Using Census Data to Explore the Spatial Distribution of Human Development

IÑAKI PERMANYER [*]

Centre d'Estudis Demogràfics, Bellaterra, Spain

Summary. — The Human Development Index (HDI) has been criticized for not incorporating distributional issues. We propose using census data to construct a municipal-based HDI that allows exploring the distribution of human development with unprecedented geographical coverage and detail. Moreover, we present a new methodology that allows decomposing overall human development inequality according to the contribution of its subcomponents. We illustrate our methodology for Mexico's last three census rounds. Municipal-based human development has increased over time and inequality between municipalities has decreased. The wealth component has increasingly accounted for most of the existing inequality in human development during the last 20 years.
© 2012 Elsevier Ltd. All rights reserved.

Key words — Human Development Index, measurement, spatial distribution, inequality, census, Mexico

1. INTRODUCTION

Since it was first introduced in the 1990 Human Development Report (HDR), the Human Development Index (HDI) has attracted a great deal of interest in policy-making and academic circles alike. As stated in Klugman, Rodríguez, and Choi (2011): "Its popularity can be attributed to the simplicity of its characterization of development—an average of achievements in health, education and income—and to its underlying message that development is much more than economic growth". Among other things, the HDI has been criticized for the arbitrariness of its weighting scheme (see Cherchye, Ooghe, & Van Puyenbroeck, 2008; Foster, McGillivray, & Seth, 2009; Permanyer, 2011a, 2011b), the limited well-being dimensions incorporated in the analysis (see, for instance, Anand & Sen, 1992; Neumayer, 2001 or Ranis, Stewart, & Samman, 2007) and for the neglect of distributional issues in its conceptualization (see Sagar & Najam, 1998; Foster, Lopez-Calva, & Szekely, 2005; Grimm, Harttgen, Klasen, & Misselhorn, 2008). More specifically, the HDI has been rightly criticized for only giving an average value at the country level that might hide large inequalities. It is in this context that this paper aims to make a contribution: using widely available census data and a simple methodology we suggest estimating human development indicators at municipal level to uncover their distribution with unprecedented geographical coverage and detail.

There have been different attempts to incorporate inequality in the assessment of human development levels, particularly in the last few years. More specifically, several papers have tried to define alternative versions of the HDI in a way that aggregate achievement values are penalized for unequal distributions within a given country (e.g., Alkire & Foster, 2010; Foster et al., 2005; Hicks, 1997; Seth, 2009). It is important to highlight that these contributions are focused on the construction of nationally representative distribution-sensitive human development indices rather than on the estimation of human development levels for certain population subgroups. In order to fill this important gap, alternative but methodologically analogous versions of the HDI have been recently defined for specific population subgroups. Grimm, Harttgen, Klasen, and Misselhorn (2008) and Grimm et al. (2010) present an

HDI for the different income quintiles and Harttgen and Klasen (2011a) calculate the HDI separately for internal migrants and for non-migrants. More recently, Harttgen and Klasen (2011b) propose a similar methodology to construct a household-based Human Development Index. As stated by the authors, these approaches are particularly attractive as they open up the possibility of performing many kinds of comparative analysis that were not previously available (e.g., comparisons between and within population subgroups based on a wide range of socio-demographic and economic characteristics).

Notwithstanding the undisputable advantages that the choice of those subgroup-specific HDI methodologies entails, there are important shortcomings that are worth pointing out. First, the different approaches attempt to forcibly replicate the original HDI variables which were originally defined at the country level [1] and estimate them for the new units of analysis. While this ensures comparability with the original HDI—as long as the variables are meaningful for the new units of analysis—it can be conceptually problematic when the units that are being compared represent extremely small population subgroups (e.g., in the case of households without children it becomes particularly questionable to define something like a household-based life expectancy, a gross enrolment ratio or the expected years of schooling). As a result, the authors need to rely on many imputations and complex technical assumptions that are hard to verify—to say the least.

Second, it is not possible to know which of the three HDI components (i.e., standard of living, education or health) accounts for most of the existing inequality levels in human development. Using the household-based approach presented in Harttgen and Klasen (2011b) it is possible to determine the different component-specific inequality levels. However, these approaches do *not* allow knowing the specific contribution of the three components to *overall* inequality in the human development distribution. Even if such decomposition analysis could be very useful to guide policy makers in any attempt to reduce disparities in human development within a country, we are not aware of any methodology providing that information.

Third, as these HDI indicators are constructed on the basis of household surveys alone, it is generally not possible to estimate their distribution in such a way that they are statistically representative for sub-national geographical units (e.g., state, province, municipality and so on) because of large sampling variation. Yet, this more detailed spatial information is crucial for a variety of purposes ranging from academic research to the design of development policies. The design of fine-tuned policy instruments can be particularly useful to deal with clusters of poverty or underdeveloped regions that are otherwise concealed under national averages.[2]

In order to overcome the aforementioned limitations we propose to use census data and a straightforward methodology to construct a "HDI-like multidimensional Human Development Index" defined at the municipal level. Our approach allows exploring the distribution of human development with unparalleled geographical coverage and detail, so it has the potential of being extremely useful for academics and policymakers alike. Among other things, it opens up the possibility of monitoring the evolution of key welfare indicators at very low aggregation levels and complementing that wealth of information with a vast array of Geographic Information System (GIS) tools commonly used by regional planners (see, for instance, Anselin, Sridharan, & Gholston, 2007). It is important to emphasize at the outset that the municipal-level Human Development Index proposed in this paper *resembles* the "classical" HDI, but these measures are not exactly the same because they are based on different indicators. In this sense, one should bear in mind that our "HDI-like" indices are not strictly comparable with the official HDI published yearly in the HDRs, even if they all proxy the same underlying concept (i.e., human development). Lastly, we also present a simple method that allows decomposing human development inequality according to the corresponding contribution of each of its three subcomponents. This kind of decomposition could be particularly useful to identify the components that might deserve priority attention in the attempts to reduce inequality in human development distributions.

The municipal-based HDI-like index proposed in this paper is an attempt to unfold the spatial dimension in the human development distribution, so it should be seen as a complement to the aforementioned subgroup-specific HDIs recently proposed in the literature. However, the attempt to estimate human development at the municipal level is certainly not new. For instance, the Mexican "Consejo Nacional de Población" (CONAPO) used the poverty mapping methodology (see Endnote #2) to generate a municipality-based Human Development Index (CONAPO, 2001). Analogously, United Nations Development Program (UNDP) Brazil used similar methodologies to generate a Human Development Atlas that estimates a municipal HDI for the year 2000 (see http://www.pnud.org.br/atlas/). Other UNDP national offices have generated their own HDI estimates at sub-national levels

(e.g., Bolivia, El Salvador). However, these exercises typically use country-specific methodologies that render international comparisons particularly difficult—if feasible at all. The main aim of this paper is to propose a simple methodology that can be easily reproduced in a wide variety of settings to generate human development estimates at very low aggregation levels on the basis of census data alone. The simplicity of the methods presented here allows replicating our results for any country with census data satisfying some minimal requirements, so they can potentially be the catalyst for future research on within- and between-country inequality in human development. In this respect, Africa is the continent that could benefit the least from the methodology presented in this paper given the scarcity of census data in that region of the world.

The rest of the paper is organized as follows. Section 2 presents the methodology we have followed to construct our municipal-based HDI. Section 3 applies that methodology to illustrate the evolution of human development in Mexico during 1990–2010. We conclude in Section 4.

2. METHODOLOGY

In this section we present the methodology used to estimate human development levels at municipal scale using census data. As is well known, the HDI has three components: health, education and standard of living. Rather than coarsely mimicking the original HDI and using exactly the same variables initially defined at the national level—that approach would force us to rely on estimation and imputation methodologies—we find it more appropriate to adapt the methodology by picking other variables that are more meaningful at municipal level. The choice of municipality as unit of analysis has been basically determined by data constraints. Ideally, one would like to have indicators at the lowest possible aggregation level—i.e., the individual—but census data have certain limitations in this respect. While it is possible to construct reasonably good education indicators at individual level and standard of living indicators at the household level, census data just allow constructing reasonably good health indicators at municipal level (see below). An adverse implication of working at municipal level is that intra-municipal variability in human development is lost.[3] On the positive side, the exhaustiveness of census data allows estimating the spatial distribution of human development levels with unprecedented geographical coverage.

(a) *Health*

This is by far the most difficult component to estimate at the individual or household level, since direct health information is typically unavailable for most census data. While there might be some country-specific exceptions, there are no health questions routinely collected in the census questionnaires that can serve the purpose of obtaining estimates in such detail. Similar difficulties have also been encountered by other attempts to construct household-based health indicators (see Harttgen and Klasen (2011b) in the context of the HDI and Alkire and Santos (2010) for multidimensional poverty indices). At municipal level and for larger geographical units, there are well-known indirect estimation techniques based on two questions routinely collected in census questionnaires concerning child survivorship[4] that can be used to generate health estimates. These methods, which were largely developed by William Brass, basically use information on child survivorship to estimate probabilities of dying at age x ($q(x)$) which can