Soc Indic Res (2012) 109:395–411
DOI 10.1007/s11205-011-9908-4

A Household-Based Distribution-Sensitive Human Development Index: An Empirical Application to Mexico, Nicaragua and Peru

Luis F. Lopez-Calva · Eduardo Ortiz-Juarez

Accepted: 17 July 2011 / Published online: 17 August 2011

Abstract In measuring human development, one of the main concerns relates to the inclusion of a measure that penalizes inequalities in the distribution of achievements across the population. Using indicators from nationally representative household surveys and census data, this paper proposes a straightforward methodology to estimate a household-based distribution-sensitive human development index aggregated through generalized means. The evidence shows that the losses in human development due to inequality reach up 22, 29 and 57% in Mexico, Peru and Nicaragua, respectively. Among dimensions, the loss in the income index reaches up 61% in Nicaragua, while the education index appears as the most sensitive in the case of Mexico and Peru, with a percentage of loss between 38 and 48%. The importance of household-level calculations is highlighted when we compare the indices computed from the entire distribution with those existing indices computed for quintiles of the distribution, which minimizes the losses due to inequality. Overall, the estimations evidence a higher sensitivity of the index to inequality, and therefore an important space for public action to reduce inequality that could involve positive development returns.

Keywords Human development index · Distribution-sensitive inequality · Generalized means · Cross-sectional data

1 Introduction

In light of the increasing availability of better data, and on the basis of the conceptual contributions by Amartya Sen, multidimensionality has attracted the interest of scholars in

L. F. Lopez-Calva · E. Ortiz-Juarez
The World Bank, 1818 H St. NW, Washington, DC 20433, USA
e-mail: lflopezcalva@worldbank.org

E. Ortiz-Juarez (✉)
UNDP, Regional Bureau for Latin America and the Caribbean, Montes Urales No. 530 4th Floor, Lomas de Chapultepec, 11000 Mexico City, Mexico
e-mail: eduardo.ortiz@undp.org

reexamining welfare measures to be axiomatically more robust, and of policy makers in implementing strategies focused on several other dimensions equally as important as income in the assessment of welfare. One of the most relevant contributions, multidimensional in nature, is the human development index (HDI) that assesses achievements in the population's educational attainment and health status, as well as in the standard of living, measured by income. The HDI calculation is straightforward and such simplicity has strong policy content both for ease in communicating the achievements and for the evaluation of policies and budget allocation. Methodologically, however, the index has been widely criticized for using an equal weighting system in the aggregation of dimensions and therefore for the impossibility to penalize inequalities in the distribution of human development across the population.

Recent literature offers several ways to capture the inequality in the distribution of human development in an overall measure. For example, Anand and Sen (1995) developed an index that assesses the achievements in human development between men and women, and then discounts the inequality between both groups using generalized means; Hicks (1997) and Foster et al. (2005) proposed indices that discount the level of inequality in each dimension of human development using the Gini index and the Atkinson inequality measure, respectively; and more recently, Seth (2009) proposed an index capturing not only the inequality in the distribution of human development, but also the interaction between dimensions.

This document is aimed at proposing a straightforward methodology to calculate a human development index that is distribution-sensitive using the class of indices developed by Foster et al. (2005). Due to data constraints, as the analysis is carried out at the household level, we use a variation of the traditional indicators from micro data for Mexico, Nicaragua and Peru: a variable that captures the years of schooling instead of the gross enrollment rate, and the child survival rate instead of the life expectancy at birth. We observe that our household-based indices using these alternative variables are not so different than those calculated using aggregated data and the traditional indicators.

The main findings show that inequality causes an important loss in human development in all three countries as measured by our household-based distribution-sensitive index. The most dramatic loss is that of Nicaragua, where inequality leads a percentage of loss that reaches up 57%; by contrast, the losses in Mexico and Peru are about 21 and 29%, respectively. Among dimensions, the loss in the income index reaches up 61% in Nicaragua, while the education component is the most sensitive in the case of Mexico and Peru (losses at around 38 and 48%, respectively). These results suggest the necessity of strategies to reduce inequality, which could involve positive development returns.

This paper is organized as follows. Section 2 describes the basic properties of the traditional HDI, as well as the main contributions that incorporate inequality in the distribution of achievements. Section 3 develops the methodology proposed in this paper, while Sect. 4 presents the empirical illustration. Finally, Sect. 5 concludes.

2 Human Development and Inequality

2.1 The Human Development Index

Over recent years, a large amount of literature recognizes that additional dimensions are equally important as income in the assessment of welfare. One of the most influential approaches has been Amartya Sen's theory of functionings and capabilities, which focuses on the achievement of "beings and doings", but more importantly, on people's freedom to choose among such achieved functionings. Although the empirical implementation of this approach hasn't achieved a widespread consensus, the HDI has been a successful attempt.

Published since 1990, the HDI was originally proposed for international comparisons of achievements in three dimensions of human development measured by four specific indicators: income, measured by the log of GDP per capita; education, measured by literacy and enrollment rates; and health, measured by life expectancy. All of these indicators are first required to be normalized and then aggregated using the arithmetic mean and equal weights, so that no dimension has a higher relative importance in the index than others. The traditional index can be expressed as:

$$\mathrm{HDI}_T = \mu[\mu(yi), \mu(ei), \mu(hi)] \tag{1}$$

where $\mu(\cdot)$ denotes the arithmetic mean, and *yi*, *ei*, and *hi* the dimensions of income, education and health, respectively.

One the most important features of the index is the ease of calculation, as it reveals the multidimensionality of welfare in the policy sphere. However, some criticisms suggest that the index suffers from important methodological flaws. On one hand, a number of studies suggest that linear aggregation and equal weighing are arbitrary, and show how the rankings are robust to changes in the weights assigned to each dimension (Cherchye et al. 2008; Foster et al. 2009). On the other hand, the index fails to incorporate both inequality between dimensions (uneven development) and inequality among people (unequal distribution of development), which could be relevant for policy purposes. Let's consider a region *x* where education's quality is unequal between public and private schools—and probably uneven in future achievements of the population—, but where the overall access to education is equitable, so that a high proportion of the population is enrolled. If the level of human development in *x* is measured as in (1), the high enrollment rate will bias the index upwards and hide educational inequalities among the population. Therefore, the design of a measure that penalizes such inequalities is highly advantageous to achieve better social outcomes.

In spite of previous criticisms, from an axiomatic point of view the index fulfills a set of desirable properties. One of the most important is *path independence*, which allows for the possibility of aggregate first across persons and then across dimensions, or vice versa, yielding the same level of human development. Another property is *continuity*, whereby the overall index does not vary due to changes in any of the three dimensions. A related property, *normalization*, indicates that if all normalized dimensions are equal, the index should not be affected; however, *linear homogeneity* indicates that if all dimensions change in the same proportion, the overall index should also change in the same proportion.

Additional properties indicate that the index remains unchanged if the population of a group is replicated several times (*replication invariance*); the index will increase if the achievement of any one person in any single dimension increases (*monotonocity*) and; the change in the index within a subgroup is associated with the corresponding change for the population as a whole (*subgroup consistency*). Finally, *symmetry in people* and *symmetry in dimensions* indicate, respectively, that each person and each dimension are equally important in measuring human development.

2.2 Inequality in the Measurement of Human Development

In measuring human development, a seminal contribution to reflect inequality is the gender-related development index (GDI) proposed by Anand and Sen (1995). The index first calculates the level of human development for men and woman as in (1), and then aggregates across the two groups using generalized means to discount for inequality.[1] It is well known that this index ignores within group inequality and violates the property of subgroup consistency. Nonetheless, the GDI has been a turning point in the development of more robust indices. In this line, Hicks (1997) proposed a distribution-sensitive index that discounts the mean of each dimension by its level of inequality, and then averages across dimensions using the arithmetic mean. The 'Hicks index' can be expressed as:

$$\text{HDI}_H = \mu[S(yi), S(ei), S(hi)] \tag{2}$$

where $S(\cdot) = \mu(\cdot)[1 - G(\cdot)]$ is the Sen's (1976) welfare measure, with G being the Gini coefficient and $\mu(\cdot)$ the arithmetic mean. In addition to fulfilling most of the abovementioned properties, the Hicks index also satisfies the *Kolm transfer principle*.[2] However, the index violates the path-independence property, since the value of the overall index depends on the order of aggregation, and also violates subgroup consistency, given that the S and G measures also violate such property.[3]

As an attempt to overcome these failures, Foster et al. (2005) proposed a class of indices (FLS from here onwards) that satisfy all the properties using a distribution-sensitive generalized mean to summarize the dimension-specific level of human development. The generalized mean can be referred as $\mu_\alpha(\cdot)$, where $\alpha = 1 - \varepsilon$, with $\varepsilon \geq 0$ being an inequality aversion parameter that measures the sensitivity to inequality of the relatively lower values of a distribution. For a given population of size n, the generalized mean is commonly expressed as:

$$\mu_\alpha(\cdot) = \left[\frac{1}{n}\sum_{i=1}^{n}(\cdot)^\alpha\right]^{\frac{1}{\alpha}} \tag{3}$$

Notice that when $\alpha = 1$ the expression (3) is the arithmetic mean, and when $\alpha = -1$ it is the harmonic mean. However, when $\alpha = 0$ the generalized mean takes a multiplicative form to yield the geometric mean as:

$$\mu_0 = \left[\prod_{i=1}^{n}(\cdot)\right]^{\frac{1}{n}} \tag{4}$$

In measuring the achievements in human development, the expressions (3) and (4) lower the mean of each dimension in accordance with its level of inequality by transforming the distribution of each dimension. Then, the aggregation of dimensions requires averaging the transformed distributions using the arithmetic mean, and then applying the reverse transformation, such that the FLS class of indices can be seen as a generalized mean of the generalized means of each dimension. The FLS class of indices can be expressed thus as:

[1] Technical details are presented in UNDP (1995).

[2] This principle establishes that for two equal distributions, namely x and y, and a bistochastic matrix **C**, the welfare in x will be lower than in y whether y becomes smoother than x ($x < y\mathbf{C}$). See Kolm (1977).

[3] See for example Foster et al. (1984, 2005). The Gini coefficient violates subgroup consistency when, for instance, inequality in groups x and y decrease between periods 1 and 2, but the overall inequality goes up. Formally, $g_{x1} > g_{x2}$ and $g_{y1} > g_{y2}$, such that $G_{x,y1} > G_{x,y2}$.

$$\text{HDI}_{\text{FLS}} = \mu_\alpha[\mu_\alpha(yi), \mu_\alpha(ei), \mu_\alpha(hi)] \quad (5)$$

As suggested, when $\alpha = 1$, the expression (5) is equivalent to the traditional index in (1) that aggregates dimensions using the arithmetic mean. Overall, as the value of α decreases (because the value of the parameter ε increases) it penalizes the inequality across dimensions and among people; therefore, the higher the inequality the lower the value of (5).[4]

Finally, Seth (2009) recently proposed a class of indices that also uses generalized means to capture not only the distribution-sensitive inequality, but also the association-sensitive inequality. The latter suggests that interactions between dimensions may alter the overall level of inequality.[5] Following the same notation, the 'Seth index' can be expressed as:

$$\text{HDI}_S = \mu_\alpha[\mu_\beta(yi), \mu_\beta(ei), \mu_\beta(hi)] \quad (6)$$

where $\alpha = 1 - \varepsilon$, as defined previously, and β is a parameter that measures the degree of substitutability between dimensions. When $\beta = 1$ the dimensions are perfect substitutes, but the higher the β the lower the substitutability; therefore, at the limit the dimensions are perfect complements. The expression (6) is not sensitive to any form of inequality if $\alpha = \beta = 1$, being equivalent to the index in (1), and it is strictly distribution-sensitive if $\beta = \alpha \leq 1$, being equivalent to the class of indices in (5). However, if $\alpha, \beta \leq 1$ and $\alpha \neq \beta$, the index is both distribution-sensitive and association-sensitive. The Seth index fulfills most of the basic properties, except path-independence, and it also fails in calculating the contribution of each dimension to the overall index.

This paper aims at calculating a household-based distribution-sensitive human development index (HB-IAHDI from here onwards) using a measure that is analytically simple and axiomatically solid, and not to question the validity of other existing indices.[6] For that purpose, we use the class of indices in (5) since it is the only one that satisfies a broad set of desirable properties according to our purposes. The two inherent characteristics of our proposal (calculation at the household level and inequality adjustment) require a data source containing the distribution of income, education, and health indicators across the entire population. The latter, however, constitutes a challenge due to two main factors. First, there is no availability of the indicators used in the traditional HDI in the same household survey, so that different but related variables must be included. Second, household surveys suffer from missing data on education (school enrollment) and health indicators (life expectancy) as these strongly depend on the presence and age of household members. In the next section we describe the methodology to address these constraints and to calculate the HB-IAHDI based on recent micro data for Mexico, Nicaragua and Peru.

[4] The inequality discounted using generalized means is equivalent to the level discounted using the Atkinson's (1970) unidimensional inequality measure I_ε, so that $\mu_\alpha(\cdot)$ can be also expressed as: $\mu_\alpha(\cdot) = \mu(\cdot)[1 - I_\varepsilon(\cdot)]$, with $I_\varepsilon(\cdot) = 1 - [\mu_\alpha(\cdot)/\mu(\cdot)]$, where a larger gap between the generalized mean and the arithmetic mean implies a higher inequality. By doing so, the overall index resulting in (5) is interpreted as the traditional index in (1) discounted by the level of inequality as measured by the Atkinson measure (Foster et al. 2005 show how to extend the Atkinson's measure to the multidimensional space).

[5] The association between dimensions is related to the notions of complementarity and substitutability, broadly analyzed by Bourguignon and Chakravarty (2003). They proposed a two-dimensional poverty measure based on a CES function that allows a degree of substitution between dimensions.

[6] The 2010 HDR proposed a new measure of human development that also penalizes the inequalities in the distribution of achievements. As the new index was named inequality-adjusted human development index (IAHDI), we decided to include the official acronym to our household-based index, thus being HB-IAHDI.

3 A Household-Based Distribution-Sensitive Human Development Index

A large amount of empirical literature has attempted to calculate a distribution-sensitive index at the household level. For example, Arim and Vigorito (2009) combined surveys with census data to estimate the class of indices in (5) for twelve Latin American countries.[7] They calculate the income index using household per capita income without rescaling to national GDP, and the education index using the average years of schooling for adults with or above the age of 25 instead of literacy and enrollment. As for the health dimension they use mortality rates at the state level, which are merged with surveys for households living in the corresponding states.

Following a regression-based approach, Grimm et al. (2008) and Harttgen and Klasen (2010) calculate the index in (1) at the household level and for all five quintiles. They combine income and education indicators from traditional household surveys with mortality data from the Demographic and Health Surveys (DHS). In the case of the income index, they estimate the correlation between the household income and a set of household characteristics in both surveys; then they use this correlation to predict a log-distributed variable in the DHS based on an asset index. In the case of the education index, as it is composed by two indicators, they use the traditional information on literacy, and for enrollment they scale the average rate in surveys to the official enrollment rate to estimate a stochastic regression for predicting missing values in households without individuals aged for school. Finally, in the case of health they regress child mortality data on socioeconomic variables, and then calculate the household average mortality rate using the predicted values.

While previous approaches provide a useful way to calculate a household-based index, one can argue that they may suffer from important drawbacks. For example, the analysis by Arim and Vigorito (2009) avoids missing data and zero values on enrollment and literacy, respectively, by using the average years of schooling for adults aged 25 or above. However, they exclude information on education for individuals below that threshold. Moreover, by using health data at the state level they suppress within-state variations and some of the between-state inequality, which bias the index upwards due to the small number of states compared with the large number of households in a survey. In the case of the regression-based approach, strong assumptions in predicting education and health indicators are the major shortcomings. For instance, the prediction for mortality is likely to be highly discontinuous due to the existence of households without children. Even if the household data is collapsed for all five quintiles and the latter constraint is avoided, all the intra-quintile inequalities will be suppressed if we calculate the index as in Eq. 5 (see Sect. 4 for an empirical exercise of this).

Previous drawbacks, however, are mainly due to the availability of household indicators for all three dimensions, as also occur in this paper. Therefore, whichever methodology depends on the existence of data and hence the resulting index can be very sensitive to several methodological decisions, which may imply, of course, a lack of comparability across countries.

The methodology proposed in this paper estimates a household-based distribution-sensitive index based on the expression (5) using income and education indicators from nationally representative household surveys, and mortality indicators from census data at

[7] The 2002 HDR for Mexico also calculated this class of indices by states using a sample from the census (UNDP 2003); the 2009 HDR for Argentina estimates the same indices using microdata (UNDP 2009a); and based on an earlier version of this paper the Latin American and the Caribbean HDR 2010 calculate the indices for 18 countries using standardized household surveys (UNDP 2010).

Table 1 Descriptive statistics from households surveys (household averages)

	Literacy rate[a]	Schooling rate[b]	Years of schooling	Household size	GDP per capita[c]	Weightned sample[d]	Distribution of population[e]
Mexico							
Urban	95.9	65.8	8.5	3.9	$22,437.2	17.7	66.2
Rural	85.5	52.1	5.7	4.3	$9,899.9	9.0	33.8
National	92.4	61.1	7.6	4.0	$18,200.4	26.7	100.0
Peru							
Urban	94.1	70.0	9.1	4.0	$12,376.4	5.1	71.4
Rural	74.7	48.1	5.0	4.3	$4,294.9	2.0	28.6
National	88.6	63.8	7.9	4.1	$10,060.9	7.1	100.0
Nicaragua							
Urban	89.1	55.4	6.8	5.0	$3,574.2	0.58	58.4
Rural	68.6	34.8	3.2	5.5	$1,939.7	0.41	41.6
National	80.6	46.8	5.3	5.2	$2,893.8	0.99	100.0

Source: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

[a] Percentage of population aged 15 and above

[b] Percentage of achievement of the accumulated years of schooling for age

[c] GDP expressed at 2005 PPP

[d] Millions of households

[e] Percentage of urban and rural population

the municipal level. In the case of Mexico, we use data from the 2008 National Household Consumption and Expenditure Survey (ENIGH) and from the National Population Council (CONAPO); in the case of Peru, calculations are based on the 2008 National Household Survey (ENAHO) and from UNDP; and for Nicaragua data is taken from the 2005 Household Living Standards Survey (EMNV) and from the National Institute for Development Information (INIDE). The representativeness of the surveys and descriptive statistics of the variables used are shown in Table 1.

3.1 Income Index

In calculating this dimension, we first construct the annual total current income for each household in the survey. Given the differences between surveys and aggregated data, we then rescale the household income to national accounts based on an adjustment factor (2.84 for Mexico, 1.19 for Nicaragua and 2.17 for Peru) calculated as the ratio between the GDP in national accounts and the total income in surveys.[8] Finally, the rescaled income is then

[8] Rescaling suggests a net difference between the income in surveys and national accounts that should be owned by both private and public sectors and used, for instance, in savings, investments in physical and human capital, and redistribution through public goods or social protection policies. The way in which the difference is adjusted—mainly related to income under-reporting in surveys—suggests different distributive effects, and some methods of adjustment have been proposed in literature in order to obtain figures closer to reality (see for example the work by Altimir 1987). In this paper, we prefer to take the view that adjustment to national accounts is distributional neutral as the methods developed are somewhat and unavoidably arbitrary. This is also supported by the idea that the distributional net effect of taxation and public expenditure and public goods provision in Latin America is very low (ECLAC 2010).

divided by the household size and expressed at 2005 prices in Purchasing Power Parity (PPP) terms for comparability purposes.

Once these adjustments are done, we compare the household i's per capita income y_i^{pc} with the minimum ($100) and maximum ($40,000) reference values established by UNDP for the calculation of the traditional HDI. The traditional formula to calculate the income index for household i (yi_i) is thus expressed as:

$$yi_i = \frac{\log(y_i^{pc}) - \log(\min)}{\log(\max) - \log(\min)} \tag{7}$$

The minimum and maximum reference values were originally proposed to calculate the HDI based on national aggregated data. However, the calculation at the household level implies that in some cases the income in the survey could be lower or higher than such reference values. In order to adjust for these variations, in this paper we use a strictly positive income, so that negative or zero values are adjusted to the minimum value of $100 to avoid the loss of observations. In the case of the upper threshold, we decide not cap income to avoid the right-truncation of the distribution, and therefore, the index for a given household could be over 1.[9]

3.2 Education Index

Literacy and enrollment indicators traditionally compose this index. As noted, however, the household-based calculation imposes the problem of missing data in households without children, as enrollment depends on the presence of individuals aged for school. The way we adopt to calculate the education index is by maintaining the information on literacy, but replacing enrollment with a continuous variable capturing the years of schooling for individuals of or above the age of 7 (the age required to complete the first year of primary education). By using this variable, we avoid missing values by imputing the household i's average schooling to children aged 6 or below, under the assumption that children could achieve at least such average over the course of their life (UNDP 2009a).

For each household member, we calculate an indicator of the years of schooling and compare it with a minimum value of zero and a maximum value that depends on age. For instance, a person aged 7 must have 1 year of schooling as maximum; a person aged 8 must have 2 years of schooling as maximum, and so on up to a maximum of 16 years of schooling which corresponds to individuals aged 22 or above.[10] Notice that if a person

[9] There are significant conceptual problems in adjusting household income to minimum and maximum reference thresholds, and the results could be sensitive to these adjustments, as many authors have pointed out (see for example the discussion by Alkire and Foster 2010). First, in dealing with survey data we face the presence of negative or zero values, and thus the calculation of the income index will suffer the loss of observations as these are being considered as missing values. In searching the best way to address this problem we decided to transform the observed negative and zero data to a strictly positive threshold of $100 as proposed by UNDP; however, at this stage appears a second problem: the presence of observations strictly positive but below the threshold. The number of these observations for the countries analyzed in this paper is lower, and there are no significant differences in the resulting income indices if we adjust to $100 or to a different value, so that we prefer to maintain the official threshold.

[10] For comparative purposes, the maximum value of 16 years of schooling corresponds to at least four completed years of tertiary education. For instance, the norm indicates that Mexican individuals have to complete 6 years of primary, 3 years of lower secondary, 3 years of upper secondary, and between 4 and 5 years of tertiary education. This implies a maximum schooling of about 16–17 years. In Peru, however, the norm indicates that population have to complete 6 years of primary, 5 years of secondary and 5 years of tertiary education, so that the maximum schooling is 16 years.

aged 7 has 2 or more years of schooling, the value is fixed up to 1; if a person aged 8 has 3 or more years of schooling, it is fixed up to 2, and so on.

Based on previous norms, the schooling index for individual j in household i is calculated as $s_{ij} = \frac{s_j - \min}{\max - \min}$, with s_j being the observed years of schooling for individual j, and min and max the reference values. The average of the individual indices is then calculated and imputed to children aged 6 or below. Notice that the schooling index for household i is therefore the average of schooling for all the individuals in the household. In order to not underestimate the index, all the zero values are replaced by 0.5 under the assumption that individuals have accumulated some learning and experience throughout their lives, regardless of if they have attended school or not.

In the case of literacy, if an adult with or above the age of 15 declared to be able to read and write a message he or she is considered as literate. Household literacy rate is then calculated as $\text{literacy}_i = \frac{1}{n}\sum_{j=1}^{m} l_j$, with n being the total number of adults in household i, m the total number of literate adults, and l_j an indicator taking the value of 1 if the adult j is literate, and 0 otherwise. Notice that the previous rate is equivalent to the literacy index. For the same reasons outlined in the case of schooling, a minimum level of 0.05 is attached instead of 0 in those households with illiterate adults.

The education index for household i (ei_i) is then aggregated using the traditional weights of 2/3 for literacy, and 1/3 for schooling, as proposed by UNDP, following the formula:

$$ei_i = 2/3(\text{literacy}_i) + 1/3(\text{schooling}_i) \tag{8}$$

3.3 Health Index

The health component of the traditional HDI uses data on life expectancy at birth. This indicator, however, is not available at the household level, so that the easiest way we adopt is to use the child survival rate at the municipality level. Some studies have justified the use of this indicator as it has a significant impact on life expectancy.[11] Moreover, although this aggregated indicator may suppress within-municipality variations, we acknowledge that it allows us to capture some of the between-municipality inequality since the number of municipalities in a survey is considerably higher than, for example, the number of states. Unless household level data become available for the dimension of health, this constraint will remain.

The child survival rate is obtained as $csr_i = 1 - cmr_i$, with cmr_i being the child mortality rate in municipality i, defined as number of deaths per thousand, and it is compared with maximum and minimum reference values to obtain the mean achievement in the health dimension. In the case of the maximum, we use the Japan's child survival rate of 0.997.[12] To calculate the minimum value, we use the formula:

$$lei_{\text{nat}} = \frac{csr_{\text{nat}} - csr_{\min}}{csr_{\max} - csr_{\min}} \tag{9}$$

where csr_{nat} is the national child survival rate, calculated as the population weighted average of the child survival rates in all municipalities; $csr_{\max}$ is Japan's child survival rate; and lei_{nat} is the life expectancy index in each country.[13] Solving by $csr_{\min}$, the above

[11] See for example Gakidou and King (2000).

[12] The higher historically observed value according to the MDG indicators.

[13] See UNDP (2007) for the life expectancy index in Mexico and Nicaragua in 2005, and UNDP (2009b) for the index in Peru in 2007.

expression yields the minimum value: 0.914 for Mexico, 0.922 for Peru and 0.872 for Nicaragua.

The municipal health index is then calculated by following the expression:

$$hi_i = \frac{csr_i - csr_{\min}}{csr_{\max} - csr_{\min}} \tag{10}$$

The resulting values are then imputed to those households in each survey living in the corresponding municipalities.

Finally, the aggregation of the individual indices obtained in (7), (8) and (10) is carried out by using the FLS class of indices shown in Eq. 5. Notice that the latter formula calculates first the mean adjusted-achievement for each dimension, and then aggregate across households to obtain the overall index from the entire distribution. However, due to the path-independence property it is possible to apply first the generalized mean across dimensions to obtain the household i's level of development as $d_i = \left[\mu\left(yi_i^{\alpha}, ei_i^{\alpha}, hi_i^{\alpha}\right)\right]^{\frac{1}{\alpha}}$; therefore, the index in (5) can be rewritten as $\mathrm{HDI}_{\mathrm{FLS}} = \mu_{\alpha}(d_1, d_2, \ldots, d_n)$.[14] As noted, with $\varepsilon = 0$ (or $\alpha = 1$) the previous expression yields the traditional index in (1); however, as the value of ε increases it penalizes the lower values in the distribution, and hence the inequality.

4 Empirical Illustration

Table 2 presents a comparison between our HB-IAHDI and its components calculated from household surveys, with $\varepsilon = 0$, and the traditional indices calculated using national aggregated data. As shown, the differences are not dramatic in relative terms. The overall index for Mexico (0.819) represented 96% of the index reported by UNDP (0.854), while for Peru and Nicaragua the numbers are 95 and 94%, respectively. Among dimensions, we observe similar figures. In the case of the education index, the value for all three countries was around 91% of that reported by UNDP, which seems high considering the use of years of schooling for the entire population instead of the enrollment rate only for individuals aged for school. Meanwhile, the income index represented 93%, also high considering the level of disaggregation, and the health index accounted for 99.7%. The similarity of the health indices could be a result of the level of data aggregation, but also of the high correlation between life expectancy and child survival, suggesting that the used indicator is a good proxy for health.

Figure 1 draws the losses in human development due to inequality, and shows the effect of increasing the inequality aversion parameter ε from 0 to 3 in our HB-IAHDI. As the value of ε increased there was a dramatic and consistent widening of the gap between the three countries. For example, the index of Mexico was 7% higher than that of Peru for a parameter 0, but the difference increased to 19% for a parameter 3. By comparing Mexico and Nicaragua the differences were dramatically higher: 25% for a parameter 0, and 130% for a parameter 3. The same occurred when comparing Peru and Nicaragua: for a parameter 0 the index of Peru was 17% higher than that of Nicaragua, but rose to 93% for a parameter 3.

These comparisons reveal that the distribution of human development is more egalitarian in Mexico than in the other two countries, and also reveal that the Nicaraguan

[14] In the case of the multiplicative form of the generalized mean, when $\varepsilon = 1$ (or $\alpha = 0$), the household i's level of development is expressed as: $d_i = [\prod yi_i, ei_i, hi_i]^{\frac{1}{3}}$.

Table 2 Human development index and its components (calculations from micro and aggregated data)

	Mexico 2008	Peru 2008	Nicaragua 2005
Household-based indices			
Education index	0.820	0.803	0.693
Income index	0.785	0.690	0.479
Health index[a]	0.853	0.795	0.789
Human development index	0.819	0.763	0.654
	Mexico	Peru	Nicaragua
Indices from aggregated data[b]			
Education index	0.886	0.891	0.760
Income index	0.826	0.728	0.542
Health index	0.850	0.800	0.795
Human development index	0.854	0.806	0.699

Source: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys, and UNDP (2009b)

[a] Based on municipal data for 2005 in Mexico and Nicaragua, and for 2007 in the case of Peru

[b] Based on national data for 2007

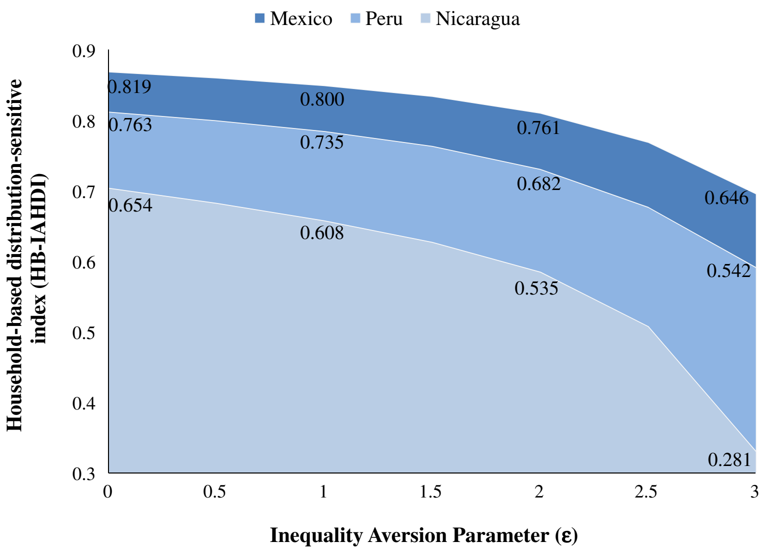

Fig. 1 Change in HB-IAHDI due to inequality. *Source*: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

distribution is more sensitive to inequality as ε increases, especially for households in the bottom being evident by the left-bulge of the bells shown in Figs. 2, 3. In these figures, inequality considerably changed the shape of the distribution, which also moved dramatically between countries.

Previous results are reproduced in Fig. 4 in terms of the percentage of loss in human development due to inequality. As shown, the loss was dramatically higher in Nicaragua,

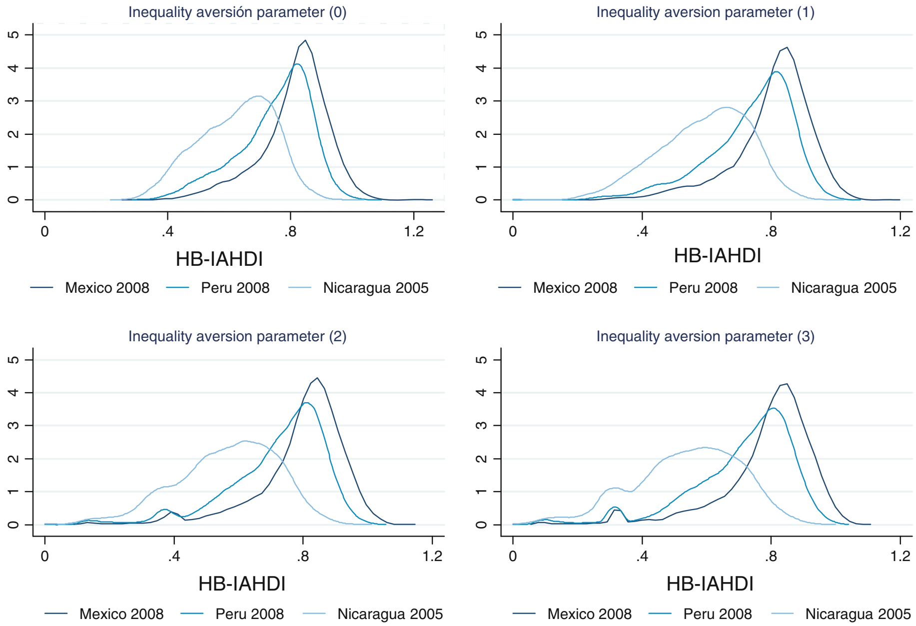

Fig. 2 Distribution of human development for different values of ε. Household distribution not truncated for values more than 1. *Source*: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

where the unequal distribution leads a loss in human development that reached up 57% when going from $\varepsilon = 0$ to $\varepsilon = 3$. By contrast, the losses in Mexico and Peru, although high, were about half that of Nicaragua. These results are particularly important in policy terms. For example, in Nicaragua the percentage of loss implies that the value of the index decreased from 0.654 to 0.281 and hence strategies to reduce income inequality and improve coverage and quality of education and health services would involve positive development returns.

The results also show that the income index was the most sensitive in the case of Nicaragua, where it decreased by 61.1% when going from $\varepsilon = 0$ to $\varepsilon = 3$. In Mexico and Peru the losses reached up 6 and 12%, respectively. In these countries, the education index was the most sensitive with percentages of loss of about 38 and 48%, respectively, being similar to that also shown for Nicaragua (47.6%). These numbers suggest that the use of a variable reflecting the years of schooling presumably reveals that a significant proportion of the population have failed to achieve the educational standards required for their age. Finally, the losses in the health index due to inequality were between 2.5 and 4%, which is an expected result due to the suppression of some within-municipality variations.[15]

[15] Due to the lack of child survival data at the household level, UNDP (2010) followed a regression approach to estimate the probability of access to water and sanitation, based on several socioeconomic indicators. The estimated probabilities were averaged and the final value was interpreted as a proxy for health. The results for 18 Latin American countries show that the distribution of the scores was highly sensitive to inequality, particularly in Central America. In the case of the countries analyzed here, the results show that when going from $\varepsilon = 0$ to $\varepsilon = 2$ Nicaragua has the highest loss in the health dimension (64%), followed by Peru (26%) and Mexico (12%).

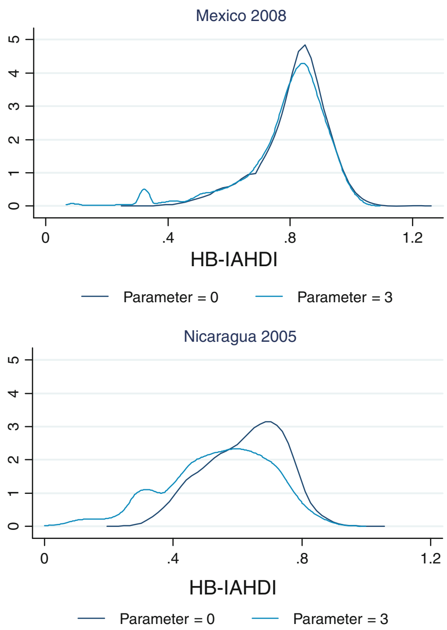

Fig. 3 Change in the distribution of human development from ε = 0 to ε = 3. Household distribution not truncated for values more than 1. *Source*: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

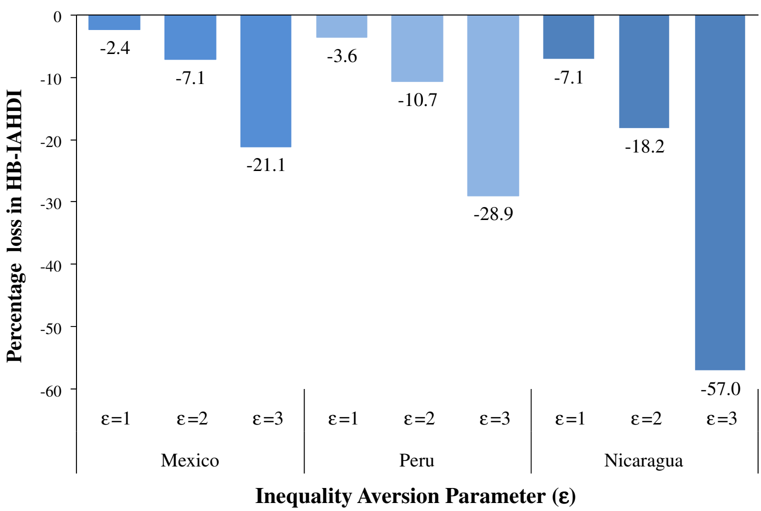

Fig. 4 Losses in HB-IAHDI due to inequality. Percentage change from ε = 0 to ε = 1, 2, 3. *Source*: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

Based on the above findings, Fig. 5 draws the variations of the overall index and of those dimensions calculated at the household level (income and education). In this figure, both a high sensitivity to inequality and a high disparity in achievements across countries is

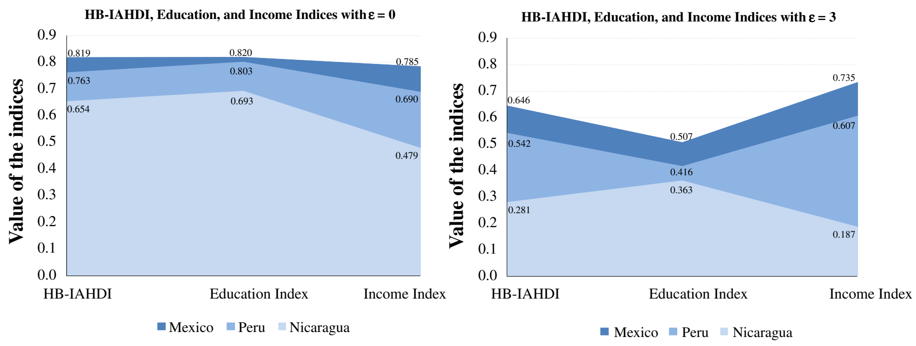

Fig. 5 Between-country inequalities in human development. *Source*: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

Table 3 Dimensions of human development for different values of ε (household-based indices)

Inequality aversion parameter (ε)	Weighted calculations								
	Income index			Education index			Health index		
	Mexico	Nicaragua	Peru	Mexico	Nicaragua	Peru	Mexico	Nicaragua	Peru
0.0	0.785	0.479	0.690	0.820	0.693	0.803	0.853	0.789	0.795
0.5	0.777	0.467	0.681	0.805	0.665	0.783	0.850	0.787	0.791
1.0	0.769	0.455	0.671	0.784	0.629	0.754	0.847	0.784	0.786
1.5	0.761	0.442	0.661	0.753	0.581	0.709	0.844	0.781	0.781
2.0	0.753	0.418	0.648	0.701	0.518	0.638	0.841	0.778	0.776
2.5	0.744	0.351	0.633	0.619	0.442	0.534	0.837	0.775	0.770
3.0	0.735	0.187	0.607	0.507	0.363	0.416	0.832	0.771	0.763

Source: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

evident, hence revealing an important space for public action. The values of the indices for different values of the inequality aversion parameter (ε) are shown in Table 3.

The importance of household-level calculations is highlighted when comparing the use of the whole distribution, and the use of the distribution collapsed by quintiles. In general, if the overall index is calculated for quintiles of the distribution, the percentage of losses due to inequality showed before are minimized, decreasing from 21.1 to 2.5% in Mexico, from 28.9 to 4.6% in Peru, and from 57 to 14.6% in Nicaragua (see Table 4). Moreover, household-based calculations also allow us to observe the variation of the index by different strata or socioeconomic status for which the survey is representative, and hence to design better policies. For example, in the case of Mexico the percentage of loss when going from $\varepsilon = 0$ to $\varepsilon = 2$ was about 13% for extreme poor households living in rural areas, while it was about 3% for non-poor households in urban ones. A more dramatic result occurred when we compared the index of Nicaragua by income deciles. While the percentage of loss in decile 1 reached up 35%, it reached only 5.4% in decile 10, and the ratio between these deciles increasing from 1.62 for $\varepsilon = 0$ to 2.37 for $\varepsilon = 2$. Additional comparisons are shown in Table 5.

Table 4 HB-IAHDI for different values of ε (household-based and quintile-based indices)

Inequality aversion parameter (ε)	Weighted calculations			Unweighted calculations					
	HB-IAHDI			HB-IAHDI			Quintile-based HB-IAHDI		
	Mexico	Nicaragua	Peru	Mexico	Nicaragua	Peru	Mexico	Nicaragua	Peru
0.0	0.819	0.654	0.763	0.817	0.625	0.748	0.817	0.625	0.748
0.5	0.811	0.633	0.751	0.808	0.602	0.735	0.814	0.613	0.743
1.0	0.800	0.608	0.735	0.796	0.576	0.719	0.810	0.600	0.738
1.5	0.784	0.577	0.714	0.780	0.543	0.697	0.807	0.585	0.732
2.0	0.761	0.535	0.682	0.756	0.494	0.664	0.804	0.569	0.726
2.5	0.719	0.458	0.628	0.716	0.391	0.609	0.800	0.552	0.720
3.0	0.646	0.281	0.542	0.645	0.202	0.524	0.796	0.533	0.713

Source: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

Table 5 HB-IAHDI for different values of ε (household-based indices by poverty status in Mexico, and by deciles across countries)

Inequality aversion parameter (ε)	National	Rural	Urban	Inequality aversion parameter (ε)	Mexico		
	Total population				Decile 1	Decile 10	Ratio
ε = 0	0.819	0.737	0.861	ε = 0	0.658	0.966	1.47
ε = 1	0.800	0.711	0.849	ε = 1	0.627	0.959	1.53
ε = 2	0.761	0.660	0.825	ε = 2	0.578	0.949	1.64
% of loss	−7.1	−10.4	−4.3	% of loss	−12.2	−1.7	
	Extreme poor households				Nicaragua		
ε = 0	0.690	0.649	0.757	ε = 0	0.499	0.810	1.62
ε = 1	0.659	0.617	0.733	ε = 1	0.423	0.796	1.88
ε = 2	0.606	0.563	0.693	ε = 2	0.323	0.767	2.37
% of loss	−12.2	−13.3	−8.5	% of loss	−35.3	−5.4	
	Non-poor households				Peru		
ε = 0	0.868	0.797	0.893	ε = 0	0.580	0.921	1.59
ε = 1	0.856	0.779	0.885	ε = 1	0.537	0.915	1.70
ε = 2	0.826	0.735	0.864	ε = 2	0.471	0.904	1.92
% of loss	−4.8	−7.8	−3.2	% of loss	−18.8	−1.8	

Source: Authors' calculations based on data from ENIGH 2008, ENAHO 2008 and EMNV 2005 household surveys

5 Concluding Remarks

This paper proposes a straightforward methodology to calculate a household-based distribution-sensitive human development index (HB-IAHDI) using income, literacy and schooling indicators from recent micro data, and child survival rates by municipality in three Latin American countries.

The inclusion of inequality in the measurement of human development show that the percentage of loss in human development, as measured by our HB-IAHDI, can reaches up 21, 29 and 57% in Mexico, Peru and Nicaragua, respectively, as the inequality aversion parameter increases. Among dimensions, the income index is the most sensitive in Nicaragua where the percentage of loss reached up 61%, while in Mexico and Peru the education index is the most sensitive with a percentage of loss of 38 and 48%, respectively. Overall, calculations reveal a higher sensitivity of the index to inequality, and therefore an important space for public action to reduce inequality, which could involve positive development returns. In addition, the importance of household-level calculations is highlighted when comparing the use of the whole distribution and the use of the distribution collapsed by quintiles, municipalities, regions or states, which minimizes the losses due to inequality.

Different methodologies for household-based calculations have been proposed in several studies without consensus in the choice of indicators, and making strong assumptions. However, it should be noted that common drawbacks are due to the availability of household indicators for all three dimensions, and therefore the index can be very sensitive to different methodological decisions. Despite this, we highlight the importance of household-based estimations for policy purposes as they allow us to observe the variation by different strata, and hence to design better policies aimed at achieve better social outcomes.

Acknowledgments The authors would like to thank Sabina Alkire, James Foster, participants at the 2010 Annual HDCA Conference in Amman, Jordan, the HDRO-UNDP Meeting on Multidimensional Poverty and Inequality held in New York, and anonymous referees for very useful comments. We also want to thank the support of the World Bank through the Summer Initiative for Research on Poverty, Inequality and Gender (SiRPiG) in the Poverty and Gender Unit, PREM-LAC. The views expressed here do not necessarily represent those of the World Bank or the United Nations Development Programme.

References

Alkire, S., & Foster, J. (2010) *Designing the inequality-adjusted human development index (HDI)*. OPHI working paper no. 37, University of Oxford.

Altimir, O. (1987). Income distribution statistics in Latin America and their reliability. *Review of Income and Wealth*, *33*(2), 111–155.

Anand, S., & Sen, A. K. (1995). *Gender inequality in human development: Theories and measurement*. HDRO occasional paper no. 19. New York: UNDP. http://hdr.undp.org/en/reports/global/hdr1995/papers/sudhir_anand_amartya_sen.pdf.

Arim, R., & Vigorito, A. (2009). *Human development and inequality: The adjusted human development index*. Background paper for the Regional Human Development Report for Latin America and the Caribbean 2010.

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

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

Cherchye, L., Ooghe, E., & Puyenbroeck, T. V. (2008). Robust human development rankings. *Journal of Economic Inequality, 6*(4), 287–321.

ECLAC. (2010). *Economic Survey of Latin America and the Caribbean 2009–2010*. Santiago: United Nations.

Foster, J., Greer, J., & Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica, 52*(3), 761–766.

Foster, J., Lopez-Calva, L. F., & Szekely, M. (2005). Measuring the distribution of human development: Methodology and an application to Mexico. *Journal of Human Development, 6*(1), 5–25.

Foster, J., McGillivray, M., & Seth, S. (2009). *Rank robustness of composite indices*. OPHI working paper no. 26, University of Oxford. http://www3.qeh.ox.ac.uk/pdf/ophiwp/OPHIwp26.pdf.

Gakidou, E., & King, G. (2000). *An individual-level approach to health inequality: Child survival in 50 countries*. Discussion paper series no. 18. Geneva: World Health Organization. http://polmeth.wustl.edu/media/Paper/king00c.pdf.

Grimm, M., Harttgen, K., Klasen, S., & Misselhorn, M. (2008). A human development index by income groups. *World Development, 36*(12), 2527–2546.

Harttgen, K., & Klasen, S. (2010). *A household-based human development index*. Human Development Research Paper 2010/22. New York. http://hdr.undp.org/en/reports/global/hdr2010/papers/HDRP_2010_22_rev.pdf.

Hicks, D. A. (1997). The inequality-adjusted human development index: A constructive proposal. *World Development, 25*(8), 1283–1298.

Kolm, S.-C. (1977). Multidimensional egalitarianisms. *Quarterly Journal of Economics, 91*(1), 1–13.

Sen, A. K. (1976). Poverty: An ordinal approach to measurement. *Econometrica, 44*(2), 219–231.

Seth, S. (2009). Inequality, interactions and human development. *Journal of Human Development and Capabilities, 10*(3), 375–396.

UNDP. (1995). *Human development report 1995. Gender and human development*. New York: Oxford University Press.

UNDP. (2003). *Informe sobre Desarrollo Humano México 2002*. Mexico: Mundi-Prensa.

UNDP. (2007). *Human development report 2007/2008. Fighting climate change: Human solidarity in a divided world*. New York: Palgrave Macmillan.

UNDP. (2009a). *Aportes para el desarrollo humano en Argentina 2009*. Buenos Aires: PNUD.

UNDP. (2009b). *Human development report 2009. Overcoming barriers: Human mobility and development*. New York: Palgrave Macmillan.

UNDP. (2010). *Human development report for Latin America and the Caribbean 2010. Acting on the future: Breaking the inequality transmission of inequality*. Costa Rica: Editorama.