Technical note 1. Calculating the Human Development Index

The Human Development Index (HDI) is a summary measure of human development. It measures the average achievements in a country in three basic dimensions of human development: a long and healthy life, access to knowledge and a decent standard of living. The HDI is the geometric mean of normalized indices measuring achievements in each dimension.

Data sources

- Life expectancy at birth: UNDESA (2009d)
- Mean years of schooling: Barro and Lee (2010)
- Expected years of schooling: UNESCO Institute for Statistics (2010a)
- Gross national income (GNI) per capita: World Bank (2010g) and IMF (2010a)

Creating the dimension indices

The first step is to create subindices for each dimension. Minimum and maximum values (goalposts) need to be set in order to transform the indicators into indices between 0 and 1. Because the geometric mean is used for aggregation, the maximum value does not affect the relative comparison (in percentage terms) between any two countries or periods of time. The maximum values are set to the actual observed maximum values of the indicators from the countries in the time series, that is, 1980–2010. The minimum values will affect comparisons, so values that can be appropriately conceived of as subsistence values or "natural" zeros are used. Progress is thus measured against minimum levels that a society needs to survive over time. The minimum values are set at 20 years for life expectancy, at zero years for both education variables and at $163 for per capita GNI. The life expectancy minimum is based on long-run historical evidence from Maddison (2010) and Riley (2005).[1] Societies can subsist without formal education, justifying the education minimum. A basic level of income, is necessary to ensure survival: $163 is the lowest value attained by any country in recorded history (in Zimbabwe in 2008) and corresponds to 45 cents a day, just over a third of the World Bank's $1.25/day poverty line.

Goalposts for the HDI in this Report

Dimension	Observed maximum	Minimum
Life expectancy	83.2 (Japan, 2010)	20.0
Mean years of schooling	13.2 (United States, 2000)	0
Expected years of schooling	20.6 (Australia, 2002)	0
Combined education index	0.951 (New Zealand, 2010)	0
Per capita income (PPP US$)	108,211 (United Arab Emirates, 1980)	163 (Zimbabwe, 2008)

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

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

$$\text{Dimension index} = \frac{\text{actual value} - \text{minimum value}}{\text{maximum value} - \text{minimum value}} \qquad (1)$$

For education, equation 1 is applied to each of the two subcomponents, then a geometric mean of the resulting indices is created and finally, equation 1 is re-applied to the geometric mean of the indices using 0 as the minimum and the highest geometric mean of the resulting indices for the time period under consideration as the maximum. This is equivalent to applying equation 1 directly to the geometric mean of the two subcomponents.

Because each dimension index is a proxy for capabilities in the corresponding dimension, the transformation function from income to capabilities is likely to be concave (Anand and Sen 2000c). Thus, for income the natural logarithm of the actual minimum and maximum values is used.

Aggregating the subindices to produce the Human Development Index

The HDI is the geometric mean of the three dimension indices:

$$\left(I_{Life}^{1/3} \cdot I_{Education}^{1/3} \cdot I_{Income}^{1/3}\right) \qquad (2)$$

Expression 2 embodies imperfect substitutability across all HDI dimensions. It thus addresses one of the most serious criticisms of the linear aggregation formula, which allowed for perfect substitution across dimensions. Some substitutability is inherent in the definition of any index that increases with the values of its components.

Example : China

Indicator	
Life expectancy at birth (years)	73.5
Mean years of schooling (years)	7.5
Expected years of schooling(years)	11.4
GNI per capita (PPP US$)	7,263

Note: Values are rounded.

Life expectancy index = $\frac{73.5-20}{83.2-20}=0.847$

Mean years of schooling index = $\frac{7.5-0}{13.2-0}=0.568$

Expected years of schooling index = $\frac{11.4-0}{20.6-0}=0.553$

$$\text{Education index} = \frac{\sqrt{0.568 \cdot 0.553} - 0}{0.951 - 0} = 0.589$$

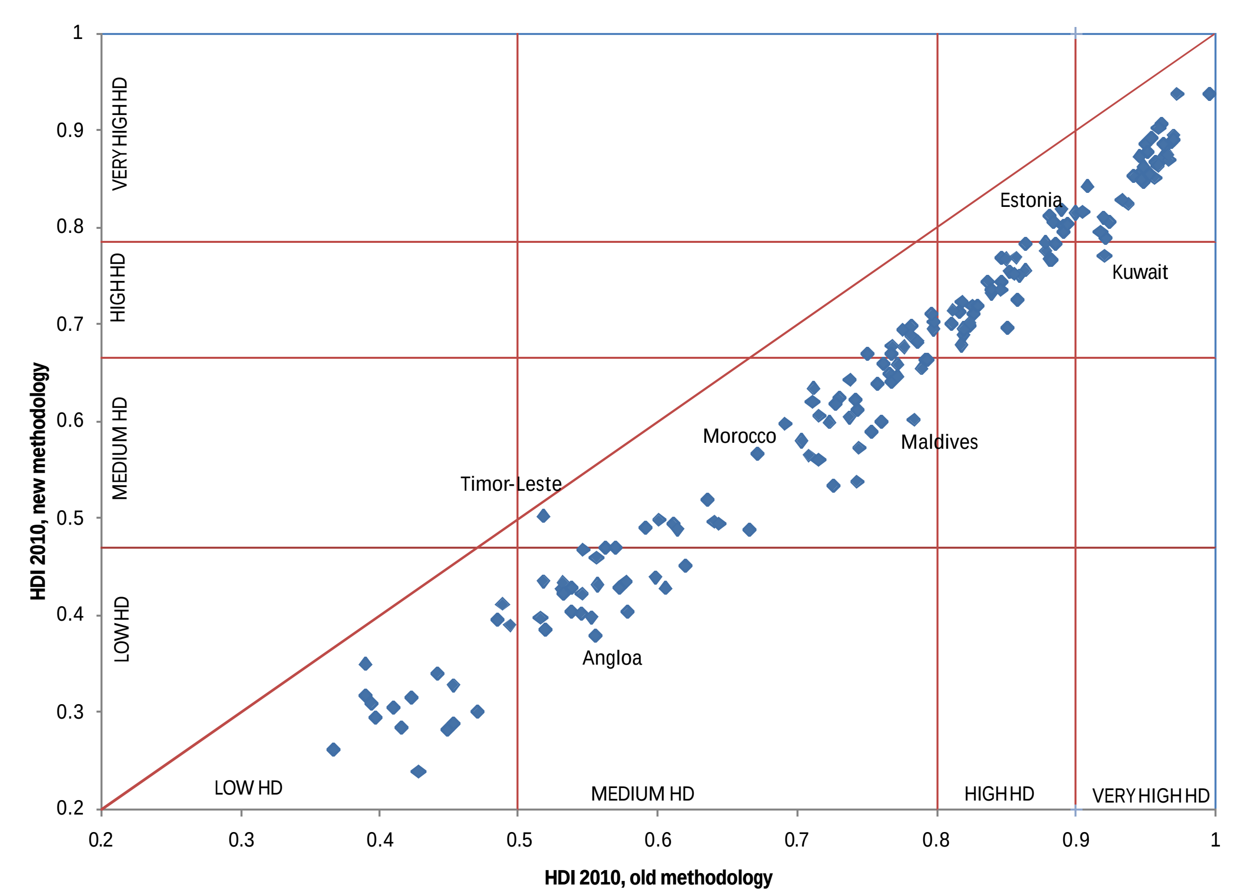
$$\text{Income index} = \frac{\ln(7{,}263) - \ln(163)}{\ln(108{,}211) - \ln(163)} = 0.584$$

$$\text{Human Development Index} = \sqrt[3]{0.847 \cdot 0.589 \cdot 0.584} = 0.663$$

Overall effects of the methodological improvements

The methodological improvements in the HDI, using new indicators and the new functional form, result in substantial changes (figure T1.1). Adopting the geometric mean produces lower index values, with the largest changes occurring in countries with uneven development across dimensions. The geometric mean has only a moderate impact on HDI ranks. Setting the upper bounds at actual maximum values has less impact on overall index values and has little further impact on ranks.

Figure T1.1 Human Development Index 2010: New and old methodology



Source: HDRO calculations using data from the HDRO database.

Analysis of historical trends in this Report

The analysis of historical trends in chapters 2 and 3 uses a different version of the HDI, the hybrid HDI, which applies the same aggregation formula as the new HDI to the set of indicators and sources used in previous Reports (since 1995) in order to allow more extensive analysis over time. Linear interpolation was used to fill missing values when both earlier and later values were present. When unavailable for the whole time period, gross enrolment ratios were projected using the last available value (for forward projections) and the first available value (for backward projections). A sensitivity analysis showed that the results of the analysis were robust to alternative extrapolation techniques. See Gidwitz et al. (2010) for further details on the construction of this data set.

The analysis in chapters 2 and 3 also uses the deviation from fit criterion to comparatively evaluate changes over time in the hybrid HDI. This measure evaluates the progress of countries in comparison to the average progress of countries with a similar initial HDI level. It is calculated as the residual of a second degree fractional polynomial regression of the annual percentage growth rate of the HDI on the logarithm of its initial HDI value. Statistical table 2 reports the country rank in the deviation from fit for the HDI for the 1980-2010 period. See Royston and Altman (1994) for a description of regression models based on fractional polynomial functions of a continuous covariate

Technical note 2. Calculating the Inequality-adjusted Human Development Index

The Inequality-adjusted Human Development Index (IHDI) adjusts the HDI for inequality in distribution of each dimension across the population. It is based on a distribution-sensitive class of composite indices proposed by Foster, Lopez-Calva, and Szekely (2005), which draws on the Atkinson (1970) family of inequality measures. It is computed as a geometric mean of geometric means, calculated across the population for each dimension separately (for details, see Alkire and Foster 2010). The IHDI accounts for inequalities in HDI dimensions by 'discounting' each dimension's average value according to its level of inequality. The IHDI will be equal to the HDI when there is no inequality across people, but falls further below the HDI as inequality rises. In this sense, the IHDI is the actual level of human development (accounting for this inequality), while the HDI can be viewed as an index of 'potential' human development (or the maximum level of IHDI) that could be achieved if there was no inequality. The 'loss' in potential human development due to inequality is given by the difference between the HDI and the IHDI, and can be expressed as a percentage.

Data sources

Since the HDI relies on country-level aggregates such as national accounts for income, the IHDI must draw on alternative sources of data to obtain the distribution of each dimension. The distributions have different units—income and years of schooling are distributed across individuals, while expected length of life is distributed across age intervals. Available distributional data are not necessarily for the same individuals or households.

The inequality in distribution of the HDI dimensions is estimated for:

- Life expectancy, which uses data from abridged life tables provided by UNDESA (2009d). This distribution is available across age intervals (0–1, 1–5, 5–10, ... , 85+), with the mortality rates and average age at death specified for each interval.
- Years of schooling and household income (or consumption), which use household survey data harmonized in international databases: the Organisation for Economic Co-operation and Development's Luxembourg Income Study, Eurostat's European Union Survey of Income and Living Conditions, the World Bank's International Income Distribution Database, the United Nations Children's Fund's Multiple Indicators Cluster Survey, the U.S. Agency for International Development's Demographic and Health Survey, the World Health Organization's World Health Survey, and the United Nations University's World Income Inequality Database.
- The inequality in standard of living dimension, which uses disposable household income per capita, household consumption per capita or income imputed based on an asset index matching methodology (Harttgen and Klasen 2010).

For a full account of data sources used for estimating inequality, see Kovacevic (2010a).

Computing the Inequality-adjusted HDI

There are three steps to computing the IHDI.

Step 1. Measuring inequality in underlying distributions

The Inequality-adjusted HDI (IHDI) draws on the Atkinson (1970) family of inequality measures and sets the aversion parameter e equal to one.[2] In this case the inequality measure is $A = 1- g/\mu$, where g is the geometric mean and μ is the arithmetic mean of the distribution. This can be written:

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

where $\{X_1, \ldots, X_n\}$ denotes the underlying distribution in the dimensions of interest. A_x is obtained for each variable (life expectancy, years of schooling and disposable income or consumption per capita) using household survey data and the life tables.[3]

The geometric mean in equation 1 does not allow zero values. For mean years of schooling one year is added to all valid observations to compute the inequality. For income per capita outliers—extremely high incomes as well as negative and zero incomes—were dealt with by truncating the top 0.5 percentile of the distribution to reduce the influence of extremely high incomes and by replacing the negative and zero incomes with the minimum value of the bottom 0.5 percentile of the distribution of positive incomes.

For more details on measuring inequality in the distribution of the HDI indicators, see Alkire and Foster (2010).

Step 2. Adjusting the dimension indices for inequality

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

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

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

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

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

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

---

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

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

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

Step 3. Computing the Inequality-adjusted HDI

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

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

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

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

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

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

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

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

which is equivalent to

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

Notes on methodology and limits

The IHDI is based on an index that satisfies subgroup consistency. This ensures that improvements or deteriorations in distribution of human development within a certain group of society (while human development remains constant in the other groups) will be reflected in changes in the overall measure of human development. This index is also path independent, which means that the order in which data are aggregated across individuals, or groups of individuals, and across dimensions yields the same result—so there is no need to rely on a particular sequence or a single data source. This allows estimation for a large number of countries.

Although the IHDI is about human development losses from inequality, the measurement of inequality in any dimension implicitly conflates inequity and inequality due to chance, choice and circumstances. It does not address the ethical and policy-relevant issues around whether these aspects should be distinguished (see Roemer 1998 and World Bank 2005b for applications in Latin America).

The main disadvantage is that the IHDI is not association sensitive, so it does not capture overlapping inequalities. To make the measure association-sensitive, all the data for each individual must be available from a single survey source, which is not currently possible.

Example : Slovenia

	Indicator	Dimension index	Inequality measure (A1)	Inequality-adjusted index
Life expectancy	78.8	0.930	0.043	(1–0.043) · 0.930 = 0.890
Mean years of schooling	9	0.682		
Expected years of schooling	16.7	0.811		
Education index		0.782	0.040	(1–0.040) · 0.782 = 0.751
Logarithm of GNI	10.16	0.780		
GNI	25,857	0.238	0.122	(1–0.122) · 0.238 = 0.209

	Human Development Index	Inequality-adjusted Human Development Index	Loss
HDI with unlogged income	$\sqrt[3]{0.930 \cdot 0.782 \cdot 0.238} = 0.557$	$\sqrt[3]{0.892 \cdot 0.738 \cdot 0.209} = 0.516$	1–0.519/0.557 = 0.068
HDI	$\sqrt[3]{0.930 \cdot 0.782 \cdot 0.780} = 0.828$	$(0.519 / 0.557) \cdot 0.828 = 0.772$	

Note: Values are rounded.

Technical note 3. Calculating the Gender Inequality Index

The Gender Inequality Index (GII) reflects women's disadvantage in three dimensions-reproductive health, empowerment, and the labour market-for as many countries as data of reasonable quality allow. The index shows the loss in human development due to inequality between female and male achievements in these dimensions. It varies between 0 – when women and men fare equally – and 1, where women fare as poorly as possible in all measured dimensions.

It is computed using the association-sensitive inequality measure suggested by Seth (2009). The index is based on the general mean of general means of different orders—the first aggregation is by the geometric mean across dimensions; these means, calculated separately for women and men, are then aggregated using a harmonic mean across genders.

Data sources

- Maternal mortality ratio (*MMR*): UNICEF (2010c)
- Adolescent fertility rate (*AFR*): UNDESA (2009d)
- Share of parliamentary seats held by each sex (*PR*): Inter-parliamentary Union's Parline database (2010)
- Attainment at secondary and higher education (*SE*) levels: Barro and Lee (2010)
- Labour market participation rate (*LFPR*): ILO (2010d)

Computing the Gender Inequality Index

There are five steps to computing the GII.

Step 1. Treating zeros and extreme values

The maternal mortality ratio is truncated symmetrically at 10 (minimum) and at 1,000 (maximum). The maximum of 1,000 is based on the normative assumption that countries where maternal mortality ratios exceed 1,000 are not different in their ability to create conditions and support for maternal health. Similarly, it is assumed that countries with 1–10 deaths per 100,000 births are essentially performing at the same level.

The female parliamentary representation of countries reporting zero is coded as 0.1% because the geometric mean cannot have zero values and because these countries do have some kind of political influence by women.

Step 2. Aggregating across dimensions within each gender group, using geometric means

Aggregating across dimensions for each gender group by the geometric mean makes the GII association sensitive (see Seth 2009).

For women and girls, the aggregation formula is

$$G_F = \sqrt[3]{\left(\frac{1}{MMR}\cdot\frac{1}{AFR}\right)^{1/2}\cdot(PR_F\cdot SE_F)^{1/2}\cdot LFPR_F}\ ,$$

and for men and boys the formula is

$$G_M = \sqrt[3]{1\cdot(PR_M\cdot SE_M)^{1/2}\cdot LFPR_M}\ .$$

Step 3. Aggregating across gender groups, using a harmonic mean

The female and male indices are aggregated by the harmonic mean to create the equally distributed gender index

$$HARM(G_F, G_M) = \left[\frac{(G_F)^{-1}+(G_M)^{-1}}{2}\right]^{-1}.$$

Using the harmonic mean of geometric means within groups captures the inequality between women and men and adjusts for association between dimensions.

Step 4. Calculating the geometric mean of the arithmetic means for each indicator

The reference standard for computing inequality is obtained by aggregating female and male indices using equal weights (thus treating the genders equally) and then aggregating the indices across dimensions:

$$G_{\overline{F},\overline{M}} = \sqrt[3]{\overline{Health}\cdot\overline{Empowerment}\cdot\overline{LFPR}}$$

where $\overline{Health} = \left(\sqrt{\frac{1}{MMR}\cdot\frac{1}{AFR}}+1\right)/2$, $\overline{Empowerment} = \left(\sqrt{PR_F\cdot SE_F}+\sqrt{PR_M\cdot SE_M}\right)/2$, and

$$\overline{LFPR} = \frac{LFPR_F + LFPR_M}{2}.$$

$\overline{Health}$ should not be interpreted as an average of corresponding female and male indices but as a half the distance from the norms established for the reproductive health indicators—fewer maternal deaths and fewer adolescent pregnancies.

Step 5. Calculating the Gender Inequality Index

Comparing the equally distributed gender index to the reference standard yields the GII, $1-\frac{Harm(G_F, G_M)}{G_{\overline{F},\overline{M}}}$.

Example : Brazil

	Health		Empowerment		Labour market
	Maternal mortality ratio	**Adolescent fertility rate**	**Parliamentary representation**	**Attainment at secondary and higher education**	**Labour market participation rate**
Female	110	75.6	0.094	0.488	0.640
Male	na	na	0.906	0.463	0.852
(F+M)/2	$\left[\sqrt{(1/110)\cdot(1/75.6)}+1\right]/2=0.505$		$\left(\sqrt{0.094\times 0.488}+\sqrt{0.906\times 0.463}\right)/2=$		(0.640 + 0.852) / 2 = 0.746

Note: na is not applicable.

Using the above formulas, it is straightforward to obtain:

G_F $$0.115=\sqrt[3]{\sqrt{\left(\frac{1}{110}\cdot\frac{1}{75.6}\right)}\cdot\sqrt{0.094\cdot 0.488}\cdot 0.640}$$

G_M $$0.820=\sqrt[3]{1\cdot\sqrt{0.906\cdot 0.463}\cdot 0.852}$$

$Harm(G_F, G_M)$ $$0.201=\left[\frac{1}{2}\left(\frac{1}{0.115}+\frac{1}{0.820}\right)\right]^{-1}$$

$G_{\overline{F,M}}$ $$0.546=\sqrt[3]{0.505\cdot 0.431\cdot 0.746}$$

GII 1–0.201/0.546 = 0.632.

Technical note 4. Calculating the Multidimensional Poverty Index

The Multidimensional Poverty Index (MPI) identifies multiple deprivations at the individual level in education, health and standard of living. It uses micro data from household surveys, and—unlike the Inequality-adjusted Human Development Index—all the indicators needed to construct the measure must come from the same survey.

Each person in a given household is classified as poor or not depending on the number of deprivations their household experiences. These data are then aggregated into the national measure of poverty.

Methodology

Each person is assigned a score according to his or her household's deprivations in each of the 10 component indicators, (*d*). The maximum score is 10, with each dimension equally weighted (thus the maximum score in each dimension is $3\frac{1}{3}$). The education and health dimensions have two indicators each, so each component is worth 5/3 (or 1.67). The standard of living dimension has six indicators, so each component is worth 5/9 (or 0.56).

The education thresholds are having no household member who has completed five years of schooling and having at least one school-age child (up to grade 8) who is not attending school. The health thresholds are having at least one household member who is malnourished and having had one or more children die. The standard of living thresholds relate to not having electricity, not having access to clean drinking water, not having access to adequate sanitation, using "dirty" cooking fuel (dung, wood or charcoal), having a home with a dirt floor, and owning no car, truck or similar motorized vehicle, and owning at most one of these assets: bicycle, motorcycle, radio, refrigerator, telephone or television.

To identify the multidimensionally poor, the deprivation scores for each household are summed to obtain the household deprivation, c. A cut-off of 3, which is the equivalent of one-third of the indicators, is used to distinguish between the poor and non-poor.[4] If c is 3 or greater, that household (and everyone in it) is multidimensionally poor. Households with a deprivation count between 2 and 3 are vulnerable to or at risk of becoming multidimensionally poor.

The MPI value is the product of two measures: the multidimensional headcount ratio and the intensity (or breadth) of poverty.

The headcount ratio, H, is the proportion of the population who are multidimensionally poor: $H = \frac{q}{n}$, where q is the number of people who are multidimensionally poor and n is the total population.

The intensity of poverty, A, reflects the proportion of the weighted component indicators, d, in which, on average, poor people are deprived. For poor households only, the deprivation scores are summed and divided by the total number of indicators and by the total number of poor persons:

[4] Technically this would be 3.33. Because of the weighting structure, the same households are identified as poor if a cut-off of 3 is used.

$A = \frac{\sum_{1}^{q} c}{qd}$, where c is the total number of weighted deprivations the poor experience and d is the total number of component indicators considered (10 in this case).

Example using hypothetical data

Indicators	**Households**				**Weights**
	1	**2**	**3**	**4**	
Household size	4	7	5	4	
Education					
No one has completed five years of schooling	0	1	0	1	5/3=1.67
At least one school-age child not enrolled in school	0	1	0	0	5/3=1.67
Health					
At least one member is malnourished	0	0	1	0	5/3=1.67
One or more children have died	1	1	0	1	5/3=1.67
Living *conditions*					
No electricity	0	1	1	1	5/9=0.56
No access to clean drinking water	0	0	1	0	5/9=0.56
No access to adequate sanitation	0	1	1	0	5/9=0.56
House has dirt floor	0	0	0	0	5/9=0.56
Household uses “dirty” cooking fuel (dung, firewood or charcoal)	1	1	1	1	5/9=0.56
Household has no car and owns at most one of: bicycle, motorcycle, radio, refrigerator, telephone or television	0	1	0	1	5/9=0.56
Weighted count of deprivation, c (sum of each deprivation multiplied by its weight)	**2.22**	**7.22**	**3.89**	**5.00**	
Is the household poor ($c > 3$)?	**No**	**Yes**	**Yes**	**Yes**	

Note: 1 indicates deprivation in the indicator; 0 indicates non-deprivation.

Weighted count of deprivations in household 1: $\left(1 \cdot \frac{5}{3}\right) + \left(1 \cdot \frac{5}{9}\right) = 2.22$.

Headcount ratio (H) = $\left(\frac{7+5+4}{4+7+5+4}\right) = 0.800$ (80% of people live in poor households)

Intensity of poverty (A) = $\frac{(7.22 \cdot 7) + (3.891 \cdot 5) + (5.00 \cdot 4)}{(7+5+4) \cdot 10} = 0.5625$ (the average poor person is deprived in 56% of the weighted indicators).

MPI = $H \cdot A = 0.450$.