Title: Human Development Index-like Small Area Estimates for Africa computed from IPUMS-International integrated census microdata

Authors: Iñaki Permanyer[1], Albert Esteve-Palos[1], Joan Garcia[1], and Robert McCaa[2]
[1]Demographic Studies Center, Autonomous University of Barcelona
[2]University of Minnesota Population Center. Corresponding author: rmccaa@umn.edu

Abstract

This paper analyzes 24 African census samples from 13 countries available via the African Integrated Census MicroData website (http://ecastats.uneca.org/aicmd) to illustrate how microdata may be used to assess development and pinpoint basic human needs at local administrative levels over time. We calculate a Human Development Index-like measure for small administrative areas, where much of the responsibility lies for executing policies related to health, education and general well-being. The methodological proposals introduced in this paper are particularly pertinent for the case of Africa. While it is true that data for much of Africa is not appropriate for economic growth rates or per capita income estimates, the analysis in this paper demonstrates that they are good enough for many other purposes. Indeed, a major aggravating problem that contributes to the 'African statistical tragedy' is the lack of accessibility to existing census microdata. This paper aims to illustrate the usefulness of census microdata – which are vastly underutilized in Africa – and hopefully contribute to make them more transparent and freely accessible.

Keywords: Human Development Index, census microdata, measurement, growth, inequality, Africa

1. Introduction

In the last few years, the measurement of living conditions in Africa has been the object of a vigorous debate. In the economic sphere, it is acknowledged that the statistical foundations of recent growth in GDP per capita or poverty reduction estimates are quite weak – so much so that some scholars speak about a "statistical tragedy" (Devarajan 2013) or about the "poverty" of such numbers (Jerven 2013). Even if some researchers have attempted to circumvent this problem using asset indices that are commonly available in household surveys as a proxy for economic performance or income poverty (Sahn and Stifel 2000, 2003, Young 2012), it is not yet entirely clear whether such methodologies are appropriate to fulfill the goals for which they were designed. While Sahn and Stifel (2000, 2003) and Young (2012) use asset indices to show that poverty reduction and economic growth in Africa might be much higher than suggested by the figures in international income statistics (*e.g.*: the Penn World Tables or the World Bank's World Development Indicators), Harttgen et al (2013) claim that such results might be biased because the relationship between growth in assets and growth in incomes or consumption is extremely weak. [[[Endnote#1]]] As suggested by Harttgen et al (2013), asset indices might not be suitable proxies for income trends but might indeed be more useful as ingredients for a broader multidimensional index of well-being – an approach that will be followed in this paper.

Despite the huge importance of getting the economic performance estimations right, it is nowadays widely agreed that other dimensions should be taken into account as well in the assessment of living conditions – not only in Africa but elsewhere. In this respect, the United Nations Development Program (UNDP) yearly publishes since 1990 the values of the Human Development Index (HDI): a composite statistic of life expectancy, education and income levels attempting to give an overall picture of countries' well-being levels that goes beyond the strictly economic dimension. [[[Endnote#2]]] Unfortunately, global indices like the HDI are only representative at the national level, and might thus actually hide important inequalities within countries. In order to address this problem, Permanyer (2013) recently proposed a methodology to construct a Human Development Index-like measure for small areas (typically municipalities, henceforth denoted as MHDI). Unlike UNDP's classic HDI, Permanyer's measure is computed solely from census microdata and therefore, when the data are accessible, may be easily calculated for small administrative areas, where much of the responsibility lies for executing policies related to health, education and general well-being. Summarizing the UNDP's HDI at the national level has its attractions, but the MHDI exposes inequalities that exist within country at the same time that it can offer a summary statistic for an entire country. [[[Endnote#3]]] Although somewhat different from the classic HDI, the MHDI attempts to construct human development indicators defined below the country level, using a single source produced by the National Statistical Office, census microdata. [[[Endnote#4]]]

One of the most attractive features of the use of census microdata is the possibility of disaggregating national-level averages and exploring the distribution of human development and its components with unprecedented geographical detail – something which is not feasible using household surveys. In particular, the availability of census microdata allows pinpointing those administrative units leaping ahead or lagging behind in the pace of progress toward well-being. Therefore, the MHDI methodology can be particularly useful for policy-makers in need of a detailed indicator computed from a single source that is comparable in space and time. [[[Endnote#5]]]

This paper analyzes 24 African census samples from 13 countries available via the African Integrated Census MicroData website (http://ecastats.uneca.org/aicmd) to illustrate how microdata may be used to assess development and pinpoint basic human needs at local administrative levels over time. The MHDI that will be used here is a composite with three components: health (proportion surviving of live-born children), education (a composite of literacy and primary education completion), and standard of living (amenities or assets, such as potable water, waste disposal and electricity). Our empirical findings suggest that the human development distribution is not only very heterogeneous across the three components but also between and within the countries included in this study. In addition, for countries with two or more suitable sets of census microdata, we are able to make comparisons over time. The dynamic study of the evolution of the three components reveals that the corresponding growth rates are not only very different among them but also quite distinct with respect to well-known economic performance indicators reported in international income statistics, a result which is in line with the findings of Harttgen et al (2013).

The results presented in this paper include 13 African countries only, so they are by no means representative of the whole continent. Unfortunately, even if in each decade since the 1970s at least 80% of the continent's population was censused, much of the census microdata are not available for scientific or policy research. In this context, the current paper identifies yet another problem that contributes even further to the "African statistical tragedy" announced by Devarajan (2013): the lack of accessibility to *existing* census microdata.

The paper is structured as follows. In section 2 we present the definitions, the data and the methodology that has been used to construct the MHDI for 24 census samples in 13 African countries. The empirical results of our analysis are shown in section 3. We discuss the implications of our results in section 4. We conclude with a discussion of methodological, theoretical, and policy implications as well as an appeal to African statistical agencies to facilitate access to census microdata.

2. Methodology

2.1. Data

Our analysis is based on harmonized census microdata samples entrusted by the respective National Statistical Offices to the Integrated Public Use Microdata Series (IPUMS) International database (Minnesota Population Center 2012) disseminated via the AICMD portal. The dataset used here contains 24 samples from 13 countries. These samples are drawn from censuses between 1982 and 2008 (see Table 1 for details on the countries and years included in the dataset). Unfortunately, some census samples available in the IPUMS database—such as Egypt 1996 and 2006, Guinea 1983, and Senegal 1988—could not be included in our analysis because they lacked the child survivorship variables necessary to compute the health component of the MHDI.

The geographical detail available for each country is not uniform, as it depends on the density of the sample size (typically between 5% and 10%), the distribution of the population and the way in which administrative units are defined for each country (see Table 1). For the case of Rwanda the microdata are coded only at the first administrative level (i.e.: the Province level), while for Mali and South Africa indicators can be computed at the third administrative level (i.e.: districts and municipalities, respectively). For the remaining samples, indicators can be computed at the second administrative level (the specific name varies with each country). In cases where the corresponding statistical agencies permits access to complete census microdata files, it would be possible to extend the analysis presented in this paper to even lower levels with increasingly greater geographical detail.

[[[Table 1 around here]]]

It should be noted that census microdata are available for a non-random sample of countries. As rightly suggested by Jerven (2013) and Harttgen et al (2013), the countries that are more successful in economic terms (which in turn tend to be less prone to experience episodes of conflict or civil wars) are more likely to conduct large scale household surveys or censuses – an important issue that might bias the assessments of economic performance of *the whole* African continent (see Young 2012) because of the lack of available data. However, this is not the purpose of the paper. As is shown below, rather than attempting to offer a comprehensive or exhaustive picture of the whole region, we have maximized the use of currently available data to illustrate the usefulness of census microdata to uncover existing inequalities in crucial well-being dimensions within the corresponding countries. Ultimately, it is precisely the very non-random distribution of countries with available census microdata that motivates this paper – which can be seen as an attempt to contribute to making such data more transparent and accessible to an increasing number of countries.

In his recent contribution, Jerven (2013) also argues that some of the population counts that are used to generate the national figures of income, consumption or expenditure per capita might be of poor quality, therefore leading to biased estimates of these indicators. More specifically, he contends that political contestation about total population data (one of the most important variables for national policy makers) might render the data very difficult to use. While extremely important when properly estimating the national accounts, we contend that this issue is not likely to affect our results in an important way. As will be shown below, the indices we are using to construct the MHDI are defined in relative terms, that is: they count *shares* of the population satisfying different conditions (e.g.: being illiterate, having surviving children and so on). Since we are not concerned about the corresponding absolute numbers, there are no reasons to believe that our indicators will give biased results in those regions of a country that might be over or under enumerated in the census. The simplicity of the indicators we are working with (see Table 2) makes them quite robust to measurement errors (see more details in section 2.2). Census microdata have been successfully used by the

World Bank to generate detailed poverty maps of different African countries (see Endnote#5), and by the United Nations to: (i) monitor demographic trends across the continent (Gerland 2013), (ii) generate some of the indicators included in the official (i.e.: country-level) HDI. [[[Endnote #6]]]

2.2. Some basic definitions

In this section we describe the methodology used to define the MHDI. Following Permanyer (2013), the MHDI for administrative unit '*i*' is an average of the health, education and wealth components (denoted as H_i, E_i and W_i respectively), the construction of which is described as follows.

Health

The health indicator for administrative unit '*i*' is the percentage of surviving children (unconditional on surviving up to any given age) born to women in that administrative unit between ages 20-39, which is denoted by P_i. The Health Index H_i simply normalizes the values of P_i between 0 and 1. It is defined as $H_i=(P_i - P_{min})/(P_{max} - P_{min})$, where P_{min}, P_{max} are the minimal and maximal benchmark values. This is the standard normalization methodology used in the construction of the classic HDI. In our empirical results, we have chosen P_{min}=50 and P_{max}=100. The choice of P_{max}=100 is quite uncontroversial, as it is the natural upper bound that would be observed in the absence of child mortality. The choice of P_{min}=50 is slightly arbitrary but it is grounded on the following reasons: i) It is a simple rounded number that involves no truncation of the distribution of the different P_is; (ii) Lower rounded bounds like 25 or 0 would be theoretically feasible but are too far away from the actual values observed in the distribution of the P_is. Analogous criteria have been used in the construction of the HDI when normalizing life expectancy values in the health component of the index. The health index H_i is particularly suitable to estimate health conditions for small size populations. Among others, it has been used to describe socio-demographic conditions of scattered indigenous populations in Latin America (ECLAC 2010).

Education

In the original HDI definition, the education component is defined as the weighted average (2/3)·ALR+(1/3)·GER, where ALR is the Adult Literacy Rate (defined as the percentage of individuals aged 15 or more who are able to read and write) and GER is the Gross Enrolment Ratio (defined as the number of students enrolled in primary, secondary and tertiary levels of education, regardless of age, expressed as a percentage of the population of theoretical school age for the three levels). While the former indicator focuses on all adults, the latter focuses on the population in school ages. Unfortunately, it is not always possible to compute the values of ALR and GER for all countries because of data limitations. In an attempt to construct other indicators which are similar to the original ones we have made the following decisions:

i) Whenever ALR was not available in a given country, we have used an alternative definition using the variable 'Years of Schooling'. More specifically, we considered those individuals with less than five years of schooling as being illiterates. This approach has been used among others by Grimm et al (2008, 2010). In order to validate the reasonableness of this approach, we have performed a couple of consistency checks.

Consistency check #1 (macro level check): For each administrative unit where both indicators were available, we have compared their values. It turns out that the correlation coefficient between them within each census sample is extremely high, almost always above 0.95. To illustrate, we show the results at the country level for the 18 samples where both indicators were available at the same time. The results – shown in Figure 1 – indicate that both indicators are highly consistent, with a correlation coefficient equal to 0.96.

[[[Figure 1 around here]]]

Consistency check #2 (micro level check): For each administrative unit where both indicators were available, we have cross-checked the classification of individuals according to both criteria (i.e.: literate/illiterate vs. less than five years of schooling/five years of schooling or more). In most cases, the percentage of agreement between both criteria is above 90%, with a few countries having agreement rates around 85%.

ii) In order to compute GER, the ‘School Enrollment’ variable is necessary. Unfortunately, this indicator is unavailable for 5 of the 24 samples considered in this paper. Unlike the previous case in which it was possible to present an alternative way of defining literacy (see (i)), there is no clear cut way of presenting alternative definitions of GER with the available data. For this reason, and in order to maximize the geographical coverage of our analysis, we have opted for an alternative solution defining a new indicator that is somewhat similar in spirit to GER. If we define $PR+_{15\text{-}24}$ as the population aged 15-24 having at least completed primary education and $POP_{15\text{-}24}$ as the population aged 15-24, we can define the index

$$PR = \frac{PR+_{15-24}}{POP_{15-24}} \qquad [1]$$

While GER compares the number of enrolled students with respect to the population of theoretical school age, PR is simply interpreted as the proportion of population aged between 15 and 24 that has at least completed primary education in the administrative area we are dealing with. Similarly to GER, the new indicator focuses on the young population (as opposed to ALR). There are several reasons why we have opted for such an indicator. First, it is very straightforward and simple to understand. Second, it can be computed in all our samples except for one. [[[Endnote#7]]] Third, it is an indicator that is directly related to the achievement of Millennium Development Goal #2 (ensure universal primary schooling). In those countries where both GER and PR are available, we have compared their values for the different administrative units we are working with. It turns out that the correlation coefficient between them is very high, in most cases around 0.9. [[[Endnote#8]]] Therefore, the administrative unit rankings that are derived from the values of both indicators are very similar. To illustrate, Figure 2 shows the country level values of GER and PR for those census samples where both indicators are available. Again, both indicators are relatively similar and the correlation coefficient is very high (0.88).

[[[Figure 2 around here]]]

Summing up, the education index for administrative unit '*i*' used in this paper can be written as the following weighted average: $E_i=(2/3)\cdot ALR_i+(1/3)\cdot PR_i$.

Standard of living

The standard of living index for administrative unit '*i*' (W_i) is an average of a household asset index defined for all households belonging to '*i*'. Our asset indices are constructed at the household level (h) using the following aggregation formula:

$$A_h = \frac{a_{h1} + \ldots + a_{hk}}{k} \qquad [2]$$

where A_h is the asset index for household h, the $a_{hj} \in \{0,1\}$ refer to the absence/presence of asset j in household h and k is the number of assets we are taking into account. Whenever a given household '*h*' owns all assets included in the list $A_h = 1$, and when it owns none, $A_h = 0$. After computing the asset index A_h for each household in the census, our wealth index (W_i) is computed for each administrative unit '*i*' as a weighted arithmetic mean of the asset indices of the households belonging to '*i*' (each household weighted by its population share within the corresponding administrative unit).

The use of asset indices is becoming increasingly popular during the last decade (particularly after the work of Filmer and Pritchett 2001 and Filmer and Scott 2008), and their advantages and disadvantages have been widely discussed (see, among others, McKenzie 2005, Stifel and Christiaensen 2007, Filmer and Scott 2008, Harttgen et al 2013). On the negative side, the following factors have been identified. First, given the discrete nature of those indicators it might happen that observations are clustered around certain values, thus complicating the task of estimating the underlying welfare distribution (McKenzie 2005). In the context of this paper, however, this problem is unlikely to be severe because our basic units of analysis are different administrative units (not the households included in them). As shown in the third column of Figure 5, the distribution of the wealth index (W_i) across administrative units tend to be quite smooth because each observation is obtained as an average of thousands of household-level asset index observations (A_h). Second, asset indices might not correctly capture the differences between rural and urban areas (see Harttgen and Klasen 2011a, 2011b): many assets are cheaper, more easily available and more desirable in urban areas, so urban households might appear to be wealthier than their rural counterparts. Again, the fact of working at different levels of administrative units may reduce that problem to a certain extent because of the heterogeneity of households within administrative units (many of which are separated in urban and rural areas). Third, asset indices are not appropriate to measure consumption levels across heterogeneous settings, and particularly to measure trends in consumption over time. To start with, the list of assets used in the construction of A_h might not include many of the goods and services that are generally available to high-income households, therefore limiting the discriminating power of the corresponding asset index. In this context, Harttgen et al

(2013) argue that changes in preferences, prices and the public provision of certain assets through government policies (often at highly subsidized rates) are important factors that may make asset indices inappropriate for portraying consumption profiles. In addition, these authors argue that it is highly problematic to proxy a flow variable (consumption) with a stock variable (asset ownership).

On the positive side, the use of asset indices has important advantages. To begin with, reporting assets at the household level is less vulnerable to measurement errors than reporting income or consumption levels (McKenzie 2005). In addition, some studies report that asset indices are very valuable as an explanatory variable or as a means of mapping economic welfare to other living standards and capabilities such as health and nutrition (Sahn and Stifel 2003). Other authors have suggested that the gradient of the outcomes of asset indices closely follow the results obtained with per capita expenditures (Filmer and Scott 2008). As a validation exercise, Figure 3 shows a scatterplot with the values of countries' real consumption per capita together with the national-level average values of an asset index for a group of 102 developing countries all over the world. [[[Endnote#9]]] As can be seen, asset indices are quite effective in discriminating countries along the income spectrum: higher income countries clearly tend to have a higher national-level average asset index and vice-versa, so both measures tend to rank countries in a quite consistent way (both in Africa and in the rest of the world). Moreover, asset indices are good proxies of long-term living standards as opposed to current income and/or consumption because the former are less vulnerable to short term economic shocks and fluctuations over time – something which is clearly in line with the conceptual foundations of UNDP's HDI. Finally, as is suggested by Harttgen et al (2013:S60), the possession of assets is welfare enhancing, so asset indices can be an extremely useful ingredient to be incorporated in multidimensional well-being indices. While inappropriate for capturing current micro or macroeconomic levels and trends, asset indices seem to offer a viable – though imperfect – way of assessing long-term living standards.

[[[Figure 3 around here]]]

The availability of household assets or amenities questions varies widely across African censuses, an issue that has imposed serious challenges for developing comparable measures of standard of living across time and space. Given the aforementioned questionnaire variability, we have opted for a two-pronged strategy to maximize the use of data. On the one hand, and in order to ensure international comparability, we have defined a standard of living indicator that included all assets that were available in the different questionnaires *at the same time*. This has produced a crude asset index consisting of three components only: access to clean drinking water, access to electricity and ownership of an improved sanitation facility. [[[Endnote#10]]] This simple asset index will be referred to as 'core standard of living index' or 'core wealth index' and will be denoted as W_i^{Core}. On the other hand, and given the crudeness of the core standard of living index, we have introduced a country specific asset index – denoted as W_i^{CS} – that might better proxy the wealth distribution in that country. This index is used for within-country comparisons only. The items introduced in each country specific standard of living index are shown in Table 2. In the construction of the different asset indices we have chosen an equal weighting scheme (as done by many others, e.g.: Montgomery et al 2000, Case, Paxson and Ableidinger 2004, Hohmann and Garenne 2010, Permanyer 2013). This way, the meaning of the indices is crystal clear: they simply count the proportion of owned assets.

For the sake of completeness, Figure 4 plots the joint values of W_i^{Core} and W_i^{CS} computed at the country level. Interestingly, it seems that both indices tend to rank order African countries included in this study in a quite similar way (the correlation coefficient equals

0.95). Therefore, despite the crudeness of its definition, the values of the core wealth index may not be overly misleading when estimating the underlying wealth distribution.

[[[Figure 4 around here]]]

[[[Table 2 around here]]]

The municipal-based HDI

After computing the three components of the index, the MHDI for administrative unit '*i*' is finally defined as the arithmetic mean $(H_i + E_i + W_i)/3$. It should be noted that since 2010, the official HDI is calculated using the geometric mean $\sqrt[3]{H_i \cdot E_i \cdot W_i}$. Both approaches have their advantages and disadvantages. On the one hand, the multiplicative HDI was introduced to reward countries with balanced (*i.e.*: similar) distributions across components and penalize those countries with unequal achievements. However, the multiplicative index drops to zero whenever any of its components is equal to zero – regardless of the value of the other two. This problem is more likely to be found when the units of analysis are very small, as it becomes increasingly possible that some components of the index equal zero. On the other hand, the additive HDI is insensitive to the extent to which achievements across dimensions are balanced or not. However, it does not have the boundary problems of its multiplicative version and – importantly for the purposes of this paper – it allows knowing the contribution of the different components to overall inequality in human development. Borrowing techniques from the economics literature that attempted to answer questions like 'what is the contribution of different income sources to total income inequality?' (*e.g.*: Shorrocks 1982, Lerman and Yitzhaki 1985), this paper uses a methodology that allows making statements like ‘In country *X*, the health component accounts for a *Y*% of the observed inequality in human development’ (the technical details of such methodology are presented in Appendix I).

3. Empirical results

In this section we present the empirical findings of the paper regarding the MHDI distribution across 13 African countries. We start exploring distributions within countries first and then proceed with comparisons between countries. In addition, we decompose inequality by the factor components methodology presented in Appendix I.

3.1. Within country analysis.

When the health, education, wealth and human development indicators (i.e.: the H_i, E_i, W_i and $MHDI_i$) are available for each administrative unit we are working with, it becomes possible to explore their distribution across the entire country with great geographical detail. In addition, when more than one census is available for the same country, it is particularly interesting to examine the evolution of the human development distribution and its three components over time. In our study, there are eight countries with more than one census (Kenya, Malawi, Mali, Morocco, Rwanda, South Africa, Tanzania and Uganda). The results are shown in Figure 5: for each of those countries we plot the distributions of human development together with the health, education and wealth components across the corresponding administrative units. Since comparisons in this subsection are within countries we use the country-specific standard of living indices (W_i^{CS}), which are expected to better capture the underlying wealth distribution than the one derived from the values of the core indicator W_i^{Core}. [[[Endnote#11]]] Therefore, it is important to highlight that while the distribution of the health and education components shown in Figure 5 are comparable across countries, the wealth and overall human development distributions are not. The corresponding comparable results across countries will be presented in section 3.2.

[[[Figure 5 around here]]]

Inspecting the shape of the density functions shown in Figure 5, one can see that there are important variations across countries. Rather than observing the traditional unimodal and highly skewed shapes that characterize income distributions, many of the distributions shown in Figure 5 have several local modes. This suggests that the levels of inequality and polarization in those countries can be very large, an issue that will be explored in more detail below. We hypothesize that the existence of local modes in those distributions might be partly attributable to the urban - rural divide: urban households tend to own more assets and their inhabitants are more likely to enjoy the benefits of nearby health and education facilities (this effect, however, might be partially blurred by the fact that the administrative units we are working with might contain both rural *and* urban households). Finally, if one compares the spread of the distributions within countries over time, no substantial changes seem discernible at first sight. At the end of the section 3.2, we will quantify the extent of inequality in all those distributions.

As can be seen from Figure 5, most distributions tend to shift to the right (*i.e.*: improve over time), an encouraging result for the corresponding countries. However, there are important exceptions to these trends. The Rwandan health distribution deteriorates from 1991 to 2002, a phenomenon that can be attributed to the mass killings that took place in the country in 1994. As a consequence, the overall MHDI distribution does not show signs of clear improvement either (despite improvements in the EI and WI distributions). In addition, the health distribution reported for South Africa also deteriorates from 2001 to 2007, a result that is in line with official figures of declining life expectancy reported in that country and which can be attributed to a large extent to the high prevalence of HIV/AIDS.

It is important to highlight that when examining such directions of change over time in Figure 5, one is only looking at the *overall* position of the corresponding distributions. In this way, however, one looses track of the administrative unit level changes that have occurred between the two moments in time. As a matter of fact, any change in the overall shape of a distribution can be the result of many different combinations of individual (*i.e.*: administrative unit level) changes. [[[Endnote#12]]] For this reason, it is essential to keep track of the changes observed for each administrative unit one is working with. Figures 6.1 to 6.5 show the distribution of the administrative units' annual growth rates for the different MHDI components we are considering in this paper (H_i, E_i and W_i) for the cases of Malawi, Morocco, Rwanda, Mali and South Africa. [[[Endnote#13]]] As can be seen, the administrative units' growth in the three components of the index has been quite heterogeneous and the patterns are highly country-specific. For all countries considered here, the growth rates patterns of the different components are quite different from one another: the shapes and positions of the corresponding density functions are quite different and they do not necessarily overlap. To illustrate: in the cases of Rwanda and South Africa there are clear deteriorations in the health component but moderate and large improvements in the education and wealth components respectively. Alternatively, in Malawi (resp. Mali) we observe deteriorations in the wealth (resp. education) component but improvements in the other two. Lastly, in Morocco there are generalized improvements in the three components but the patterns are also quite different from one another. In order to frame our results within the recently debated relationship between economic growth and household assets growth (*e.g.*: Younger 2012, Harttgen 2013), Figures 6.1 – 6.5 also plot the corresponding levels of annual growth in GDP per capita reported in the Penn World Tables

8.0 for comparative purposes. For the small sample of countries considered in this study, economic growth does not seem to be related to national level assets' growth in a clear way: for Mali, Morocco and South Africa the results are roughly similar while for Malawi (resp. Rwanda), the value of the former is clearly above (resp. below) the latter. While it is clear that such results cannot be generalized to the African continent because of the limited number of countries, the apparent lack of relationship between both indicators seems to be consistent with findings reported in Harttgen et al (2013).

[[[Figures 6.1 – 6.5 around here]]]

3.2. Across country analysis.

We will now compare the distribution of human development and its components across countries. For that purpose, we will make use of the 'core' wealth index W_i^{Core} that includes the same assets for all the countries included in this study – thus ensuring cross-country comparability. We start by examining the population weighted country-level average of our MHDI indicator and its health, education and standard of living components, which are shown in Table 3. The country average MHDI values differ greatly, ranging from 0.17 (observed in Mali 1987) to 0.84 (South Africa 2007). Being an average of three different components, it is also important to explore them separately. Interestingly, the three components of the MHDI behave quite differently. The values of the country average health index range from 0.31 (Mali 1987) to 0.90 (South Africa 2001), those of the education index from 0.17 (Mali 1998) to 0.88 (South Africa 2007) and those of the core wealth index from an appallingly low 0.02 (Uganda 1991) to 0.75 (South Africa 2007). Taking into account the fact that the theoretical range of these indicators is the interval [0,1], the observed range of variation for each case is considerably large. This illustrates the heterogeneity that is observed among the African countries included in the analysis.

Since the MHDI is simply the arithmetic mean of the HI, EI and WI indices, it is straightforward to compute the contribution of each of these subcomponents to the aggregate value of the index. To illustrate: the percentage contribution of the health component to the aggregate MHDI value is simply computed as 100·HI/(HI+EI+WI). As is shown in Figure 7, the contribution of the three components varies greatly across countries. Figure 7 shows that as the country-level MHDI values decrease, the relative contribution of the wealth index tend to decrease as well while the contribution of the health component tends to increase. As can be seen, the percentage contribution of the three components is balanced (*i.e.*: around 33% each) only for those countries with the largest MHDI values (South Africa and Morocco). For the other countries, the MHDI values tend to be overwhelmingly accounted for by the health and education components, that is: for less developed countries, the wealth component is by large the one that is comparatively smaller with respect to the other two.

[[[Table 3 around here]]]

[[[Figure 7 around here]]]

In order to compare the results of our methodology with the official HDI results reported yearly in the Human Development Reports, the latter are also shown in Table 3. Figure 8 plots the country-level values of our MHDI indicator against UNDP's HDI. The results of this validation exercise are quite encouraging: the values of the MHDI and HDI are closely related in a linear fashion, and no large discrepancies are observed (the correlation coefficient for the values plotted in Figure 8 is very high: 0.94). The country that differs the most from the predicted linear model is Rwanda 1991, perhaps because it is the only country where the education component has been calculated using a slight variation with respect to the others and thus artificially inflated its 'true' value. As can be seen, the values of MHDI tend to be higher than those of the official HDI (the dots are mostly below the dashed equality line). This issue is not particularly troubling since neither the HDI nor the MHDI have a specific unit of measurement. Therefore, what is especially relevant is the *ordinal* information (*i.e.*: the rankings) that is derived from the values of those indices, rather than the cardinal values themselves.

[[[Figure 8 around here]]]

The country-level MHDI values shown in Table 3 and Figures 7, 8 are the result of averaging MHDI values across a large number of administrative units defined at sub-national level. Despite the interest that such country-level averages might have, the main rationale for introducing the MHDI methodology is to uncover the inequalities that are hidden behind those aggregate numbers. In order to compare not only the average value of the MHDI distribution but also its spread within the corresponding country, Figure 9 plots the density functions of the MHDI distributions for all African countries included in the analysis except for the case of Sudan (its 2008 census was conducted about a decade later than the others, an issue that seriously compromises its comparability). For those countries with several observations, we have chosen those belonging to the 2000 census round. As can be seen, there are large variations not only in the average MHDI values but also in the spread of the human development distributions within countries. [[[Endnote#14]]] Again, many of the distributions have several local modes, suggesting that the levels of inequality and polarization in the corresponding countries must be very high (this is particularly the case for Guinea, Malawi, Mali, Rwanda and Sierra Leone). At the other extreme, countries like Kenya, Morocco and South Africa have relatively smooth distributions with a relatively wide range of variation. In this context, it is particularly interesting for policy making purposes to identify the administrative units that are located in the lower and upper tails of the corresponding MHDI distributions. In the following section, we quantify more precisely the extent of inequality observed in these distributions.

[[[Figure 9 around here]]]

Inequality in human development

Table 4 shows the values of the Gini index for the MHDI, HI, EI and WI distributions. For the case of the MHDI distribution, its values range from 0.06 (observed in Rwanda 1991) to 0.27 (Mali 1987). This range of variation is relatively similar to the one observed for the Gini index of the health distributions: from 0.03 (observed in South Africa 2001) to 0.21 (Mali 1987). However, the range of variation is larger for the Gini index of the education distributions (from 0.04 in South Africa 2007 to 0.40 in Mali 1998) and even larger for the wealth distributions (from 0.16 in South Africa 2007 to 0.82 in Mali 1987). Taking into account the fact that the Gini index can only range from 0 to 1 (0 denoting complete equality and 1 complete inequality), the observed Gini values for many of the wealth distributions are appallingly high. [[[Endnote#15]]] This suggests that the three assets included in the core wealth index (access to clean drinking water, access to electricity and ownership of an improved sanitation facility) are very unevenly distributed within these countries (most probably concentrated in metropolitan urban areas). All in all, the country with highest human development inequality levels in our sample is Mali (in 1987). At the other extreme, South Africa is among the countries with lowest observed levels of inequality in human development.

Table 4 also shows the results of the inequality decomposition by factor components presented in Appendix I. As can be seen, the contributions of the three components to MHDI inequality do not show clear cut patterns. An inspection of the values of the component-specific Gini indices on the one hand and the contribution of that component to the overall inequality in human development on the other reveals that both magnitudes do not necessarily run in the same direction. In other words: high values of a component-specific

Gini index do not necessarily imply that the corresponding contribution to overall inequality in human development is also high (e.g.: Mali has a very high Gini index for the wealth component but the percentage contribution of that component to overall inequality in human development barely reaches 20%). This apparently surprising fact is attributable to the weak correlation structure of the data (i.e.: the administrative units' rankings within a given country can be quite different depending on the specific indicators that are used to rank them, see Shorrocks 1982 and Lerman and Yitzhaki 1985 for details).

As is shown in Table 4, the health component is the one that tends to contribute the least to observed MHDI inequality levels, but there are important exceptions (Mali 1987 and 1998, Rwanda 1991 and 2002). The education and wealth components tend to dominate the contribution to overall MHDI inequality, but, again, it is difficult to discern simple patterns in the data. Comparing the results of Table 4 with those of Table 3, it turns out that the contribution of the wealth component to overall human development inequality tends to be larger for countries with higher human development *levels*. For instance, South Africa and Morocco, the countries with highest human development levels in our study, are the countries where the contribution of the wealth component to overall inequality is the largest (above 50%). At the other extreme, in Mali, Rwanda and Uganda (the countries with lowest human development levels in our study), the contribution of the wealth component is at its lowest values (around 20%). These results, together with the examination of the component-specific distribution graphs shown in Figure 5, lead us to hypothesize that as countries progress towards higher human development levels, the education and health distributions tend to become more homogeneous, therefore increasing the contribution of the wealth component to overall human development inequality. [[[Endnote#16]]] However, this is a challenging issue that is beyond the scope of this paper and should be explored in future research.

[[[Table 4 around here]]]

4. Discussion and concluding remarks

In this paper we use new measurement techniques recently proposed by Permanyer (2013). The measures add human development indices at small aggregation levels to an operational toolkit that can be used by scholars, researchers, practitioners, national and international institutions and policy makers alike. As argued in the paper, access to census microdata is extremely important for a variety of purposes ranging from academic research to the design of development policies. On the academic side, the lack of reliable data at sub-national levels is a major hurdle that critically undermines the possibility of (i) assessing the large, unmeasured heterogeneity within countries; and (ii) empirically testing alternative theoretical efforts proposed in different disciplines of the social sciences that aim to establish formal linkages and interactions between variables operating at the micro and macro aggregation levels. From the policy-making perspective, there is a need for more accurate information that can be used for the design and evaluation of public policy and to reduce the risk of falling into the ecological fallacy trap. The design of fine-tuned policy instruments can be particularly useful to identify and monitor the evolution of small administrative units that are otherwise concealed by national averages.

In this respect, our empirical findings clearly illustrate the necessity of going beyond national-level averages: they allow visualizing and quantifying the extent of inequality in human development and the corresponding contribution of its different subcomponents – an important issue that so far remains largely unexplored. Among other things, our results indicate that the distributions of the three basic components of the MHDI across administrative units are quite heterogeneous and the corresponding inequality levels are high for all countries considered in this study – particularly for the wealth component. Roughly speaking, the contribution of that component to overall human development inequality tends to be higher for those countries with higher human development levels. The sizable amount of human development inequality observed in our samples confirms the rationale for constructing inequality-sensitive well-being indices that penalize overall achievement levels for unequal distributions at sub-national levels (as suggested, for instance, by Alkire and Foster 2010 with the definition of the Inequality-adjusted Human Development Index). For those countries with more than one available census, we have been able to explore the dynamics of the different MHDI components over time. Again, the patterns of growth are highly heterogeneous and depend very much on the component and the country one is dealing with: in most cases, some components tend to grow over time while others decline during the same period. Even if our findings are by no means representative of the whole continent, they are consistent with the weak relationship between income growth and asset growth reported by Harttgen et al (2013). In this context, our results suggest that improvements in well-being and improvements in economic performance do not necessarily go hand in hand – at least for the countries and periods considered here. This important issue, however, remains hypothetical and should be explored in much more detail in future research.

The methodology presented in this paper is far from perfect and various caveats should be borne in mind. First, some might argue that the indicators used in the construction of the MHDI are too crude and simple, so they might fail to give a faithful portrait of the well-being distribution in a given country. While it is true that other variables typically available in detailed household surveys can be more useful in that endeavor, the use of simple variables included in population censuses has other advantages that can outweigh such problem. On the one hand, the universal geographical coverage of census data allows a detailed administrative units' level analysis that is not feasible when using household surveys because of extreme sampling variability. On the other hand the variables included in censuses are good enough to measure and monitor the levels of some basic human capabilities that are needed to lead a minimally decent life. Indeed, the three components used in this paper greatly discriminate across the administrative units we have been working with: as shown in our empirical results the welfare distributions derived from the values of our indices are not clustered around a certain point, but tend to be quite disperse across the range of admissible values. Regarding data quality issues, we should highlight that the simplicity of the variables we are working with makes them much less prone to the measurement errors that usually afflict economic performance indicators like reported income / expenditures or GDP per capita. In addition, in a context where ordinal information is all that is needed to identify the administrative units that are doing better or worse in terms of well-being, the existence of measurement errors is less likely to overly influence our results. Last but not least, the fact that our basic indicators are constructed on a relative basis (as opposed to an absolute one, *i.e.*: we are measuring the *percent* of surviving children, adult illiterates, young people completing primary education and so on, not their absolute numbers) makes them much less vulnerable to the eventual problem of over or under enumerating the population counts of the administrative units we are working with (a problem which might have occurred to a certain extent in the censuses considered here).

The methodological proposals introduced in this paper are particularly pertinent for the case of Africa. While Devarajan (2013) and Jerven (2013) are correct when they conclude that data for much of Africa is not appropriate for economic growth rates or per capita income estimates, the MHDI analysis in this paper demonstrates that they are good enough for many other purposes. As a matter of fact, a major aggravating problem that contributes to the ‘African statistical tragedy’ is the lack of accessibility to existing census microdata (indeed, this is the main reason why our results cannot yet be extended to the whole continent). This paper aims to illustrate the usefulness of census microdata – which are vastly underutilized in Africa (Alderman et al 2003, p.193) – and hopefully contribute to make them more transparent and freely accessible.

In 1999, the Minnesota Population center began a global initiative, IPUMS-International, offering free, internet access (www.ipum.org/international) to integrated census microdata for researchers world-wide under a single license agreement with National Statistical Office partners. Microdata for 74 countries, totaling 540 million person records (234 samples), are accessible for research (June 2013). IPUMS-International disseminates microdata encompassing 80% of the world's population, but the coverage for Africa, at 42%, is barely half. Africa is under-represented in the database, not only due to a slow start out of deference to the African Census Analysis Project, which began a census microdata initiative for the continent somewhat earlier (Zuberi and Bangha 2012), but also because African statistical offices are exceedingly reluctant to allow outside access to the microdata. Nonetheless, as this study has shown, microdata for fifteen African countries (29 censuses, 55 million person records) are currently available. Integration work is underway for another eleven countries, but some very important nations—Nigeria, Algeria, Cote d'Ivoire, Zimbabwe, Burundi, etc.—are not yet participating in the initiative (see Appendix II, table A1). [[[Endnote#17]]] Africa has a treasure trove of census microdata. It is unfortunate, unlike other regions of the world, such a large fraction of the continent's census microdata remains inaccessible.

Endnotes

Endnote#1: *Inter alia*, Harttgen et al (2013) point to changes in preferences, prices and the public provision of certain assets through government policies (often at highly subsidized rates) as possible sources of the aforementioned bias.

Endnote#2: In the same line, the UNDP also publishes since 2010 the new Multidimensional Poverty Index (MPI): an attempt to capture countries' poverty levels on the basis of deprivations in the education, health and economic dimensions.

Endnote#3: In recent years, there have been different attempts to construct a national level HDI that is sensitive to internal inequality and/or association between dimensions (e.g.: Foster, Lopez Calva and Szekely 2005, Seth 2009, Alkire and Foster 2010). The approach taken in this paper is different: rather than summarizing detailed information into an aggregate measure, we have emphasized the importance of exploring the distribution of human development at low aggregation levels.

Endnote#4: Other conceptually related approaches are those of Grimm et al (2008, 2010) who present an HDI for different income quintiles; Harttgen and Klasen (2011a), who calculate the HDI separately for internal migrants and for non-migrants and Harttgen and Klasen (2011b), who define a household-based Human Development Index. Since these indicators are constructed on the basis of household surveys alone, it is 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 sampling variability.

Endnote#5: In an attempt to have high-precision welfare estimates at very low aggregation levels the World Bank has been recently using the so-called "poverty mapping" techniques (see Elbers, Lanjouw and Lanjouw (2003), Bedi, Coudouel and Simler 2007), which have been criticized by Tarozzi and Deaton (2009) and Tarozzi (2011) because the underlying assumptions upon which they are based might not be necessarily satisfied in practice. While both the approach presented in this paper and the poverty mapping methodology attempt to construct welfare indicators with high geographical detail, some differences are worth pointing out. On the one hand, poverty mapping techniques require having a census *and* a household survey carried out the same year, while the methodology presented here is based on census data alone. On the other hand, poverty mappings generate estimates of income or consumption levels, while our methodology includes other non-monetary dimensions like education and health. All in all, these complementary approaches make up a valuable new component for the contemporary analysts' and policy-makers' toolbox that is extremely useful to help inform current development debates.

Endnote#6: As is later shown in Figure 8, a calibration exercise comparing the national-level average of our measure against the conventional HDI shows that both measures are highly consistent and tend to rank the different countries considered in this study in the same way.

Endnote#7: In the case of Rwanda 1991, that information is not available. In order to include that country in our sample, we have defined a simpler version of the education index, in which we only included the ALR. Since this compromises the comparability of that specific country, special caution should be exercised when interpreting the corresponding results.

Endnote#8: The big exception to that rule was found in South Africa, which has the highest levels of school attendance, primary education completion and human development among the countries included in this paper. In that case the correlation coefficient is around 0.3.

Endnote#9: We have computed an asset index for the 102 countries with a Demographic and Health Survey (DHS) that included the following list of assets: 1. Electricity: The household has electricity; 2. Sanitation (toilet facility): The household sanitation facility is improved and not shared with other households; 3. Water: the household does have access to clean drinking water, or clean water is less than 30 minutes walking from home; 4. Floor: The household has no dirt, sand or dung floor; 5. Radio: The household has a Radio; 6. TV: The household has a TV; 7. Telephone: The household has a Telephone; 8. Refrigerator: The household has a Refrigerator; 9. Bike: The household has a Bike; 10. Motor vehicle: The household has a Motor vehicle (Motorbike, Car, Truck). The same weights have been used for all variables, but analogous results arise when we use alternative systems of weights. In this validation exercise we have chosen the DHS for the great comparability of questionnaires across countries.

Endnote#10: As can be seen in Table 2, there are a couple of exceptions to that rule. For the case of Mali there is no information regarding the water supply and for the case of Malawi 1987 there is no information on access to electricity. In those cases, the corresponding asset index is based on the remaining two components only. Since this compromises the comparability of those specific countries, special caution should be exercised when interpreting the corresponding results.

Endnote#11: It should be pointed out that for the cases of Malawi and South Africa, we only show the results corresponding to their last two censuses. Both countries have a third census which, unfortunately, does not have the same list of household assets as the other two, so they are not strictly comparable. Morocco is the only country with the same list of variables for the three available censuses, so the corresponding results are shown for all of them.

Endnote#12: To illustrate: for any given distribution, *any* permutation of its values across the different administrative units one is working with will result in another distribution that coincides exactly with the original one. This illustrates the fact that many possibly different changes at the *individual* level might lead to the same observed changes for the *overall* distribution.

Endnote#13: Kenya, Tanzania and Uganda have not been included in this analysis because the number of administrative units of these countries has greatly increased between the two available censuses, a matter that greatly difficults the estimation of the different components' growth at sub national levels.

Endnote#14: It is important to highlight that the level of geographical disaggregation is not the same for all countries (see Table 1). Therefore, countries with greater geographical detail are likely to exhibit larger spread in their MHDI distribution. This should be borne in mind when comparing distributions' spread.

Endnote#15: The same observation as those presented in footnote #14 apply.

Endnote#16: Albeit in a completely different geographical context, the results shown in Permanyer (2013) for Mexico are in line with this hypothesis.

Endnote#17: The Integrated Household Survey Network has supported capacity building in National Statistical Offices to facilitate access to microdata. Unfortunately, while the project appears to be successful—in Africa websites hosted by national statistical offices number in the dozens—as a matter of fact few are functional. An exhaustive 2009 study reported that only four African statistical offices provided relatively obstacle free access to microdata (Woolfrey 2009).

Appendix I

Inequality decomposition by factor components

Following Permanyer (2013), we briefly present the methodology used in this paper to compute the contribution of the different components to overall inequality in human development. For each administrative unit '*i*' let Y_i, H_i, E_i and W_i be the corresponding human development, health, education and wealth indices. In case of additive human development indices we have that

$$Y_i = \frac{H_i}{3} + \frac{E_i}{3} + \frac{W_i}{3} \qquad [3]$$

The distribution of human development, health, education and wealth indices will be denoted as Y, H, E and W respectively. According to Shorrocks (1982:195), if the human development distribution is ordered so that $Y_1 \leq Y_2 \leq \ldots \leq Y_n$, then the corresponding Gini inequality index can be written as

$$G(Y) = \frac{2}{n^2 \mu_y} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) Y_i \qquad [4]$$

where n is the number of administrative units we are taking into account and μ_y is the mean of the human development distribution. Plugging equation [3] into equation [4] we have

$$G(Y) = \frac{2}{n^2 3\mu_y} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) (H_i + E_i + W_i) = \frac{\mu_h}{\mu_y} \overline{G}(H) + \frac{\mu_e}{\mu_y} \overline{G}(E) + \frac{\mu_w}{\mu_y} \overline{G}(W) \qquad [5]$$

where μ_h, μ_e and μ_w are the means of the health, education and wealth distributions and

$$\left. \begin{aligned} \overline{G}(H) &= \frac{2}{3n^2 \mu_h} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) H_i \\ \overline{G}(E) &= \frac{2}{3n^2 \mu_e} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) E_i \\ \overline{G}(W) &= \frac{2}{3n^2 \mu_w} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) W_i \end{aligned} \right\} \qquad [6]$$

which are known as the pseudo-Ginis for factors H, E and W respectively (see Shorrocks 1982:196 and Lerman and Yitzhaki 1985:152). Equation [5] shows a natural additive decomposition of the Gini index where the contribution of the H, E and W components is clearly established.

Appendix II

Table A1

Is Africa's statistical tragedy due to the absence of census data, or to the lack of access to microdata?
Current status of census microdata entrusted to the IPUMS-Africa project

	IPUMS-Africa		Census Round (bold = microdata exist)					
Country	**Entrusted**	**Need**	**2005-14**	**1995-04**	**1985-94**	**1975-84**	**1965-74**	**1955-64**
A. Microdata for 17 countries (47 censuses; 35 integrated and disseminating)								
Burkina-Faso	**3**	Complete	**2006**	**1996**	**1985**	1975		
Cameroon	**3**	Complete	**2005**			**1987**	**1976**	
Egypt	**3**	Integrate 1986	**2006**	**1996**	**1986**	1976	1966	
Ghana	**3**	Integrating 2010	**2010**	**2000**		**1984**	1970	
Guinea	**2**	Complete	2013	**1996**		**1983**		
Kenya	**5**	Complete	**2009**	**1999**	**1989**	**1979**	**1969**	
Malawi	**3**	Complete	**2008**	**1998**	**1987**	1977	1966	
Mali	**4**	Integrating 2009	**2009**	**1998**	**1987**	**1976**		
Morocco	**3**	Complete	2014	**2004**	**1994**	**1982**	1971	1960
Rwanda	**2**	Complete	2013	**2002**	**1991**	1978		
Senegal	**3**	Integrate 1976	2013	**2002**	**1988**	**1976**		
Sierra-Leone	**1**	Complete	2014	**2004**	1985		1974	1963
South-Africa	**3**	Need 2011	**2011/7**	**2001**	**1996**	**1980, 5**	**1970**	1960
South-Sudan	**1**	Complete	**2008**					
Sudan	**4**	Integrate 1973-93	**2008**		**1993**	**1983**	**1973**	
Tanzania	**2**	Need 2012	2012	**2002**		**1988**	1978	1967
Uganda	**2**	Complete	2013	**2002**	**1991**	1980		1969
B. 11 Countries (29 datasets) entrusted to IPUMS-International. Integrating (bold); awaiting most recent census								
Benin	**3**	Integrating	2013	**2002**	**1992**		**1979**	1961
Botswana	**3**	Need 2011	**2011**	**2001**	**1991**	**1981**	1971	1964
Chad	**1**	Need 2009	**2009**		**1993**			1962
Ethiopia	**3**	Integrating	**2007**		**1994**	**1984**		
Lesotho	**2**	Integrating	**2006**	**1996**	1986	1976	1966	
Liberia	**2**	Integrating	**2008**				**1974**	
Madagascar	**1**		2013		**1993**		1975	1966
Mauritius	**2**	Need 2010	**2010**	**2000**	**1990**	**1983**	1972	1962
Mozambique	**2**	Integrating	**2007**	**1997**		1980	1970	1960
Niger	**2**	Need 2009	**2009**	**2001**	**1988**	1977		
Nigeria GHS (NBS)	**5**	Need 2012	**2006-11**	1996	1990/2			
Zambia	**3**	Integrating	**2010**	**2000**	**1990**	1980		1969**?**
C. Agreement signed, but no microdata sets entrusted: 4 countries								
Cape Verde	**0**	Need all	**2010**	**2000**	**1990**	1980		
Central-African- Rep.	**0**	Need all	2013	**2003**		**1988**	1975	1960
Cote-d'Ivoire	**0**	Need all	2013	**1998**	**1988**	1975		
Guinea-Bissau	**0**	Need all	**2009**		1991		1979	1960
D. No agreement signed: 16 countries								
Algeria	**0**	Need all	**2008**	**1998**	**1987**	1977	1966	
Burundi	**0**	Need all	**2008**		1990	1979	1970	
Comoros	**0**	Need all	2013	**2003**	**1991**	1980		
Congo-Republic	**0**	Need all	**2007**	1996		1984	1974	1960/61
Djibouti	**0**	Need all	**2009**					
Equatorial Guinea	**0**	Need all	2013	**2002**	1994	1983		
Gabon	**0**	Need all	2013	2003	1993			
Gambia-The	**0**	Need all	2013	**2003**	**1993**	**1983**	**1973**	1963
Libya	**0**	Need all	2005	1995		1984	1973	1964
Mauritania	**0**	Need all	2013	**2000**	**1988**	1977		
Namibia	**0**	Need all	**2011**	**2001**	**1991**	1981	1970	1960
Nigeria: Census (NPC)	**0**	Need all	**2006**		**1991**	1980	1973	1963
São Tomé and Principe	**0**	Need all	2012	2001	1991	1981		
Swaziland	**0**	Need all	**2007**	1997	**1986**	**1976**	1966	

Togo	**3**	Need 2010	**2010**			**1981**	**1970**	**1958**
Tunisia	**1**	Need pre-2004	2014	**2004**	**1994**	**1984**	**1975**	1966
Zimbabwe	**0**	Need all	2012	**2002**	**1992**	1982		1969
E. No agreement signed; no census microdata until next census is conducted: 4 countries								
Angola	**0**		2013			1984	1970	1960
Congo-Democratic Rep.	**0**		2014			1984	1970	
Eritrea	**0**			1998				
Somalia	**0**				1987	1975		

Note: "Entrusted" indicates the number of sets of census microdata entrusted to the IPUMS-Africa project.
"Need" indicates the microdata sets not yet entrusted to the project, as of this writing.

References

Alderman, H., Babita, M., Demombynes, G., Makhatha, N., and Özler, B. (2003) 'How Low Can You Go? Combining Census and Survey Data for Mapping Poverty in South Africa', *Journal of African Economies*, *11*(2), pp. 169-200.

Alkire, S. and Foster, J. (2010) 'Designing the Inequality-Adjusted Human Development Index (IHDI)', *Working Paper No. 37,* Oxford Poverty and Human Development Initiative.

Bedi, T., Coudouel, A. and Simler, K. (eds.) (2007) *More than a pretty picture: Using poverty maps to design better policies and interventions,* Washington DC: The World Bank.

Case, A., Paxson C., and Ableidinger, J. (2004) 'Orphans in Africa: Parental Death, Poverty, and School Enrollment', *Demography, 41*(3), pp. 483-508.

Devarajan, S. (2013) 'Africa's Statistical Tragedy', *Review of Income and Wealth* (DOI: 10.1111/roiw.12013).

ECLAC (2010). *Pueblos Indígenas y Afrodescendientes*. Publication series by ECLAC's Population Division.

Elbers C., Lanjouw J. and Lanjouw, P. (2003) 'Micro-level estimation of poverty and inequality', *Econometrica*, *71*(1), pp. 355-364.

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

Filmer, D. and Scott, K. (2008). Assessing asset indices. World Bank Policy Research Working Paper No. 4605, The World Bank. Foster, J., Lopez-Calva, L. and Szekely, M. (2005) 'Measuring the distribution of Human Development: methodology and an application to Mexico', *Journal of Human Development*, *6*(1), pp. 5-25.

Gerland, P. (2013), 'Uses of Census Microdata by the United Nations Population Division', paper presented for the KOSTAT pre-event to XXVII IUSSP International Population Conference, Busan, South Korea.

Grimm, M., Harttgen, K., Klasen, S. and Misselhorn, M. (2008) 'A Human Development Index by Income Groups', *World Development, 36* (12), pp. 2527-2546.

Grimm, M., Harttgen K., Klasen, S., Misselhorn, M., Munzi, T. and Smeeding, T. (2010) 'Inequality in Human Development: An Empirical Assessment of 32 Countries', *Social Indicators Research*, 97, pp. 191-211.

Harttgen, K. and Klasen, S. (2011a) 'A Human Development Index by Internal Migration Status', *Journal of Human Development and Capabilities*, *12*(3), pp. 393-424.

Harttgen, K. and Klasen, S. (2011b) 'A Household-based Human Development Index', *World Development*, *40*(5), pp. 878-899.

Harttgen, K., Klasen, S. and Vollmer, S. (2013) 'An African growth miracle? Or: What do asset indices tell us about trends in economic performance?, *Review of Income and Wealth* 59, Special Issue October 2013.

Hohmann, S. and Garenne, M. (2010) ‘Health and wealth in Uzbekistan and sub-Saharan Africa in comparative perspective’, *Economics and Human Biology*, 8(3), pp. 346-360.

Jerven, M. (2013) *Poor Numbers: How We Are Misled by African Development Statistics and What To Do About It*, Ithaca: Cornell University Press.

Lerman, R. and Yitzhaki, S. (1985) ‘Income inequality effects by income sources: a new approach and applications to the United States’, *The Review of Economics and Statistics*, 67 (1), pp. 151-156.

McKenzie, D. (2005). Measuring inequality with asset indicators. *Journal of Population Economics*, 18, pp. 229-260.

Minnesota Population Center. 2012. Integrated Public Use Microdata Series, International; Version 6.0 [Machine-readable database]. Minneapolis, University of Minnesota.

Montgomery, M., Gragnolati, M., Burke, K., and Paredes, E. (2000) ‘Measuring Living Standards with Proxy Variables’, *Demography* 37(2), pp. 155-174.

Permanyer, I. (2013) ‘Using Census Data to Explore the Spatial Distribution of Human Development’, *World Development*, 46, pp. 1-13.

Sahn, D. and Stifel, D. (2000) ‘Poverty comparisons over time and across countries in Africa’, *World Development*, 28, pp. 2123-2155.

Sahn, D. and Stifel, D. (2003) ‘Exploring alternative measures of welfare in the absence of expenditure data’, *Review of Income and Wealth*, 49, pp. 463-489.

Seth, S. (2009) ‘Inequality, Interactions and Human Development’, *Journal of Human Development and Capabilities*, 10, pp. 375-396.

Shorrocks, A. (1982) ‘Inequality decomposition by factor components’, *Econometrica*, 50(1), pp. 193-211.

Stifel, D. and Christiaensen, L. (2007) ‘Tracking Poverty Over Time in the Absence of Comparable Consumption Data’, *The World Bank Economic Review*, 21(2), pp. 317–41.

Tarozzi, A. (2011) ‘Can census data alone signal heterogeneity in the estimation of poverty maps?’ *Journal of Development Economics*, 95, pp. 170-185.

Tarozzi, A. and Deaton, A. (2009) ‘Using census and survey data to estimate poverty and inequality for small areas’, *Review of Economics and Statistics* 91 (4), pp. 773-792.

UNDP (2010). ‘Human Development Report 2010. The Real Wealth of Nations: Pathways to Human Development’, New York.

Young, A. (2012), ‘The African Growth Miracle’, *Journal of Political Economy*, 120(4), pp. 696-739.

Woolfrey, L. (2009). ‘African microdata access survey’, Paper presented at the *International Statistical Institute World Congress* (Session 100), Durban, South Africa.

Zuberi, T. and Bangha, M.W. (2012[2006]). The history and future of the African Census Analysis Project. In: *Les Systemes D’information en Demographie et en Sciences Sociales: Nouvelles Questions, Nouveaux Outils? Chair Quetelet 2006.* Louvain: Presses Univ., pp. 47-70.

Tables

Country	ISO 3166 Code	Year	Population	Sample density	Administrative Unit Name	Levels	Number
Ghana	GH	2000	18,941,330	10%	Districts	2	110
Guinea	GN	1996	7,290,710	10%	Prefectures	2	34
Kenya	KE	1989	21,481,960	5%	Districts	2	31
Kenya	KE	1999	28,150,940	5%	Districts	2	69
Malawi	MW	1987	7,986,690	10%	Districts	2	24
Malawi	MW	1998	9,913,930	10%	Districts	2	26
Malawi	MW	2008	13,419,770	10%	Districts	2	31
Mali	ML	1987	7,853,840	10%	Communes	3	221
Mali	ML	1998	9,913,300	10%	Communes	3	221
Morocco	MA	1982	20,257,460	5%	Province/prefecture	2	63
Morocco	MA	1994	25,880,520	5%	Province/prefecture	2	60
Morocco	MA	2004	29,654,400	5%	Province/prefecture	2	60
Rwanda	RW	1991	7,429,180	10%	Province	1	11
Rwanda	RW	2002	8,433,920	10%	Province	1	12
Senegal	SN	2002	9,945,620	10%	Department	2	34
Sierra Leone	SL	2004	4,942,980	10%	Districts	2	14
South Africa	ZA	1996	40,578,357	10%	Municipalities	3	284
South Africa	ZA	2001	44,769,106	10%	Municipalities	3	225
South Africa	ZA	2007	47,173,595	2%	Municipalities	3	225
Sudan	SD	2008	38,206,344	15%	Districts/counties	2	202
Tanzania	TZ	1988	23,145,678	10%	Districts	2	39
Tanzania	TZ	2002	33,505,374	10%	Districts	2	56
Uganda	UG	1991	16,598,197	10%	Districts	2	113
Uganda	UG	2002	24,974,490	10%	Districts	2	129

Table 1. Basic information on the samples included in the analysis.
Source: http://ecastats.uneca.org/aicmd/en-us/samples.aspx.

		Core vars.			Country-specific variables													
Country	Year	El	WS	Tl	Sw	FC	FH	FL	RO	WL	RF	TV	RD	PC	PH	AU	CL	Tot
Ghana	2000	X	X	X		X		X	X	X								7
Guinea	1996	X	X	X														3
Kenya	1989	X	X	X	X	X		X	X	X								8
Kenya	1999	X	X	X	X	X		X	X	X								8
Malawi	1987		X	X		X							X					4
Malawi	1998	X	X	X		X							X					5
Malawi	2008	X	X	X		X							X					5
Mali	1987	X		X		X		X	X	X								6
Mali	1998	X		X		X		X	X	X								6
Morocco	1982	X	X	X														3
Morocco	1994	X	X	X														3
Morocco	2004	X	X	X														3
Rwanda	1991	X	X	X		X		X	X	X			X					8
Rwanda	2002	X	X	X		X		X	X	X			X					8
Senegal	1988	X	X	X	X			X	X	X	X	X			X			10
Senegal	2002	X	X	X	X			X	X	X	X	X			X			10
S.Leone	2004	X	X	X		X		X	X	X	X	X		X	X	X	X	13
S.Africa	1996	X	X	X		X	X				X	X	X	X	X		X	11
S.Africa	2001	X	X	X		X	X				X	X	X	X	X		X	11
S.Africa	2007	X	X	X		X	X				X	X	X	X	X		X	11
Sudan	2008	X	X	X		X					X	X	X	X	X	X	X	11
Tanzania	1988	X	X	X														3
Tanzania	2002	X	X	X														3
Uganda	1991	X	X	X		X		X	X	X								7
Uganda	2002	X	X	X		X		X	X	X								7

Table 2. Variables used in the construction of the country-specific wealth indices W_i^{CS}. El= ‘Electricity’, WS= ‘Water Supply’, Tl=’Toilet’, Sw= ‘Sewage’, FC= ‘Cooking fuel’, FH= ‘Heating fuel’, FL=’Floor’, RO= ‘Roof’, WL= ‘Wall’, RF= ‘Refrigerator’, TV= ‘Television’, RD= ‘Radio’, PC= ‘Personal Computer’, PH = ‘Phone’, AU= ‘Autos’, CL= ‘Cell Phone’. Source: https://international.ipums.org/international-action/variables/group

Country	Year	HI	EI	WI	MHDI	UNDP's HDI
Ghana	2000	0.60	0.60	0.31	0.50	0.43
Guinea	1996	0.46	0.21	0.11	0.26	NA
Kenya	1989	0.63	0.70	0.17	0.50	0.44
Kenya	1999	0.68	0.66	0.18	0.51	0.42
Malawi	1987	0.38	0.38	0.14*	0.30	0.27
Malawi	1998	0.45	0.52	0.10	0.36	0.34
Malawi	2008	0.51	0.61	0.10	0.41	0.37
Mali	1987	0.31	0.19	0.03*	0.17	0.17
Mali	1998	0.42	0.17	0.06	0.21	0.25
Morocco	1982	0.67	0.32	0.39	0.46	0.35
Morocco	1994	0.72	0.41	0.54	0.56	0.45
Morocco	2004	0.73	0.52	0.67	0.64	0.53
Rwanda	1991	0.49	0.54*	0.08	0.37	0.22
Rwanda	2002	0.42	0.50	0.10	0.34	0.30
Senegal	2002	0.68	0.41	0.44	0.51	0.37
S. Leone	2004	0.43	0.35	0.10	0.29	0.29
S. Africa	1996	0.69	0.78	0.63	0.70	0.63
S. Africa	2001	0.90	0.81	0.70	0.80	NA
S. Africa	2007	0.84	0.88	0.75	0.82	0.59
Sudan	2008	0.74	0.44	0.20	0.46	0.37
Tanzania	1988	0.54	0.65	0.15	0.45	0.33
Tanzania	2002	0.60	0.69	0.16	0.48	0.35
Uganda	1991	0.49	0.52	0.02	0.34	0.28
Uganda	2002	0.70	0.63	0.07	0.47	0.37

Table 3. Official HDI, country-level MHDI and its health, education and standard of living components calculated from IPUMS census samples. The values of EI for Rwanda 1991 and the values of WI for Mali and Malawi 1987 are not strictly comparable, as slightly different definitions of these indices have been used for those censuses (they have been marked with an asterisk to distinguish them from the other values). NA = ‘Not available’. Source: Authors’ calculations using census microdata disseminated by IPUMS-International.

Country	Year	G_MHDI	G_H	G_E	G_W	%C_H	%C_E	%C_W
Ghana	2000	0.16	0.06	0.18	0.41	9.7	42.6	47.7
Guinea	1996	0.22	0.10	0.33	0.69	23.3	36.1	40.6
Kenya	1989	0.13	0.10	0.13	0.49	29.6	36.5	34.0
Kenya	1999	0.13	0.10	0.14	0.51	26.4	38.6	35.1
Malawi	1987	0.15	0.12	0.16	0.39	30.3	39.2	30.5
Malawi	1998	0.11	0.07	0.12	0.36	21.5	53.2	25.3
Malawi	2008	0.12	0.09	0.11	0.55	25.4	41.6	33.0
Mali	1987	0.27	0.21	0.39	0.83	42.0	42.7	15.2
Mali	1998	0.24	0.16	0.40	0.61	42.2	36.3	21.5
Morocco	1982	0.19	0.07	0.28	0.35	14.6	33.2	52.2
Morocco	1994	0.14	0.05	0.21	0.25	8.9	35.0	56.1
Morocco	2004	0.11	0.05	0.15	0.16	12.1	37.0	50.9
Rwanda	1991	0.06	0.05	0.05	0.32	34.4	36.5	29.0
Rwanda	2002	0.07	0.07	0.06	0.34	34.2	37.7	28.1
Senegal	2002	0.20	0.08	0.35	0.37	16.7	29.9	53.4
S.Leone	2004	0.18	0.10	0.22	0.49	24.6	47.5	27.9
S.Africa	1996	0.13	0.09	0.08	0.27	20.6	20.4	59.0
S.Africa	2001	0.09	0.03	0.07	0.19	12.1	27.5	60.4
S.Africa	2007	0.07	0.03	0.04	0.16	13.1	20.0	66.9
Sudan	2008	0.19	0.07	0.30	0.57	10.0	47.2	42.8
Tanzania	1988	0.09	0.09	0.08	0.40	25.7	33.5	40.8
Tanzania	2002	0.12	0.09	0.11	0.52	20.1	40.2	39.7
Uganda	1991	0.11	0.08	0.12	0.80	27.6	49.2	23.2
Uganda	2002	0.09	0.05	0.11	0.63	17.5	53.5	28.9

Table 4. Gini index for the MHDI, HI, EI and WI distributions. Percent contribution of the health, education and standard of living components to inequality in the MHDI distribution. Source: Authors' calculations using census microdata disseminated by IPUMS-International.

Figures

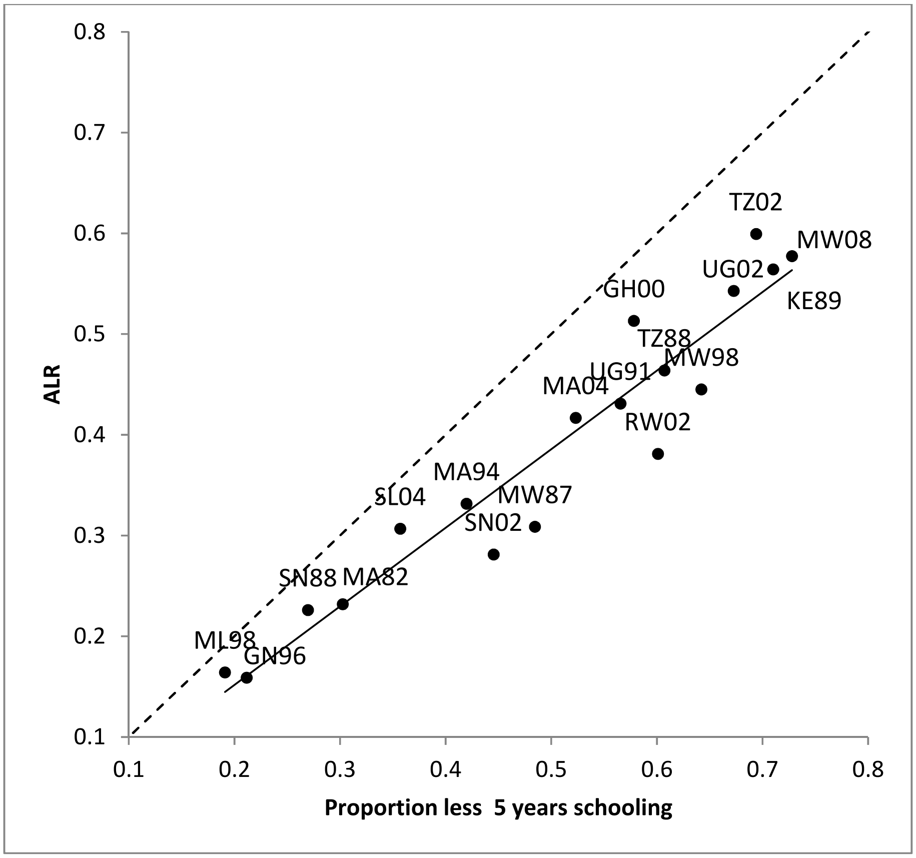

Figure 1. Comparison of ALR (vertical axis) and proportion of adults with less than five years of schooling at the country level (horizontal axis). The dashed line is the 45º equality line. The solid line is the best linear fit line. Countries are labeled with the ISO 3166 codes plus the last two digits of the year in which the corresponding census was conducted. Authors’ calculations using census microdata disseminated by IPUMS-International.

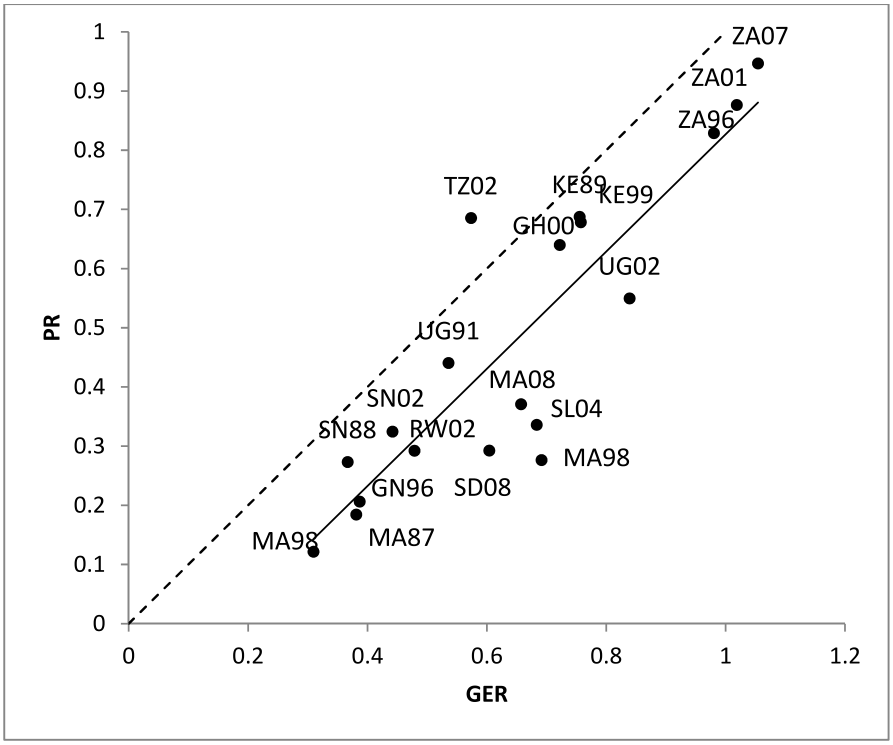

Figure 2. Comparison of Gross Enrolment Ratio (horizontal axis) and Primary completion rates (vertical axis). The dashed line is the 45º equality line. The solid line is the best linear fit line. Countries are labeled with the ISO 3166 codes plus the last two digits of the year in which the corresponding census was conducted. Authors' calculations using census microdata disseminated by IPUMS-International.

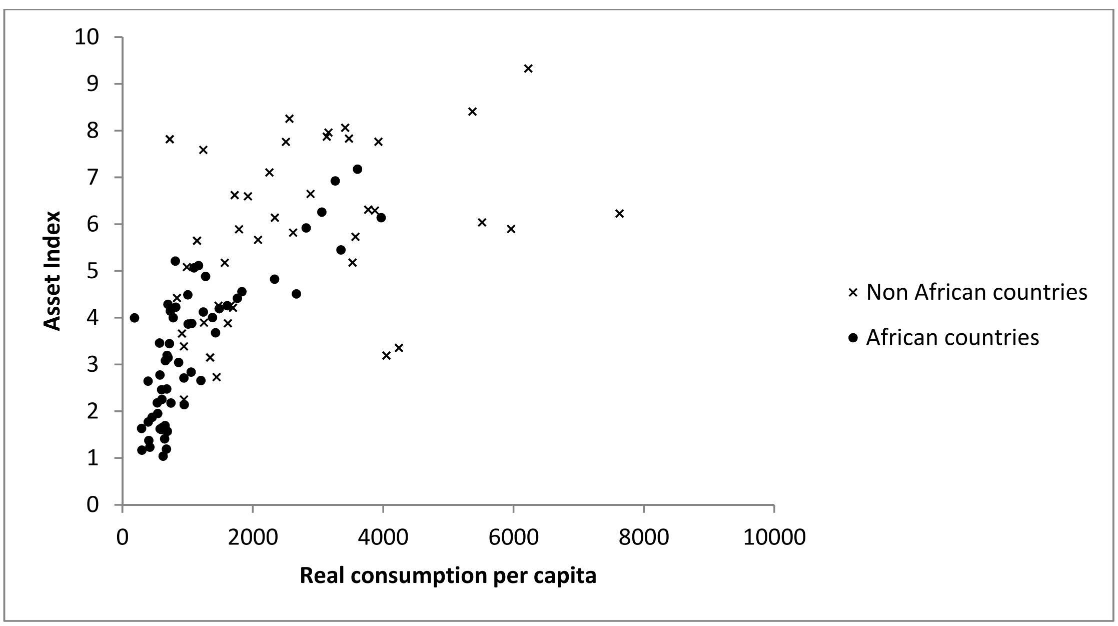

Figure 3. Real consumption per capita vs country-level average of an asset index in 102 developing countries. Source: Authors' calculations using DHS data and the World Penn Tables (version 8.0).

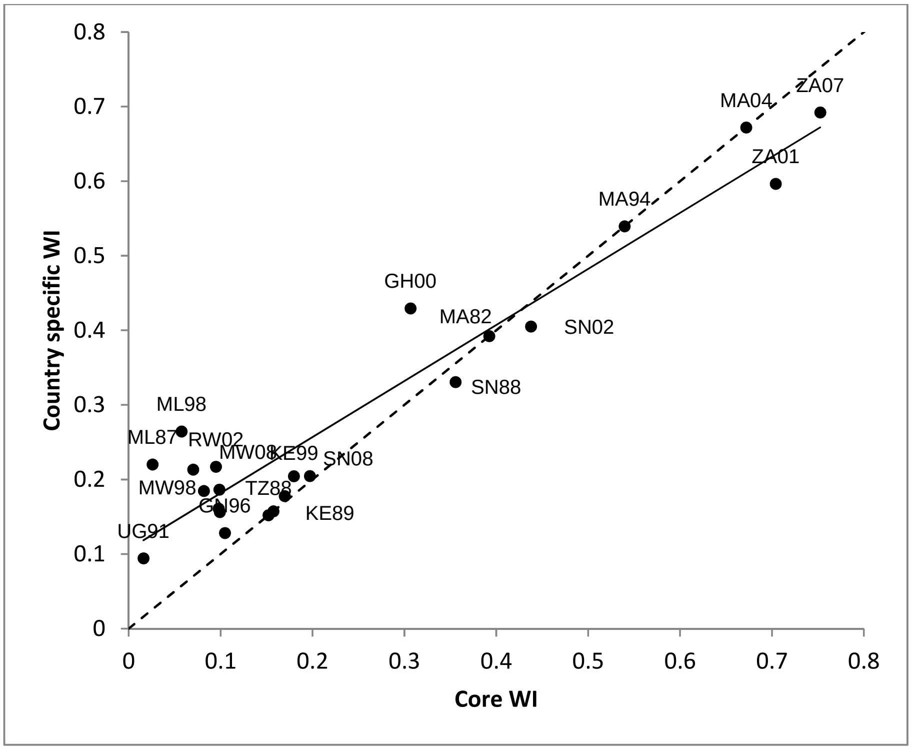

Figure 4. Comparison of the core Wealth Index (horizontal axis) versus the country specific Wealth Index (vertical axis). The dashed line is the 45º equality line. The solid line is the best linear fit line. Countries are labeled with the ISO 3166 codes plus the last two digits of the year in which the corresponding census was conducted. Authors' calculations using census microdata disseminated by IPUMS-International.

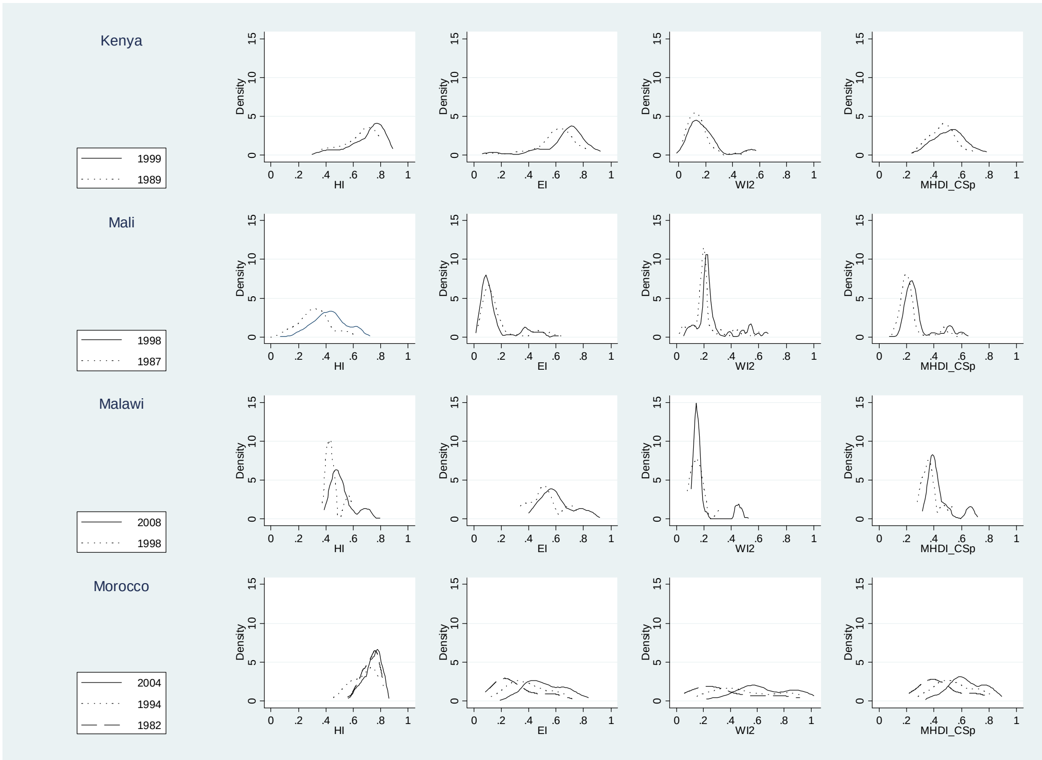

Figure 5. Density functions of the health, education, wealth and human development distributions for Kenya, Mali, Malawi and Morocco. The wealth and human development distributions have been constructed using country specific definitions. Authors' calculations using census microdata disseminated by IPUMS-International.

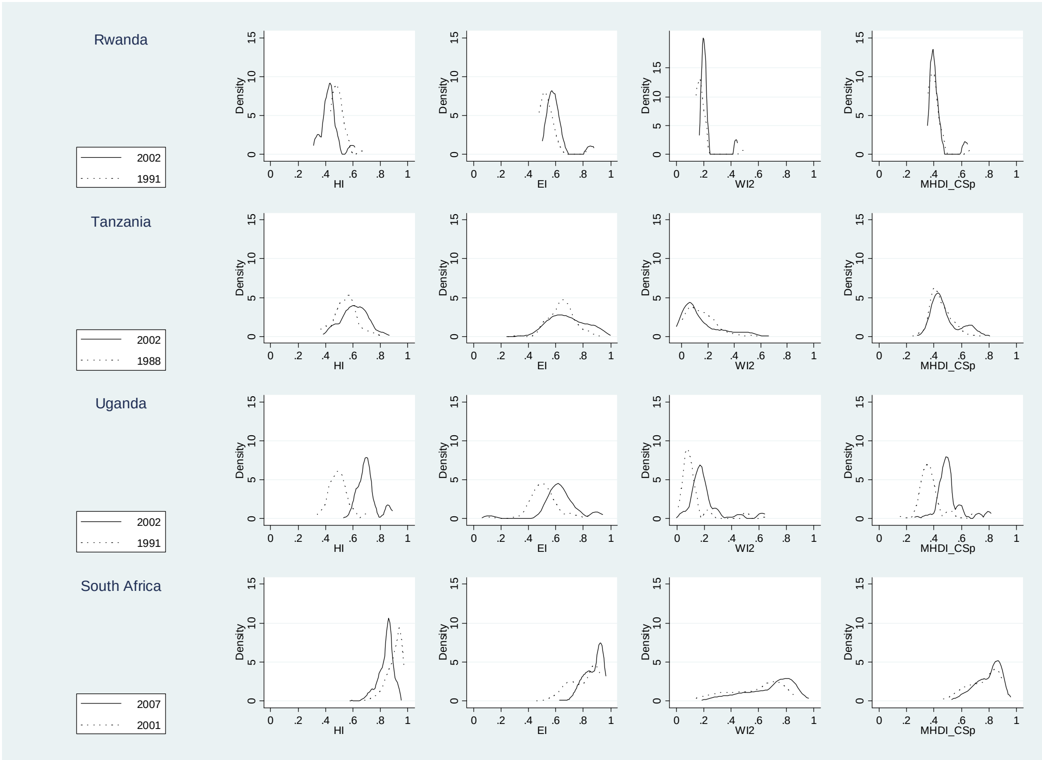

Figure 5 (Continued). Figure 4. Density functions of the health, education, wealth and human development distributions for Rwanda, Tanzania, Uganda and South Africa. The wealth and human development distributions have been constructed using country specific definitions. Authors' calculations using census microdata disseminated by IPUMS-International.

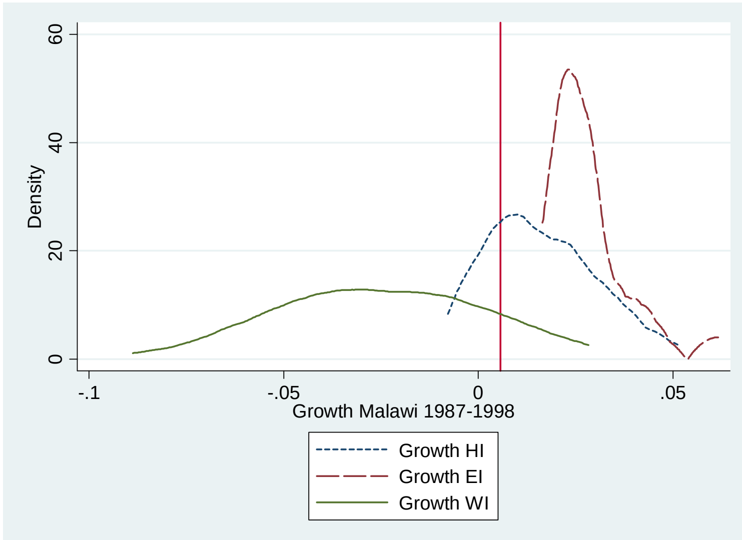

Figure 6.1: Administrative units' distribution of the annual growth rates of the health, education and wealth component for Malawi, 1987-1998. The vertical line indicates the real GDP per capita growth from the Penn World Tables 8.0. Authors' calculations using census microdata disseminated by IPUMS-International.

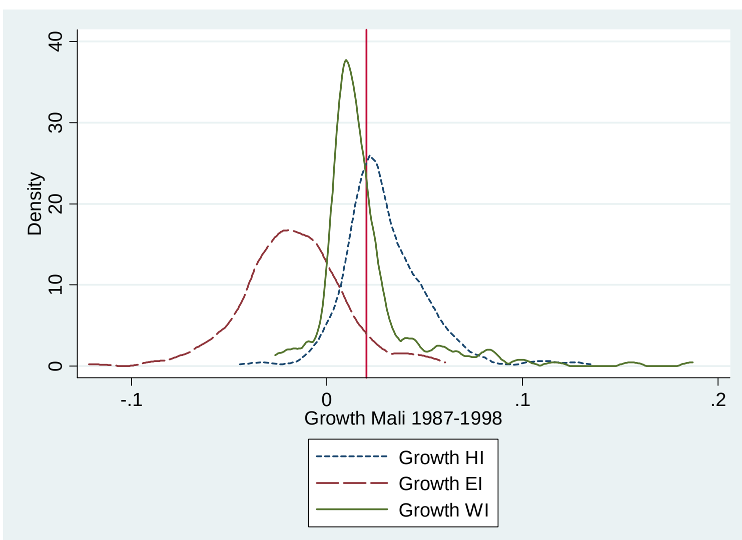

Figure 6.2: Administrative units' distribution of the annual growth rates of the health, education and wealth component for Mali, 1987-1998. The vertical line indicates the real GDP per capita growth from the Penn World Tables 8.0. Authors' calculations using census microdata disseminated by IPUMS-International.

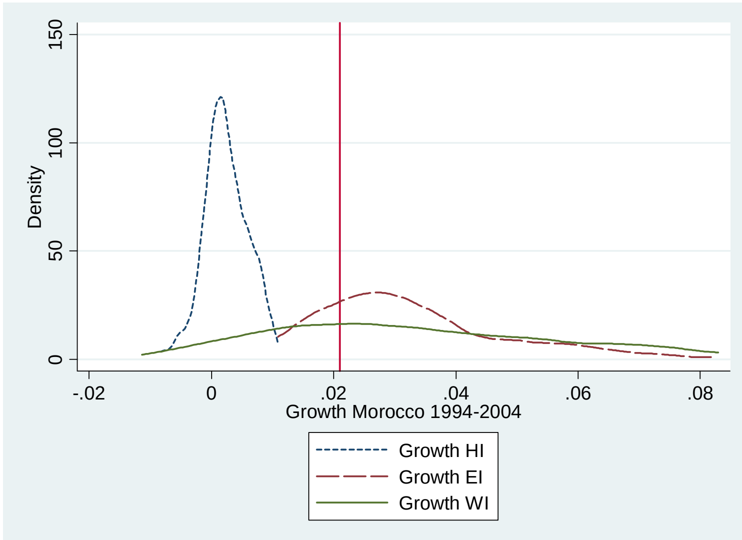

Figure 6.3: Administrative units' distribution of the annual growth rates of the health, education and wealth component for Morocco, 1994-2004. The vertical line indicates the real GDP per capita growth from the Penn World Tables 8.0. Authors' calculations using census microdata disseminated by IPUMS-International.

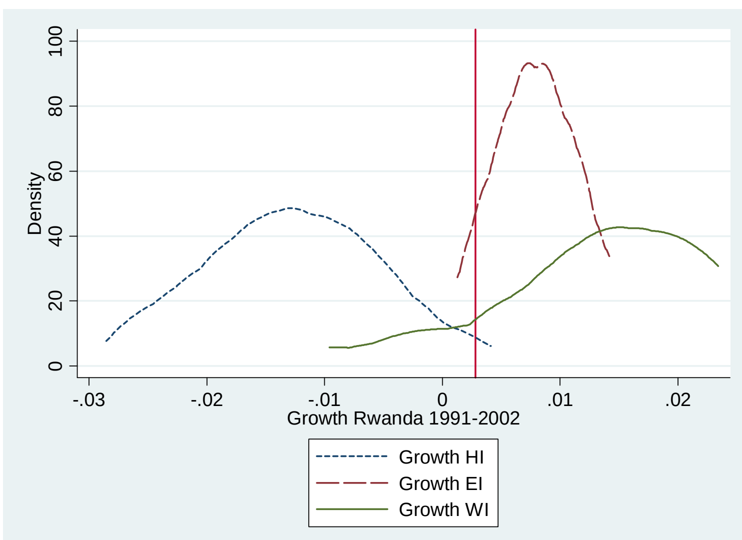

Figure 6.4: Administrative units' distribution of the annual growth rates of the health, education and wealth component for Rwanda, 1991-2002. The vertical line indicates the real GDP per capita growth from the Penn World Tables 8.0. Authors' calculations using census microdata disseminated by IPUMS-International.

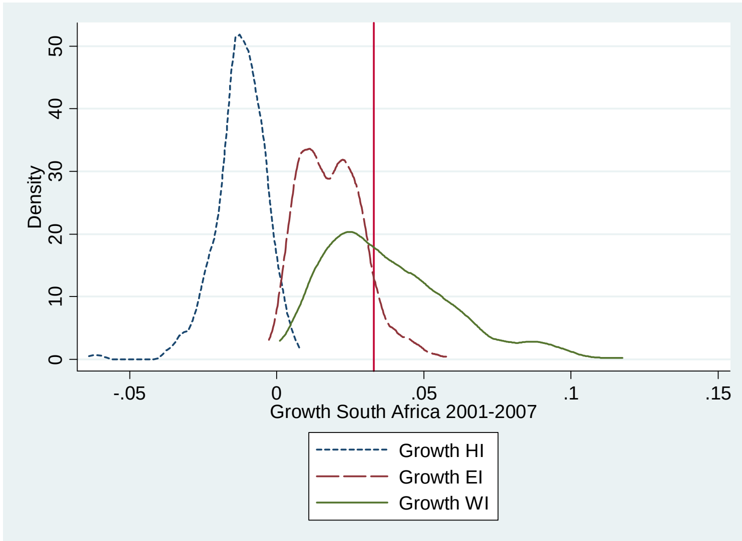

Figure 6.5: Administrative units' distribution of the annual growth rates of the health, education and wealth component for South Africa, 2001-2007. The vertical line indicates the real GDP per capita growth from the Penn World Tables 8.0. Authors' calculations using census microdata disseminated by IPUMS-International.

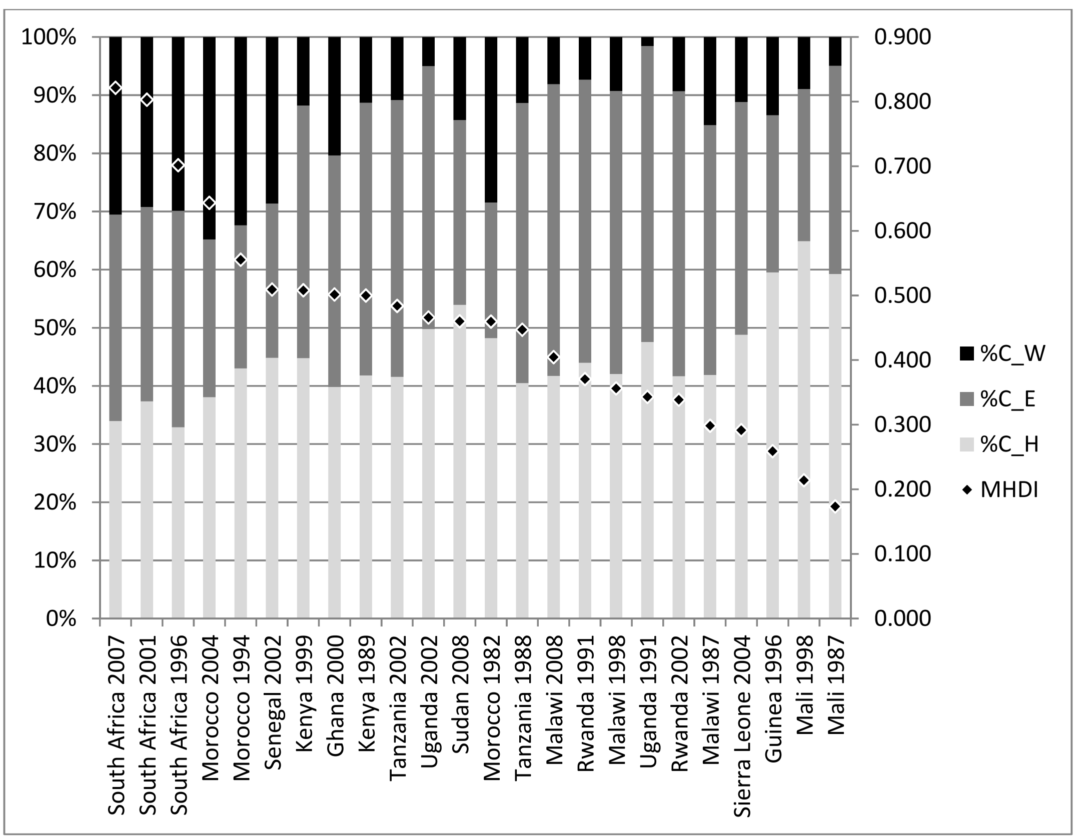

Figure 7. Country-level MHDI values (right vertical axis) with the corresponding percent contributions of the health, education and standard of living components (left vertical axis). Source: Authors' calculations using census microdata disseminated by IPUMS-International.

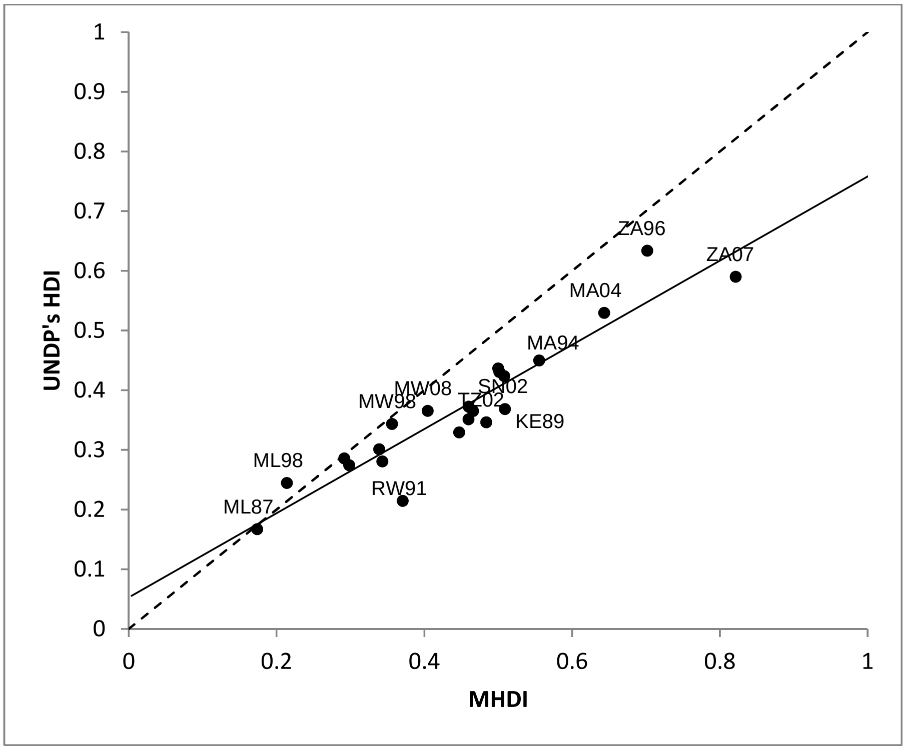

Figure 8. Country level MHDI (horizontal axis) vs UNDP's HDI values (vertical axis). The dashed line is the equality line. The solid one is the best linear fit line. Countries are labeled with the ISO 3166 codes plus the last two digits of the year in which the corresponding census was conducted. Source: Authors' calculations using census microdata disseminated by IPUMS-International and HDRs data.

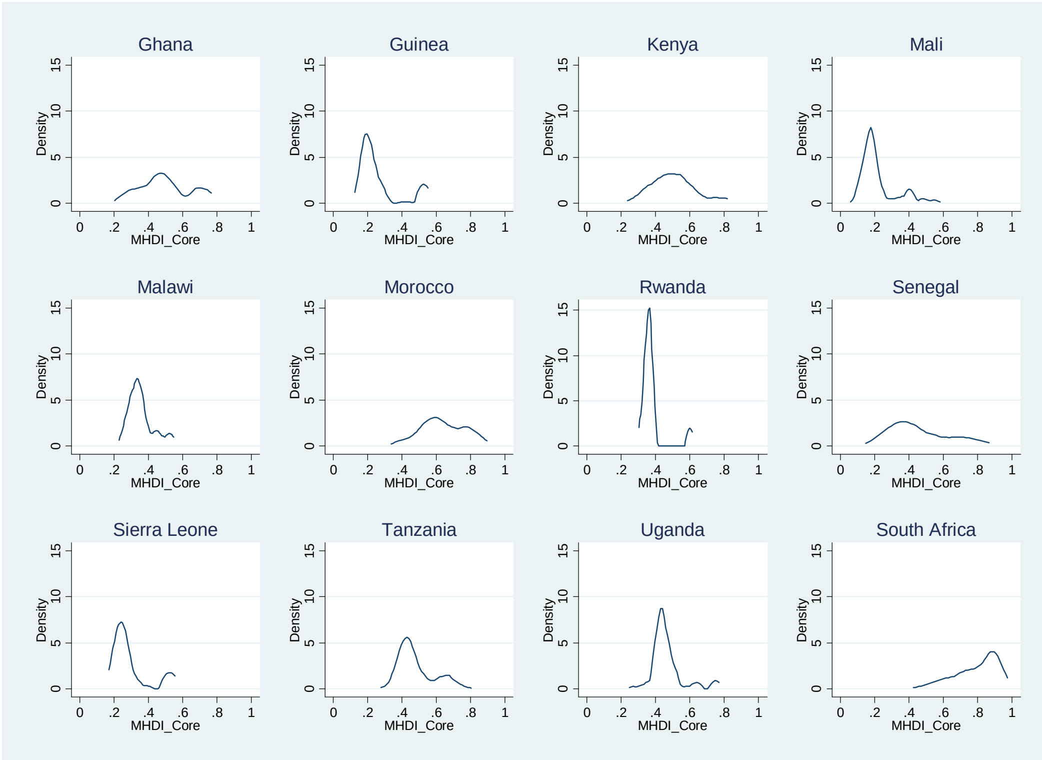

Figure 9. Density functions of the MHDI distributions for the 12 African countries included in the analysis around year 2000 (Ghana 2000, Guinea 1996, Kenya 1999, Malawi 1998, Mali 1998, Morocco 2004, Rwanda 2002, Senegal 2002, Sierra Leone 2004, South Africa 2001, Tanzania 2002, Uganda 2002). Source: Authors' calculations using census microdata disseminated by IPUMS-International.