Vietnam (2002)	0.338	0.656	1.940	0.379	0.631	1.667	0.193	0.562	2.919	0.253	0.578	2.287	0.793	0.867	1.093
Zambia (2002)	0.161	0.543	3.378	0.197	0.511	2.590	0.054	0.527	9.803	0.243	0.709	2.918	0.317	0.427	1.348
Aaverage	0.299	0.716	2.878	0.326	0.663	2.347	0.189	0.619	4.176	0.269	0.817	6.936	0.605	0.745	1.259

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table 4: HDI by education of household head and age of household head

Country	HDI By education of household head No education	Prim	Sec.	Ratio sec. to no edu.	By age of household head Agegroup 20-29	60+	Ratio oldest to youngest
Armenia (2005)	0.591	0.619	0.645	1.093	0.606	0.678	1.119
Burkina Faso (2003)	0.180	0.433	0.657	3.660	0.220	0.242	1.100
Bolivia (2003)	0.344	0.514	0.728	2.114	0.538	0.603	1.120
Egypt (2007)	0.496	0.674	0.675	1.359	0.708	0.727	1.027
Ethiopia (2005)	0.141	0.312	0.448	3.182	0.194	0.216	1.112
India (2005)	0.299	0.457	0.542	1.816	0.359	0.481	1.340
Indonesia (2003)	0.381	0.494	0.608	1.596	0.429	0.507	1.181
Kyrgyz R. (1997)	0.463	0.468	0.540	1.166	0.341	0.471	1.381
Nicaragua (2000)	0.356	0.533	0.710	1.992	0.468	0.537	1.147
Nigeria (2003)	0.209	0.427	0.501	2.393	0.317	0.354	1.115
Pakistan (2007)	0.333	0.490	0.585	1.757	0.372	0.467	1.257
Peru (2005)	0.457	0.625	0.778	1.702	0.586	0.720	1.228
Senegal (2005)	0.313	0.532	0.653	2.082	0.355	0.406	1.142
Vietnam (2002)	0.353	0.448	0.564	1.597	0.410	0.497	1.212
Zambia (2002)	0.206	0.309	0.424	2.060	0.267	0.302	1.132
Average	0.342	0.489	0.604	1.971	0.411	0.481	1.174

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table 5: HDI and sub-components by country rural and urban areas

	HDI			GNI index			Education index			Health index		
Country	Urban	Rural	Ratio rural/urban	Urban	Rural	Ratio rural/urban	Urban	Rural	Ratio rural/urban	Urban	Rural	Ratio rural/urban
Armenia (2005)	0.709	0.611	0.862	0.578	0.428	0.740	0.735	0.640	0.871	0.839	0.834	0.994
Burkina Faso (2003)	0.530	0.189	0.356	0.457	0.253	0.555	0.574	0.056	0.098	0.569	0.474	0.833
Bolivia (2003)	0.738	0.390	0.529	0.584	0.302	0.517	0.923	0.323	0.350	0.745	0.609	0.818
Egypt (2007)	0.678	0.575	0.849	0.604	0.501	0.830	0.620	0.489	0.789	0.832	0.777	0.934
Ethiopia (2005)	0.500	0.222	0.445	0.327	0.194	0.595	0.600	0.112	0.187	0.637	0.502	0.789
India (2005)	0.537	0.419	0.781	0.475	0.353	0.742	0.473	0.316	0.668	0.690	0.662	0.961
Indonesia (2003)	0.583	0.491	0.843	0.490	0.406	0.828	0.510	0.385	0.755	0.791	0.758	0.958
Kyrgyz Republic (1997)	0.626	0.519	0.830	0.435	0.309	0.711	0.724	0.636	0.878	0.777	0.711	0.914
Nicaragua (2000)	0.639	0.402	0.630	0.512	0.315	0.617	0.611	0.280	0.458	0.836	0.739	0.884
Nigeria (2003)	0.484	0.341	0.704	0.437	0.294	0.671	0.554	0.343	0.619	0.468	0.393	0.840
Pakistan (2007)	0.584	0.427	0.731	0.498	0.387	0.776	0.513	0.285	0.556	0.779	0.706	0.907
Peru (2005)	0.803	0.577	0.719	0.688	0.445	0.647	0.833	0.568	0.682	0.904	0.761	0.842
Senegal (2005)	0.553	0.281	0.508	0.472	0.313	0.664	0.562	0.140	0.250	0.636	0.503	0.791
Vietnam (2002)	0.603	0.504	0.837	0.496	0.352	0.710	0.510	0.436	0.854	0.864	0.835	0.966
Zambia (2002)	0.454	0.297	0.655	0.425	0.231	0.543	0.557	0.352	0.632	0.395	0.323	0.819
Average	0.601	0.416	0.685	0.498	0.339	0.676	0.620	0.357	0.576	0.717	0.639	0.883

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table 6: Inequality decomposition of the HDI by subgroups

	Theil index of HDI						
		By income deciles		By urban and rural areas		By Education of household head	
	Total	Within group	Between group	Within group	Between group	Within group	Between group
Armenia (2005)	0.343	0.338	0.004	0.341	0.002	0.341	0.002
Burkina Faso	0.907	0.551	0.356	0.648	0.259	0.586	0.319
Bolivia	0.787	0.678	0.109	0.691	0.096	0.741	0.044
Egypt (2007)	0.587	0.505	0.082	0.580	0.007	0.545	0.043
Ethiopia (2005)	0.709	0.591	0.118	0.629	0.080	0.379	0.329
India (2005)	0.572	0.516	0.056	0.562	0.010	0.448	0.125
Indonesia (2003)	0.286	0.272	0.014	0.281	0.005	0.273	0.013
Kyrgyz Republic (1997)	0.295	0.284	0.011	0.293	0.003	0.292	0.003
Nicaragua (2000)	0.414	0.338	0.076	0.378	0.036	0.349	0.063
Nigeria (2003)	0.630	0.497	0.133	0.605	0.025	0.422	0.207
Pakistan (2007)	0.568	0.490	0.078	0.549	0.018	0.454	0.113
Peru (2005)	0.245	0.223	0.022	0.234	0.011	0.230	0.015
Senegal (2005)	0.824	0.636	0.188	0.674	0.150	0.686	0.139
Vietnam (2002)	0.227	0.209	0.018	0.225	0.001	0.211	0.015
Zambia (2002)	0.410	0.328	0.082	0.382	0.029	0.363	0.047

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Appendix

Table A.1: OLS Regression of household literacy and enrolment - Burkina Faso

VARIABLES	enrolment rate
Age of household member	0.0349***
	(0.00136)
Age of household member^2	-0.0576***
	(0.00258)
Age of household member^3	0.00300***
	(0.000156)
Urban (=1)	-0.00473
	(0.00541)
Female headed household (=1)	-0.000989
	(0.00415)
Household size	-0.00762***
	(0.000437)
Number of children at home	0.0123***
	(0.000694)
Sex of household member (1=male)	0.0220**
	(0.00866)
Head has no education (=1)	-0.0781***
	(0.00305)
sex*age	-0.000109
	(9.72e-05)
sex*urban	-0.00743
	(0.00456)
urban*age	2.45e-05
	(3.73e-05)
Literacy (mean per cluster)	-0.00388
	(0.0337)
Enrolment (mean per cluster)	0.984***
	(0.0126)
Years of education (mean per cluster)	-0.00516
Constant	-0.565***
	(0.0254)
Observations	57866
R-squared	0.415

Note: Also controlled for within-country regions.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table A.2: Regression of child mortality - Burkina Faso

Under 5 mortality	Coef.	Std. Err.	z
Urban (=1)	-0.192	0.123	-1.560
Female headed household (=1)	-0.408	0.167	-2.450
Household size	-0.003	0.002	-1.680
Number of children at home	-0.339	0.008	-42.500
Age of mother	-0.056	0.003	-21.970
Total children ever born	0.222	0.007	33.940
Mother currently pregnant (=1)	-0.014	0.039	-0.360
Mother currently breast feeding (=1)	-0.306	0.032	-9.600
Mother wants another child (=1)	0.174	0.029	6.010
Mother works (=1)	-0.134	0.116	-1.150
Mother catholic (=1)	-0.059	0.028	-2.150
Mother is not literate (=1)	0.191	0.091	2.110
Mother has primary education (=1)	-0.050	0.081	-0.620
Mother is married (=1)	-0.164	0.029	-5.620
Mother received tetanus after last birth (=1)	-0.419	0.102	-4.100
Mother received prenatal care after last birth (=1)	-1.482	0.065	-22.890
Mother received professional assistance during delivery (=1)	-0.319	0.058	-5.530
Urban*age	0.004	0.002	1.890
Urban*fhh	-0.161	0.144	-1.120
Sex*age	0.006	0.003	1.780
Literacy rate (cluster mean)	0.322	0.346	0.930
Primary education (cluster mean)	-0.019	0.329	-0.060
Constant	0.851	0.091	9.360
N	9046		

Note: Discrete time proportional hazard model with piece-wise constant baseline hazard. In addition, dummy variables were included for 12 month time intervals.
Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table A.3: Correlation between indicators

	income / edu years	income / exp. edu years	income / life expec.	life expec. / edu years	life expec. / exp. edu years	edu years exp. edu years	edu index/ health index	edu index/ GNI index	GNI index/ health index	HDI new/ HDI old
Armenia (2005)	0.39	0.07	0.05	0.07	0.30	0.12	0.33	0.15	0.02	0.97
Burkina Faso (2003)	0.63	0.47	0.23	0.13	0.21	0.40	0.14	0.50	0.18	0.95
Bolivia (2003)	0.54	0.25	0.24	0.29	0.18	0.25	0.24	0.46	0.37	0.97
Egypt (2007)	0.58	0.20	0.25	0.20	0.30	0.22	0.10	0.09	0.24	0.93
Ethiopia (2005)	0.64	0.37	0.19	0.13	0.23	0.32	0.20	0.56	0.18	0.93
India (2005)	0.54	0.10	0.11	0.07	0.21	0.20	0.18	0.41	0.11	0.94
Indonesia (2003)	0.51	0.15	0.18	0.11	0.21	0.20	0.23	0.36	0.18	0.95
Kyrgyz Republic (1997)	0.38	0.13	0.24	0.22	0.32	0.17	0.36	0.16	0.23	0.99
Nicaragua (2000)	0.55	0.25	0.19	0.27	0.29	0.38	0.32	0.54	0.25	0.94
Nigeria (2003)	0.53	0.29	0.22	0.23	0.32	0.40	0.28	0.52	0.26	0.96
Pakistan (2007)	0.54	0.28	0.19	0.14	0.24	0.35	0.21	0.49	0.21	0.95
Peru (2005)	0.47	0.07	0.23	0.31	0.20	0.11	0.31	0.41	0.33	0.95
Senegal (2005)	0.47	0.17	0.26	0.19	0.18	0.38	0.21	0.41	0.29	0.95
Vietnam (2002)	0.51	0.11	0.20	0.23	0.24	0.22	0.31	0.37	0.18	0.96
Zambia (2002)	0.57	0.29	0.28	0.20	0.26	0.28	0.31	0.47	0.26	0.97

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table A.4: Descriptive statistics

	Income simulated	Income scaled	Years of educ.	Exp . years of educ.	Life expect.	Life expect. (small)	Sample size	Sample size (small)
Armenia (2005)	989	4471	8.84	10.98	79.74	84.04	5643	2175
Burkina F. (2003)	545	1100	0.93	3.74	57.32	60.41	9028	1058
Bolivia (2003)	2185	4778	6.73	8.06	65.92	66.83	19028	4740
Egypt (2007)	1336	5149	6.67	10.47	73.75	77.65	7111	1533
Ethiopia (2005)	537	595	1.53	6.91	55.37	57.58	13636	1861
India (2005)	490	1788	3.75	9.58	76.33	80.26	108443	22029
Indonesia (2003)	615	3206	6.55	11.17	73.06	78.40	31059	6649
Kyrgyz R. (1997)	696	1364	10.06	10.13	68.99	71.21	3647	883
Nicaragua (2000)	1748	2805	4.64	10.44	69.67	73.19	11244	1458
Nigeria (2003)	472	1656	4.70	10.36	55.39	57.42	7111	1533
Pakistan (2007)	822	2696	3.51	9.52	63.37	65.55	13483	1185
Peru (2005)	2386	7297	7.77	12.74	69.14	70.79	28354	6843
Senegal (2005)	914	1963	1.82	6.87	60.11	61.71	7154	664
Vietnam (2002)	635	2058	6.79	12.41	77.35	79.12	6987	1358
Zambia (2002)	561	1186	6.35	8.25	57.79	61.82	6995	960
Average	995	2807	5.38	9.44	66.89	69.73	18595	3662

Note: Years of education refers to the mean years of education of per household of adults aged 25+. Expected years of education are based on enrolment by age at all levels of education and population of official school age for each level of education (UNESCO 2010). Household income per capita is expressed in USD PPP. Scaled income refers to the household per capita income that is scaled to the national GNI per capita (USD PPP) for the respective country and year taken from the Human Development Report. The index 'small' refers to the sample of households with school aged children.
Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table A.5: HDI components by education of the household head

	GNI index				Education index				Health index			
Country	No education	Prim	Sec.	Ratio sec. to no edu.	No education	Prim	Sec.	Ratio sec. to no edu.	No education	Prim	Sec.	Ratio sec. to no edu.
Armenia (2005)	0.410	0.447	0.502	1.223	0.575	0.631	0.644	1.121	0.874	0.841	0.832	0.951
Burkina Faso (2003)	0.269	0.350	0.460	1.711	0.045	0.450	1.059	23.713	0.482	0.515	0.582	1.208
Bolivia (2003)	0.329	0.409	0.561	1.707	0.196	0.507	0.925	4.713	0.634	0.656	0.744	1.174
Egypt (2007)	0.488	0.526	0.563	1.154	0.330	0.741	0.669	2.026	0.759	0.787	0.816	1.075
Ethiopia (2005)	0.194	0.211	0.268	1.378	0.028	0.290	0.613	22.125	0.517	0.494	0.547	1.057
India (2005)	0.343	0.381	0.417	1.213	0.118	0.370	0.566	4.801	0.657	0.675	0.676	1.028
Indonesia (2003)	0.394	0.425	0.481	1.220	0.185	0.372	0.590	3.193	0.758	0.763	0.791	1.043
Kyrgyz Republic (1997)	0.309	0.318	0.332	1.072	0.462	0.486	0.652	1.410	0.693	0.663	0.728	1.051
Nicaragua (2000)	0.342	0.411	0.529	1.544	0.181	0.465	0.773	4.267	0.729	0.791	0.875	1.200
Nigeria (2003)	0.276	0.352	0.398	1.443	0.087	0.505	0.713	8.145	0.381	0.439	0.444	1.166
Pakistan (2007)	0.391	0.420	0.455	1.162	0.135	0.384	0.581	4.309	0.701	0.729	0.760	1.084
Peru (2005)	0.448	0.523	0.637	1.423	0.293	0.588	0.833	2.845	0.728	0.795	0.887	1.219
Senegal (2005)	0.354	0.413	0.488	1.378	0.161	0.625	0.892	5.551	0.541	0.585	0.639	1.181
Vietnam (2002)	0.300	0.337	0.396	1.320	0.184	0.329	0.529	2.880	0.798	0.812	0.856	1.072
Zambia (2002)	0.209	0.253	0.367	1.758	0.127	0.351	0.560	4.416	0.330	0.330	0.372	1.125
Average	0.337	0.385	0.457	1.380	0.207	0.473	0.707	6.368	0.639	0.658	0.703	1.109

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table A.6: HDI components by age of the household head

	GNI index			Education index			Health index		
Country	20-29	30-39	Ratio old/ young	20-29	60+	Ratio old/ young	20-29	60+	Ratio old/ young
Armenia (2005)	0.528	0.495	0.938	0.540	0.740	1.371	0.784	0.854	1.089
Burkina Faso (2003)	0.259	0.290	1.119	0.090	0.092	1.021	0.455	0.531	1.167
Bolivia (2003)	0.443	0.504	1.139	0.580	0.616	1.061	0.607	0.706	1.163
Egypt (2007)	0.530	0.511	0.964	0.991	0.995	1.004	0.676	0.756	1.118
Ethiopia (2005)	0.206	0.205	0.994	0.102	0.094	0.918	0.347	0.523	1.507
India (2005)	0.356	0.403	1.132	0.219	0.427	1.951	0.596	0.648	1.089
Indonesia (2003)	0.413	0.452	1.094	0.305	0.390	1.281	0.628	0.738	1.174
Kyrgyz R. (1997)	0.346	0.322	0.931	0.213	0.493	2.314	0.537	0.657	1.223
Nicaragua (2000)	0.368	0.453	1.231	0.383	0.444	1.160	0.727	0.769	1.057
Nigeria (2003)	0.322	0.323	1.002	0.298	0.325	1.090	0.333	0.422	1.268
Pakistan (2007)	0.397	0.429	1.078	0.232	0.352	1.517	0.558	0.678	1.214
Peru (2005)	0.518	0.639	1.234	0.545	0.709	1.301	0.714	0.825	1.155
Senegal (2005)	0.383	0.387	1.010	0.241	0.307	1.277	0.485	0.561	1.155
Vietnam (2002)	0.313	0.400	1.280	0.307	0.382	1.242	0.716	0.803	1.121
Zambia (2002)	0.280	0.252	0.900	0.253	0.320	1.264	0.268	0.341	1.274
Average	0.378	0.404	1.070	0.353	0.446	1.318	0.562	0.654	1.185

Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

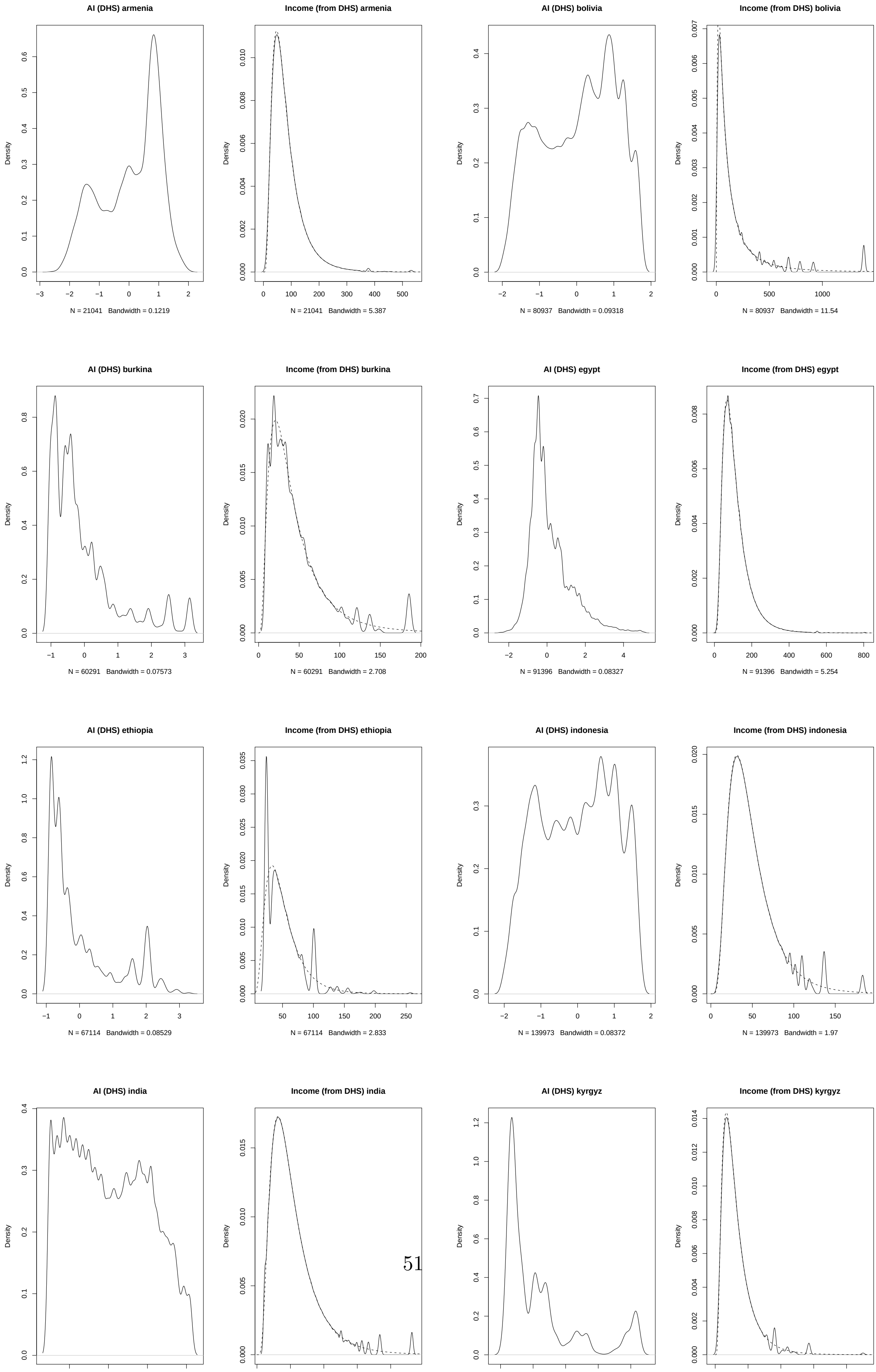
Table A.7: Inequality decomposition of the HDI components by subgroups

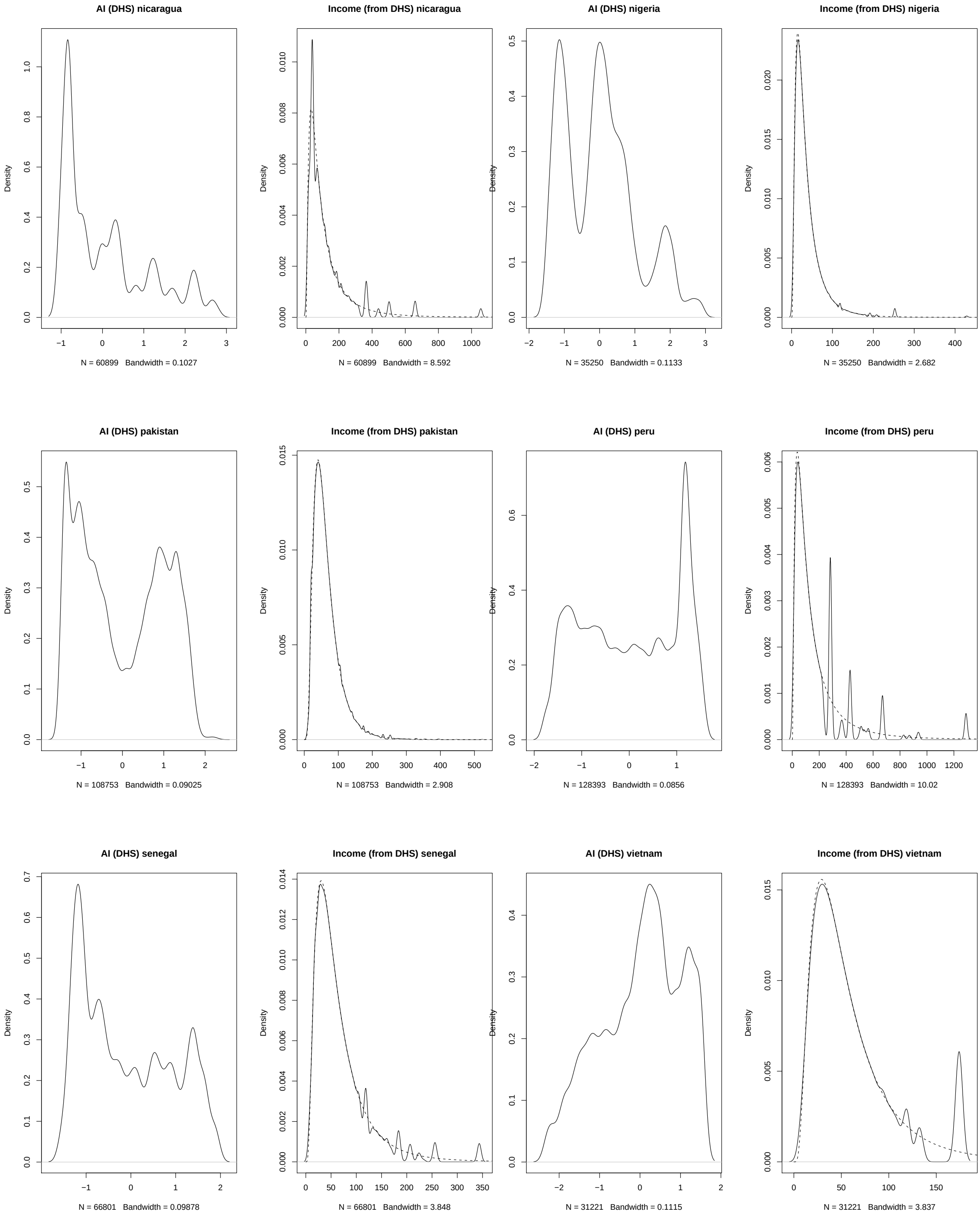
	Theil index of GNI index							Theil index of education index							Theil index of health index						
		By income deciles		By urban/ rural areas		By edu of hhhead			By income deciles		By urban/ rural areas		By edu of hhhead			By income deciles		By urban/ rural areas		By edu of hhhead	
	Tot.	With.	Betw.	With.	Betw.	With.	Betw.	Tot.	With.	Betw.	With.	Betw.	With.	Betw.	Tot.	With.	Betw.	With.	Betw.	With.	Betw.
Armenia	0.024	0.002	0.022	0.013	0.011	0.021	0.003	0.380	0.379	0.001	0.380	0.000	0.379	0.001	0.026	0.026	0.000	0.026	0.000	0.026	0.000
Bolivia	0.104	0.004	0.101	0.077	0.028	0.092	0.013	0.731	0.425	0.306	0.493	0.238	0.413	0.316	0.135	0.133	0.003	0.133	0.002	0.134	0.001
Burkina F.	0.141	0.018	0.123	0.141	0.132	0.117	0.023	0.726	0.688	0.038	0.680	0.046	0.695	0.029	0.099	0.091	0.008	0.095	0.005	0.095	0.003
Egypt	0.014	0.000	0.013	0.009	0.005	0.011	0.002	0.836	0.756	0.080	0.833	0.003	0.791	0.044	0.036	0.032	0.004	0.034	0.002	0.035	0.002
Ehtiopia	0.060	0.003	0.057	0.045	0.014	0.052	0.007	0.873	0.764	0.109	0.794	0.079	0.466	0.407	0.143	0.140	0.003	0.140	0.003	0.142	0.001
India	0.037	0.002	0.035	0.028	0.009	0.031	0.005	0.679	0.628	0.051	0.671	0.008	0.517	0.162	0.050	0.049	0.000	0.049	0.000	0.049	0.000
Indonesia	0.019	0.001	0.018	0.015	0.004	0.016	0.003	0.376	0.366	0.010	0.372	0.004	0.353	0.023	0.051	0.051	0.001	0.051	0.000	0.051	0.000
Kyrgyz R.	0.055	0.006	0.049	0.043	0.012	0.052	0.003	0.347	0.346	0.002	0.347	0.000	0.344	0.002	0.057	0.055	0.002	0.056	0.001	0.056	0.001
Nicaragua	0.097	0.002	0.095	0.070	0.027	0.080	0.017	0.502	0.443	0.060	0.465	0.038	0.418	0.082	0.064	0.061	0.002	0.062	0.002	0.061	0.002
Nigeria	0.125	0.032	0.093	0.107	0.019	0.108	0.017	0.732	0.643	0.090	0.719	0.014	0.489	0.242	0.189	0.180	0.009	0.186	0.003	0.186	0.003
Pakistan	0.025	0.001	0.024	0.018	0.007	0.022	0.004	0.695	0.625	0.070	0.680	0.014	0.562	0.132	0.099	0.097	0.002	0.098	0.001	0.098	0.001
Peru	0.055	0.005	0.050	0.034	0.020	0.045	0.010	0.256	0.247	0.009	0.250	0.006	0.239	0.017	0.067	0.062	0.004	0.063	0.003	0.064	0.003
Senegal	0.073	0.007	0.066	0.052	0.021	0.066	0.007	0.831	0.679	0.152	0.695	0.136	0.669	0.162	0.118	0.111	0.007	0.112	0.007	0.118	0.002
Vietnam	0.044	0.004	0.040	0.036	0.008	0.038	0.006	0.323	0.310	0.013	0.322	0.000	0.296	0.027	0.030	0.029	0.001	0.029	0.000	0.029	0.000
Zambia	0.168	0.002	0.166	0.123	0.045	0.138	0.030	0.508	0.473	0.035	0.496	0.013	0.456	0.052	0.115	0.109	0.007	0.111	0.004	0.113	0.003

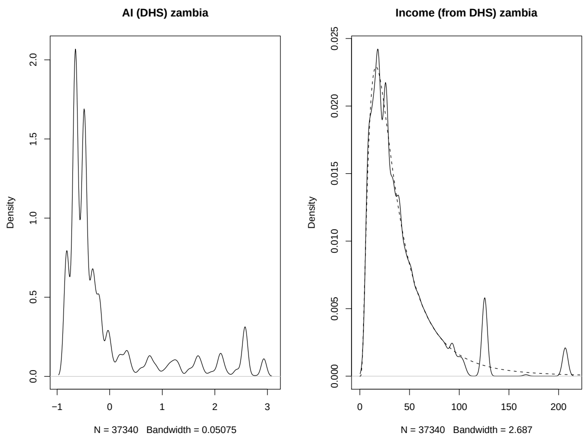
Note: HDI is based on multiple imputation for enrolment rates to calculate the expected years of education and on the regression-based approach for under five mortality to calculate the life expectancy.

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Figure A.1: Distributions of the asst index and income




AI (DHS) nicaragua
N = 60899 Bandwidth = 0.1027
Income (from DHS) nicaragua
N = 60899 Bandwidth = 8.592
AI (DHS) nigeria
N = 35250 Bandwidth = 0.1133
Income (from DHS) nigeria
N = 35250 Bandwidth = 2.682
AI (DHS) pakistan
N = 108753 Bandwidth = 0.09025
Income (from DHS) pakistan
N = 108753 Bandwidth = 2.908
AI (DHS) peru
N = 128393 Bandwidth = 0.0856
Income (from DHS) peru
N = 128393 Bandwidth = 10.02
AI (DHS) senegal
N = 66801 Bandwidth = 0.09878
Income (from DHS) senegal
N = 66801 Bandwidth = 3.848
AI (DHS) vietnam
N = 31221 Bandwidth = 0.1115
Income (from DHS) vietnam
N = 31221 Bandwidth = 3.837
Density


AI (DHS) zambia
N = 37340 Bandwidth = 0.05075
Income (from DHS) zambia
N = 37340 Bandwidth = 2.687
Density