ELSEVIER
www.elsevier.com/locate/worlddev

World Development Vol. 40, No. 5, pp. 878–899, 2012

0305-750X/$ - see front matter

doi:10.1016/j.worlddev.2011.09.011

A Household-Based Human Development Index

KENNETH HARTTGEN and STEPHAN KLASEN *
University of Göttingen, Germany

Summary. — One of the most serious weaknesses of the human development index (HDI) is that it does not take into account the distribution of human development within a country. All attempts to capture inequality in the HDI have used aggregate information and there exists no HDI at the household level. We provide a method for proxying the HDI at the household level. This allows the immediate analysis of human development by population subgroups and/or socioeconomic characteristics as well as analysis of inequality in human development across population subgroups and households characteristics. We illustrate our approach for 15 developing countries. Inequality in the HDI is stunningly large for some countries, particularly in countries with low overall human development, driven mostly by very high inequality in the education and income components of the HDI.

Key words — human development index, income inequality, differential mortality, inequality in education

1. INTRODUCTION

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

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

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

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

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

There are more closely related precursors to our work. Grimm *et al.* (2008, 2009) produce an HDI for different income quintiles. Based on a method and computations described in detail in Grimm *et al.* (2006), the HDR 2006 presented an HDI for all five income quintiles for a sample

* 1We would like to thank Sabina Alkire, Joachim von Braun, Francois Bourguignon, James Foster, Jeni Klugman, Milorad Kovacevic, Francisco Rodriguez, three anonymous referees as well as participants at workshops and conferences in Oxford, Hannover, New York, and Göttingen for helpful comments on earlier versions of the paper. Funding from UNDP in support of this work is gratefully acknowledged Final revision accepted: August 29, 2011.

of 11 OECD countries and 21 developing countries. Harttgen and Klasen (2011) calculate the HDI separately for internal migrants *versus* nonmigrants.[5] Both of these papers calculate the HDI separately for the specified groups. One cannot use these results to calculate unconditional inequality measures such as the Gini coefficient for the HDI or its components, or calculate it for other groups.

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

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

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

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

The objective of this paper is to illustratively demonstrate the feasibility of such an approach. We illustrate our approach for 15 developing countries using DHS survey data. Because of the comprehensive availability of the DHS data, the household-based HDI can be calculated for more than 50 countries for several time periods. The remainder of this paper is organized as follows. Section 2 presents our methodology. Section 3 presents the sample of countries and presents the results. Section 4 concludes.

2. METHODOLOGY

(a) *Calculating the GNI index*

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

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

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

The literature has also examined whether it can serve as an appropriate proxy for income or expenditure. For example, Stewart and Simelane (2005) validate the use of the asset index as a proxy for income to predict child mortality in South Africa. They find a very close relationship between income and the asset index. The recent paper by Filmer and Scott (2008) provides an excellent validation of the use of various asset index methods by comparing how asset index outcomes match to results using *per capita* expenditures with respect to the ranking of households and with respect to inequality analysis outcomes in education, health care use, fertility, and child mortality. They show that inferences about inequalities in education and health are robust to the use of the asset index. The gradient of the outcomes of the asset index closely follows the outcome using *per capita* expenditures.[10] Using Mexico's national income and expenditure survey, McKenzie (2005) also shows that assets provide a reliable proxy for inequality in education, after controlling for household income and demographic characteristics. However, Howe, Hargreaves, Gabrysch, and Huttly (2009) provided a systematic review of