Consequences of Data Error in Aggregate Indicators: Evidence from the Human Development Index

Hendrik Wolff
University of Washington

Howard Chong
University of California, Berkeley

Maximilian Auffhammer
University of Californi, Berkeley

Consequences of Data Error in Aggregate Indicators: Evidence from the Human Development Index

Abstract

This paper examines the consequences of data error in data series used to construct aggregate indicators. Using the most popular indicator of country level economic development, the Human Development Index (HDI), we identify three separate sources of data error. We propose a simple statistical framework to investigate how data error may bias rank assignments and identify two striking consequences for the HDI. First, using the cutoff values used by the United Nations to assign a country as 'low', 'medium', or 'high' developed, we find that currently up to 45% of developing countries are misclassified. Moreover, by replicating prior development/macroeconomic studies, we find that key estimated parameters such as Gini coefficients and speed of convergence measures vary by up to 100% due to data error.

Consequences of Data Error in Aggregate Indicators
Evidence from the Human Development Index

Hendrik Wolff
University of Washington

Howard Chong
University of California, Berkeley

Maximilian Auffhammer
University of California, Berkeley

January 2008

Abstract

This paper examines the consequences of data error in data series used to construct aggregate indicators. Using the most popular indicator of country level economic development, the Human Development Index (HDI), we identify three separate sources of data error. We propose a simple statistical framework to investigate how data error may bias rank assignments and identify two striking consequences for the HDI. First, using the cutoff values used by the United Nations to assign a country as 'low', 'medium', or 'high' developed, we find that currently up to 45% of developing countries are misclassified. Moreover, by replicating prior development/macroeconomic studies, we find that key estimated parameters such as Gini coefficients and speed of convergence measures vary by up to 100% due to data error.

Keywords: Measurement Error, International Comparative Statistics
JEL Codes: O10, C82

For correspondence contact Hendrik Wolff, Department of Economics, University of Washington, 524 Condon Hall, Box 353330, Seattle, WA 98195-3330. Tel: (510) 220-7961. Email: hgwolff@u.washington.edu. We are indebted to Alison Kennedy from UNDP for helpful correspondence and providing the "revised" HDI statistics. We thank Jenny Aker, David Albouy, Richard Carson, Maria Damon, Alain DeJanvry, Levis Kochin, James Rauch, Elisabeth Sadoulet and George Wright for helpful comments. We gratefully acknowledge generous funding provided by a University of California's Institute on Global Conflict and Cooperation faculty research grant. All errors in this manuscript are the authors'.

"Perhaps the greatest step forward that can be taken, even at short notice, is to insist that economic statistics be only published together with an estimate of their error."

Oskar Morgenstern, 1970

1. Introduction

A large number of social and economic indices are used to create policy relevant rankings of countries. Examples of popular indicators include the *Gross National Income* (GNI) measure (World Bank), the *Index of Economic Freedom* (Wall Street Journal), the *Political Risk Index* (Business Environment Risk Intelligence), the *Corruption Perceptions Index* (Transparency International), and the *Press Freedom Index* (Reporters Sans Frontières). In some cases, the policy relevance of these ordinal rankings is obvious as for example, the GNI determines a countries' eligibility for borrowing from various loan programs managed by the World Bank. In other cases the rank assignments have no direct legal consequence, and rather reveal their significance in fueling policy debates.

Despite the substantial use of international comparative statistics, their data quality is often considered dissatisfying; however, to our knowledge, no formal study measures the magnitude of the data error and reveals how poor data quality may bias rank assignments of countries.[1] In this paper, we propose a simple statistical framework to analyze such indicators which enables us to calculate country-specific variances of the noise distributions. We pick a popular index to show how three different sources of data error affect its cardinal values and ordinal rankings. Then, by re-estimating key parameters of selected published development/macroeconomic studies, we analyze the sensitivity of these parameters and find that coefficients can vary by up to 100% due to data error.

[1] Chay et al. (2005) analyze the consequences of data noise due to 'mean reversion' of student test-scores and show that this is problematic for small class sizes. Our paper differs from this in that we estimate country level specific probability measures of misclassification with respect to three sources of data error.

In particular, we apply our analysis to the Human Development Index (HDI) which has become the most widely used measure to communicate a country's development status. Compared to the Gross Domestic Product (GDP), the HDI is a broader measure of development, since it captures not only the level of income, but also incorporates measures of health and education (Srinivasan, 1994; Anand and Sen, 2006). Depending on the HDI score, a country is classified into one of the following three rank categories: 'low human development', 'medium human development' or 'high human development'. Although these categories are not formally tied to official development aid or imply any other direct legal consequence, today, these three mutually exclusive development categories are utilized widely. They are used to define the term *developing country*, to study health outcomes across countries (Guindon and Boisclair, 2003), and are used in academic studies in communications (Hargittai, 1998; Keiser et al. 2004), development economics (Kelley, 1991; Noorbakhsh, 1998; Baliamoune, 2004), and macroeconomics (Mazumdar, 2002; Noorbakhsh, 2006). Further, the indicator is frequently invoked to structure discussions in development-political debates (United Nations, 1997; HDR 1999 to 2006; Geneva Global, 2007).

Despite extensive use of the HDI statistics, the drastic changes in the distribution of HDI scores for developing countries, as displayed in Figure 1 below, have gone unnoticed in the academic and policy literature. When the HDI was first published in 1990, the cross country-distribution appears to be approximately uniformly distributed between zero (least developed) and one (most developed). Today, however, the distribution is twin-peaked with two sharp spikes around the values of 0.5 and 0.8, which are the cut-off values for categorizing countries of 'low', 'medium' and 'high' human development.

In this paper, we investigate the role of data error on the published HDI and the consequences for its use in statistical analysis. We address these questions by exploiting (1) the originally published HDI time series, (2) the subindicator variables used to construct the HDI, (3) changes to the HDI formula, and (4) documented data revisions. We identify three major sources of data error: measurement error due to data revisions, data error due to formula updating and misclassification due to inconsistent cut-off values, each of which is discussed in more detail in section 3. Based on these errors we estimate country specific variances of the HDI scores. We show that the HDI contains data error ranging from 0.04 standard deviations (Algeria) to 0.11 standard deviations (Niger), which is significant given the scale of 0 to 1. Mapping these cardinal noise measures onto the ordinal dimension, we find that 12%, 24% and 45% of developing countries can be interpreted as currently misclassified due to the three sources of data error, respectively.

Moreover, our results have direct implications for the academic literature. The HDI has been used to analyze the evolution of the world's distribution of well being, to explore issues of inequality, polarization, foreign direct investment, development aid and to econometrically test various convergence hypotheses in macroeconomics (*e.g.* Pillarisetti, 1997; Ogwang, 2000; Jahan, 2000; Globerman, Shapiro, 2002; Mazumdar, 2002; Neumayer, 2003; Arcelus et al., 2005; Noorbakhsh, 2006; Prados de la Escosura, 2007). By replicating some of these studies and carrying out sensitivity analysis, we find that key parameters, such as estimated Gini coeffients and speed of convergence parameters, vary by up to 100% in their values, simply due to the measurement error we directly observe in the published HDI series.

As a consequence of our findings, we suggest that the United Nations should discontinue the practice of classifying countries into the three bins. Based on our analysis, we view the cut-off values as arbitrary. The classification does not add any substantial informational value but rather has the potential to severely misguide users of the HDI statistics. Further, the analysis in this paper may be of broader interest since the same variables used to construct the HDI (education, health and income purchasing power statistics) serve as inputs to many international comparative statistics used e.g. by OECD, UNESCO, WHO, and World Bank.

The remainder of the paper is structured as follows. Section 2 outlines the data, section 3 measures the misclassification due to formula changes and data revisions, section 4 discusses empirical examples of how the HDI is used today and how measurement error affects prior analysis. We conclude with policy recommendations in section 5.

2. Data

The HDI is a composite indicator measuring a country's level of development along three dimensions: health, education and income. These dimensions are expressed as unit free and double bounded subindicators y_1, y_2, y_3, each taking values between zero and one. The subindicators themselves are functions of data $\boldsymbol{x}$ on primary and secondary school enrollment statistics, life expectancy and per capita purchasing power (PPP). Finally, the HDI is calculated as a simple average of the three subindicators, HDI = $1/3\Sigma_k y_k(\boldsymbol{x})$, which is then used for ordinal and cardinal comparisons. The HDI is published in the Human Development Reports (HDR) by the United Nations Development Program (UNDP), which are available for the years 1990 to 2006 (HDR, 1990 to 2006).

2.1. Original versus Revised Data

In our analysis we exploit the fact that the original historical data matrix $\boldsymbol{x}_t$ used by the UNDP in year t, does not correspond to the at a later date s revised matrix $\boldsymbol{x}^{\mathbf{R},s}_t$ which is used by the UNDP at time s. The original $\boldsymbol{x}_t$ is available for the years t =1999 to 2006, whereas the revised data $\boldsymbol{x}^{\mathbf{R},s}_t$ are available for all years of the analyses, t = 1990 to 2006 and s = 2006. In this paper, $\boldsymbol{x}^{\mathbf{R}}_t$ refers to the variables for year t kindly provided to us in the fall of 2006 by the UNDP office, except stated otherwise. $\boldsymbol{x}_t$ refers to the data that we hand-copied[2] from the t^{th} year Human Development Report (HDR, 1990 to 2006).

2.2. The HDI Formulas and Computation of Counterfactuals

Since 1990, the UNDP has made three major updates to the formula used to construct the HDI. For each year t and country i denote the HDI formula by

$$HDI_{it} = h_f(\mathbf{x}_{it}).$$

The formula h changed thrice as indexed by $f \in \{A, B, C\}$ which corresponds to the time periods 1990, 1995-1998 and 1999-2006, respectively.[3] The three formulas are explained in the HDR technical appendices (HDR, 1990 to 1999) and in Jahan (2000). Combining data updating and formula changes, we construct three 'counterfactuals' denoted by $h_A(\mathbf{x}^{\boldsymbol{R}}_{it})$, $h_B(\mathbf{x}^{\boldsymbol{R}}_{it})$, and $h_C(\mathbf{x}^{\boldsymbol{R}}_{it})$. Hence, for the entire time series we recalculate what the HDI would have been if the alternate formulas had been in place, using the most recent available historical data on the subindicators. In the analysis we exploit exactly these differences between the "original" HDI generated by the formula that was active at time t compared to the HDI generated by the other two formulas that were not active in that particular year t.

2.3. The Sample

For comparability of the yearly HDI distributions it is important that the number of countries be constant over time so that the distributions are based on a consistent sample. We construct a balanced panel from 1990 to 2006. Whether a country is included in the panel is

[2] Copying statistics from the original HDRs is time intensive. Hand copying may produce data errors. Since the purpose of this study is to measure the error of the HDI statistics (and not our own data entry error) the data were hand-copied separately by two of the authors. Only after verifying that the two hand-copied data sets are 100% identical, we proceeded with the analysis. Data are available upon request.

[3] Note that period A refers to the year 1990 only. There were two minor changes to the formula in the year 1991 and 1994. However, these formulas require data that are not available any longer and could not be replicated by the authors. In particular the variable 'mean year of schooling' and 'world average income' could not be precisely replicated in a way the UN had used those variables in the years 1991 to 1994.

determined by the following three conditions: (a) the country exists continuously between 1990 and 2006 (*e.g.*, Croatia is dropped); (b) for each country and subindicator, not more than five data points are missing over the period of the analysis[4]; and (c) it is not an industrialized country[5]. In this way we obtain a panel of HDI scores for 72 non-industrialized countries which we also, more conventionally, denote as the sample of 72 developing countries.

3. Sources of Data Error and Results

In the following, we provide a detailed discussion of the three sources of data error: measurement error due to data revisions, data noise due to formula updating and misclassification due to inconsistent cut-off values. We propose a useful, yet simple, statistical framework to analyze these sources of errors, which will allow us to calculate country specific variances and confidence intervals and simulate country specific probabilities of misclassification.

3.1. First Source of Data Error: Measurement error

To obtain a first measure of the randomness of the HDI data, we exploit the following exogenous changes to the data over time: The data $\boldsymbol{x}_t$ (as used by the UNDP for the HDR at year t) are in general not the same data as the UNDP publishes in year s for the same data year t. Hence, as revised statistics become available, the UNDP updates the original data matrix $\boldsymbol{x}_t$ at year s, $s \geq t$, which we then denote $\boldsymbol{x}^{\mathbf{R}\,s}_{t}$.

This implies that whenever an analyst/researcher uses UNDP data, the same analysis run at a later date, will result in different estimates due to a changed data matrix. Hence, when the HDI for a given year t is released in year t, the value must be understood as an inexact value

[4] If we would require that all data points were available, then our sample would drop considerably.
[5] We drop all industrialized countries from the data set which are essentially all countries in the OECD and the former Soviet Union and Eastern Europe. The exact listing of the industrialized countries is given in the HDR report of 1991 Table 1.1.

subject to future data revisions. This problem is what we refer to as measurement error from data updating.

To parameterize this measurement error, assume that the relationship between the observed HDI score of country i and the true (but unknown) subindicators, denoted by y^*_{itk}, can be expressed as

$$\mathrm{HDI}_{it} = 1/3\Sigma_k y^*_{itk} + \varepsilon_{itk}$$

where ε_{itk} is orthogonal to y^*_{itk} and is distributed with mean m_{kti} (not necessarily equal to zero) and country specific variance s^2_{kti}. The relationship between the observed HDI score of country i and the true HDI^* consequently is $\mathrm{HDI}_{it} = \mathrm{HDI}_{it}^* + e_{it}$ with e_{it} being the composite error term distributed with mean $1/3\Sigma_k m_{kti}$ and country specific variance σ^2_i that is determined by the countries' covariance structure of the measurement error of the subindicators.

Exploiting the original $\boldsymbol{x}_t$ and revised $\boldsymbol{x}^{\mathbf{R}}_t$, we now are in the position to calculate country specific variances of the measurement error due to data (D) updating given by

$$\sigma^2_{\mathrm{D},i} = \Sigma_t(h_t(\boldsymbol{x_{it}})-h_t(\boldsymbol{x_{it}}^{\mathbf{R}}))^2/T \quad \text{for } t = 1990 \ \& \ 1995,1996,...,2005. \qquad (1)$$

with h_t denoting the formula which was active at time t. Hence, the variance of the data-updating measurement error is based on the difference between the original HDI as published in the HDR at year t and the reconstructed HDI for year t using revised data available to us today, $\mathrm{HDI}^{\mathrm{R}}$.[6]

3.2. Second Source: Changes in HDI Formula

In an effort to improve the HDI statistics, after being criticized on methodological and statistical grounds (e.g. Desai, 1991; McGillivray, 1991; Srinivasan, 1994, Noorbakhsh, 1998),

[6] We do not compute the variance using the data of 2006, since for 2006 the revised HDI is by definition equivalent to the originally published HDI. We also do not use the data of the years 1991 to 1994 (see footnote 3).

the UNDP has made three major updates to the formula used to construct the HDI. These three changes are clearly visible in the empirical distribution of the HDI displayed in Figure 2.

In particular, different distributional characteristics occur for the following subperiods A (1990), B (1995-1998) and C (1999-2006) that correspond to the three formula regimes $h_A(\boldsymbol{x}_{it}^{R})$, $h_B(\boldsymbol{x}_{it}^{R})$, and $h_C(\boldsymbol{x}_{it}^{R})$, respectively. We exploit this variation of the HDI scores across the counterfactual formulas to calculate country specific variances due to the formula (F) updates that is

$$\sigma^2_{F,i} = \Sigma_t \Sigma_f (h_f(\boldsymbol{x}_{it}^{R}) - h_C(\boldsymbol{x}_{it}^{R}))^2/(T\text{x}2) \text{ for } t = 1990 \text{ \& } 1995,1996,...,2005$$

where f is the index to sum over the three formula indices A, B and C. Hence the variance $\sigma^2_{F,i}$ is based on the country specific differences of the HDI generated by the most recent and improved formula h_C compared to the HDI counterfactuals generated by the other two formulas h_B and h_A. We do acknowledge that the formula revisions were undertaken to improve the HDI statistics and hence one interpretation of $\sigma^2_{F,i}$ is to understand it as a measure of *historic* noise due to the formula updates. Alternatively, the country specific measures $\sigma^2_{F,i}$ can be interpreted as a *present* measure of noise, if the UNDP will similarly continue to change the formula in the future and the rankings today would have to be understood as subject to those future formula revisions.

3.3. Third Source of Misclassification: Arbitrariness of the Cut-off Values

The third measure of misclassification is due to the arbitrariness of the two cut-off values used to categorize countries into ‘low’, ‘medium’ and ‘high’ development countries. Despite the fact that changes made to the HDI formula did have considerable impacts on the HDI distributions as displayed in Figure 2, surprisingly the UNDP has used the *same* cut-off values (0.5 and 0.8) since 1990. Since the original cutoff-values are supposed to distinguish three *qualities* of human development, with each formula change the UNDP could and should have adjusted the cut-off values in such a way that the new adjusted thresholds again reflect these same value judgment for the levels of quality. Hence, our procedure to obtain revised threshold values—that would be consistent with the initial 1990 value judgment of classifying quality and consistent with the entire history of formula changes—is as follows. In 1990, Morocco and Egypt were the two countries closest around the original cut off value of 0.5 (with HDI scores of 0.49 and 0.50, respectively). On the counterfactual distribution of formula h_c applied to 1990, these two countries take on the values 0.54 and 0.56. Taking the mean (0.55) provides the revised threshold for separating between the low and medium human development groups. Similarly we proceed with the cut off value 0.8 and obtain the revised value 0.70.

3.4. Simulation: The expected number of misclassified countries

For the first two sources of data error, for each country we can calculate the exact probability of being misclassified. Given the parameterization of the measurement error as $HDI_{i2006}^{*} = HDI_{i2006} - e_{i2006}$ and $e_{i2006} \sim N(0,\sigma^2_{.,i})$, normally distributed with mean zero[7] and variance $\sigma^2_{.,i}$ (as calculated by $\sigma^2_{F,i}$ or $\sigma^2_{D,i}$) we analytically calculate for each country the probability of being misclassified as

$$\begin{aligned}
&\int_{0.5}^{1} p(\widehat{HDI}_i)\, d\widehat{HDI}_i \quad \forall\, i \text{ with } HDI \in [0.0, 0.5) \quad , \\
&\int_{0.0}^{0.5} p(\widehat{HDI}_i)\, d\widehat{HDI}_i + \int_{0.8}^{1} p(\widehat{HDI}_i)\, d\widehat{HDI}_i \quad \forall\, i \text{ with } HDI \in [0.5, 0.8) \quad , \qquad (2) \\
&\int_{0.0}^{0.8} p(\widehat{HDI}_i)\, d\widehat{HDI}_i \quad \forall\, i \text{ with } HDI \in [0.8, 1.0] \quad ,
\end{aligned}$$

[7] In this section, we assume that the country specific means of the data error distribution are zero. In section 4.3, we find, however, an upward bias for most of the countries. If we were taking into account these asymmetries, then the misclassification measures reported in section 3.5. would lead to even larger values.

where $p(\widetilde{HDI})$ is the probability density function of the estimated HDI_i* distributions. Hence, for countries reported to be of ‘low development’, we calculate the probability of being classified as a medium or a high development country; similarly, for the ‘medium’ countries we calculate the probability of being low or high, and for the ‘high’ development countries the probability of being low or medium. Finally, adding these integrals over all countries provides the expected number of misclassified countries.

3.5 Results

If one followed Oskar Morgenstern’s (1970) advice given in the introduction, an alternative way for UNDP to report HDI scores would be to report country specific noise measures. To do so, we display country specific standard errors in table 1 below. We find that the standard errors due to the measurement error $\sigma_{D,i}$ range between a minimum value of 0.01 (Malaysia) and a maximum value of 0.07 (Syria). The estimated $\sigma_{F,i}$ due to the formula updates range between a minimum value of 0.01 (Algeria) and a maximum value of 0.11 (Niger). Given that the HDI is an average over three subindicators, whereby positive and negative deviations in the subindicators could on average cancel out,[8] and given that the HDI is scaled from of 0 to 1, these standard deviations are large and significant. These estimated standard errors $\sigma_{D,i}$ and $\sigma_{F,i}$ reflect noise measures of the cardinal scale of the HDI. Since the HDI is, however, primarily used as an ordinal measure, we now turn to the impact of these cardinal measures on the ordinal dimension. To illustrate, Figure 3 below displays the case of the “average” country with HDI = 0.65 using the average standard deviation over *all* developing countries due to data revisions, σ_D=0.03 and due to formula updates σ_F=0.08. Figure 3 shows that substantial probability mass is

[8] The correlation between the three subindicator error terms ε_{itk}, $k = \{1,2,3\}$ is close to zero, such that the three subindicator error variables can be viewed as distributed approximately independent. Hence the average standard deviation of the subindicator errors s^2_k must be larger in magnitude, compared to the standard deviation of the HDI, $\sigma_{D,I}$ Section 4.3, in which we analyze the structure of the compound error term in more detail confirms this.

spread over all three development categories. In table 1, the category specific probabilities are displayed for all developing countries in columns 5-7 and columns 10-12. For example, as of 2006, Mongolia, India, Honduras, Bolivia and others have non-zero probabilities of belonging to all three categories simultaneously. Even a high human development country, such as Costa Rica with HDI of 0.84, can still be a 'low' with 0.1% probability and yet be 'medium' to 35%. Finally, columns 8 and 13 display the total probability of a particular country being misclassified by using formula (2). The sum over these column probabilities show that currently, in expectation, 8.4 countries are misclassified due to data updating measurement error and 17.6 countries are misclassified due to formula updates; these numbers translate into, 12% and 24% of the developing countries being misclassified. For these calculations, we assumed that the mean of the error distributions is zero. In fact, the mean over all countries is an insignificant -0.0005.

Turning now to the third measure of misclassification, the adjustment of the cut-off value. If the UNDP had adjusted the cut-off values in a manner consistent with the 1990 classification, since 1999 (the year of the last formula update), the thresholds should be at the values 0.55 and 0.70, as opposed to 0.5 and 0.8. This lack of adjustment of the cutoff values results in 45% of the countries being misclassified today.[9] With such a high percentage, statements such as '*over the last decade x% of African countries successfully moved from the 'low' to the 'medium' human development category*'—as expressed in numerous policy papers and news reports (United Nations,1997; People's Daily, 2001; Daily Times, 2005) become useless at best, if not blatantly misleading. The listing of the misclassified countries due to this source of error as of 2006 is provided in Table 2.

[9] The percentage of countries misclassified is calculated as the number of countries that have HDI scores in the ranges [0.5, 0.55) and [0.70, 0.8) divided by the total number of countries in our sample (72).

We interpret the misclassification of 12% due to data updating as *conservative* because $\sigma^2_{D,i}$ is just based on "short term" differences between $\boldsymbol{x}_t$ and $\boldsymbol{x}^{\mathbf{R}}_t$, based on the years from 1990 to 2006.[10] There, however, also exists "long term" data updating error, which taking into account, may increase σ^2_D as $\|\boldsymbol{x}_t\ \boldsymbol{x}^{\mathbf{R}\,s}_t\|$ increases with s. While we cannot capture this long term effect by formula (1) (due to the lack of published original data prior to the HDR of 1990), we are able however to illustrate the magnitude of such "long term" drift effects: since 1999, the UNDP publishes historic HDI scores for the year 1975, HDI_{1975}. Figure 4 displays HDI_{1975} scores as they are reported in each of the HDR reports from 1999 to 2006. In every year, between 1999 and 2006, substantial data revisions took place for the *same* 1975 HDI score. For example, while in 2000 Portugal was reported to have a historic HDI_{1975} of 0.73 in 2000 (that was below the HDI_{1975} of Venezuela), by 2006 the Portugal HDI_{1975} significantly increased and is now substantially above the 2006 reported HDI_{1975} of Venezuela. On average over all countries the updating bias is 0.003 with $\sigma_{1975} = 0.012$. Given that the data updates took place after a quarter of a century, we consider 0.012 as a sizable standard deviation. Instead, in a world of good data quality, after a quarter of a century σ_{1975} should be close to zero.

4. Discussion of the results

The HDI is frequently used in development/political debates and in the academic literature. Given, however, that the HDI is subject to a considerable amount of measurement error, the use of the HDI and its triple bin classification system leads to serious interpretability problems. The following examples shed some light on these issues.

[10] $\sigma_{D,i}$ is based on the "short term" differences between the original and the revised time series provided in 2006. The minimum short term difference is hence one year (the 2005 data updated in 2006) and the maximum is seventeen years (the 1990 data updated in 2006). The "long term" data updating error is based on the fact that even after a quarter of a century, the historic 1975 data are updated in every year from 1999 to 2006.

4.1 The HDI as a definitional measure

The definition of the term "developing country" is often directly linked to the HDI, as being a country with low to moderate development status. In fact, the first hit on *Google* for the search term 'Developing Country' leads to a site that displays a world map of HDI scores. Here it is common to differentiate development status using three different colors. In Figure 5, we recreate such a map by displaying the HDI scores for 2006. To demonstrate the impact of misclassification of non-industrialized countries in our sample, we reclassify the non-industrialized countries using the updated thresholds of 0.55 and 0.70 as discussed in section 3.4. The visual impact of this reclassification is striking, especially in South America, Southeast Asia and Africa. This misclassification is particularly problematic, if organizations/institutions use these categories to design particular policies or rules.

4.2 The HDI and Foreign Development Aid:

Although, to our knowledge, the HDI is not formally used by any development agency as the sole index used to determine the distribution of development funds, there is a clear indication that the HDI does play a significant role in governmental institutions and NGOs when debating over the need for foreign aid allocation.[11] In 2000, the Deputy Director of the UNDP Selim Jahan exemplified this debate by stating:

> *"At the global level, issues are now being explored as to whether bilateral aid can be allocated on the basis of HDI, or the core funds of multilateral agencies can be based on the index […]"* (p. 10, Jahan, 2000).

In fact, 'charity scorecards' are increasingly used as a tool for helping individuals decide which countries to donate money to. Here the HDI can be used to construct such a score. For

[11] For a related discussion see Alesina and Dollar, 2000; Alesina and Weder, 2002; Arcelus et al. 2005; Bandyopadhyay and Wall, 2006; Easterly et al., 2004.

example, on the start homepage of the most prominent charity scorecard organization (http://www.charityscorecard.org/) a world map of HDI scores is displayed, similar to the one shown in Figure 4. The use of the HDI in this context may explicitly and implicitly steer users of these scorecards to "misclassified countries". Further, the triple bin classification is often used for report writing purposes to describe donor activities (United Nations, 1997; HDR 2001 to 2007; Geneva Global, 2007). For example, Geneva Global (2007), which holds investments of 60 million client dollars in development projects, structures its funds according to the three HDI categories. Also the United Nations (HDR 2001 to 2006) analyzes development aid data in the domain of the three human development categories. Table 3 shows that, across all years, countries in the 'low' category obtained 3.4 times the official development assistance (ODA) per capita as compared to the medium development countries, which we do not claim is a causal effect but rather an interesting correlation.

4.3 Structure of the Measurement Error

4.3.1. Measurement Error with Respect to the HDI

In the following we analyze the structure of the measurement error due to data revisions for the most recent years 1999 to 2006, period C.[12] Figure 6 displays the relationship between the country specific measurement error due to the data revisions, $\sigma_{D,i}$ and the countries' HDI score (as of 2006). Clearly, we see that as countries become more developed, the data updating variance declines, which could be an indication that richer countries have better statistical agencies. Looking at the graph in more detail we also note that the group of countries with HDI scores close to the threshold value of 0.5 has a larger than average variance of $\sigma^2_{D,i}$, which can exarbate the missclassification problem.

[12] We restrict this section to period C, when the formula h_C has remained constant over time and the quality of the subindicator data has improved considerably compared to period A and B.

Figure 7 displays the empirical densities of the updating error by year, $-e_{tD}$, that are calculated by differencing the originally reported HDI and the revised HDI^R. The updating has the smallest mean in the most recent year for which updated data are available - 2005 data revised in year 2006. This is intuitive, as not enough time has passed to more substantially revise the data. For all other years (1999-2004), the average updating implies a structural *upward bias* by about +0.01 (see Table 4) and this bias consistently positive since 1999 for every single year. This is in contrast to the bias in the nineties, when for some years the bias is positive and for some negative (the empirical mean over all years is 0.0005, see section 3.5). To investigate this further, zooming to the +/- 0.05 HDI range around the threshold 0.5, we find that 36% of these countries were reclassified in the period 1999-2004 and that 82% of the reclassifications countries ex-post were assigned to the next *higher* category. Hence many countries originally reported to be of 'low' development in year t < 2006, were in 2006 ex-post revised to have been in fact of 'medium' development status in given year t. As an example, Laos had an HDI of 0.485 in the year 2000. In 2006, however, the $HDI_{2000Laos}^R$ is now reported as 0.523 for data year 2000.

4.3.2. Measurement Error with Respect to the Subindicators

Thus far, we analyzed the data error for the *overall* HDI. Since the same variables used to construct the HDI serve as inputs to many international comparative statistics (used e.g. by OECD, UNESCO, WHO, and World Bank and in the academic literature), it is worthwhile to analyze the subindicators pertaining to health, education and purchasing power in more detail.

The first five columns of Table 4 display basic summary statistics of the subindicator updating error $\boldsymbol{\varepsilon}$ and the overall HDI updating error e for our sample of 72 non-industrialized countries. In general, the standard deviations of the health and education indexes are larger than the standard error of the income statistics. It is interesting to note, however, that the main driver for the HDI upward bias stems from the change to the purchasing power index (m_{income}=0.02).[13] Instead, the errors on the health and the education indices show distributions that are centered around zero. Note, however, that the min/max columns in table 4 still reveal enormous changes due to the data updating; for example, the income index changed by 15% and the education index even by 25% of the total scale from 0 to 1.

One may ask whether the three subindicator updating errors are correlated. An analysis of the year by year correlation matrices of the errors does not show any systematic co-movement, as the correlation coefficients are all close to zero in all years. This suggests that the statistical adjustments on the three dimensions are independent of each other (and indicates that the respective national statistical offices responsible for health, education, and income statistics have no systematic contemporaneous responses). Furthermore, statistical independence of the three subindicator error variables ε_k implies that their errors must be on average larger than the variance of the HDI error e, which is confirmed by table 4. Hence, while the three subindicator errors offset each other with respect to the HDI,[14] when working with the variables of education, income and health, one faces even larger data error.

Although this paper focuses on developing countries, one also may ask, what role measurement error plays for the industrialized world. Table 4 shows a comparison of means of the updating errors and shows the ratio of standard deviations between the industrialized countries and the developing countries. What we find is not flattering for the industrialized world. The industrialized countries have on average *larger* updating bias on all three

[13] Statistically this upward bias with a standard deviation of 0.02 is not significantly different from zero
[14] Under the assumption of independence, the standard deviation for the composite HDI error, e, is given by std(e)=SQRT[($\Sigma_k s_k^2$/9)], which, after replacing s_k by s_hat$_k$, then equals to std(e)= 0.0163. The estimated standard deviation of the HDI measurement error by formula (1) (applied to period C) is 0.0158 (see table 4), hence, in fact, very close to this theoretical result.

subindicators compared to the non-industrialized countries. Only the variability of these updates is less pronounced, as shown by the lower ratio of standard deviations in the last column, confirming the downward trend of Figure 6.

4.4 Use of the HDI statistics in the academic literature:

The HDI has been increasingly employed in the academic literature to describe the evolution of the world's "welfare" distribution in terms of various measures of inequality, such as the Gini coefficient, and to discuss the path of polarization (*e.g.* Pillarisetti, 1997; Ogwang, 2000; Mazumdar, 2002; Noorbakhsh, 2006; Prados de la Escosura, 2007). The results published in these studies, however, can differ largely depending on which year the researcher collected the data. To illustrate, in Figure 8 we display HDI Gini coefficients using the formulas h_A, h_B and h_C for data covering the years 1990 to 2006. The values produced by h_A are about 50% *higher* and the time trend *steeper* compared to the time series generated by formula h_C. This substantial difference would lead to different conclusions or policy recommendations by the analyst. For a recent discussion on the relevance of levels and gradients of Gini estimates see for example Sala-i-Martin (2006) and Prados de la Escosura (2007).

Further we find that a number of recent studies are very sensitive to random selection of countries that is due to the "arbitrariness" of the cut-off values: For example in the macroeconomic literature, Mazumdar (2002) and Noorbakhsh (2006) use the triple bins to analyze the existence of convergence clubs (Quah, 1996) by testing the beta and the sigma conditional convergence hypothesis (originally discussed in Barro and Sala-i-Martin, 1992). In particular, Noorbakhsh (2006) runs beta-convergence regressions of the form

$$\ln(hdi_{it+T}/hdi_{it})/T = \alpha + \beta \ln(hdi_{it}) + \varepsilon_{it} \quad (3)$$

conditional on the country belonging to the ‘low’ development bin. The dependent variable is the annualized growth of the HDI variable for country i over the period t to $t+T$ and hdi_{it} is the ratio of HDI in the i^{th} country to the average for the sample.[15] The regression is then repeated for the bins ‘medium’ and ‘high’ and the comparison of the β estimates is used to analyze the existence of convergence clubs.

To illustrate the consequences of the random selection, we first rerun the convergence regression (3) conditional on the HDI being in the interval [0.5, 0.8) as specified in Noorbakhsh (2006, p. 10, table 3). Then we perform the same regression with the adjusted cut-off values [0.55, 0.70], which we motivated in section 3.4. The results are displayed in Table 3. Comparing the main parameter of interest, β, the estimate of the second regression is about 100% off the first regression, as it is almost exactly twice that of the first regression which would imply a much faster speed of convergence. Also note that the β estimates are statistically very different for the [0.5,0.8) and [0.55,0.70) sample respectively. This example demonstrates that regression results based on the reported HDI are very sensitive to changes of the HDI triple bin classification system.

4.5 Implications of the results in statistical analysis

Econometrically speaking, the average error measures σ_D and σ_F calculated in section 3.3 imply that there is a 3% and 19% downward attenuation bias in a ordinary least squares (OLS) regression $y = \beta_1 + \beta_2\text{HDI}^* + \varepsilon$, if the observed HDI—instead of the “true” (but unknown)

[15] A value of β in the range of (-1, 0) would imply β-convergence of the countries in the sample. A β of zero means no convergence and a positive value for β indicates divergence, with the speed of convergence/divergence the higher the absolute value of β.

HDI*—is used as the regressor variable (for any variable y of interest). The bias of the OLS estimate b_2 is given by[16]

$$\text{plim } b^D{}_2 = [1\text{-}\sigma^2{}_{\text{D}}/(\sigma^2{}_{\text{D}}+\sigma^2{}_{\text{HDI*}})]\beta_2 \approx 0.97\beta_{2,}$$

and

$$\text{plim } b^F{}_2 = [1\text{-}\sigma^2{}_{\text{F}}/(\sigma^2{}_{\text{F}}+\sigma^2{}_{\text{HDI*}})]\beta_2 \approx 0.81\beta_{2,}$$

This is important since in many econometric cross country studies the HDI is used as a regressor and regressand (see for example Arcelus et al., 2005, Globerman, Shapiro, 2002; Jahan, 2000; Mazumdar, 2002; Neumayer, 2003; Noorbakhsh, 2006; Ogwang, 2000; Pillarisetti, 1997; Prados de la Escosura, 2007; Sanyal and Samanta, 2004). This is even more crucial when working with the individual subindicator variables, since (as shown in section 4.3.2) their average standard deviation of the measurement error is larger than the error of the HDI. Figure 9 displays the relationship between the attenuation bias and the standard deviations of the error variables for the range of noise measures as displayed in Table 1, with the lowest attenuation for Algeria and the highest for Niger.

5. Conclusions

Frequently social and economic indicators on a country are collapsed into a single, unit free and often double bounded index which forms the basis for cross country comparisons. Such indexes are used to assess country investment risk, political stability, development status, to name but a few. The objective of this paper is to show some of the consequences if indicators are subject to data error. In our empirical analysis we examine the United Nations' Human Development Index (HDI) which has become the most widely used measure to communicate the

[16] $\sigma^2{}_{\text{HDI*}}$ is approximated by the empirical analogue of the 2006 HDI scores, $\hat{\sigma}^2{}_{\text{HDI*}} = 0.027$.

state of a country’s development status. The HDI is currently further applied to differentiate between countries of ‘low’, ‘medium’ and ‘high’ development status. Institutions as well as the academic literature explicitly and implicitly accept the HDI values of 0.5 and 0.8 to separate countries into these triple bins.

We identify three sources of HDI data error and make the following three empirical contributions. First, we calculate country specific noise measures due to measurement error and formula choice/inconsistencies in the cut-off values. Second, we calculate the misclassification measures with respect to these three sources of data error by simulating the probabilities of being misclassified and sensitivity analysis of the cut-off values. Third, we reproduce prior academic studies and again apply sensitivity analysis with respect to the three sources of data error. Regarding our first contribution we find that the HDI statistics contain a substantial amount of noise on the order of 0.01 to 0.11 standard deviations. Secondly, we show that up to 45% of the developing countries are misclassified due to failure to update the cutoff values. The continuous HDI score jointly with this framework of the discrete classification system is vulnerable when many countries are close to the thresholds, as is the case in the most recent years. Third, we discuss various empirical examples from the prior macroeconomic/development literature where the HDI has been employed (Gini coefficients, convergence regressions and foreign aid) and find that its use is very problematic as key parameters of the past academic literature vary by up to 100% in their values.

Our results raise serious concerns about the triple-bin classification system and we suggest that the United Nations should discontinue the practice of classifying countries into these bins of human development. In our view the cut-off values are arbitrary, can provide incentives for strategic behavior in reporting official statistics, and have the potential to misguide politicians, investors, charity donators and the public at large.

This paper did not investigate the drivers of why in the early years of the HDI—when its political role was still uncertain—its distribution as displayed in Figure 1 looked so different from today's. However, we should caution future private investors, donor organizations and users of the charity scorecards not to take the triple bin system as a tool for investments (Arcelus et al. (2005) and the allocation of foreign aid (Neumayer, 2003). The relationship between the availability of development aid as a direct function of the HDI might potentially provide perverse incentives for a developing country to manipulate the subindicator variables, if it has realized the comparative advantage of being i.e. 0.49 vs. a 0.51 country. In fact, announcements such as the statement by Jahan (2000) (discussed in section 4.2) might have just created these incentives. We refer to Oskar Morgenstern (1970):

> *"Governments, too are not free from falsifying statistics. This occurs, for example, when they are bargaining with other governments and wish to obtain strategic advantages or feel impelled to bluff [...]. A special study of these falsified, suppressed, and misrepresented government statistics is greatly needed and should be made."*

References

Alesina, A. and Dollar, D. (2000). Who Gives Foreign Aid to Whom and Why? *Journal of Economic Growth* 5(1), 33-63.

Alesina, A. and Weder, B. (2002). Do Corrupt Governments Receive Less Foreign Aid? *American Economic Review* 92(4), 1126-1137.

Anand, S. and A. Sen. (2000): The Income Component of the Human Development Index. *Journal of Human Development*, Vol. 1(1), 83-106.

Arcelus, Francisco J., Basu Sharma and Gopalan Srinivasan (2005): Foreign Capital Flows and the Efficiency of the HDI Dimensions. *Global Economy Journal*. Vol. 5, Issue 2, Article 4.

Bandyopadhyay, Subhayu and Howard J. Wall (2006): The Determinants of Aid in the Post-Cold War Era. Working Paper 2006-021B, Federal Reserve Bank of St. Louis.

Baliamoune, M. (2003): On the Measurement of Human Well-Being: Fuzzy Set Theory and Sen'S Capability Approach. *World Institute for Development Economic Research* (UNU-WIDER), Research paper 2004/16.

Barro, Robert J., and Xavier Sala-i-Martin (1992): Convergence, *Journal of Political Economy*, C, 223–251.

Chay, K., McEwan, P. J., and Urquiola, M. (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review*, 95(4), 1237-58.

Daily Times (2005): Pakistan ranked at 135th in human development. September 15.

Desai, M. (1991): Human Development: Concepts and Measurement. *European Economic Journal*, 35: 350-7.

Easterly, W., Levine, R. and Roodman, D. (2004). Aid, Policies, and Growth: Comment. *American Economic Review* 94(3), 774-780.

Geneva Global (2007): Geneva Global Performance Philanthropy. 15550 Liberty Ridge Drive Wayne, PA 19087, USA. http://www.genevaglobal.com/sector-priorities/high-hd/.

Globerman, Steven and Daniel Shapiro (2002): National Political Infrastructure and Foreign Direct Investment. Working Paper Number 37. Industry Canada Research Publications Program.

Guindon, G.E. and D. Boisclair (2003): Past, current and future trends in tobacco use. *Tobacco Control*. WHO Tobacco Control Papers, TRENDS2003.

Hargittai, Eszter (1998): Holes in the Net: The Internet and International Stratification Revisited. *Proceedings of the Internet Society's Internet Summit meetings*.

HDR (1990 to 2006): Human Development Report, UNDP, Oxford University Press.

Keiser, Jennifer, Jürg Utzinger, Marcel Tanner, and Burton H Singer (2004): Representation of authors and editors from countries with different human development indexes in the leading literature on tropical medicine: survey of current evidence. *British Medical Journal* 328:1229-1232.

Kelley, Allen C (1991): The Human Development Index: "Handle with Care". *Population and Development Review*, Vol. 17, No. 2, pp. 315-324.

Jahan, Selim (2000): Measurement of Human Development: Seven Questions. Lecture Reading by Deputy Director of the Human Development Report Office, UNDP.

Mazumdar, Krishna (2002): A Note on Cross-Country Divergence in Standard of Living. *Applied Economic Letters* 9(2), 87-90.

McGillivray, M. (1991): ‘Redundant Composite Development Indicator’. *World Development*, 19 (10): 1461-9.

Morse, S. (2003): Greening the United Nations’ Human Development Index. *Sustainable Development*, 11, 183–198.

Morgenstern, Oskar (1970): On the Accuracy of Economic Observations. Second edition, Princeton University Press.

Neumayer, E. (2003). The Determinants of Aid Allocation by Regional Multilateral Development Banks and United Nations Agencies, *International Studies Quarterly* 47(1), 101-22.

Noorbakhsh, Farhad (1998): The human development index: some technical issues and alternative indices *Journal of International Development*, Volume 10, Issue 5, pp.589–605.

Noorbakhsh, Farhad (2006): International Convergence or Higher Inequality in Human Development? Evidence for 1975 to 2002. World Institute for Development Economic Research (UNU-WIDER). Research Paper No. 2006/15.

Ogwang, Tomson (2000): Inter-Country Inequality in Human Development Indicators. *Applied Economic Letters*, 7(7), pp. 443-46.

Quah, Danny (1996): Twin Peaks: Growth and Convergence in Models of Distribution Dynamics, *Economic Journal*, CVI, 1045–1055.

People's Daily (2001): Ghana's Human Development Performance Recommended by UNDP. July 11.

Pillarisetti, J. Ram (1997): An Empirical Note on Inequality in the World Development Indicators. *Applied Economic Letters*, 4 (3), pp. 145-47.

Prados de la Escosura, Leandro (2007): International Inequality and Polarization in Living Standards, 1870-2000: Evidence from the Western World. Universidad Carlos III de Madrid, Working Papers in Economic History, WP 07-05.

Sala-i-Martin, Xavier (2006): The World Distribution of Income: Falling Poverty and…Convergence, Period. *Quarterly Journal of Economics.* 121(2), pp. 351-97.

Sanyal, Rajib N. and Subarna K. Samanta (2004): Determinants of bribery in international business. Thunderbird International Business Review, Volume 46, Issue 2, pp133–148.

Srinivasan, T. N. (1994): Human Development: A New Paradigm or Reinvention of the Wheel? *American Economic Review*, Vol. 84, No. 2, pp. 238-243.

United Nations (1996): Africa Recovery, UN-NADAF Midterm Review, September 1996.

Table 1: Country *i* specific standard deviations and probabilities of belonging to development category *j*

			Measures based on formula updates (F)					Measures based on measurement error due to data revisions (D)				
Country *i*	2006 reported human development status	2006 HDI	$\sigma_{F,i}$	Pr{*i*='low'}	Prob{*i*='mid'}	Prob{*i*='high'}	Prob{*i*=mis-classified}	$\sigma_{D,i}$	Pr{*i*='low'}	Prob{*i*='mid'}	Prob{*i*='high'}	Prob{*i*=mis-classified}
Niger	'low'	0.31	0.11	95.5	4.5	0.0	4.5	0.03	100.0	0.0	0.0	0.0
Mali	'low'	0.34	0.10	94.4	5.6	0.0	5.6	0.03	100.0	0.0	0.0	0.0
Burkina Faso	'low'	0.34	0.10	94.9	5.1	0.0	5.1	0.02	100.0	0.0	0.0	0.0
Chad	'low'	0.37	0.09	92.2	7.8	0.0	7.8	0.04	100.0	0.0	0.0	0.0
Ethiopia	'low'	0.37	0.09	91.3	8.7	0.0	8.7	0.03	100.0	0.0	0.0	0.0
Burundi	'low'	0.38	0.10	88.6	11.4	0.0	11.4	0.02	100.0	0.0	0.0	0.0
Mozambique	'low'	0.39	0.10	86.4	13.6	0.0	13.6	0.03	100.0	0.0	0.0	0.0
Malawi	'low'	0.40	0.11	81.8	18.1	0.0	18.1	0.01	100.0	0.0	0.0	0.0
Zambia	'low'	0.41	0.07	89.8	10.2	0.0	10.2	0.04	98.8	1.2	0.0	1.2
Côte d'Ivoire	'low'	0.42	0.08	84.5	15.5	0.0	15.5	0.02	100.0	0.0	0.0	0.0
Benin	'low'	0.43	0.09	79.8	20.2	0.0	20.2	0.03	99.3	0.7	0.0	0.7
Tanzania	'low'	0.43	0.07	83.0	17.0	0.0	17.0	0.02	99.9	0.1	0.0	0.1
Nigeria	'low'	0.45	0.09	71.3	28.7	0.0	28.7	0.04	88.4	11.6	0.0	11.6
Senegal	'low'	0.46	0.07	70.4	29.6	0.0	29.6	0.02	99.6	0.4	0.0	0.4
Mauritania	'low'	0.49	0.08	57.3	42.7	0.0	42.7	0.03	67.1	32.9	0.0	32.9
Kenya	'low'	0.49	0.07	54.8	45.2	0.0	45.2	0.02	64.8	35.2	0.0	35.2
Zimbabwe	'low'	0.49	0.06	56.2	43.8	0.0	43.8	0.03	62.8	37.2	0.0	37.2
Lesotho	'low'	0.49	0.07	53.5	46.5	0.0	46.5	0.02	59.8	40.2	0.0	40.2
Togo	'low'	0.50	0.07	52.8	47.2	0.0	47.2	0.04	55.2	44.8	0.0	44.8
Uganda	'medium'	0.50	0.08	49.1	50.9	0.0	49.1	0.02	46.0	54.0	0.0	46.0
Cameroon	'medium'	0.51	0.07	46.5	53.5	0.0	46.5	0.04	44.3	55.7	0.0	44.3
Madagascar	'medium'	0.51	0.07	45.0	55.0	0.0	45.0	0.03	38.9	61.1	0.0	38.9
Sudan	'medium'	0.52	0.07	40.6	59.4	0.0	40.6	0.03	31.6	68.4	0.0	31.6
Congo	'medium'	0.52	0.07	38.7	61.3	0.0	38.7	0.05	34.7	65.3	0.0	34.7
Pap. N. Guinea	'medium'	0.52	0.06	34.5	65.5	0.0	34.5	0.04	26.9	73.1	0.0	26.9
Nepal	'medium'	0.53	0.08	36.3	63.6	0.0	36.3	0.02	9.5	90.5	0.0	9.5
Bangladesh	'medium'	0.53	0.07	34.2	65.8	0.0	34.2	0.02	6.6	93.4	0.0	6.6
Ghana	'medium'	0.53	0.07	31.6	68.4	0.0	31.6	0.04	19.6	80.4	0.0	19.6
Pakistan	'medium'	0.54	0.07	27.5	72.5	0.0	27.5	0.03	9.8	90.2	0.0	9.8
Lao Peoples	'medium'	0.55	0.07	23.0	77.0	0.0	23.0	0.06	17.8	82.2	0.0	17.8
Botswana	'medium'	0.57	0.05	6.4	93.6	0.0	6.4	0.04	2.9	97.1	0.0	2.9
India	'medium'	0.61	0.06	3.1	96.8	0.1	3.1	0.01	0.0	100.0	0.0	0.0
Morocco	'medium'	0.64	0.04	0.1	99.9	0.0	0.1	0.02	0.0	100.0	0.0	0.0
Guatemala	'medium'	0.67	0.05	0.0	99.5	0.5	0.0	0.02	0.0	100.0	0.0	0.0
Honduras	'medium'	0.68	0.07	0.3	95.5	4.1	0.3	0.02	0.0	100.0	0.0	0.0
Mongolia	'medium'	0.69	0.08	1.0	89.9	9.2	1.0	0.06	0.1	96.7	3.3	3.4

Country i	2006 reported human development status	2006 HDI	Measures based on formula updates (F) $\sigma_{F,i}$	Pr{i='low') in %	Prob{i='mid') in %	Prob{i='high') in %	Prob{i=mis-classified) in %	Measures based on measurement error due to data revisions (D) $\sigma_{D,i}$	Pr{i='low') in %	Prob{i='mid') in %	Prob{i='high') in %	Prob{i=mis-classified) in %
Bolivia	'medium'	0.69	0.06	0.2	95.1	4.8	5.0	0.02	0.0	100.0	0.0	0.0
Nicaragua	'medium'	0.70	0.05	0.0	97.0	3.0	3.0	0.04	0.0	99.4	0.6	0.6
Egypt	'medium'	0.70	0.04	0.0	99.1	0.9	0.9	0.03	0.0	99.8	0.2	0.2
Vietnam	'medium'	0.71	0.09	0.9	83.9	15.2	16.1	0.02	0.0	100.0	0.0	0.0
Indonesia	'medium'	0.71	0.07	0.1	90.8	9.1	9.2	0.03	0.0	99.9	0.1	0.1
Syria	'medium'	0.72	0.07	0.1	89.1	10.9	11.0	0.07	0.1	89.6	10.3	10.4
Jamaica	'medium'	0.72	0.07	0.1	85.1	14.8	14.9	0.02	0.0	100.0	0.0	0.0
Algeria	'medium'	0.73	0.04	0.0	97.4	2.6	2.6	0.04	0.0	97.9	2.1	2.1
El Salvador	'medium'	0.73	0.06	0.0	89.3	10.7	10.7	0.05	0.0	91.9	8.1	8.1
Iran	'medium'	0.75	0.05	0.0	86.9	13.1	13.1	0.02	0.0	98.5	1.5	1.5
Dominican R.	'medium'	0.75	0.06	0.0	80.9	19.1	19.1	0.02	0.0	99.9	0.1	0.1
Sri Lanka	'medium'	0.76	0.09	0.2	69.0	30.7	30.9	0.02	0.0	97.1	2.9	2.9
Turkey	'medium'	0.76	0.06	0.0	72.8	27.1	27.1	0.01	0.0	93.8	6.2	6.2
Paraguay	'medium'	0.76	0.07	0.0	75.1	24.9	24.9	0.03	0.0	100.0	0.0	0.0
Tunisia	'medium'	0.76	0.05	0.0	71.9	28.1	28.1	0.02	0.0	90.0	10.0	10.0
Jordan	'medium'	0.76	0.07	0.0	78.1	21.9	21.9	0.03	0.0	96.8	3.2	3.2
Philippines	'medium'	0.76	0.07	0.0	71.5	28.5	28.5	0.03	0.0	91.4	8.6	8.6
Peru	'medium'	0.77	0.05	0.0	74.2	25.8	25.8	0.02	0.0	97.4	2.6	2.6
China	'medium'	0.77	0.08	0.0	66.5	33.5	33.5	0.02	0.0	95.4	4.6	4.6
Lebanon	'medium'	0.77	0.06	0.0	67.0	33.0	33.0	0.04	0.0	75.7	24.3	24.3
Saudi Arabia	'medium'	0.78	0.06	0.0	64.7	35.3	35.3	0.02	0.0	87.9	12.1	12.1
Thailand	'medium'	0.78	0.08	0.0	57.9	42.1	42.1	0.02	0.0	80.9	19.1	19.1
Venezuela	'medium'	0.78	0.08	0.0	58.4	41.6	41.6	0.02	0.0	80.7	19.3	19.3
Colombia	'medium'	0.79	0.08	0.0	55.2	44.8	44.8	0.02	0.0	72.9	27.1	27.1
Brazil	'medium'	0.79	0.07	0.0	54.6	45.4	45.4	0.02	0.0	63.1	36.9	36.9
Mauritius	'high'	0.80	0.08	0.0	50.0	50.0	50.0	0.01	0.0	50.0	50.0	50.0
Malaysia	'high'	0.81	0.08	0.0	47.6	52.4	47.6	0.01	0.0	23.5	76.5	23.5
Trinidad/Tobago	'high'	0.81	0.09	0.0	45.4	54.6	45.4	0.01	0.0	39.5	60.5	39.5
Panama	'high'	0.81	0.08	0.0	46.1	53.9	46.1	0.03	0.0	25.6	74.4	25.6
Mexico	'high'	0.82	0.09	0.0	40.7	59.3	40.7	0.01	0.0	3.3	96.7	3.3
Costa Rica	'high'	0.84	0.11	0.1	35.0	65.0	35.1	0.01	0.0	0.0	100.0	0.0
Uruguay	'high'	0.85	0.09	0.0	28.2	71.8	28.2	0.01	0.0	0.0	100.0	0.0
Chile	'high'	0.86	0.09	0.0	26.6	73.4	26.6	0.01	0.0	0.0	100.0	0.0
Argentina	'high'	0.86	0.07	0.0	18.2	81.8	18.2	0.01	0.0	0.0	100.0	0.0
Korea	'high'	0.91	0.06	0.0	3.9	96.1	3.9	0.02	0.0	0.0	100.0	0.0
Hong Kong	'high'	0.93	0.05	0.0	0.4	99.6	0.4	0.02	0.0	0.0	100.0	0.0
Expected # of countries misclassified					17.6					8.4		

Table 2: As of 2006, countries misclassified due to the arbitrary cut off points

Countries with $HDI_{2006} \in$ [0.5 and 0.55)	Countries with $HDI_{2006} \in$ [0.7 and 0.8)
Bangladesh	Brazil
Cameroon	China
Congo	Colombia
Ghana	Dominican Republic
Madagascar	Algeria
Nepal	Egypt
Pakistan	Indonesia
Papua New Guinea	Iran, Islamic Rep. of
Sudan	Jamaica
Uganda	Jordan
	Lebanon
	Sri Lanka
	Peru
	Philippines
	Paraguay
	Saudi Arabia
	El Salvador
	Syrian Arab Republic
	Thailand
	Tunisia
	Turkey
	Venezuela
	Vietnam

Table 3: Official development assistance (ODA) received in US dollar per capita by year and human development category

	2006	2005	2004	2003	2002	2001
‘medium’	7.2	6.5	6.5	5.7	5.9	6.6
‘low’	30.1	27.9	24.2	18.4	14.9	14.5

Data are from the Human Development Reports 2001 to 2006.

Table 4: Updating error summary statistics for the period 1999 to 2004

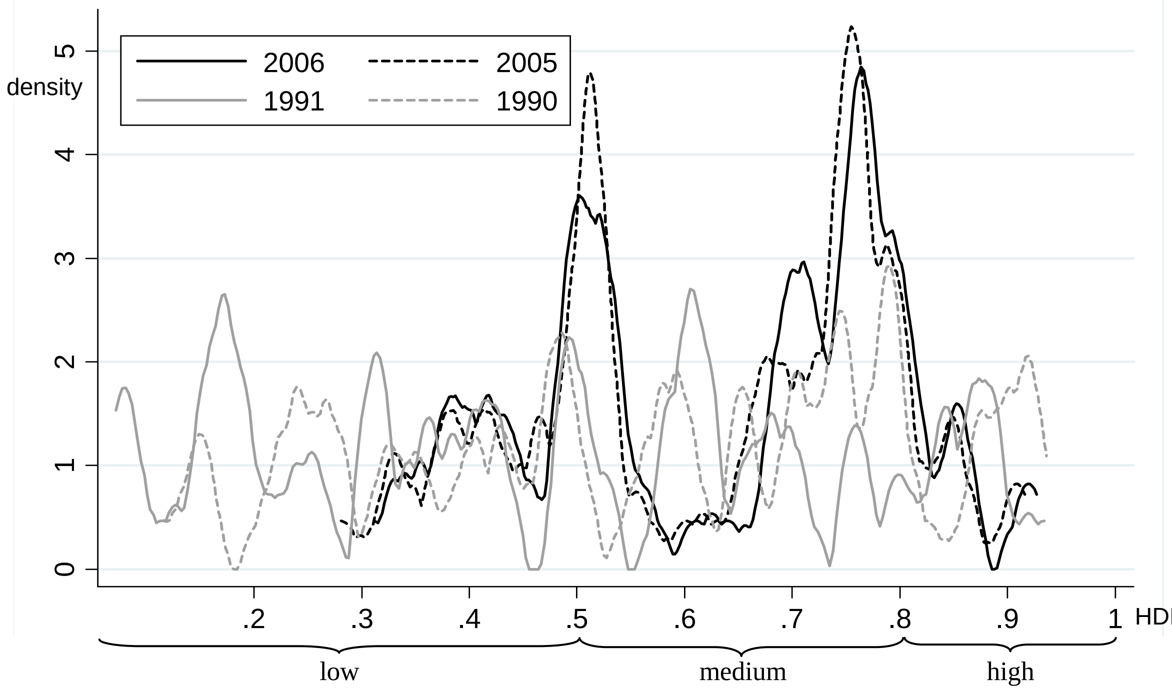
	Developing Countries				Industrialized Countries				Industrial vs. Developing Countries	
Indicators	Mean	std. dev.	min	max	mean	std. dev.	min	max	Difference in means	Ratio of std.dev.s
HDI	0.01	0.02	-0.06	0.08	0.01	0.01	-0.03	0.05	0.006	0.493
Health	0.00	0.04	-0.14	0.11	0.00	0.02	-0.11	0.06	0.004	0.424
Education	0.00	0.03	-0.11	0.25	0.00	0.02	-0.13	0.08	0.00	0.646
Income	0.02	0.02	-0.07	0.15	0.03	0.02	-0.05	0.13	0.011	1.062

Table 3: Convergence club regression results for medium development category

Sample conditional on	$HDI_{2006} \in [0.5,0.8)$	$HDI_{2006} \in [0.55,0.70)$
constant α	-.02556 (-56.69)	-.02847 (-35.36)
slope β	-.01380 (- 6.74)	-.02667 (-4.59)
adjusted R^2	.53	.74

t statistics in parentheses.

Figure 1: Historical HDI scores for Developing Countries in 1990/91 and 2005/06[17]



[17] On the horizontal axis we display the HDI, which ranges from 0 to 1. 1990/91 are the first and 2005/06 are last two years for which the HDI scores originally have been made available (HDR, 1990, 1991, 2005, 2006). To make the HDI-distributions comparable across years we use the balanced panel of 72 developing countries that have been evaluated by the UNDP for all years. Countries that existed for a subset of years only (e.g. Croatia) are not considered. All densities are estimated by the Epanechnikov kernel method with bandwidth 0.02.

Figure 2: Density of HDI as published by the HDR reports

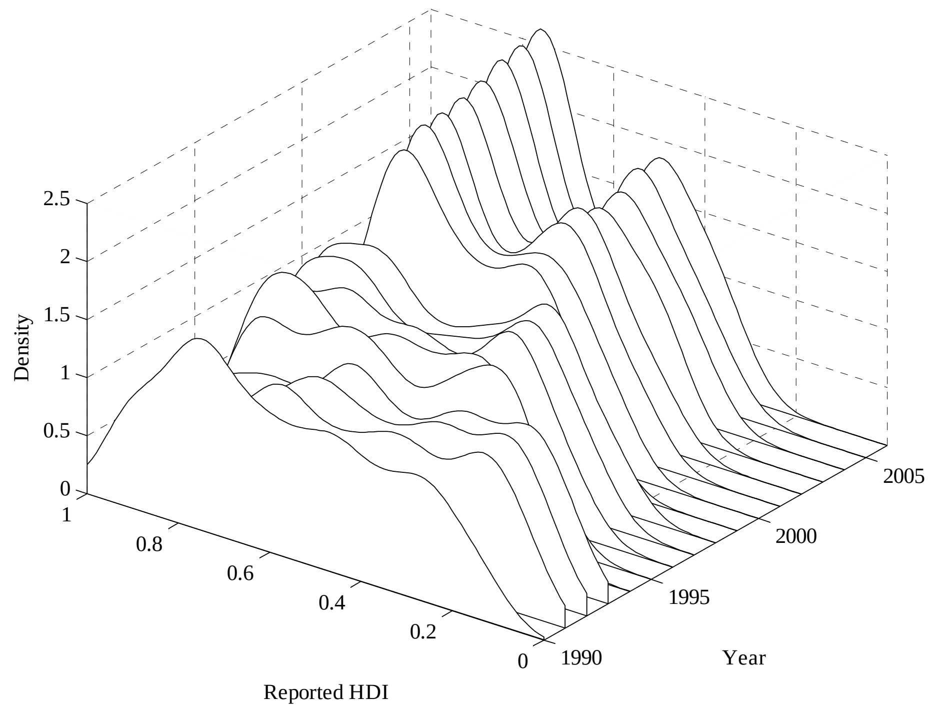


Figure 3: Representation of data error of a country with HDI = 0.65

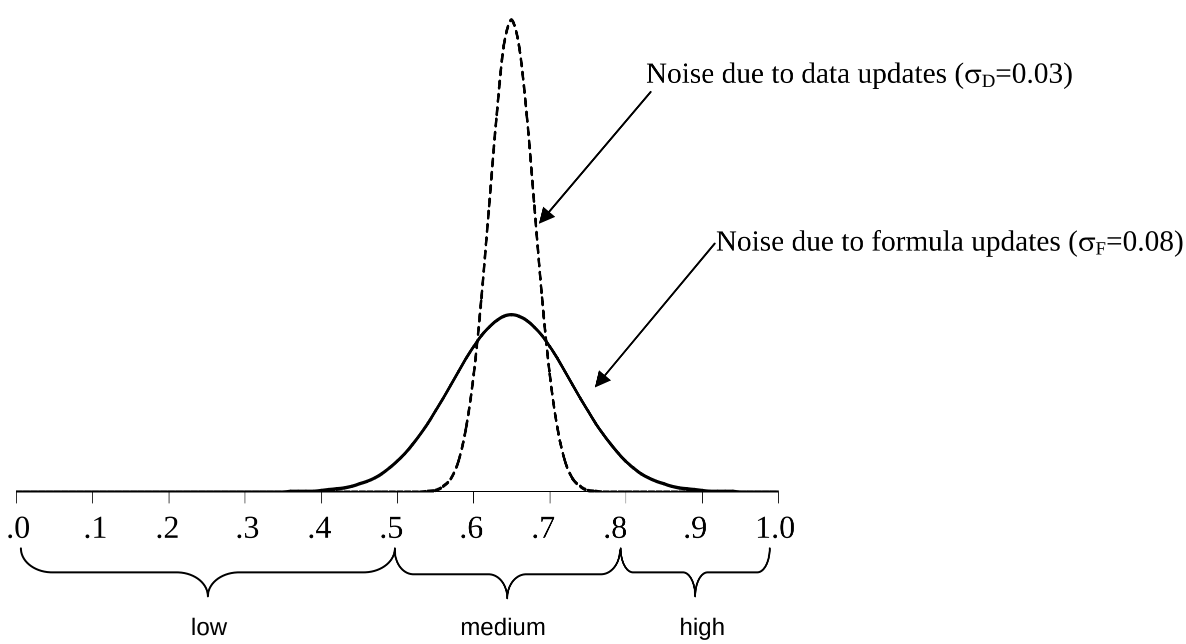


Figure 4: HDI of 1975 of Portugal and Venezuela as reported in the years 1999 to 2006

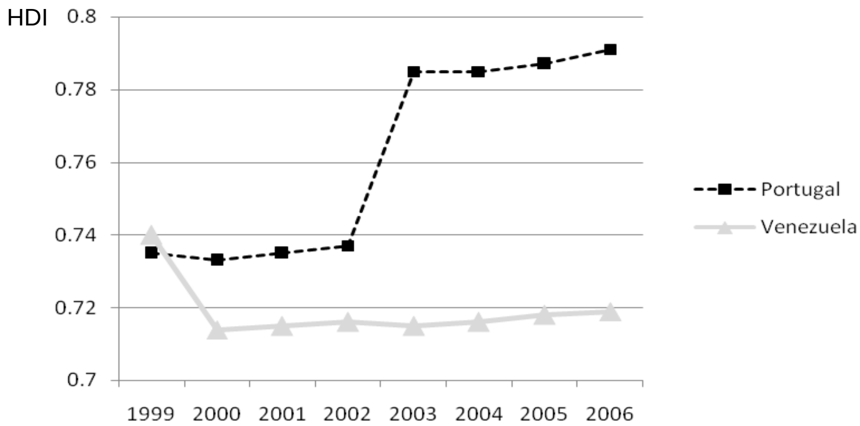
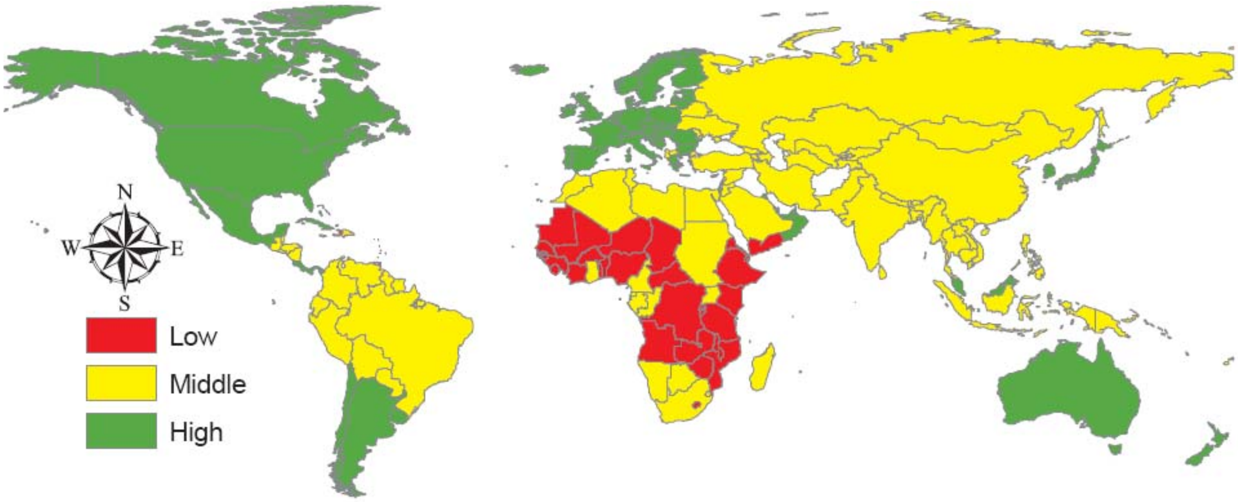
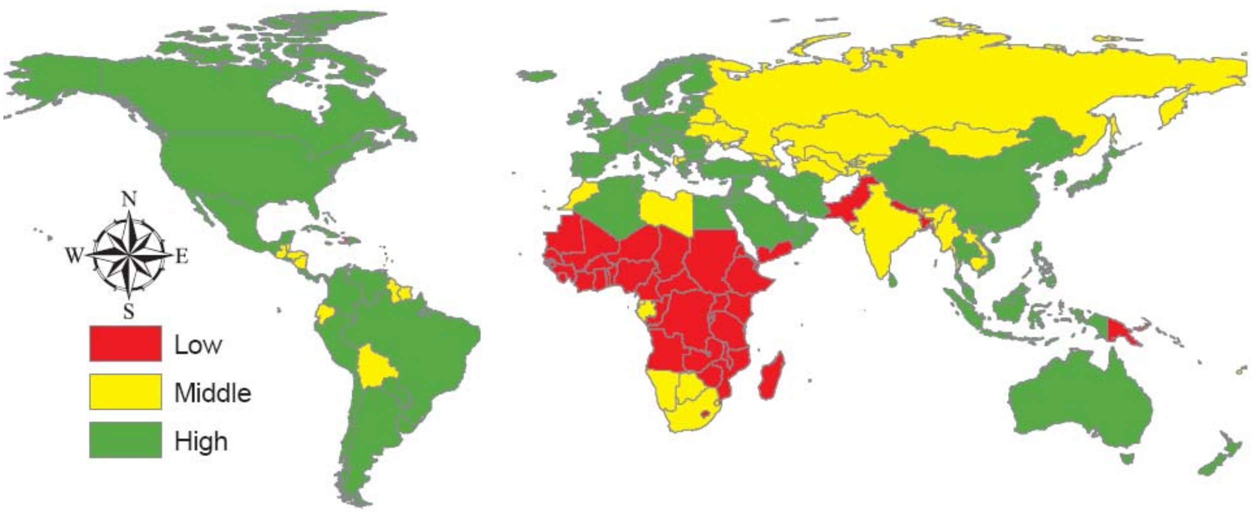


Figure 5: World map of the Human Development Index (2006)

Panel (a): Reported Human Development Index 2006



Panel (b): Adjusted Human Development Index 2006



Note: Panel (a) displays the classification using the actually reported HDI Index for the year 2006 for all reported countries (industrialized and non-industrialized). Countries in white have no reported data. Panel (b) displays the same classification for industrialized countries as in panel (a). For the 72 non-industrialized countries, the classification is based on the revised thresholds that we calculate in section 3.4. if the UNDP had consistently updated the cutoff values for classification.

Figure 6: Relationship between countries' development status and the standard deviations due to measurement error generated by data updates.

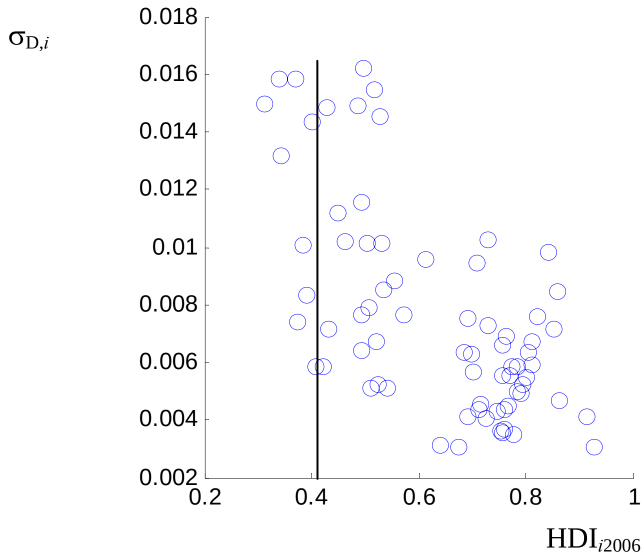


Figure 7: Densities of the HDI data updating error for the years 1999 to 2005

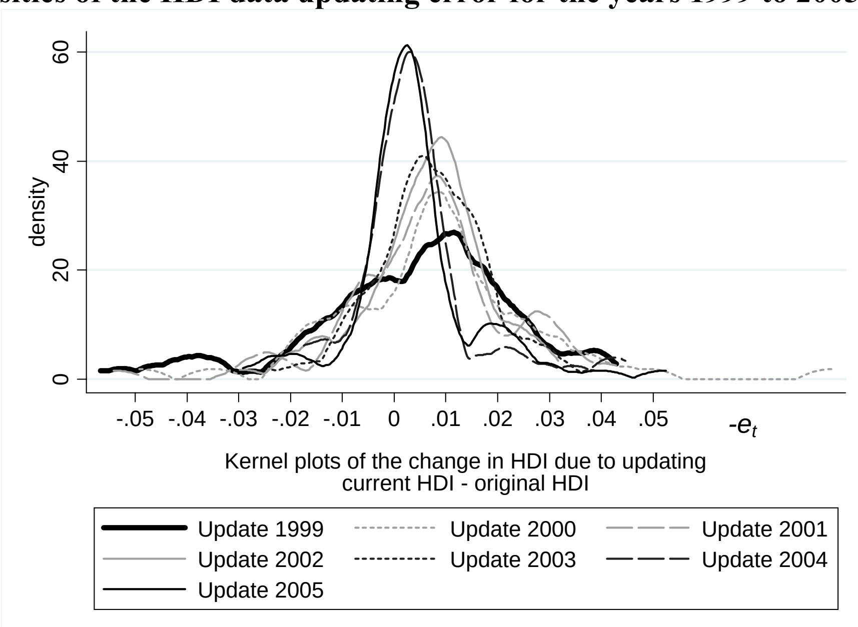


Figure 8: Gini Coefficients computed by the HDI formulas A, B and C

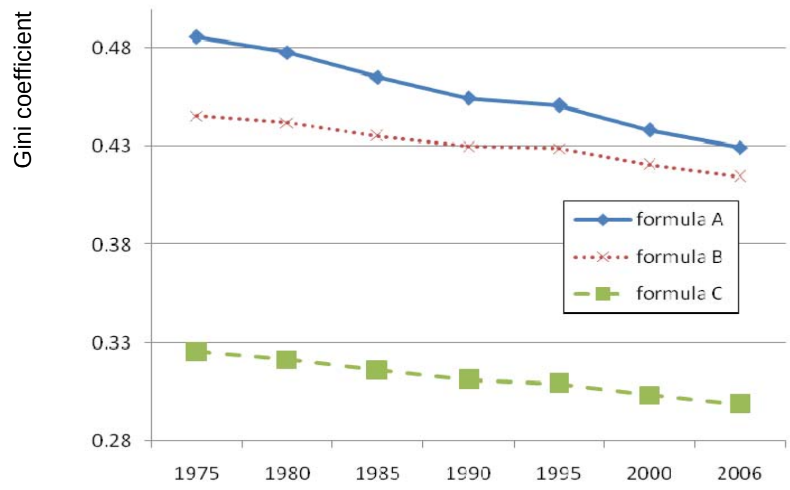


Figure 9: Attenuation bias as function of the error variable standard deviation