UNDP

Human Development Report **2006**

Human Development Report Office
OCCASIONAL PAPER

A Human Development Index by Income Groups

Michael Grimm, Kenneth Harttgen, Stephan Klasen, and Mark Misselhorn

2006/7

A Human Development Index
by Income Groups

Michael Grimm, Kenneth Harttgen, Stephan Klasen
and Mark Misselhorn*

University of Göttingen, Department of Economics

July 25, 2006

Abstract

One of the most frequent critiques of the HDI is that is does not take into account inequality within countries in its three dimensions. We suggest a relatively easy and intuitive approach which allows to compute the three components and the overall HDI for quintiles of the income distribution. This allows to compare the level in human development of the poor with the level of the non-poor within countries, but also across countries. An empirical illustration for a sample of 13 low and middle income countries and 2 industrialized countries shows that inequality in human development within countries is indeed high. The results also show that the level of inequality is only weakly correlated with the level of human development itself.

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

1 Introduction

The Human Development Index (HDI) is a composite index that measures the average achievement in a country in three basic dimensions of human development: a long and healthy life, as measured by life expectancy at birth; knowledge, as measured by the adult literacy rate and the combined gross enrollment ratio for primary, secondary and tertiary schools; and a decent standard of living, as measured by GDP per capita in purchasing power parity US dollars (UNDP, 2005). Based on available statistics UNDP was able to provide an HDI for 175 countries in the latest Human Development Report. The HDI is today widely used in academia, the media and in policy circles to measure and compare progress in human development between countries and over time.

Despite its popularity, which is among other things due to its transparency and simplicity, the HDI is criticized for several reasons.[1] First, it neglects several other dimensions of human well-being, such as for example human rights, security and political participation (see e.g. Anand and Sen (1992), Ranis, Stewart and Samman (2006)). Second, it implies substitution possibilities between the three dimension indices, e.g. a decline in life expectancy can be off set by a rise in GDP per capita.[2] Related to that critique is the third point, which charges that the HDI uses an arbitrary weighting scheme (see e.g. Kelley (1991), Srinivasan (1994) and Ravallion (1997)). For instance, why should education be worth as much as income or health? Finally and fourth, the HDI is often criticized because it only looks at average achievements and, thus, does not take into account the distribution of human development within a country (see e.g. Sagar and Najam (1998)). The implications of this problem are nicely illustrated by the UNDP itself in last year's Human Development Report (UNDP, 2005).

UNDP computed for a set of countries an alternative HDI, where the income dimension does not measure the national average income, but the average income of the poorest 20%. The two other dimensions indices where for simplicity left unchanged. However, only that change led for some highly unequal countries to a sharp drop in the HDI rank. For instance, Brazil is ranked at the 63rd position using the global HDI and only at the 115th position, when the income of the poorest quintile is used. In fact the situation of the poorest 20% in Brazil is comparable to that in countries such as Guatemala, Honduras and Mongolia (UNDP, 2005). Among the other countries with large differences in HDI ranks for the poorest are, as one can expect, many Latin-American countries, for instance Mexico, Chile and Argentina. But even for rather equal and highly developed countries such

---

[1]For a critical review, see e.g. Sagar and Najam (1998).

[2]Moreover, if poor people face higher mortality, their deaths would increase per capita incomes of the survivors, generating a further distortion, particularly in HDI trends over time.

as Sweden, UNDP found, that the poorest 20% would rank at position 25, compared with position 6 when average incomes are used. That exercise is illustrative, but did not suggest any adjustment for health and education, and therefore underestimated the impact of inequality on the HDI for the poorest income quintile as the people in that quintile are likely to suffer from lower life expectancy and poorer educational opportunities as well.

When constructing distribution-sensitive measures of human development, limited data availability on the distribution of human development achievements seriously constrains analyses. Household income surveys are today widely undertaken and, hence provide data on income distribution, but it is much more difficult to get data on inequality in life-expectancy, educational achievements and literacy. Inequality in these dimensions seems, at least in developing countries, also to be very high. There is also broad empirical evidence that mortality as well as educational attainment vary with income and wealth in both rich and poor countries (see e.g. Cutler, Deaton and Lleras-Muney (2005) and Filmer and Pritchett (1999)).

In the past several attempts have been made to integrate inequality into the human development index. 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.[3] One reason might be that it is not easy to compute the Gini coefficient for education and life-expectancy due to data limitations and conceptual problems. Another reason might be that it is not clear how to interpret the interaction between the Gini coefficient and the average achievement in a component.

The gender related development index, or GDI, was another attempt in that direction. Its motivation was the 1995 Human Development Report's emphasis on gender inequalities. The GDI adjusts the HDI downward by existing gender inequalities in life-expectancy, education and incomes. The GDI calculates each dimension index separately for men and women and then combines both by taking the harmonic mean, penalizing differences in achievement between men and women. The overall GDI is then calculated by combining the three gender-adjusted dimension indices by taking the arithmetic mean. This concept could of course also be applied using other segmentation variables than gender, such as different ethnic or income groups. However for gender in particular, it is not clear how gender related inequality in income can reasonably be measured.[4] In most cases men and

---

[3]See Grün and Klasen (2006) for an analysis of a Gini-adjusted GDP measure.

[4]Generally, the GDI uses information on earned income of males and females, based on sex-specific labor force participation rates and earnings differentials (UNDP, 2005).

women pool incomes in households. Usually not much information is available how the pooled income is then allocated among household members. That and other critical issues related to the GDI are discussed in detail by Klasen (2006a, 2006b).

Another attempt was undertaken by Foster, López-Calva and Székely (2003). They chose an axiomatic approach to derive a distribution sensitive HDI. They suggest a three-step procedure. First, each dimension index is calculated on the lowest possible aggregation level, given the data availability. For instance, income at the level of households and life-expectancy at the level of municipalities (taken from census data). Second, for each dimension an overall index is computed by taking the generalized mean μ_q. The formula for the generalized mean is $\mu_q = [(x_1^q + \ldots + x_n^q)/n]^{1/q}$. For $q = 1$ μ yields the arithmetic mean, but for negative values for q, μ gives more emphasis on lower levels of x. The higher the absolute value of q, the more weight is given to low levels of x. Third, the overall HDI is computed by taking again the generalized mean instead of the simple arithmetic mean. The advantage of this approach is its axiomatic foundation. For instance, the index is decomposable by sub-groups, which is not the case for the Gini-corrected HDI. The problem with this approach is, however, that the generalized mean may not seem very intuitive for many users of the HDI. It obviously also raises the question of how to determine the 'right' inequality aversion parameter q. An additional problem is, that again no generally applicable methodology is suggested, which could help to compute the three dimension indices on the lowest disaggregation level.

The approach chosen in this paper differs from the others in that, first, we focus of inequality in human development across the income distribution and, second, we do not try to incorporate the aggregate well-being costs associated with existing inequalities, but rather generate a separate HDI for different segments of the income distribution. More precisely, we take household income and demographic data to compute the three dimension indices for quintiles of the income distribution. This allows on the one hand to track the progress in human development separately for 'the poor' and 'non-poor' and on the other hand to compare the level of human development of the poor to the level of the average population and the level of the non-poor. In contrast to previous attempts, we also present, at least for developing countries, a clear methodology how the three dimension indices for different segments of the income distribution can be calculated with commonly available data sources. Applying our methodology to developed countries entails some data availability and comparability problems which we discuss below. Due to these problems, we are only able to provide rough estimates for two developed countries.

The objective of this paper is first of all illustrative. We will show that our methodology has also some shortcomings, and, hence, all presented re-

sults should be interpreted with caution and in the light of our assumptions. The reminder of this paper is organized as follows. Section 2 presents our methodology. Section 3 presents the sample of countries for which we illustrate it. Section 4 discusses the results. Section 5 shows how our approach could be applied to industrialized countries. Section 6 offers a critical assessment of our methodology. Section 7 concludes.

2 Methodology

2.1 General idea and overview

The basic idea of our method is to use disaggregated data to calculate the three dimension indices which constitute the HDI for different segments of the income distribution. This will allow us to get an idea of the heterogeneity and inequality in human development which exists within a country. As data sources, we use household surveys. As segments of the income distribution, we define income quintiles.

Since the early nineties, two types of surveys are undertaken in almost all developing countries. First, there are so-called *Living Standard Measurement Surveys* (LSMS) or a shorter version of it called *Priority Surveys* (PS). Even in countries were none of these two surveys are available, there exist normally at least some other type of living standard survey. These surveys provide, apart from information on household and individual characteristics, data on educational achievement, school enrollment and household income or household expenditure. In what follows, we call this type of survey simply 'household income survey' or 'HIS'. Second, there are so called 'Demographic and Health Surveys' or 'DHS' in short. These surveys are undertaken by the *Macro International Inc., Calverton, Maryland* (usually in cooperation with local authorities and funded by USAID) and provide among other things detailed information on child mortality, health, and fertility. How to proceed for industrialized countries, where usually other types of surveys are undertaken, will be discussed later.

Hence, we will use the HIS to calculate the quintile specific education and GDP indices and the DHS to calculate the quintile specific life expectancy index. The main problems in proceeding so, are that both surveys do not interview the same households (or if so, these households can at least not be matched) and that the DHS does not contain any information on household income or household expenditure, i.e. it is not possible to sort directly the DHS households and individuals by income quintiles.

To solve this problem, we use a simple variant of so-called data matching techniques. The principle of this technique is to estimate the correlation between income and a set of household characteristics which are available in the HIS and the DHS and then to use this correlation pattern to predict income for the households covered by the DHS. This will allow to calculate

5

the life expectancy index also by income quintiles, even if income is not directly available in the DHS. To ensure that the data matching method gives acceptable results, both surveys should of course have been undertaken within a relative narrow time period and should be representative for the same (at best the total) population.

However, given that the quality of such a matching process depends heavily on the data quality and data consistency of both types of surveys, we present a second and alternative approach where we use a so-called 'asset index' as segmentation variable. This measure is often used to get an idea of the living standard of households interviewed in the DHS. In the DHS, households are asked regarding the ownership of various consumer durables, such as telephone, radio, television, fridge and about housing characteristics and educational endowments. This information can then be aggregated into one single metric index and can be used as a proxy for income (see below).

Once the three dimension indices are calculated, we simply calculate the quintile specific HDI, which we name QHDI, by taking the arithmetic average of the three dimension indices. In what follows, each step of our method is explained in detail and illustrated by an example.

2.2 Imputing income for DHS households

2.2.1 A regression based approach

To impute incomes for the households covered by the DHS, we use techniques similar to those used in the poverty mapping literature (see e.g. Elbers, Lanjouw and Lanjouw (2003)).[5]

The HIS provides information about household income and/or household expenditure. If income is used, the aggregate should contain earned (e.g. wages and profits) as well as unearned income (e.g. transfers). If expenditures are used, the aggregate should contain the expenditure for all items purchased plus the value of the self-produced consumption. According to usual practice in poverty analysis (see e.g. Deaton and Zaidi (2002) expenditures on durables should be excluded. For house owners, hypothetical rents should be imputed. Regional variations in the cost of living should be eliminated using appropriate price deflators. Once the welfare aggregate is calculated, we divide it by household size to receive a per capita measure. We do not use any particular equivalence scale to ensure consistency with the general HDI, which also uses a per capita measure for the income index. In what follows, our per capita welfare measure is denoted y_h, where the index h stands for households $h = 1, \ldots, K$.

Once, y_h is calculated, the common set of variables Ω_h in the HIS and the DHS has to be identified. The variables of Ω have to be correlated with

---

[5]Grosse, Klasen and Spatz (2005) recently also used such a technique to match HIS and DHS data for Bolivia.

y_h and should at least contain (i) some characteristics of the household head such as age and educational achievement, (ii) characteristics of the household like the number of children, the number of male and female adults in working age and regional variables (such as urban vs. rural, region or province of residence), and (iii) housing conditions like materials of the floor, the roof and the walls, type of electricity and water connection and possibly the number of rooms per person.

Once all these variables are calculated, y_h is regressed in logarithmic form on this set of variables using OLS estimators:

$$\ln y_h^{HIS} = \beta^{HIS} \Omega^{HIS} + u_h, \tag{1}$$

where β^{HIS} is a vector of parameters and u_h is the residual.

Using the vector of estimated parameters $\hat{\beta}^{HIS}$, hypothetical incomes for the households covered by the DHS can be calculated by:

$$\hat{\ln} y_h^{DHS} = \hat{\beta}^{HIS} \Omega^{DHS}. \tag{2}$$

Given that regressions as in Equation (1) rarely explain more than half or three quarter of the total variance in $\ln y_h$, one could generate residuals to account for the unobserved determinants of y_h. We think that would be important, when the objective was to calculate any inequality measure. However, given our objective, we think it is sufficient to assume that the included variables contain enough information on the true income quintile and that in contrast hypothetical residuals may well preserve the natural variance in the data, but at the price of a higher probability of missclasifications over income quintiles.[6] One may also argue that drawing residuals would help to prevent ties, i.e. that households with an identical set Ω will have the same imputed income. However, if Ω is large enough and contains also continuous variables that problem will not arise.

Once the hypothetical incomes for the DHS are imputed, it is possible for both surveys to calculate the cumulative distribution functions of income (person weighted) $F(\ln y_h^{HIS})$ and $F(\hat{\ln} y_h^{DHS})$. Using these distributions it can be determined for each household in which income quintile ($Q = 1, 2, ..., 5$) it is situated.

However, what could pose a problem is, first, that household expenditure may in some cases not be a good proxy of permanent income due to measurement error and limited possibilities of households to smooth consumption, and, second that in some cases the comparability of the HIS and DHS is not high enough, and, hence predicted incomes in a DHS give a biased impression of the distribution of income. Therefore, we present, as

---

[6]Moreover, when imputing residuals for the DHS households, one would in addition have to take into account that the HIS and DHS have generally different sample sizes and a different regional stratification. Hence, the unobserved determinants of y_h will not be distributed identically (see Elbers, Lanjouw and Lanjouw, 2002).

mentioned above, a second alternative to classify households in the DHS by income quintile which is based on an asset index approach.

Example: Imputing income for DHS households in Indonesia

First, we regress the logarithm of household expenditure per capita $(\ln y_h^{HIS})$ on the set of variables Ω_h.

Table 1
Correlates of the logarithm of household per capita expenditure
in Indonesia (2000) $(\hat{\beta}^{HIS})$

	Coeff.	S.E.
Age of head	-0.019***	0.008
Age of head^2/100	0.035***	0.017
Age of head^3/1000	-0.002***	0.001
Urban (=1)	0.050***	0.014
Head women (=1)	-0.039***	0.017
Household size	-0.021***	0.003
Number of children ($age \leq 1$)	-0.075***	0.019
Number of children ($1 < age \leq 5$)	-0.115***	0.013
Number of children ($5 < age \leq 15$)	-0.088***	0.007
Head has no education	-0.381***	0.023
Head has primary education	-0.412***	0.019
Head has secondary education	-0.180***	0.020
Head has tertiary education	(Ref.)	
Possesses TV (=1)	0.183***	0.017
Possesses appliances (=1)	0.071***	0.018
Access to piped drink. water (=1)	0.084***	0.015
Having flush toilet (=1)	0.205***	0.015
Having electricity (=1)	-0.027	0.023
Inferior floor material (=1)	-0.108***	0.022
Inferior wall material (=1)	-0.117***	0.035
Cons	15.366***	0.117
R^2	0.29	
n	9921	

Notes: $*$ significant at the ten percent level. $**$ significant at the five percent level. $***$ significant at the one percent level.
Source: Third wave of the Indonesian Family Life Survey (IFLS) 2000; estimations by the authors.

The regression results in Table 1 show that household expenditures per capita are in average higher, as one can expect, in households living in urban areas, where the head is a male and is well educated and where are fewer children in the household. Moreover, the possession of durables such as TV and better housing conditions are a good signal for higher household per capita expenditures. The set of variables Ω explains roughly 30 percent of the total variance of log household expenditures.

Second, we use the coefficients of Table 1 to impute for each household in the DHS hypothetical expenditures $\hat{\ln y}_h^{DHS}$ (hypothetical income in what

8

follows). For instance for household $n = 123$ (where the household head is 45 years old, ..., the house is built with inferior wall material), we calculate:

$$(-0.019) \times 45 + (0.035) \times (45^2)/100 + ... + (-0.117) \times 1 + 15.366 = 13.802.$$

Third, we calculate for the HIS and the DHS, the person weighted cumulative income distribution. Table 2 shows the quintile specific average incomes per capita in USD PPP of 2000 (base year 1996, further details, see below). The data matching approach works quite well. The quintile specific incomes imputed in the DHS are very close to the incomes calculated in the HIS. Only in the fifth quintile we seriously underestimate the average income. This stems from the fact that extreme high incomes in the HIS which influence strongly the average income in the fifth quintile are not reproduced by the observable determinants in our regression approach. This entails however no serious problem, given that we need for the computation of the life expectancy index only the classification of households into quintiles and not the exact quintile specific average incomes.

Table 2
Quintile specific average household incomes per capita
for Indonesia (2000) (PPP)

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
HIS	617.11	1042.28	1478.91	2180.85	5048.03	1918.81
DHS	796.86	1144.49	1433.01	1802.51	2841.08	1531.54

Source: Third wave of the Indonesian Family Life Survey (IFLS) 2000, Demographic and Health Survey (DHS) 2003; estimations by the authors.

2.2.2 An asset index based approach

In order to construct an asset index for DHS households, first, a set of household assets has to be identified. We used the ownership of a radio, TV, refrigerator, bicycle, motorized vehicle, floor material of housing, type of toilet, type of water source and some other assets depending on the country. Second, these assets have to be aggregated into one single metric index for each household using principal component analysis, or, alternatively, the closely related factor analysis (see Filmer and Pritchett (2001) and Sahn and Stifel (2000, 2003)). We used principal component analysis. Once the asset index is built, one can construct, similar to the regression-approach, the cumulative distribution function of the asset index and, hence, households in the DHS can be classified into asset quintiles. Under the assumption that the ownership of assets is a good proxy for income, it can be assumed that the asset quintiles yield a consistent classification to that obtained via observed income in the HIS. Hence, in that case matching between both surveys using these quintiles is possible.

We will use both approaches, the regression based approach and the asset index approach. In principle, the regression based approach is to be favored as income is one of the three components of the HDI and therefore it is consistent to use that approach. Moreover the asset index is sometimes biased, because it reflects not correctly differences in income between rural and urban areas, due to usually huge differences in prices and the supply of such assets as well as differences in preferences for assets between both areas. On the other hand, the income regression approach yields biased results whenever the distribution of explanatory variables in the regression is not consistent in the HIS and DHS, due either to measurement error or due to different definitions of the variables used in both surveys. As will be shown below, we suspect such a problem to exist particularly in some very poor African countries, and hence in this case it might be that the asset index is a better predictor of true income in these circumstances than predicted income using the estimated regression.

2.3 Calculating the life expectancy index by income quintiles

To calculate a life expectancy index by income quintile we combine information on child mortality with model life tables. As mentioned above, the HIS provides usually no information on mortality. The DHS provides only information on child mortality, but not on mortality by all age groups, which would be necessary to construct a life table and to calculate life expectancy directly.

In a first step, we calculate under one child mortality rates by income quintile. To do this we use the information on all children born in the five years preceding the survey. For each child i we calculate the survival time S_i expressed in months m and the survival status d_i. The status variable takes the value one if the child died at the end of S_i and the value zero, if the child was still alive at the age of one. Then we use a simple non-parametric life table estimator to estimate the survival probability for each months after birth, $p_{m,m+1}$. Through cumulative multiplication we derive for each income quintile the mortality rate q_1:

$$q_1^Q = 1 - \prod_{m=1}^{12} p_m^Q, \tag{3}$$

We also estimate q_1 over the whole sample, to be able to construct the aggregate life expectancy index.

In a next step, we use the estimated mortality rate q_1 and Ledermann model life tables to calculate quintile specific life expectancy. Ledermann (1969) used historical mortality data for many countries and periods to estimate the relationship between life-expectancy and age-specific mortality rates. He found the following relationship (note that the log function uses

10

the basis 10):

$$\log \hat{q}_j = \hat{a}_{j,0} + \hat{a}_{j,1} \log(100 - e_0), \tag{4}$$

where $\hat{q}_j$ is the predicted mortality rate for the age group j, e_0 is the life expectancy at birth and $\hat{a}_{j,0}$ and $\hat{a}_{j,1}$ are the estimated regression coefficients by Ledermann. Ledermann considered age groups defined over five-year intervals, except for the first age group, which he divided into children 0 to 1 year old and 1 to 5 years old. In principle, we could also use the Princeton model life tables (Coale and Demeny, 1983), but the problem with those tables is, that first they use not e_0 but e_{10} as entry, i.e. life expectancy at the age of 10. Obviously, it is easier to estimate e_{10} given the probably higher measurement error in child mortality, but to construct the $QHDI$ we need e_0 not e_{10}. Second, Princeton tables end already at a life expectancy of 75 years. Third, Princeton tables are defined separately for men and women, and, hence we would need to estimate child mortality rates separately for boys and girls. This would reduce the number of death events in each income quintile to extremely low levels and therefore lead to very unstable life expectancy estimates. We checked however, whether our life expectancy estimates were consistent with those one would obtain using the Princeton Life Tables 'West'. That was the case, and, hence, we are confident that our Lederman approach yields acceptable results. However, a drawback of both types of tables—Ledermann and Princeton—is that their estimation included almost no countries of today's developing world and no countries affected by the AIDS epidemic. In particular the latter omission might be problematic, given that AIDS usually affects strongly the age-mortality pattern by increasing mortality among children below the age of 5 (through mother-child transmission) and mortality among adults in age of activity.

To calculate quintile specific life expectancy, we take the inverse of Equation (4) and the regression coefficients for the age group 1 year old:

$$\hat{e}_0^Q = 100 - \left[\frac{q_1^Q}{10^{\hat{a}_{1,0}}} \right]^{\frac{1}{\hat{a}_{1,1}}} \quad \forall \, Q = 1, 2, \dots, 5. \tag{5}$$

with $\hat{a}_{1,0} = -1.98384$ and $\hat{a}_{1,1} = 2.40372$ (Ledermann, 1969).

Aggregate life expectancy can be calculated using q_1 instead of q_1^Q.

Then we calculate the quintile specific life expectancy index, L^Q, using the usual minimum and maximum values for life expectancy employed to calculate the HDI:

$$L^Q = \frac{\hat{e}_0^Q - 25}{85 - 25} \quad \forall \, Q = 1, 2, \dots, 5. \tag{6}$$

The aggregate life expectancy index L can be calculated using $\hat{e}_0$ instead of $\hat{e}_0^Q$.

In a last step, we rescale linearly L^Q and L to achieve consistency with the aggregate HDI calculated by UNDP. As rescaling factor we use the ratio between our aggregate life expectancy index L and the aggregate life expectancy index calculated by UNDP for the particular year in question.[7] Consistency is not automatic, given that our approach and UNDP's approach are based on different data sources. Given that the objective of our approach is first of all to examine the *distribution* of human development, differences in levels should not present any serious problem.

Example: Computing the quintile specific life expectancy index for Indonesia

First, we estimate the mortality rate q_1 (see Table 3). Figure 1 shows in addition the survival curves (S) for the five income quintiles. One can clearly see that in Indonesia the survival rate is positively correlated with the income quintile: Whereas in the highest income quintile only 2 percent of all children were dead after 12 months, in the lowest income quintile almost 5 percent were dead after 12 months.[8]

Figure 1

Survival curve of children five years after birth for each income quintile ($\hat{S}_i^Q$) for Indonesia (2003).
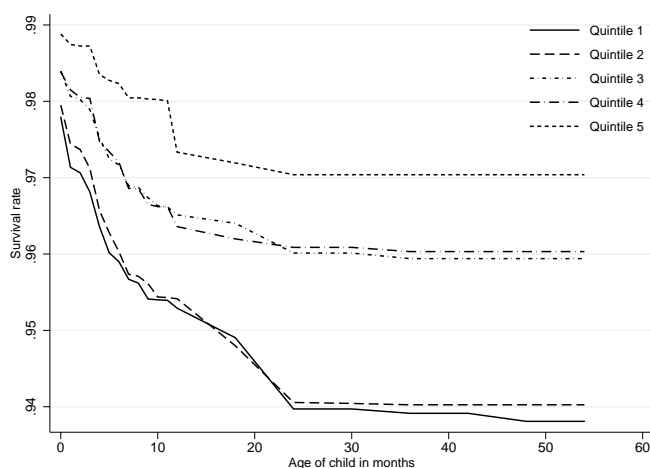
Source: Demographic and Health Survey (DHS) 2003; estimations by the authors.

---

[7]If the DHS and HIS are from different years, we re-scale to the later year.

[8]In the case of Indonesia the survival curves are indeed very similar for the first and second and the third and fourth quintile. This is however not a general result, which would apply to the other countries in our sample as well.

Second, we calculate using equation (5) and (6) the quintile specific life expectancies and life expectancy indices. The results are shown in Table 3. As one can expect, given the quintile-specific pattern of the survival curves, life expectancy and the life expectancy index increase with the income quintile. Table 3 also shows the life expectancy index by asset quintile, which will be used later alternatively. However, in the case of Indonesia, L^{QI} and L^{QA} differ not much except in the fourth quintile.

Table 3
Quintile specific child mortality rates,
life expectancies, and life expectancy indices
for Indonesia (2003)

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
q_1^Q (per 1000)	43.5	39.6	36.7	32.5	28.9	36.5
$\hat{e}_0^{QI}$ (income regression)	67.1	67.2	71.1	71.1	76.8	70.1
L^{QI} (income regression)	0.65	0.65	0.71	0.71	0.80	0.70
$\hat{e}_0^{QA}$ (asset index)	64.9	65.9	69.0	74.5	82.6	70.1
L^{QA} (asset index)	0.62	0.63	0.68	0.76	0.89	0.70

Notes: QI stands for life expectancy by income quintile. QA stands for life expectancy by asset quintile.

Source: Demographic and Health Survey (DHS) 2003; estimations by the authors.

2.4 Calculating the education index by income quintiles

To calculate the quintile specific education index, we use the information on literacy and school enrollment provided by the HIS.

2.4.1 Calculating the adult literacy index

The questions providing information about adult literacy may significantly vary from one HIS to the other. Sometimes adults are simply asked whether they are able to read and write. Other surveys are much more specific in asking whether the person is able to read a newspaper and to write a letter. This is even sometimes directly tested. In addition, in some countries one has to distinguish between having knowledge of any local language or of the official language of the country. Finally in some surveys, such information is completely missing. In this latter case, it is possible to use educational achievement as proxy for literacy. However, it is far from evident to determine after how many years of school a person is literate. This varies a lot from country to country or even within a country (for West-Africa, see e.g. Michaelowa (2001)). We proceeded as follows. If an adult declared to be able to read and/or write in any language (with or without proof), we considered him or her as literate. If that information was not available, we considered somebody as literate if he or she achieved at least a grade which

corresponds to five years of schooling. Adults are defined as persons above the age of 15.

Quintile specific adult literacy is then calculated by the following equation:

$$a^Q = \frac{1}{n^Q} \sum_{i(\forall j > 15)} I(a_i^Q > \bar{a}) \quad \forall Q = 1, 2, \dots, 5, \tag{7}$$

where n^Q is the total number of adults in quintile Q and I is an indicator function which takes the value one if literacy status of adult i, a_i is over the above defined threshold value $\bar{a}$ and zero otherwise. We calculate also the aggregate adult literacy rate a.

Then we calculate the quintile specific adult literacy index, A^Q, using again the corresponding usual minimum and maximum values employed in the HDI:

$$A^Q = \frac{a^Q - 0}{1 - 0} \quad \forall Q = 1, 2, \dots, 5. \tag{8}$$

The aggregate adult literacy index A can be calculated using a instead of a^Q.

In a last step, we rescale again linearly A^Q and A to achieve consistency with the aggregate HDI calculated by UNDP for the relevant year. As rescaling factor we use the ratio between our aggregate literacy index A and the aggregate literacy index calculated by UNDP.

2.4.2 Calculating the enrollment index

To calculate the quintile specific gross enrolment index, we calculate first the combined gross enrolment rate for each quintile. Each individual attending school or university whether general or vocational is considered as enrolled. We define this rate over all individuals of the age group 5 to 23 years old. Age is for each individual the age at the date of the interview. This yields:

$$g^Q = \frac{1}{n^Q} \sum_{i(\forall 5 \leq j \leq 23)} I(g_i^Q > 0) \quad \forall Q = 1, 2, \dots, 5, \tag{9}$$

where n^Q is the total number of individuals of age 5 to 23 in quintile Q and I is an indicator function which takes the value one if an individual i independent of age, is enrolled, i.e. $g_i > 0$. We calculate also the aggregate gross enrolment rate g.

Then we calculate the quintile specific gross enrollment index, G^Q using the usual minimum and maximum values used for the calculation of the HDI:

$$G^Q = \frac{g^Q - 0}{1 - 0} \quad \forall Q = 1, 2, \dots, 5. \tag{10}$$

The aggregate gross enrollment index G can be calculated by using g instead of g^Q. Finally, we also rescale G^Q and G to the level of the HDI enrollment index.

2.4.3 Calculating the education index

The quintile specific education index E^Q is calculated using the same weighted average as the HDI:

$$E^Q = (2/3) \times A^Q + (1/3) \times G^Q \quad \forall\, Q = 1, 2, \dots, 5. \tag{11}$$

The aggregate education index E can be calculated by using A and G instead of A^Q and G^Q.

Example: Computing the quintile specific education index for Indonesia

Table 4 shows the quintile specific literacy and enrollment rates as well as the corresponding indices and overall education index or Indonesia. Again, literacy and enrollment are clearly positively correlated with income. Whereas the education index is 0.73 in the poorest quintile, it is 0.90 in the richest quintile.

Table 4
Quintile specific literacy and enrollment rates,
and education indices
for Indonesia (2000)

	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	All
a^Q	0.76	0.82	0.85	0.89	0.94	0.85
A^Q	0.79	0.86	0.89	0.93	0.98	0.88
g^Q	0.55	0.58	0.62	0.64	0.67	0.60
G^Q	0.61	0.65	0.69	0.71	0.74	0.67
E^Q	0.73	0.79	0.82	0.85	0.90	0.81

Source: Third wave of the Indonesian Family Life Survey (IFLS) 2000; estimations by the authors.

2.5 Calculating the GDP index by income quintiles

To calculate the GDP index by income quintile, we use our income variable from the HIS. One main difference with the two other dimension indices is that mean income calculated from the HIS can be *very* different from GDP per capita derived from National Accounts data, which is used for the GDP index in the general HDI. This has two reasons: first, because of conceptual differences and, second, because of measurement error on both levels. GDP

15

measures the value of all goods and services produced for the market within a year in a given country evaluated at market prices. Income in the household survey is either measured, as mentioned above, via household expenditure (including self-consumed production) or via the sum of earned and unearned household income. Therefore, non distributed profits of enterprises, property income and so on will not be included in the household income variable. Moreover, on the household survey side, there may be measurement errors, because it is difficult to get accurate responses from households concerning wages, profits and expenditures. On the National Accounts side, while supply-side information on output and income for some sectors is based on high-quality surveys or census data for agriculture and industry, information about subsistence farmers and informal producers is harder to obtain and usually of lower quality.[9]

We proceed as follows. First, to eliminate differences in national price levels 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):

$$y_h^{PPP} = y_h \times PPP. \tag{12}$$

Second, we rescale y_h^{PPP} using the ratio between $\bar{y}^{PPP}$ and GDP per capita expressed in PPP (taken from the general HDI):

$$ry_h^{PPP} = y_h^{PPP} \times \left[\frac{GDPPC^{PPP}}{\bar{y}^{PPP}} \right]. \tag{13}$$

Once, theses adjustments are done, it is straightforward to calculate the quintile specific GDP index, again using the usual minimum and maximum values of the HDI:

$$Y^Q = \frac{\log \bar{ry}^{Q,PPP} - \log(100)}{\log(40,000) - \log(100)} \quad \forall\, Q = 1, 2, \dots, 5, \tag{14}$$

where $\bar{ry}^{Q,PPP}$ is the quintile specific arithmetic mean of the rescaled household income per capita.

It should be noted that in richer countries the GDP per capita measure for the richest quintile, $\bar{ry}^{5,PPP}$, could easily exceed 40,000 USD PPP and, hence, the index could take a value greater than 1, and this could, in extreme cases, push the overall HDI for the richest quintile also above 1, which would cause problems for interpretation.[10]

---

[9]A detailed discussion of all these problems can be found in Ravallion (2001) and Deaton (2005).

[10]An obvious 'solution' to this problem could be to widen the income range for the HDI and the quintile-specific HDI.

Example: Computing the quintile specific GDP index for Indonesia

Table 5 shows that in Indonesia GDP per capita is 59 percent above the survey mean. The adjusted mean income (or GDP per capita) is roughly 980 USD PPP in the first quintile and 8000 USD PPP in the fifth quintile, i.e. more than double. Accordingly, the GDP index increases from 0.43 to 0.83 over the five quintiles.

Table 5
Quintile specific mean incomes
and GDP indices for Indonesia (2000)

	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	All
PPP				1521.37		
$GDPPC^{PPP}/\bar{y}^{PPP}$				1.59		
$\bar{ry}^{Q,PPP}$	980.80	1656.54	2350.50	3466.12	8023.06	3049.65
Y^Q	0.43	0.53	0.60	0.67	0.83	0.59

Source: Third wave of the Indonesian Family Life Survey (IFLS) 2000, Penn World Tables 6.1, World Development Indicators 2005; estimations by the authors.

2.6 Calculating the overall HDI and the HDI by income quintiles

Once the quintile specific dimension indices have been calculated, determining the QHDI is straightforward. It is the simple average of the three dimension indices:

$$HDI^Q = (1/3) \times L^Q + (1/3) \times E^Q + (1/3) \times Y^Q$$

$$\forall\, Q = 1, 2, \dots, 5. \tag{15}$$

The aggregate HDI is as usual given by:

$$HDI = (1/3) \times L + (1/3) \times E + (1/3) \times Y. \tag{16}$$

To get a sense of the inequality in human development within a country, one may compute the ratio between the HDI for the richest quintile and the poorest quintile:

$$RQHDI^{5,1} = \frac{HDI^{Q=5}}{HDI^{Q=1}}, \tag{17}$$

or the ratio of the quintile specific HDI to the aggregate HDI:

$$RQHDI^{1,mean} = \frac{HDI^{Q=1}}{HDI} \quad \text{and} \quad RQHDI^{5,mean} = \frac{HDI^{Q=5}}{HDI}. \tag{18}$$

All these indicators can of course also be calculated for each dimension index. Hence, the QHDI cannot only be used to inform about the level of human development of the poor, the rich and the groups in-between, but also about the inequality in human development within a society. Moreover, the quintile specific indices can be compared across countries. This may lead to results where country A has a higher overall HDI than country B, but that in country B human development of the poor is on a substantial higher level than in country A.

Example: Computing the quintile specific HDI for Indonesia

Table 6 shows for our example Indonesia again the quintile specific dimension indices and the resulting quintile specific HDI, QHDI. The QHDI is 0.59 in the poorest quintile, i.e. 14 percent under the HDI and 0.87 in the richest quintile, i.e. 24 percent above average. That is a huge differential in human development across the income distribution. Put differently, the ratio in human development between the very rich and the very poor is 1.48. Indonesia fares relatively better in life expectancy and education than in income. Inequality in the income index is also higher than in the two other dimension indices.

Table 6
Quintile specific HDI

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
L^{QI}	0.65	0.65	0.71	0.71	0.80	0.70
L^{QA}	0.62	0.63	0.68	0.76	0.89	0.70
E^Q	0.73	0.79	0.82	0.85	0.90	0.81
Y^Q	0.43	0.53	0.60	0.67	0.83	0.59
HDI^Q	0.59	0.65	0.70	0.76	0.87	0.70
$RQHDI^{5,1}$			1.48			
$RQHDI^{1,mean}$			0.84			
$RQHDI^{5,mean}$			1.24			

Source: Third wave of the Indonesian Family Life Survey (IFLS) 2000, Demographic and Health Survey (DHS) 2003; estimations by the authors.

3   Sample of developing countries

We illustrate our approach for a sample of 13 developing countries: 7 countries from Sub-Saharan Africa (Burkina Faso, Côte d'Ivoire, Cameroon, Guinea, Madagascar, Mozambique, South-Africa and Zambia), 3 countries from Latin America (Bolivia, Colombia, and Nicaragua), and 2 countries from Asia (Indonesia and Vietnam). These countries are listed in Table A1. We tried to restrict the sample to countries where a HIS and DHS were undertaken within a two-year time period. For 2 countries both surveys were

undertaken in the same year. For 3 countries there is a gap of one year and for 4 countries a gap of two years. Only in three countries (Guinea, Indonesia, and Madagascar) we were not able to follow this rule and have actually a gap between both surveys of 3 to 4 years. Moreover, we tried to include countries where both surveys are not older than 5 years. This was however not possible for 4 countries (Côte d'Ivoire, Guinea, Madagascar, South-Africa), where the HIS or the DHS (or both) were undertaken at the end of the 1990s. The survey dates should also be taken into account when comparing our unscaled QHDI with the usual HDI. The published HDI in the UNDP's Human Development Report 2005 (UNDP, 2005) refers to the year 2003. But a closer look at the data sources shows that literacy rates and life-expectancy estimates were usually based on censuses or surveys conducted between 2000 and 2004. In several countries the data sources even stem from data collected in the 1990s (e.g. Belarus, Burkina Faso, Kazakhstan, Mali). Hence, time consistency between the different dimension indices and actuality of the data is not a problem specific to our approach, but rather is present for both the usual HDI and the QHDI.

The life expectancy estimates are based throughout on Demographic and Health Surveys (DHS). The HIS data is a bit more heterogenous across countries but in most cases it stems from surveys which are in design similar to the World Bank's LSMS (at least regarding the variables we included).

4 Results

Table 7a presents the QHDI and the overall HDI for our sample of countries. The differences between the HDI of the poorest quintile and the HDI of the richest quintile are in all countries substantial. In each country the QHDI strongly increases with income. Tables 8a, 9 and 10 show that this is also the case for each dimension index, except for the life expectancy index. Education and GDP per capita, the latter one of course by construction, clearly increase in income. The correlation between the life expectancy index and income is in most cases also positive, but if inspected quintile by quintile one finds in some countries the opposite (e.g. Burkina Faso, Guinea, Mozambique), i.e. the index decreases with income. Several explanations might be invoked. First of all, given that we derived life expectancy from survey based estimates of child mortality, the potential measurement error is obviously high, due to in some cases rather small sample sizes and potentially very imprecise household's declarations regarding the death date of their children. These errors might themselves be correlated with income. The life table approach introduces an additional bias given that the used tables do not account for AIDS specific age-mortality patterns. Second, as already mentioned above, the suggested method to match data from the HIS and the DHS by income quintile might pose problems when the data quality is

limited. This is in particular the case in some of the African countries. For instance, when the set of common variables Ω is rather small or when the distribution of the variables included in Ω differs in both surveys. This may arise if the variable definitions are not exactly the same in both surveys.[11]

Given this difficulties, we also calculated life expectancy by asset quintiles (Table 8b) instead of income quintiles and recalculated the QHDI. The results are shown in Tables 7b. Interestingly, using the asset index one obtains for each country the expected positive correlation between life expectancy and the asset index. This suggests that for some of the countries in our sample the asset index may be a better proxy for income than predicted income itself. The asset index is also positively correlated with the education index (numbers not presented in table). The indicated inequality in education across asset quintiles is a bit higher than across income quintiles. This is not the case for the GDP per capita index. This is in fact what we expect; the asset index is not perfectly correlated with income and hence the income differential across asset quintiles is lower than across predicted income quintiles.

Table 7a
Quintile specific HDI by country
(L^Q computed using predicted income)

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
Country						
Burkina Faso (2003/2003)	0.270	0.306	0.334	0.362	0.469	0.348
Bolivia (2002/2003)	0.561	0.638	0.699	0.749	0.820	0.690
Côte d'Ivoire (1998/1999)	0.360	0.419	0.420	0.462	0.575	0.430
Cameroon (2001/2004)	0.431	0.479	0.519	0.551	0.627	0.523
Colombia (2003/2005)	0.682	0.737	0.790	0.833	0.931	0.790
Guinea (1995/1999)	0.379	0.468	0.504	0.559	0.627	0.467
Indonesia (2000/2003)	0.604	0.658	0.710	0.746	0.844	0.701
Madagascar (2001/1997)	0.370	0.437	0.484	0.569	0.687	0.488
Mozambique (2002/2003)	0.326	0.366	0.375	0.389	0.473	0.387
Nicaragua (2001/2001)	0.549	0.614	0.677	0.719	0.815	0.667
South Africa (2000/1998)	0.564	0.637	0.701	0.756	0.853	0.691
Vietnam (2004/2002)	0.637	0.665	0.726	0.757	0.818	0.713
Zambia (2002/2002)	0.327	0.385	0.429	0.481	0.571	0.426

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

---

[11]Or interviewers coded the answers not exactly identically, although the questions have been asked in exactly the same way.

Table 7b
Quintile specific HDI by country
(L^Q computed using asset index)

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
Country						
Burkina Faso (2003/2003)	0.257	0.306	0.331	0.365	0.489	0.348
Bolivia (2002/2003)	0.550	0.640	0.704	0.741	0.863	0.690
Côte d'Ivoire (1998/1999)	0.343	0.416	0.434	0.515	0.561	0.430
Cameroon (2001/2004)	0.417	0.477	0.529	0.553	0.644	0.523
Colombia (2003/2005)	0.673	0.741	0.800	0.857	0.927	0.790
Guinea (1995/1999)	0.340	0.457	0.490	0.594	0.696	0.467
Indonesia (2000/2003)	0.593	0.651	0.700	0.764	0.874	0.701
Madagascar (2001/1997)	0.343	0.463	0.496	0.563	0.684	0.488
Mozambique (2002/2003)	0.305	0.355	0.380	0.417	0.504	0.387
Nicaragua (2001/2001)	0.531	0.629	0.678	0.720	0.830	0.667
South Africa (2000/1998)	0.627	0.680	0.718	0.765	0.828	0.713
Vietnam (2004/2002)	0.561	0.640	0.700	0.743	0.879	0.691
Zambia (2002/2002)	0.317	0.390	0.431	0.476	0.583	0.426

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

Figure 2
Correlation between the overall HDI, the QHDI for the poorest quintile and the QHDI for the richest quintile
(L^Q computed using predicted income)
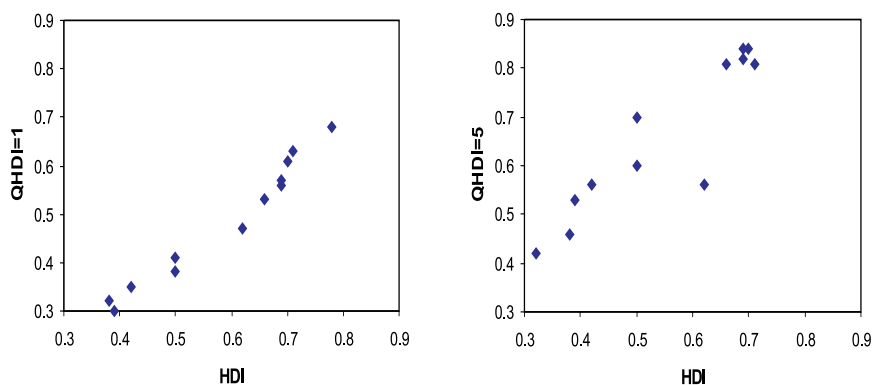
Source: Computations by the authors.

Figure 2 shows that the quintile specific HDI, here illustrated for the first and fifth quintile, is strongly correlated with the overall HDI. However, below we will show that this is not the case for inequality in human development within countries. Interestingly, the latter one seems rather unrelated to the level of human development.

Table 8a
Quintile specific life expectancy indices by country
(L^Q computed using predicted income)

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
Country						
Burkina Faso (2003/2003)	0.38	0.39	0.38	0.36	0.36	0.38
Bolivia (2002/2003)	0.63	0.62	0.67	0.69	0.68	0.65
Côte d'Ivoire (1998/1999)	0.37	0.39	0.30	0.31	0.43	0.36
Cameroon (2001/2004)	0.37	0.34	0.34	0.32	0.34	0.34
Colombia (2003/2005)	0.81	0.78	0.80	0.80	0.79	0.80
Guinea (1995/1999)	0.53	0.49	0.47	0.46	0.42	0.48
Indonesia (2000/2003)	0.65	0.65	0.71	0.71	0.80	0.70
Madagascar (2001/1997)	0.51	0.43	0.46	0.57	0.57	0.50
Mozambique (2002/2003)	0.33	0.31	0.26	0.23	0.22	0.28
Nicaragua (2001/2001)	0.73	0.70	0.75	0.76	0.78	0.74
South Africa (2000/1998)	0.42	0.47	0.53	0.54	0.52	0.48
Vietnam (2004/2002)	0.71	0.71	0.82	0.81	0.85	0.76
Zambia (2002/2002)	0.21	0.20	0.21	0.21	0.21	0.21

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

When the ranking of countries for the overall HDI is compared with the ranking of countries for the $QHDI^1$, one finds, for instance, that Mozambique fares as good as Zambia when the $QHDI^1$ ranking is used. Hence, despite a lower overall level of human development in Mozambique, the poor in Mozambique fare as well as the poor in Zambia (see Table 7a). Or, if Bolivia is compared to Indonesia one can state that both countries have a similar overall HDI of 0.70, but that in Indonesia the poorest quintile has an HDI of 0.60 and in Bolivia of only 0.56 (see Table 7a). In contrast, in Bolivia and Nicaragua the level *and* the distribution of human development is almost identical (see Table 7a). All these results also hold if the asset index instead of predicted income is used to compute quintile specific life expectancy (Table 7b). Such comparisons are one of the appealing features of the QHDI, i.e. that it allows to compare human development across countries for groups which share the same position in their country's income

Table 8b
Quintile specific life expectancy indices by country
(L^Q computed using asset index)

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
Country						
Burkina Faso (2003/2003)	0.35	0.39	0.37	0.37	0.42	0.38
Bolivia (2002/2003)	0.60	0.63	0.68	0.67	0.81	0.65
Côte d'Ivoire (1998/1999)	0.32	0.38	0.34	0.46	0.38	0.36
Cameroon (2001/2004)	0.33	0.33	0.37	0.33	0.39	0.34
Colombia (2003/2005)	0.79	0.79	0.83	0.87	0.78	0.80
Guinea (1995/1999)	0.41	0.46	0.43	0.56	0.62	0.48
Indonesia (2000/2003)	0.62	0.63	0.68	0.76	0.89	0.70
Madagascar (2001/1997)	0.43	0.51	0.50	0.56	0.57	0.50
Mozambique (2002/2003)	0.26	0.28	0.28	0.31	0.31	0.28
Nicaragua (2001/2001)	0.68	0.74	0.76	0.76	0.83	0.74
South Africa (2000/1998)	0.40	0.48	0.53	0.50	0.60	0.48
Vietnam (2004/2002)	0.68	0.75	0.80	0.83	0.88	0.76
Zambia (2002/2002)	0.19	0.22	0.21	0.20	0.25	0.21

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

distribution.

When examining the individual components, it becomes clear that the biggest effect of inequality on the quintile-specific HDI is in the income component. As Table 10 shows, in many countries the richest quintile has an income index (Y) that is often more than twice or even up to five times as high as among the poorest quintile. Here many of the Sub-Saharan African countries have the highest inequality, followed closely by the Latin American and the two East Asian countries. This may seem surprising since it is well-known that Latin American countries have, on average, (slightly) higher income inequality than Sub-Saharan African countries. The reason why this is not reflected here is that the income index uses a logarithmic transformation of incomes under the assumption that the well-being effects of higher incomes among the rich is declining with higher incomes. Thus what is being measured here is not the differential in incomes but, in line with the general treatment of the income component in the HDI, the differential in important aspects of quality of life such as nutrition, housing, clothing, and other aspects that are closely correlated with incomes. In that sense it is particularly worrying that the differential is so stark in Africa and Latin America.

Table 9
Quintile specific education indices by country

Country	Q = 1	Q = 2	Q = 3	Q = 4	Q = 5	All
Burkina Faso (2003/2003)	0.19	0.21	0.23	0.26	0.37	0.26
Bolivia (2002/2003)	0.72	0.83	0.89	0.92	0.95	0.87
Côte d'Ivoire (1998/1999)	0.37	0.42	0.45	0.49	0.55	0.44
Cameroon (2001/2004)	0.58	0.66	0.72	0.75	0.80	0.71
Colombia (2003/2005)	0.79	0.83	0.87	0.89	0.93	0.86
Guinea (1995/1999)	0.30	0.43	0.44	0.49	0.46	0.41
Indonesia (2000/2003)	0.73	0.79	0.82	0.85	0.90	0.81
Madagascar (2001/1997)	0.46	0.60	0.61	0.65	0.82	0.59
Mozambique (2002/2003)	0.43	0.46	0.46	0.46	0.52	0.47
Nicaragua (2001/2001)	0.62	0.63	0.67	0.69	0.72	0.67
South Africa (2000/1998)	0.81	0.82	0.82	0.82	0.82	0.82
Vietnam (2004/2002)	0.78	0.81	0.82	0.87	0.88	0.83
Zambia (2002/2002)	0.59	0.66	0.71	0.77	0.83	0.70

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

The differential in educational achievements (E) between the richest and the poorest quintile are also sizable, but smaller than in the income index (see Table 9). In some Sub-Saharan African countries such as Burkina Faso and Madagascar the rich have nearly twice the educational achievement of the poor. But in many other countries such as South Africa, Vietnam, Nicaragua and Colombia, the differentials are not very large reflecting substantial efforts to improve education across the entire income spectrum. One should note, however, that education is only using literacy and enrolment rates and says little about educational quality which is likely to differ much more strongly between the rich and the poor.

The differential in life expectancy achievements (L) between the richest and poorest quintile are also substantial, but generally the smallest of the three components (see Tables 8a and 8b). While some of this may be related to data quality issues and the assumptions that were made in order to derive these estimates, it appears that inequality in life expectancy is indeed smaller in the developing countries we consider than other forms of inequality. Three cautionary notes are important, however. To some extent, such smaller inequality is to be expected given that life expectancy is effectively bounded above, i.e. there are limits to life expectancy that even high income people run up against. Second, the differences in actual life expectancy (rather

Table 10

Quintile specific GDP indices by country

	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$	$Q = 5$	All
Country						
Burkina Faso (2003/2003)	0.23	0.33	0.39	0.47	0.67	0.41
Bolivia (2002/2003)	0.33	0.46	0.54	0.63	0.83	0.55
Côte d'Ivoire (1998/1999)	0.34	0.45	0.51	0.59	0.75	0.48
Cameroon (2001/2004)	0.34	0.44	0.51	0.58	0.74	0.51
Colombia (2003/2005)	0.44	0.60	0.70	0.81	1.07	0.71
Guinea (1995/1999)	0.30	0.48	0.60	0.73	1.00	0.51
Indonesia (2000/2003)	0.43	0.53	0.60	0.67	0.83	0.59
Madagascar (2001/1997)	0.14	0.28	0.38	0.48	0.66	0.37
Mozambique (2002/2003)	0.22	0.33	0.40	0.48	0.68	0.41
Nicaragua (2001/2001)	0.30	0.51	0.61	0.71	0.94	0.60
South Africa (2000/1998)	0.46	0.62	0.75	0.90	1.21	0.77
Vietnam (2004/2002)	0.42	0.48	0.53	0.59	0.73	0.54
Zambia (2002/2002)	0.18	0.30	0.38	0.46	0.67	0.37

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

than the life expectancy index) are still substantial with gaps between the poorest and richest quintile amounting to more than 10 years in 5 countries. Third, even seemingly smaller differentials in life expectancy may be seen as just as important, or even more important, than larger differentials in the other components. After all, the chance to live and be free from the fear of premature mortality is a fundamental precondition for all other aspects of life.

Table 11 presents the inequality measures formulated in Equations (17) and (18): the ratio between the HDI of the richest quintile and the HDI of the poorest quintile, the ratio between the HDI of the poorest quintile and the overall HDI, and the ratio between the HDI of the richest quintile and the overall HDI. The results reveal again very stark differences in human development between the richest and the poorest quintile. For example, in Guinea, Burkina Faso, Zambia, and Madagascar, the HDI for the richest income quintile is about twice as high as in the poorest quintile. In a second group of countries, including Bolivia, Cameroon, Nicaragua, Côte d'Ivoire, Mozambique, and South Africa, the gap between the rich and the poor is also very large, between 50% and 65%. In a third group of countries, comprising Colombia, Vietnam, and Indonesia, the differential in the HDI for the richest and poorest quintile is smaller but still substantial at about

Table 11
Ratios of the quintile specific HDI to the aggregate HDI by country
(L^Q computed using predicted income)

| Country	HDI	$RQHDI^{5,1}$	$RQHDI^{1,mean}$	$RQHDI^{5,mean}$
Burkina Faso (2003/2003)	0.35	1.74	0.78	1.35
Bolivia (2002/2003)	0.69	1.46	0.81	1.19
Côte d'Ivoire (1998/1999)	0.43	1.60	0.84	1.34
Cameroon (2001/2004)	0.52	1.45	0.82	1.20
Colombia (2003/2005)	0.79	1.37	0.86	1.18
Guinea (1995/1999)	0.47	1.66	0.81	1.34
Indonesia (2000/2003)	0.70	1.40	0.86	1.20
Madagascar (2001/1997)	0.49	1.86	0.76	1.41
Mozambique (2002/2003)	0.39	1.45	0.84	1.22
Nicaragua (2001/2001)	0.67	1.48	0.82	1.22
South Africa (2000/1998)	0.69	1.51	0.82	1.23
Vietnam (2004/2002)	0.71	1.28	0.89	1.15
Zambia (2002/2002)	0.43	1.75	0.77	1.34

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS data set, the second to the DHS data set. All indices are rescaled to UNDP's reported HDI value of the second survey year.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table A1), Human Development Reports; calculations by the authors.

30%-50%. The correlation between the level of the HDI and inequality in human development seems to be negative but only weakly so as Figure 3 illustrates. In relation to the mean, the poor fare relatively well in Côte d'Ivoire, Colombia, Vietnam and Indonesia and relatively badly in Madagascar, Burkina Faso and Zambia. The rich are particularly well off with respect to the mean in Madagascar, Côte d'Ivoire and Guinea and less so in Colombia and Vietnam.

Finally in Table 12 we compare rank positions of the different quintiles. For example, the richest quintile in Bolivia is at rank 34, i.e. among the countries with high human development, actually at the same level as Poland, whereas the poorest quintile is at rank 132. The average HDI in Bolivia was ranked 112th in last year's report. In some Sub-Saharan African countries such as Cameroon, Guinea and Madagascar the richest quintile achieves a level similar to those countries with medium human development, i.e. far above the threshold of 0.5. In contrast the poorest quintiles of these countries all rank among the 15 countries with the lowest HDI. Put differently, the differences within countries are as high as the differences between high and medium as well as medium and low human development countries.

Figure 3
Correlation between the overall HDI and the ratio between the QHDI for the poorest quintile
and the QHDI for the richest quintile
(L^Q computed using predicted income)
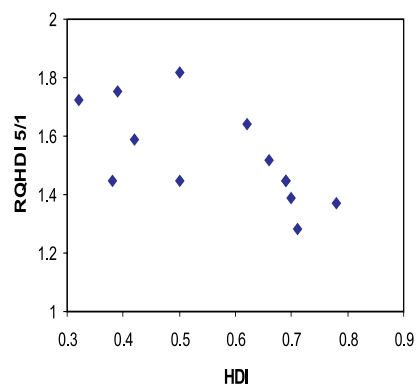
Source: Computations by the authors.

Table 12
Rank positions of the different quintiles
(L^Q computed using asset index)

	Ranking All	Ranking Q=1	Ranking Q=5
Country			
Burkina Faso (2003/2003)	172	178	151
Bolivia (2002/2003)	112	132	34
Côte d'Ivoire (1998/1999)	163	173	132
Cameroon (2001/2004)	137	165	123
Colombia (2003/2005)	66	115	22
Guinea (1995/1999)	156	174	111
Indonesia (2000/2003)	110	129	31
Madagascar (2001/1997)	152	173	114
Mozambique (2002/2003)	167	176	146
Nicaragua (2001/2001)	116	135	51
South Africa (2000/1998)	112	132	30
Vietnam (2004/2002)	108	125	51
Zambia (2002/2002)	163	175	129

Note: The years in brackets refer to the respective survey years. The first year refers to the HIS
data set, the second to the DHS data set. The ranking analysis is based on the ranking in the
2005 Human Development Report.
Source: Household Income Survey (HIS) and Demographic and Health Surveys (DHS) (see Table
A1), Human Development Reports; calculations by the authors.

5 Calculating the HDI by income quintiles for OECD countries

So far we focused in our presentation on low and middle income countries. The application of our approach to OECD countries entails some additional problems. The data availability is very different in developing and industrialized countries. Whereas for a long time access to disaggregated and harmonized income, education and health data was much better in industrialized countries than in developing countries is seems today to be the other way around. For many developing countries there exist today at least roughly comparable income, education and health data thanks to the household income surveys and Demographic and Health Surveys. In many industrialized countries, such standardized surveys are either absent or not easily accessible. Moreover, due to very low infant and child mortality levels in rich countries, we could not easily apply our methods of deducing life expectancy from infant or child mortality rates available in household survey data as the absolute number of infant and child deaths are too low in such surveys to calculate life expectancies (and its differential by income) with any reliability. Thus we will briefly discuss data availability and outline an approach to construct quintile-specific HDIs in rich countries and illustrate it for Finland and the USA. However, these calculations are not fully comparable to the calculations for developing countries and thus should be viewed as tentative.

Matters are easiest for the income component. Here we can rely on the Luxemburg Income Study (LIS), which produces harmonized micro data sets on income, demographics, labor market status and expenditures on the level of households and individuals for 30 OECD countries.[12] These data are of very high quality and probable more reliable than the income/expenditure data available in many developing countries. For our examples, Finland and the USA, we took the LIS income data for the year 2000 and simply rescaled it to fit UNDP's GDP index.

Unfortunately, the data sets contained in LIS do not have educational enrolment or adult literacy information and only provide information on educational achievements by levels of education passed. Therefore, for Finland and the USA, we assume no inequality in adult literacy and use the schooling achievement differential by income for 2000 as reported in the Luxembourg Income Study to estimate income differentials in enrolments, after which we rescale again. Alternatively, enrolment rates by income quintile could probably be generated from national household income surveys (or co-ordinated surveys such as the European Household Panel Survey) but this would mean that we rely on two different income measures to calculate the two different components (as we had to do with the HIS and the DHS for developing

---

[12]For details see: http://www.lisproject.org.

countries).

A different approach would be to use data from the 'International Adult Literacy Survey (IALS)' for the education component. This is an international comparative study designed to provide information about the skills of the adult populations. It was conducted in three phases (1994, 1996 and 1998) in 20 nations.[13] There exists also a follow up survey called 'Adult Literacy and Lifeskills (ALL) Survey' but which exists so far only in six countries. A problem with using that information would be that it is not directly comparable to the literacy and enrolment measures used for all the other countries.

By far the most difficult issues arise however with the life expectancy component. As already stated, using quintile specific child mortality to derive an estimate of quintile specific life expectancy from household surveys would not be possible as child mortality in most OECD countries is so low that no meaningful differentials by income could be identified. Moreover, child mortality in these countries is much related to premature births, genetic defects, complications during birth and due to accidents all of which not closely related to income. In fact, it is likely that existing income differentials in life expectancy in rich countries are largely due to mortality beyond childhood.

In principle, one could try to rely on census or census-like sample surveys with large numbers of observations. An alternative would be to rely on death registrations. These data sources are generally used in rich countries to calculate mortality rates and associated life expectancy statistics. But these data sources usually do not include incomes and cannot be used to calculate income differentials. Two exceptions are the USA and Finland where specialized analyses on the link between incomes and mortality were undertaken. We therefore use the results from Rogot *et al.* (1997) and Martikainen *et al.* (2001) on the life expectancy differential by incomes. These data are based on linked income survey data with vital registration data and are covering the adult mortality experience for 1979-85 for the USA, and 1991-96 for Finland. Through matching the mortality experience by income quintile with the Model Life Tables 'North' (Coale and Demeny, 1983), we derive life expectancy at birth for the two countries, after which we re-scale as described above.[14]

An alternative way would be to use similar data matching techniques that we used above to impute incomes into the DHS to impute incomes into census data and then generate life expectancy information by income quintile. That presupposes access to census data (which are not available or accessible in some countries) and a detailed analysis of the potential of such

---

[13]For details see: http://nces.ed.gov/surveys/all/.

[14]The 'income' that is referred to in these studies does not closely match annual household per capita income that we would use for the income component which causes a further complication. See also discussion below.

a method.

Given these caveats, we included only Finland and the USA in the current report and focus otherwise solely on low and middle income countries and leave the calculation of a QHDI for OECD countries for future work. Table 13 shows that among both countries, all three differentials are considerably smaller than in the considered developing countries. Income differentials (especially when expressed using the logarithmic transformation) are considerably smaller suggesting smaller differentials in income-sensitive human development achievements than elsewhere. Education differentials are, as expected, also smaller as schooling up to secondary level and thus basic literacy is near universal and only slight differentials exist at the post-secondary level. Also life expectancy differentials by income (based on cause of death information for the 1980s or early 1990s) are smaller in developing countries but remain sizable. In both the USA and Finland, the top quintiles enjoys about five more years of life than the poorest quintile. Given the wealth of these countries and the ability to provide health case to all, such differentials seem still unacceptably large. in these two countries, the differences between the quintile specific values are smaller than in developing countries, however large differences remain, particularly in the USA. If a rank analysis similar to that in Table 12 is done, the USA (followed by Finland) would top the list of human development achievements, but the poorest quintile in the US would only achieve rank 48, considerably worse off than the richest quintile in South Africa, Colombia, Bolivia, or Indonesia.

Table 13
Quintile specific HDI for industrialized countries
for Finland and USA

Finland (2000)	$Q=1$	$Q=2$	$Q=3$	$Q=4$	$Q=5$	All
L^{QI}	0.85	0.87	0.89	0.91	0.93	0.89
E^{Q}	0.97	0.97	0.98	0.99	1.02	0.99
Y^{Q}	0.82	0.88	0.92	0.96	1.04	0.94
HDI^{Q}	0.87	0.90	0.92	0.94	0.99	0.93
USA (2000)						
L^{QI}	0.82	0.86	0.88	0.89	0.90	0.87
E^{Q}	0.94	0.96	0.98	0.99	1.02	0.98
Y^{Q}	0.75	0.87	0.93	0.99	1.11	0.97
HDI^{Q}	0.84	0.89	0.93	0.96	1.01	0.94

	Finland	USA				
$RQHDI^{5,1}$	1.14	1.20				
$RQHDI^{1,mean}$	0.93	0.89				
$RQHDI^{5,mean}$	1.06	1.07				

Source: Data provided by the Luxembourg Income Study (LIS); calculations by the authors.

6 Limits and shortcomings of the suggested approach

Computing an index of well-being for different income groups is a serious challenge. The exercise is first of all constrained by data availability. In addition there is clearly a trade-off between transparency, simplicity and an intuitive interpretation on the one hand and accuracy and computational complexity on the other hand. In our approach we rather tried to elaborate an index which is relatively transparent, simple to calculate and easy to interpret. In consequence, we were forced to make many simplifications. The most important ones are discussed in what follows. Hence, the paper should first of all be seen as an illustrative exercise, which hopefully enhances the discussion and sensitizes policy makers for inequality in human development within countries. But it should not be seen by economists and demographers as an attempt to reflect accurately and exactly inequality and income differentials in health and education.

First, our segmentation variable, household income, has obviously a different temporal dimension than our indicators for life expectancy and education. Household income as measured in household surveys is clearly a period estimate, even if it is approximated by household expenditure, which could be seen as a rough measure of permanent income. Hence, assuming that people stay at this level throughout life, which is implicitly done the way we use it, is probably false and is likely to overstate income inequality. Whether this also leads to an overestimation in the income differentials of life expectancy and education is unclear. However, if such a bias exists, it would at least partly be offset by a bias in the opposite direction: If the difference between permanent income and period income is mainly driven by age and if education and life expectancy are higher among younger cohorts, then the education and life expectancy differentials by income are underestimated.

This leads directly to the second problematic point. In industrialized countries, where education at least up to some grade and basic health provision is provided costlessly to everyone, income differentials in health and education may to a large extent be driven by preferences. However, this is certainly less the case for developing countries, where health and education are often very costly. Hence, the QHDI we suggest, might have a very different interpretation in industrialized and developing countries.

Third, the matching method we use to impute incomes for the DHS is, as mentioned above, based on a couple of strong assumptions. Among other things, we assume that the distribution of unobservable factors is the same in both surveys and uncorrelated with income. Both assumptions are certainly not met and, hence, life expectancy is not exactly calculated for the same quintiles of households than education and average income.

31

Fourth, as the results show it is hard to get precise estimates of the human development index for very poor countries. This is on the one hand due to the general lower quality of data in poor countries and on the other hand in particular due to the difficulty to derive reliable estimates of life expectancy. Given the high prevalence rates of AIDS in many poor countries our life table approach provides only a very rough measure of mortality. However, the usual aggregate estimates are, at least to some extent, also affected by these problems and hence there is also uncertainty regarding the general HDI in these countries.

Fifth, in contrast to Foster *et al.* (2003), we did not follow an axiomatic approach. Hence, our HDI is not additively decomposable into the quintile-specific QHDI. This is the case, because life expectancy is not exactly additively decomposable by population sub-groups (in our case income quintiles).

7 Conclusion

One of the most often heard critiques of the HDI is that this index does not take into account inequality in its three dimensions within countries. We suggested a relatively easy, transparent and intuitive approach which allows to compute the three dimension indices and the overall HDI for quintiles of the income distribution. This allows to compare the level in human development of the poor with the level of the non-poor within countries. In addition, the approach allows to compare the level of human development for the poor and non-poor across countries. The illustration for a sample of 13 low and middle income countries, as well as 2 rich countries showed that inequality in human development within countries is indeed high. The results also showed that the level of inequality is only weakly linked to the level of human development itself. Therefore, this information is not yet contained in the overall HDI.

The implementation of our approach is obviously more time consuming and data demanding than the calculation of the usual HDI. However the necessary data—a Household Income Survey and a Demographic and Health Survey—exists now in at least most of the low and middle income countries. As discussed above, for industrialized countries getting harmonized data on education and life expectancy differentials is surprisingly a bit more problematic.

Of course our approach is not without its limits. This was discussed in detail in the previous section. However, we think it can make a useful contribution to the measurement of human development and should sensitize policy makers to inequality not only in income but also in education and life expectancy which are without any doubt two important determinants of individual well-being.

Appendix

Table A1

Data sources for developing country analysis

Country	Year	Type of Survey
Burkina Faso	2003	Demographic and Health Survey (DHS)
	2003	Enquete Burkinabe Sur Les Conditions De Vie Des Menages
Bolivia	2003	Demographic and Health Survey (DHS)
	2002	Living Standard Survey (LSMS)
Côte d'Ivoire	1999	Demographic and Health Survey (DHS)
	1998	Enquete Niveau De Vie Des Menages
Cameroon	2004	Demographic and Health Survey (DHS)
	2001	Enqute Camerounaise auprs des mnages
Colombia	2005	Demographic and Health Survey (DHS)
	2003	Encuesta de Calidad de Vida
Guinea	1999	Demographic and Health Survey (DHS)
	1995	Enquete Integrale Avec Module Budgetconsummation
Indonesia	2003	Demographic and Health Survey (DHS)
	2000	Third wave of the Indonesian Family Life Survey (IFLS)
Madagascar	1997	Demographic and Health Survey (DHS)
	2001	Enquete Aupres Des Menages (EPM)
Mozambique	2003	Demographic and Health Survey (DHS)
	2002	Inquérito Nacional Aos Agregados Familiares Sobre As Condicões De Vida
Nicaragua	2001	Demographic and Health Survey (DHS)
	2001	Encuesta Nacional de Hogares Sobre Medición de Nivel de Vida (EMNV)
South Africa	1998	Demographic and Health Survey (DHS)
	2000	Income and Expenditure Survey
Vietnam	2002	Demographic and Health Survey (DHS)
	2004	Living Standard Survey (LSMS)
Zambia	2002	Demographic and Health Survey (DHS)
	2002	Living Conditions Monitoring Survey III (LCMS III)

References

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

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

Cutler D., A. Deaton and A. Lleras-Muney (2005), The Determinants of Mortality. Mimeo, Department of Economics, Harvard University.

Deaton A. (2005), Measuring Poverty in a Growing World (or Measuring Growth in a Poor World). *Review of Economics and Statistics*, 87 (1): 1-19.

Deaton A. and S. Zaidi (2002), Guidelines for Constructing Consumption Aggregates for Welfare Analysis. Washington D.C.: World Bank.

Elbers C., J.O. Lanjouw and P. Lanjouw (2003), Micro-Level Estimation of Poverty and Inequality. *Econometrica*, 71 (1): 355-364.

Filmer D. and L. Pritchett (1999), The Effect of Household Wealth on Educational Attainment: Evidence from 35 countries. Population and Development Review, 25 (1): 85-120.

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

Foster J.E., L.F. López-Calva und M. Székely (2003), Measuring the Distribution of Human Development. Mimeo, Vanderbilt University, Nashiville.

Grün C. and S. Klasen (2006), Inequality, and Well-Being: Comparisons across space and time. Mimeo, University of Göttingen.

Grosse M., S. Klasen and J. Spatz (2005), Creating National Poverty Profiles and Growth Incidence Curves with Incomplete Income or Consumption Expenditure Data: An Application to Bolivia. Ibero America Institute for Econonomic Research (IAI) Discussion Papers No. 129, University of Göttingen.

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

Klasen S. (2006a), Guest Editor's Introduction. *Journal of Human Development*, 7(2): 145-159.

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

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

Michaelowa K. (2001), Primary Education Quality in Francophone Sub-Saharan Africa. *World Development*, 29 (10): 1699-1716.

Martikainen P., P. Mäkelä, S. Koskinen and T. Valkonen (2001), Income differences in mortality: a register-based follow-up study of three million men and women. *International Journal of Epidemiology*, 30: 1397-1405.

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.

Ravallion M (2001), Should Poverty Measures be Anchored to the National Accounts? *Economic and Political Weekly*, August 26-September 2: 3245-3252.

Rogot E., P.D. Sorlie and N.J. Johnson (1992), Life Expectancy by Employment Status, Income, and Education in the National Longitudinal Mortality Study. *Public Health Reports*, 107 (4): 457-481.

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

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

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

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

UNDP (2005), *Human Development Report 2005. International cooperation at a crossroads.* UNDP, New York.

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