NBER WORKING PAPER SERIES

INEQUALITY WITHIN COUNTRIES IS FALLING: UNDERREPORTING-ROBUST ESTIMATES OF WORLD POVERTY, INEQUALITY AND THE GLOBAL DISTRIBUTION OF INCOME

Maxim Pinkovskiy
Xavier Sala-i-Martin
Kasey Chatterji-Len
William H. Nober

Working Paper 32203
http://www.nber.org/papers/w32203

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2024

We thank Ruchi Avtar and Marie Camara for outstanding research assistance. We thank Leonardo Gasparini and Leopoldo Tornarolli for sharing with us standardized regional survey data for multiple Latin American countries through SEDLAC. We thank Arvind Subramanian for guiding us to the "junked" report of the 2017 Indian NSS. We thank numerous staff members at the Luxembourg Income Study for help using their data. We thank Christoph Lakner for sharing with us code for using the Luxembourg Income Study data. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, or the National Bureau of Economic Research. Any errors or omissions are our own.

Inequality Within Countries is Falling: Underreporting-Robust Estimates of World Poverty, Inequality and the Global Distribution of Income
Maxim Pinkovskiy, Xavier Sala-i-Martin, Kasey Chatterji-Len, and William H. Nober
NBER Working Paper No. 32203
March 2024
JEL No. C83,D31,D33

ABSTRACT

Household surveys suffer from persistent and growing underreporting. We propose a novel procedure to adjust reported survey incomes for underreporting by estimating a model of misreporting whose main parameter of interest is the elasticity of regional national accounts income to regional survey income, which is closely related to the elasticity of underreporting with respect to income. We find this elasticity to be substantial but roughly constant over time, implying a large but relatively constant correction to survey-derived inequality estimates. Underreporting of income by the bottom 50% of the world income distribution has become particularly important in recent decades. We reconfirm the findings of the literature that global poverty and inequality have declined dramatically between 1980 and 2019. Finally, we find that within-country inequality is falling on average, and has been largely constant since the 1990s.

Maxim Pinkovskiy
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
maxim.pinkovskiy@ny.frb.org

Xavier Sala-i-Martin
Department of Economics
Columbia University
420 West 118th Street, 1005
New York, NY 10027
and NBER
xs23@columbia.edu

Kasey Chatterji-Len
Federal Reserve Bank of New York
kasey.clen@gmail.com

William H. Nober
Columbia University
420 118th Street
New York, NY 10027
whn2105@columbia.edu

1 Introduction

Over the past two decades, it has become widely accepted that global poverty and inequality have fallen substantially since the 1980s. Evidence for both of these claims has generally been obtained from household surveys conducted in developing and developed countries around the world, and assembled by the World Bank as part of its Poverty and Inequality Platform (World Bank 2023; PIP, formerly PovCalNet). These data are the basis for the World Bank's poverty monitoring program that has found sustained declines in $1-a-day poverty around the world (Chen and Ravallion 2010). Research by others using these surveys has found even steeper poverty declines (Sala-i-Martin 2002a and b, 2006), and a large literature has employed these surveys and their antecedents to find that world inequality is decreasing (Bourguignon and Morrisson 1992, Bhalla 2001; Sala-i-Martin 2002a and b, 2006; Pinkovskiy and Sala-i-Martin 2009).

However, the literature has recognized considerable uncertainty in the income and distributional data coming from these surveys, with nonrandom nonresponse and misreporting of data being prime challenges to interpreting the data as true (Deaton 2005; Korinek, Mistiaen and Ravallion 2006). Comparing household survey means and GDP per capita with data on satellite-recorded nighttime lights, Pinkovskiy and Sala-i-Martin (2016) further substantiate this concern, and find that household survey means have very little predictive power for unobserved true income once GDP per capita is used. For this reason, a substantial strand of the literature (Bhalla 2001, Sala-i-Martin 2002a and b, 2006; Pinkovskiy and Sala-i-Martin 2009) have used GDP per capita instead of the survey mean to re-center country income distributions. Chen and Ravallion (2010) also recognize the fact that survey means are likely to systematically underestimate unobserved true income by considering a robustness check in which the mean of the income distribution is taken to be the geometric mean of the survey mean and GDP per capita. Nevertheless, the baseline poverty and inequality estimates in the PIP reflect survey data alone.

Now, if household surveys systematically understate aggregate income, it is likely that they also distort its distribution. Deaton (2005) recognized this possibility, however all of the literature cited above – both Chen and Ravallion (2010) and the literature rescaling country income distributions to be centered at GDP per capita – has assumed survey nonresponse and misreporting to only affect the mean and otherwise operate in a distribution-neutral manner. An alternative approach to solving this problem has been to use tax data (Piketty 2003, Piketty and Saez 2003, Alvaredo et al. 2018), which is applied to the world distribution of income in Chancel et al. 2022. The resulting dataset of individual country and global Lorenz curves, using a mix of survey and tax data as well as surveys adjusted using other countries' relationships between surveys and tax data is known as the World Inequality Database (WID). While tax data is of high quality for most developed countries, covering a large fraction of their population and national income (but see Burkhauser et al. 2015), tax data in the developing world frequently is absent completely (as in Africa) or covers very small fractions of population and income. Thus, for the world's two largest countries, Piketty, Yang and Zucman (2019) see tax data for only 0.5% of all Chinese tax units, and Chancel and Piketty (2019) see tax data for less than 2% of Indian ones. Therefore, questions on how to correct household surveys for nonresponse and misreporting remain.

In this paper, we propose a methodology to adjust household survey income distributions by exploiting regional data on national accounts income, survey means and survey inequality. Intuitively, while we cannot observe each individual's contribution to national accounts income ("true income") directly in order to compare it to their report in the household survey, we can do so at the level of subnational regions. Since household surveys are conducted at the national level, every respondent in every region should face the same questionnaire, and the relationship between their survey response and their actual income should be roughly homogeneous across regions. Then, if survey misreporting is a function of a respondent's true income, we should be able to deduce the misreporting relationship based on the empirical relationship between subnational region income in the surveys and subnational region income in the national accounts. In particular, if in household surveys, the rich underreport their income by a greater fraction than do the poor, one would expect the regional distribution of household survey means to be less unequal than the regional distribution of GDP per capita. Comparing the two distributions, we should be able to estimate the extra fraction of income that the rich underreport on household surveys compared to what the poor underreport, and adjust the reported household survey distributions.

Our main assumptions are that 1) true income and survey income have, on average, a nationally homogeneous loglinear relationship, and 2) data on regional GDP are subject to error that is not correlated with regional survey data. The first assumption is a linear approximation to any misreporting relationship in which reported survey income is a function of unobserved true income. We produce evidence that the first assumption nests the approach of the WID of using tax data to correct survey income distributions by noting the strong loglinear associations between Lorenz curves in household surveys and Lorenz curves estimated by combining surveys and tax data in the WID. The second assumption exploits the intuition developed in the previous paragraph because subnational GDP data are explicitly constructed on a regional basis, while survey instruments are generally common nationwide. In particular, we do not have to assume that the relationship between survey income and true income is constant across any particular set of countries, or within a country across years, as do Chancel et al. (2023) in constructing their tax data-adjusted income distribution for Africa.

Armed with these assumptions, we collect regional GDP per capita and regional household survey means data from a variety of sources to estimate the loglinear relationship between true income and survey income for countries accounting for up to 75% of the world's population. Some of the most important sources for the regional GDP per capita data are the dataset used by Gennaioli et al. (2014), the OECD Stats database, and the national statistical yearbooks for several large countries (in particular, China and India) for the most recent data, while for the regional survey means we employ the Luxembourg Income Study, the Socio-Economic Database for Latin American Countries (SEDLAC) and publications by national statistical offices. For each household survey with regional data, we estimate inequality adjustment parameters using data from that survey (and the corresponding year's national accounts) alone without having to make cross-country homogeneity assumptions. Extrapolating the adjustment parameters for the countries without regional data, we create a consistent world distribution of national accounts GDP per capita, in which the mean is equal to global GDP per capita, and the dispersion is given by suitably adjusting household surveys.

Our methodology does not restrict us to computing the distribution of GDP, but allows us to obtain the distribution of any income-related variable that is available at the subnational level, and for which we can assume that reported survey income is a function of that variable. In particular, it is reasonable to compute the distribution of of national accounts household final consumption expenditure (HFCE) per capita. GDP and HFCE represent two alternative measures of welfare; one that includes saving that is done on behalf of households by firms and governments, and one that focuses only on what households receive directly, or what they actually consume. Deaton (2005) considers HFCE per capita a closer proxy than GDP per capita to the disposable income concept used in most of the World Bank's household surveys for the purpose of computing poverty, and the poverty lines defined in Ravallion (2010) are typically in terms of levels of disposable income. Therefore, we construct estimates of the distribution of HFCE per capita to better compare our results to the World Bank's household survey estimates of global poverty at different poverty rates.

We obtain four main sets of findings. First, we show that the poor also misreport, and that they have been playing a more important role in aggregate misreporting over time. Whereas the bottom 50% of the global income distribution reported a higher fraction of their income in 1980 than did the top 10%, the reverse is true in 2019. Moreover, by 2019, the bottom 50% also account for a larger fraction of overall underreporting of disposable income than their survey income share (which has also been rising over this time period). Hence, the poor also underreport, and their underreported income is a significant fraction both of underreporting across the entire income distribution, as well as of the true underlying income of the poor themselves. Overall, we find that underreporting progressivity – the extent to which the rich misreport more than do the poor – experienced a slight decline between 1980 and 2019, evolving smoothly over time. In contrast, the rates of underreporting progressivity that would rationalize the estimates in the WID must have experienced a dramatic kink after 2000, having been flat for the 1980s and 1990s but rising steeply afterwards. We believe that the patterns we find in underreporting progressivity are more straightforward to explain than those that would be needed to justify the WID.

Second, we replicate the findings of the earlier literature that global poverty and inequality have declined since the 1980s, with Gini inequality in GDP per capita close to but declining faster than comparably measured inequality in the World Inequality Database, and inequality in HFCE per capita close to but declining faster than comparably measured inequality in the World Bank PIP. The global poverty rate at the $2.15 poverty line is just over 6% for our baseline estimate using HFCE per capita, but 10% using the comparable measurement based on household surveys alone in the World Bank PIP. We also find that global welfare, measured by the Sen index (Sen 1996) has increased over 20% more using either our adjusted GDP per capita or our adjusted HFCE per capita distributions relative to either the comparable estimates using the World Bank PIP or using the World Inequality Database. Therefore, even after adjusting for systematic survey misreporting, our results provide a substantially more optimistic picture of the world distribution of income than do the results of the World Bank or of the World Inequality Database. Intuitively, the poor also misreport their income on household surveys (Meyer and Sullivan 2015), and accounting for misreporting at all points on the income distribution does not fully compensate for the enormous difference between national accounts aggregates and household survey means.

Third, we find that not only overall global inequality has fallen, but this decrease was also reflected for within-country inequality. Population-weighted within-country inequality according to a range of measures, and for both the GDP per capita distribution and the HFCE per capita distribution, has declined since the mid-2000s to attain levels of the early to mid-1990s. Within-country inequality has also fallen for several large countries such as China, India and Indonesia. Therefore, the recent declines in inequality are driven by both falling across-country inequality and falling within-country inequality, rather than within-country inequality acting as a headwind to global inequality reduction. While adjusting GDP and HFCE distributions for misreporting makes the within-country inequality reduction more salient, this result does not depend on our adjustment procedure, as the reduction in population-weighted within-country inequality is present in the unadjusted household survey data, and data from the World Inequality Database for the same countries and years as the World Bank PIP are consistent with a stabilization of within-country inequality.

Finally, the world is doing much better than we thought not only at eliminating extreme poverty – reducing the \$2.15-a-day headcount ratio – but also reducing poverty at higher poverty lines. Using the World Bank's preferred poverty lines of \$3.65 a day and \$6.85 a day (the latter two reflecting the medians of lower-middle-income and upper-middle-income country poverty lines), we find that poverty rates at these higher thresholds have declined to 30% and 50%, respectively, of their 1990 levels for the HFCE per capita world distributions, which is considerably lower than estimated by the World Bank using household surveys alone. Hence, the world distribution of income is less defined by a large "precariat" modestly above an extreme poverty line but liable to fall back into destitution following a global shock, but rather increasinly by a true "global middle class" that is not poor even by upper-middle-income country standards.

The rest of the paper is organized as follows: Section 2 describes the empirical approach. Section 3 describes the data. Section 4 presents the results. Section 5 concludes.

2 Using Regional Data to Correct Survey Misreporting

Let y_i be an income measure for individual i, such as their contribution to national accounts GDP or to household final consumption expenditure (HFCE). We will refer to y_i as "true income." The variable y_i is not observable at the level of the individual i, and therefore its national Lorenz curve $L^Y(p)$ is not directly observable. However, its aggregates $E_r(Y)$ are observable at the level of regions r. These regions may be U.S. states, French regions, Chinese provinces etc.

Let x_i be the income that individual i reports on a household survey. Unlike the case for y_i, we have unit record data on the values of x_i. In particular, for each region r, we observe not just the survey mean income $E_r(X)$ but also the Lorenz curve of the regional survey income distribution $L_r(p)$.

We assume that the individual survey income is loglinearly related to their true income, specifically

$$\ln x_i = a + b \ln y_i + \eta_i \tag{1}$$

with a and b constant across regions and η_i independent and identically distributed across individuals (and hence, regions). In principle, as long as $\ln x_i = f(\ln y_i) + \eta_i$ for some function

f, we can always use a linear form as a first-order approximation for this function, and it is highly likely that individuals with higher underlying income will report higher income in a household survey, at least on average. The core of assumption 1 lies in the homogeneity of a and b across all regions r and the homogeneity of the distribution of η_i across all regions r. We believe this is a reasonable assumption because the survey instrument used to elicit x_i is typically homogeneous across all regions r of a country. We exploit this homogeneity for identification of relevant functions of the parameters of a and b to estimate the Lorenz curve of underlying income $L^Y(p)$.

We note that Assumption 1 assumes that all discrepancies between survey income and unerlying income result from misreporting (typically, underreporting) on surveys, rather than from failing to respond to surveys altogether. A richer model allowing for both misreporting and nonresponse is presented in the Appendix, where we show that estimating the nonresponse component would critically depend on differences in inequality across regions. This model is computationally unreliable to estimate in practice. Meyer and Sullivan (2015) document that substantial underreporting, as well as nonresponse, can be found in surveys of transfer payments in the U.S., and Ravallion (2018) notes that surveys in the developing world likely experience substantial underreporting, especially of capital income, in addition to nonresponse. Outside of OECD countries, nonresponse tends to be low. Moreover, nonresponse by top income earners can be modeled within the structure of Assumption 1 as a greater extent of misreporting at the top, so while we will not be able to structurally disentangle misreporting from nonresponse, our results should not be particularly affected by our modeling choice.

We can rewrite Assumption 1 as

$$\ln y_i + \xi_i = \tilde{\alpha} + \beta \ln x_i \tag{2}$$

noting that ξ_i (a multiple of η_i) is orthogonal to $\ln y_i$ but not to $\ln x_i$.

Exponentiating and integrating both sides at the level of the region r, it now follows that

$$\begin{aligned}\ln E_r(Y) &= \alpha + \ln E_r\left(X^\beta\right) \\ &= \alpha + \beta \ln E_r(X) + \ln\left(\int_0^1 \left(L_r'(p)\right)^\beta dp\right)\end{aligned}$$

where $\alpha = \tilde{\alpha} - E(\exp \xi_i)$, a parameter that is constant across regions r. Now, the quantities $E_r(Y)$, $E_r(X)$ and $L_r(p)$ are all observable to the econometrician and correspond to the national accounts aggregate, the regional survey mean, and the regional survey Lorenz curve for region r. Assuming further than national accounts aggregates are measured with loglinear error that is not correlated with survey aggregates at the regional level, we can then estimate the equation

$$\ln E_r(Y) = \alpha + \beta \ln E_r(X) + \ln\left(\int_0^1 \left(L_r'(p)\right)^\beta dp\right) + \varepsilon_i \tag{3}$$

by nonlinear least squares to get a consistent and asymptotically normal estimates of α and β.

Define $\ln \hat{y}_i = \ln y_i + \xi_i$ and let $\hat{Y}_i = \exp(\ln \hat{y}_i)$, with $L^{\hat{Y}}(p)$ being its Lorenz curve. Appendix I presents a proof showing that since ξ_i is orthogonal to $\ln y_i$, $L^{\hat{Y}}(p) \leq L^Y(p)$ for all p. Armed with the parameter β, we can estimate $L^{\hat{Y}}(p)$ straightforwardly as

$$L^{\hat{Y}}(p, \beta) = \frac{\int_0^p \left(L'(p)\right)^\beta dp}{\int_0^1 \left(L'(p)\right)^\beta dp} \tag{4}$$

We will conservatively estimate $L^Y(p)$ with $L^{\hat{Y}}(p)$ in the rest of the paper and refer to the latter as $L^Y(p)$, abusing notation.

The parameter β – the elasticity of the survey response to underlying income – is the critical parameter determining how inequality in household surveys needs to be adjusted for survey misreporting to be made comparable with the income measure Y. It is apparent from equation (4) that a higher β implies greater inequality in the Lorenz dominance sense (if $\beta_1 \leq \beta_2$, then $L^Y(p, \beta_1) \geq L^Y(p, \beta_2)$). It is also clear that if $\beta = 1$, then $L^Y(p) = L(p)$ and the household survey Lorenz curve can be combined with national accounts data on Y without adjustment. On the other hand, if $\beta > 1$, then survey inequality understates inequality in underlying income y_i and if $\beta < 1$, then survey inequality overstates inequality in y_i.

It is important to stress that the parameter β measures the distribution of misreporting across different income levels, or *misreporting progressivity*, rather than the extent of misreporting. For a constant level of misreported income, different values of β apportion it to different parts of the income distribution. If $\beta = 1$, the apportionment is proportional to a group's income share in the household survey; however if $\beta > 1$, poorer groups are assigned shares of misreported income that are smaller than their shares of reported survey income, while richer groups are assigned shares that are larger. While the way in which misreported income is apportioned depends on the survey income distribution, an illustrative example is the USA in 2000, presented in the table below.

Table I (I)

Values of Beta and Misreporting: USA in 2000		
	(1)	(2)
	Fraction of Overall Misreported Income Attributed to Group	
Beta	Bottom 50%	Top 10%
1 (Survey)	23.3	30.2
1.1	19.1	35.7
1.2	15.1	41.4
1.3	11.3	47.3
1.4	7.7	53.3
1.5	4.5	59.5
1.6	1.4	65.8
1.7	-1.3	72.1
1.8	-3.8	78.4
1.9	-6.1	84.6
2	-8.2	90.7
WID	7.3	54.1
Our Estimate	7.9	53.1

Note: Each row of the table reports the fractions of overall misreported income attributable to the bottom 50% and the top 10%, respectively, if the U.S. household survey for 2000 in the LIS (the 2000 March CPS) is assumed to follow the process in equation (2), ignoring ξ_i with the given value of β. The rows labeled "WID" and "Our Estimate" report these fractions either directly from the WID and the LIS household survey data, or using our estimate of $\hat{\beta}_{adj}$ using national accounts data on U.S. state GDP, discussed later in this section.

For $\beta = 1$, the bottom 50%, who earn 23.3% of reported survey income, also account for 23.3% of survey misreporting. However, if β is increased, this percentage falls – for a value of β as low as 1.3, the share of misreported income accruing to the bottom 50% is more than halved (to about 11%), for $\beta = 1.6$, this share becomes a nearly negligible 1.6%, and values of $\beta = 1.7$ or greater would imply that the bottom 50%, on average, *overreports* its income on the survey, an unlikely finding to be verified empirically given the results of Meyer and Sullivan (2015). Calculations using data from the WID combined with U.S. household survey data in 2000 (the 2000 March CPS as processed by the Luxembourg Income Study) suggest that the bottom 50% accounts for only 7.3% of the aggregate misreported income, while the top 10% accounts for over 50%. Our estimates (specifically $\hat{\beta}_{adj}$ using national accounts data on U.S. state GDP, to be explained later in this section) agree with the WID for this particular survey, yielding essentially identical proportions.

Equation (4) implies that our corrected inequality estimates will be part of a parametric family of curves, each determined by making $L(p)$ more or less convex using the parameter β. It is worth asking whether this parametric family is sufficiently flexible to accommodate plausible modifications to household survey inequality estimates for all countries and years (Table I shows that it is for the US in 2000). One way of answering this question is by seeing how closely we can approximate the Lorenz curves in the World Inequality Database with this functional form by estimating

$$\beta^{WID} = \arg\min_{\beta} \sum_{p=1}^{100} \left(L^{WID}(p) - \frac{\int_0^p (L'(p))^\beta \, dp}{\int_0^1 (L'(p))^\beta \, dp} \right)^2 \tag{5}$$

where $L(p)$ is the Lorenz curve of the corresponding survey in the World Bank PIP. We obtain that for the typical survey, the best-fitting parametric curve statistically explains over 99.8% of the variation of $L^{WID}(p)$ around the line of perfect equality, and for no survey does it explain less than 97.75%. The values of β^{WID} are, on average, 1.4, with a standard deviation of 0.18, suggesting that for nearly all household surveys, the World Inequality Database (which combines survey and tax data, or imputes tax data to household surveys) suggests higher inequality than the unadjusted survey estimates.

To visualize the region-based estimation procedure, we work with a version of Equation 3 that is linearized around $\beta = 1$, which is

$$\begin{aligned} \ln E_r(Y) - \ln E_r(X) &= \alpha + \tilde{\beta} \left(\ln E_r(X) + \int_0^1 (L_r'(p)) \ln (L_r'(p)) \, dp \right) + \varepsilon_r \\ \tilde{Y}_r &= \alpha + \tilde{\beta} \tilde{X}_r + \varepsilon_r \end{aligned} \tag{6}$$

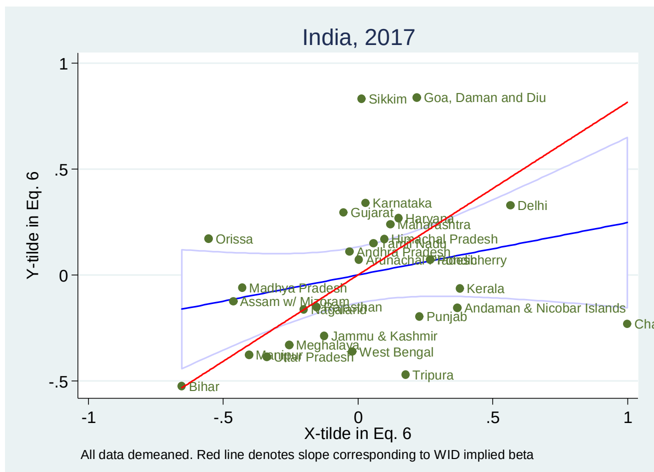
Here, $\tilde{\beta} = \beta - 1$ can be estimated by running an OLS regression of $\tilde{Y}_r = \ln E_r(Y) - \ln E_r(X)$ on $\tilde{X}_r = \ln E_r(X) + \int_0^1 (L_r'(p)) \ln (L_r'(p)) \, dp$.

For our example in Figure 1, we use regional data on 2017 Indian national accounts GDP per capita by Indian state as well as survey data from the suppressed report of the 2017 NSS (Subramnian 2019, Jha 2023[1]). The blue line shows the regression line (with slope approximately 0.24) and the light blue region presents the 95% confidence intervals. We see that the linear model in equation 6 offers a reasonable fit to the data, and that a regression line with slope 0 (implying $\beta = 1$ or no adjustment to the underlying survey data) cannot be

[1] We thank Arvind Subramanian and Somesh Jha for guiding us to the publicly available version of this report

excluded, suggesting that while most likely the NSS understated Indian inequality in 2017, we cannot reject the null hypothesis that it measured inequality accurately in that year. The red line shows the regression line that we would have expected based on the value of β^{WID} for India in 2017 obtained using equation 4, which would have implied a regression slope of about 0.81. We see that this regression line is quite different from the one we obtain using the regional data and is, in fact, rejected by the confidence intervals, though the confidence intervals are wide.

Figure 1 (1)



Note: Household survey mean and distributional regional data for India obtained from suppressed 2017 NSS report and authors' calculations. Regional GDP per capita obtained from Indian national accounts. The Y and X variables are as defined in equation (6). The blue line shows the slope $\tilde{\beta}$, where $\tilde{\beta}$ is defined in equation (6), and the faint blue lines show the 95% prediction intervals. The red line has the slope $\beta^{WID}-1$, where β^{WID} is the WID implied value of β, defined in equation (5).

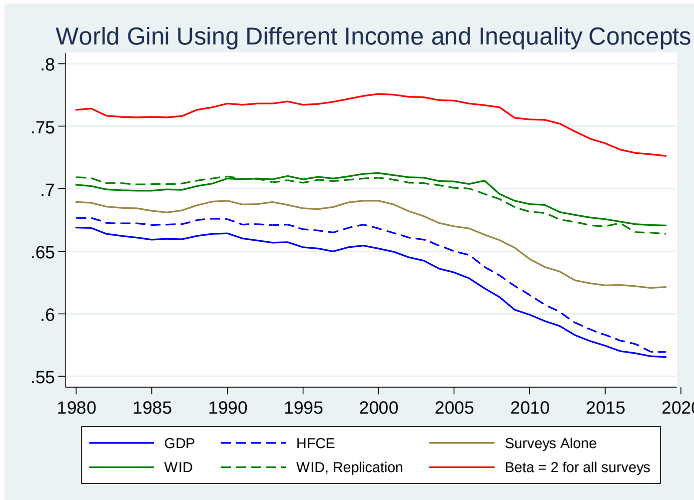
Figure 2 shows the degree to which world inequality estimates may be affected by correcting the household survey Lorenz curves through varying the parameter β. The solid blue line presents the time path of the world Gini coefficient using household surveys from World Bank PIP, rescaling their mean to equal GDP per capita at 2017 PPP from the World Development Indicators, while the solid brown line shows the world Gini coefficient using household survey data both for the mean and for the dispersion of country income distributions. We can see that the world Gini is declining regardless of the measure used, but it is lower and declining faster for the series using GDP per capita to anchor each country's income distribution. The dashed blue line presents global inequality anchoring household surveys to HFCE per capita at 2017 PPP. We see that this slightly increases global inequality, but it is considerably closer to anchoring to GDP per capita than to the household survey mean.

The solid green line presents global inequality estimated using country-level data from the World Inequality Database, which uses tax data to adjust survey inequality and centers the distribution using GDP.[2] It suggests much higher global inequality than the other three series (a Gini of 0.67 in 2019, as opposed to 0.62 for the series using PIP data alone and 0.57 for the series anchoring household surveys to GDP), and a much smaller decline in the Gini since 1980. The dashed green line presents global inequality estimated by adjusting household surveys using each survey's individual β^{WID} parameter and equation 4. We see how closely we can reproduce global inequality estimates based on the WID using the parametric families of curves defined by equation 4, implying that our methodology does not exclude the WID estimates through model choice.

Finally, the solid red line shows the upper bound of the world Gini in every year if the parameter β is allowed to vary up to a value of 2 (a value exceeded or attained by β^{WID} for only 8 of the 1934 household surveys in PIP), and if distributions are centered using GDP. It is clear that, hypothetically, inequality could be much higher than even the WID estimates suggest time paths of inequality between the solid blue line and the solid red line could be increasing or decreasing over time in a variety of ways, making estimation of the parameter β for each survey critical for understanding global inequality.

[2] This series is different from the official series for the world Gini coefficient on wid.world because 1) the official series uses national income per adult instead of GDP per capita to center each country's distribution, 2) the official series weighs each country by its adult (age 20+) population rather than by its total population, 3) we use only countries and years covered by World Bank PIP household surveys while WID uses distribution data for other countries and years, notably some income surveys from the former Soviet Union, and 4) the official series computes the integrals underlying the Gini using the trapezoid rule while we use the rectangle rule. Adopting the WID's versions of conditions 1-4 allows us to replicate their official Gini coefficient series exactly after 1990 and very closely before then. We prefer our modifications of conditions 1-3 to better compare with the household surveys, and we prefer our modification of condition 4 for transparency and ease of replication.

Figure 2 (2)



Note: Each line shows the time path of the world Gini using different approaches to compute the world distribution of income. The "GDP" series combines national accounts GDP and PIP surveys' Lorenz curves. The "HFCE" series combines national accounts HFCE and PIP surveys' Lorenz curves. The "Surveys Alone" series uses only survey data from the PIP for both the mean and the Lorenz curve of each income distribution. The "WID" series uses country-level data from the WID. The "WID, Replication" series adjusts the PIP survey data by estimated values of β^{WID} in equation (5) to approximate WID data. Finally, the "$\beta = 2$" series adjusts the PIP survey data by a value of $\beta = 2$. Procedures in Section 2.3 are used whenever necessary to interpolate and extrapolate estimates.

2.1 Accounting for Measurement Error in Household Surveys

A practical concern may be that instead of observing x_i for the universe of individuals i in the country in question, we observe it only for a random sample. In this case, instead of observing the population regional survey mean $E_r(X)$ we only observe its sample estimate $\bar{E}_r(X)$ (and similarly instead of observing $E_r\left(X^\beta\right)$ for any β, we can observe only $\bar{E}_r\left(X^\beta\right)$). With measurement error on the right hand-side, nonlinear least squares estimation of equation (3) will be biased and inconsistent. However, this concern can be resolved by using unit record data from the original survey to estimate the sampling uncertainty in moments like $\bar{E}_r\left(X^\beta\right)$ and incorporate it into our objective function.

If we had population survey moments, we could obtain β by solving the equation

$$cov\left(\ln E_r\left(Y\right)-\ln E_r\left(X^{\beta}\right),\frac{E_r\left(X^{\beta}\ln X\right)}{E_r\left(X^{\beta}\right)}\right)=0$$

Our nonlinear least squares estimate $\hat{\beta}^{NLS}$ instead satisfies

$$cov\left(\ln E_r\left(Y\right)-\ln \bar{E}_r\left(X^{\hat{\beta}_{NLS}}\right),\frac{\bar{E}_r\left(X^{\hat{\beta}_{NLS}}\ln X\right)}{\bar{E}_r\left(X^{\hat{\beta}_{NLS}}\right)}\right)=0$$

Now, it is straighforward to show that

$$\begin{aligned} & cov\left(\ln E_r\left(Y\right)-\ln E_r\left(X^{\beta}\right),\frac{E_r\left(X^{\beta}\ln X\right)}{E_r\left(X^{\beta}\right)}\right) \\ = \; & cov\left(\ln E_r\left(Y\right)-\ln \bar{E}_r\left(X^{\beta}\right),\frac{\bar{E}_r\left(X^{\beta}\ln X\right)}{\bar{E}_r\left(X^{\beta}\right)}\right) \\ & +cov\left(\left(\ln \bar{E}_r\left(X^{\beta}\right)-\ln E_r\left(X^{\beta}\right)\right),\frac{\bar{E}_r\left(X^{\beta}\ln X\right)}{\bar{E}_r\left(X^{\beta}\right)}-\frac{E_r\left(X^{\beta}\ln X\right)}{E_r\left(X^{\beta}\right)}\right) \end{aligned} \tag{6}$$

The first term on the right hand-side of (6) is the first-order condition satisfied by the nonlinear least squares estimate and is estimable for any value of β. The second term is the covariance of the differences of two functions of sample means and their population analoges, and thus should be obtainable from data on the distribution of X and the sampling procedure via the Central Limit Theorem and the delta method. Specifically, the second term can be obtained as

$$\begin{aligned} S\left(\beta\right) \quad : & = cov\left(\left(\ln \bar{E}_r\left(X^{\beta}\right)-\ln E_r\left(X^{\beta}\right)\right),\frac{\bar{E}_r\left(X^{\beta}\ln X\right)}{\bar{E}_r\left(X^{\beta}\right)}-\frac{E_r\left(X^{\beta}\ln X\right)}{E_r\left(X^{\beta}\right)}\right) \\ & \rightarrow {}^{p}V\left(\beta\right)\cdot Avar\begin{pmatrix}\bar{E}_r\left(X^{\beta}\right)\\ \bar{E}_r\left(X^{\beta}\ln X\right)\end{pmatrix}\cdot V\left(\beta\right)' \end{aligned}$$

where

$$V\left(\beta\right)=\frac{1}{\bar{E}_r\left(X^{\beta}\right)^2}\begin{pmatrix}\bar{E}_r\left(X^{\beta}\right) & 0\\ -\bar{E}_r\left(X^{\beta}\ln X\right) & \bar{E}_r\left(X^{\beta}\right)\end{pmatrix}$$

We define β^{adj}, the adjustment value of β, to be the solution of

$$cov\left(\ln E_r\left(Y\right) - \ln \bar{E}_r\left(X^{\beta}\right), \frac{\bar{E}_r\left(X^{\beta}\ln X\right)}{\bar{E}_r\left(X^{\beta}\right)}\right) + S\left(\beta\right) = 0 \tag{8}$$

In general, $S(\beta)$ is positive for $\beta = \hat{\beta}_{NLS}$, and as the first term is decreasing in β, adjusting the first-order condition by the second term implies that $\hat{\beta}_{adj} \geq \hat{\beta}_{NLS}$. We report results with β^{adj} as our baseline results; the results using nonlinear least squares estimation without adjustment are qualitatively similar, differing by a level shift but having essentially identical trends in inequality, poverty and welfare over time.

2.2 Assumptions for Countries with Missing Data

While we have regional GDP data for the largest countries, and for most countries in Latin America and the OECD, there are still many countries, particularly in Asia and Africa, without regional data. We also do not have regional data (or survey data) for every country in every year, even in the better-covered regions. Data on regional consumption data is even more sparse. We therefore use the following rules to extend our estimates to countries with missing data:

1. For countries and years with regional GDP data and unit record survey data to perform the adjustment (the countries with LIS data) we compute $\hat{\beta}^{adj}$ using equation (8).

2. For countries and years with regional GDP data but without unit record survey data to perform the adjustment (all countries with survey regional data from SEDLAC and national statistical offices) we compute $\hat{\beta}^{adj}$ as the predicted values of the regression of $\hat{\beta}^{adj}$ on $\hat{\beta}_{NLS}$ for developing countries with both LIS and non-LIS data (these are Brazil, Mexico, Colombia, Chile, Panama, Guatemala, Paraguay, China and India).

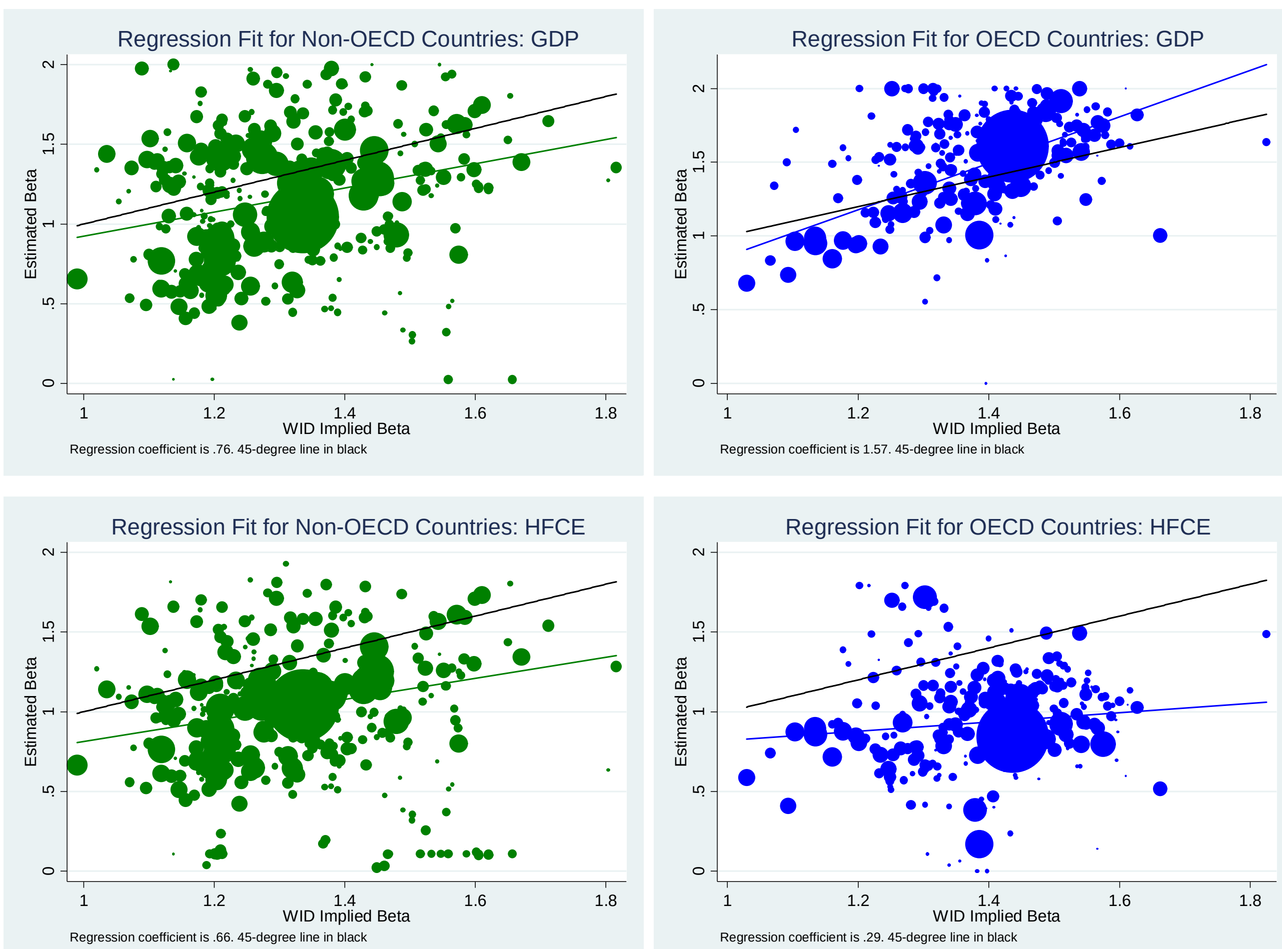
3. For countries with regional GDP data for some years but not others, we linearly interpolate the adjusted Lorenz curves for the missing years (as the convex combination of two Lorenz curves is a Lorenz curve), and use the earliest and latest Lorenz curve for years outside of the time span with regional data.

4. We use regional data on disposable income (OECD countries), primary income (Russia, Mexico, other non-OECD countries), HFCE (China) and earnings (Brazil) as proxies for regional HFCE data. For countries with regional consumption data available for fewer years than regional GDP data, we estimate β by linearly interpolating and horizontally extrapolating the series $\beta^{GDP} - \beta^{HFCE}$ across all years with regional GDP

data, and adding back β^{GDP} for years with available regional GDP data but missing regional HFCE data. We then follow step 3 to interpolate Lorenz curves for years with missing survey data.

5. For countries with missing regional data altogether, we compute β by regressing available values of β on corresponding values of β^{WID}, weighting by the standard error of β. We run two separate regressions: one for OECD countries and the other for non-OECD countries. (Using the survey Gini coefficient instead of β^{WID} generates very similar results, as β tends to be negatively correlated with the survey Gini coefficient).

Figure 3 (3)



Note: Each panel of this figure represents a regression of an estimated value of β from a household survey and either regional accounts data on GDP or on HFCE for a given survey on the left hand-side, and an estimated value of β^{WID} from equation (5) for that same survey on the right hand-side. Each observation represents a single survey in a given country and year, and has weight equal to the reciprocal of the variance estimate of β, which is represented by the size of its dot. Lines represent the lines of best fit and the 45-degree line.

Figure 3 plots the relationships between values of β estimated based on regional variation and values of β that rationalize inequality estimates in the WID. Each point represents a survey and is weighted by the variance of the region-based estimate of β. We see that the regional-based estimates of β are generally lower than the WID-based estimates, although for developing countries they tend to rise at the same rate. For OECD countries, regional-based estimates of β for GDP rise much more than one-for-one with the WID-based estimates, while regional-based estimates of β for HFCE (in this context, generally proxied by disposable income, which likely is more unequally distributed than HFCE) rise much less. This is likely because the regional-based estimates of β for GDP capture inequalities in production stemming from inequalities in public investment and government consumption, which is often disproportionately assigned to the capital region in the national accounts, with there being a much smaller degree of asymmetry in disposable income between the capital region and other regions. As we use regional data on GDP when regional data on HFCE or its proxies is not available, it is likely that we are masking similar though less pronounced trends for non-OECD countries, most of which have no regional HFCE data. It is clear from figure 3 that the subnational region-based inequality estimates will tend to be lower than those in the WID for many developing countries, though unequal GDP distribution in the OECD will provide a countervailing trend when global inequality is considered.

2.3 Estimating Adjusted Income

Once we obtain estimates of adjusted Lorenz curves ($L_Y(p)$), it is straightforward to estimate Y_p, true income for each percentile, as

$$Y_p = E(Y) L'_Y(p)$$

For some values of β, the estimate of Y_p may be smaller than the estimate of X_p, survey income at each percentile. This is inconsistent with the general finding that respondents do not overreport on surveys. We seek to adjust our estimates of Y_p by constructing a regularized series $\hat{Y}_p$ so that the adjusted series never falls below X_p, so that $\hat{Y}_p$ is monotonic increasing in p, and so that $E\left(\hat{Y}\right) = E(Y)$. We construct this series as follows:

1. If $Y_p < X_p$, then $\hat{Y}_p = X_p$

2. Let S be the set of p such that $Y_p < X_p$. Let $L = \sum_{p \in S} (X_p - Y_p)$ and $G = \sum_{p \notin S} (Y_p - X_p)$. Generally, $G > L$ (as $E(Y) > E(X)$). Let $\alpha = L/G \in [0, 1)$.

3. Then, if $Y_p > X_p$, we define $\hat{Y}_p = (1-\alpha) Y_p + \alpha X_p$. Note that for $p \in S^c$, $Y_p > \hat{Y}_p > X_p$, and $\hat{Y}_p$ is monotonic increasing in p. Also note that if S is the null set, then $L = 0 = \alpha$.

4. Then, $E\left(\hat{Y}\right) = E(Y) + L - \alpha G = E(Y)$ and $\hat{Y}_p \geq X_p$ for all p.

In practice, poverty and inequality estimates using Y_p and $\hat{Y}_p$ are very similar.

2.4 Data

To estimate β for each country and year, we need to have national accounts data on regional GDP and / or disposable income, as well as household survey data on reported income and its distribution by region. We get this information for the countries below from the following sources.

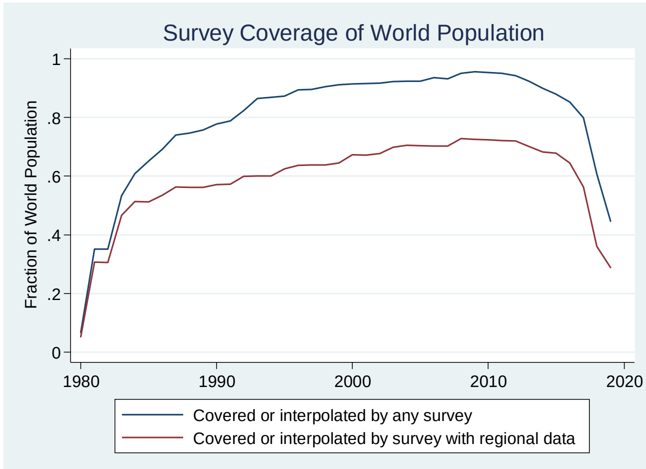
For the US, Canada, most European countries (including Eastern European countries) and Russia, we used the OECD Stats database and Gennaioli et al. (2014) for regional GDP and disposable income per capita, and the Luxembourg Income Study for regional household survey data (Luxembourg Income Study 2023). We use the household per capita income concept.

For Latin American countries (all of the Americas except for Canada and the US) we used the OECD Stats database in conjunction with Gennaioli et al. (2014) for regional GDP and (the occasional) disposable income per capita, and the SEDLAC database for regional household survey data.

For China, India, Indonesia, Pakistan, Bangladesh, Nigeria and the DRC we obtained regional accounts GDP data (and in the case of China, regional accounts HFCE data) as well as regional household survey data directly from these countries' statistical yearbooks.

Overall, countries with any regional data account for 70-80% of the world population depending on the year, while countries with available surveys account for over 90% of the world population across our time period. Considering the more demanding criterion of the fraction of the world population in each year with either a survey conducted in that year or surveys conducted in prior and subsequent years, this fraction is over 80% from the early 1990s to the mid-2010s. If we restrict to the fraction of the world population for which the surveys in question have regional data that can be matched to national accounts data, this fraction is above 60% from the mid-1990s to the mid-2010s and reaches a maximum of 72.5% in 2009. Figure 4 below shows how both these proportions vary by year.

Figure 4 (4)



3 Results

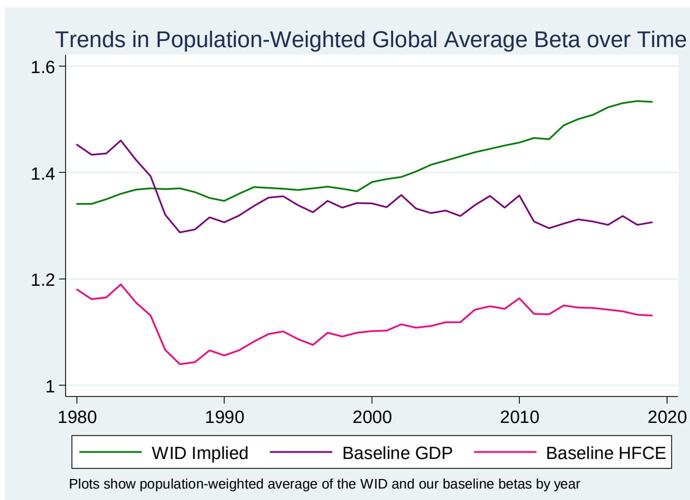
3.1 Basic Facts about Misreporting Corrections

We first present evidence on how progressive misreporting is in the world overall. Figure 5 presents time series of the population-weighted average value of β for the world as a whole for our baseline datasets of β using either GDP or HFCE as the income measure in the national accounts, as well as for the implied β based on the data in the WID. We see that our corrections for misreporting deliver alternative adjustments to income distributions than does the WID. On average, the β that rationalizes survey Lorenz curve estimates in the PIP with those in the WID is approximately flat through the 1980s and the 1990s at 1.35 (implying that for a distribution like that of the U.S., the bottom 50% underreports only half as much as under proportional misreporting). After 2000, the average WID-implied value of β rises sharply over time, exceeding 1.5 by the late 2010s, and implying that household surveys' underreporting becomes more and more tilted towards the rich over the course of the 2000s. In contrast, our estimates of β based on the regional GDP distribution start relatively high (at 1.45) and oscillate through the 1980s and early 1990s until stabilizing slightly below the level of the WID-implied average value of β by the mid-1990s.[3] They remain at this level through the late 2010s, possibly declining a little. Therefore, for the 1990s and early 2000s, our estimates agree with the WID on the overall progressivity of underreporting

[3]The high values of our estimates of β in the 1980s are driven by estimates from the Chinese household surveys, and if these are excluded, average values of our estimates of β are roughly constant over time.

in the household surveys, but subsequently they suggest that underreporting progressivity remained flat rather than experiencing a dramatic rise. It is notable that the WID requires a dramatic change in underreporting behavior around 2000 to explain its findings, while our findings are consistent with merely the continuation of the preexisting pattern of how the rich underreported relative to the poor. Our estimates of β based on the regional distribution of HFCE (or disposable income) suggest lower underreporting progressivity overall (less than 1.2, or the bottom 50% underreporting two-thirds as much as they would under proportional underreporting), with the trend of underreporting progressivity rising slightly in the 1990s and 2000s and flattening out in the 2010s. They also do not suggest the stark upward trend in underreporting progressivity that is consistent with the estimates in the WID. The stabilization of underreporting progressivity in our estimates suggests that the dynamics of within-country inequality should be better approximated by those in the household surveys than in the WID.

Figure 5 (5)

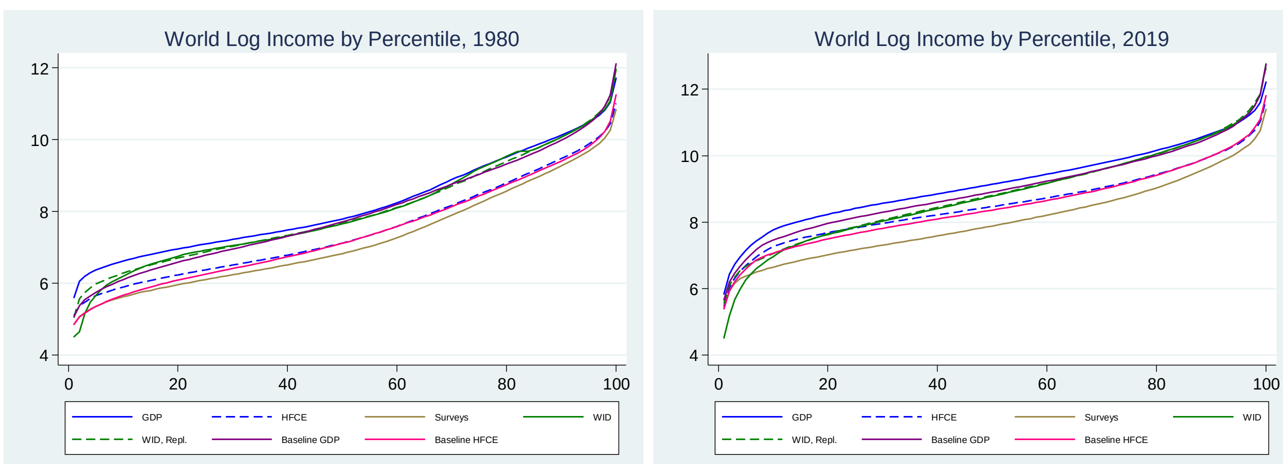


Note: Each line represents the population-weighted average of β for the given series. For the WID implied series, this is the weighted average of β^{WID} from equation (5). For the Baseline GDP series, this is the weighted average of $\hat{\beta}_{adj}$ using national accounts GDP (see equation 8). For the Baseline HFCE series, this is the weighted average of $\hat{\beta}_{adj}$ using national accounts HFCE or disposable income.

We can immediately use our estimates of β to assess the distribution of underreporting across percentiles of the world distribution of income. Figure 6 presents the value of log income according to several ways of computing the world distribution of income in 1980 (left panel) and 2019 (right panel). In addition to presenting estimates for the series in

Figure 2, we also offer estimates of log income by percentile when re-centering surveys to GDP and adjusting distributions using values of β estimated from data on regional GDP (purple), and when re-centering surveys to HFCE and adjusting distributions using values of β estimated from data on regional disposable income or HFCE (pink). We see that in 1980, our baseline GDP approach broadly coincides with the WID estimates, and in particular suggests lower incomes for the bottom 50% than if we were to re-center the unadjusted surveys to GDP. On the other hand, our baseline HFCE estimates broadly coincide with the estimates obtained from using surveys alone for the bottom 20% of the income distribution, and broadly coincide with the estimates obtained by rescaling surveys to match HFCE for the top 60% of the income distribution. In particular, correcting for underreporting revises our estimates of the incomes of the poor in 1970 downwards relative to assuming proportional underreporting. Relative to the estimates obtained using surveys alone, we see considerable underreporting in all the series representing pre-tax income or GDP (suggesting that one source of systematic underreporting is the difference in income concept) and very little underreporting by the poor accompanied by moderate underreporting above the 20th percentile in our HFCE-based series.

Figure 6 (6)



Note: Each line shows the world Lorenz Curve in either 1980 or 2019 using different approaches to compute the world distribution of income. The "GDP" series combines national accounts GDP and PIP surveys' Lorenz curves. The "HFCE" series combines national accounts HFCE and PIP surveys' Lorenz curves. The "Surveys Alone" series uses only survey data from the PIP for both the mean and the Lorenz curve of each income distribution. The "WID" series uses country-level data from the WID. The "WID, Replication" series adjusts the PIP survey data by estimated values of β^{WID} in equation (5) to approximate WID data.

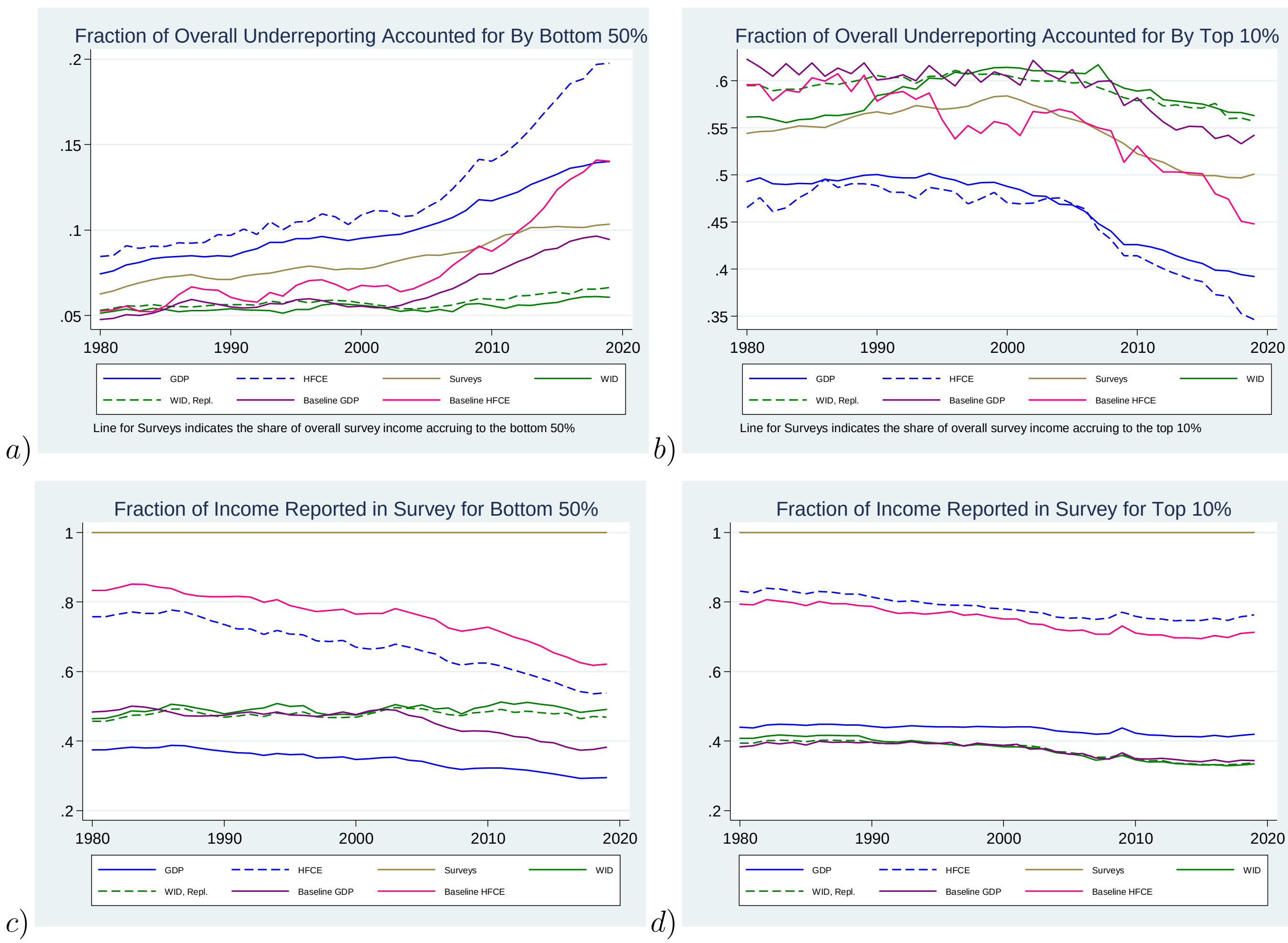
The "Baseline GDP" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts GDP. The "Baseline HFCE" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts HFCE or disposable income.

By 2019, however, the picture changes. Now, our estimates of income at each percentile for the baseline GDP series are well above the corresponding estimates in the WID for the poor up to the median of the world distribution of income, while remaining close to them for higher incomes. In turn, our estimates of income for the baseline HFCE series are above those obtained using surveys alone, even for the bottom 20%. Underreporting appears to be prevalent in the surveys relative to both the baseline GDP and the baseline HFCE series.

To see that underreporting by the poor is both present, increasing, and significant, and to compare and contrast our findings on underreporting over time, Figure 7 presents estimates of two underreporting-related measures for the bottom 50% of the world income distribution as well as for the top 10%, for various approaches of estimating the world distribution of income. Panel a) shows the fraction of overall underreporting accounted for by the bottom 50% over time. Note that if underreporting were proportional to survey income with the same proportionality constant for all countries, this would just be the income share of the bottom 50% in the surveys, which is what is presented by the brown line in this chart. We see that rescaling the surveys to match GDP or HFCE suggests that the poor account for a greater share of underreporting than they do of income; this is a consequence of poorer countries' household surveys capturing a lower proportion of their GDP or HFCE than richer countries'. Thus, underreporting by the poor is an increasingly significant part of the overall distribution of income. On the other hand, estimates from the WID (using tax data and imputation for much of the developing world) suggest that the bottom 50% account for a smaller share of world underreporting than they do of survey income, with this share roughly stable at 5% of overall underreporting and rising slightly only in the 2010s. Our results using both rescaling and adjustment of the underlying distributions paint an intermediate picture. Until the early 2000s, both our baseline GDP and baseline HFCE series indicate that underreporting by the bottom 50% is around 5% of overall misreporting (though a respectable fraction of their survey income share) and flat over time. However, starting in the early 2000s, misreporting of the bottom 50% accounts for a larger and larger fraction of overall misreporting for both our baseline GDP and HFCE series, exceeding the survey income share of the bottom 50% by 2019 for the baseline HFCE series and closely approaching it for the baseline GDP series. By 2019, over 10% of aggregate underreporting is accounted for by the bottom 50%. No such rise is observed either in the WID data or in our replication of it through constructing values of β that approximate the difference between the unadjusted survey data and the WID.

Panel b) of Figure 7 shows the fraction of overall underreporting accounted for by the top 10%. It offers much the same picture as panel a), with this fraction falling over time in line with the falling survey income share of the top 10%, but with the decline more pronounced in the baseline GDP and especially baseline HFCE series than in the WID data. We conclude that over time, the underreporting of the poor matters increasingly more relative to the underreporting of the rich. (We note that replicating the WID data by suitable choices of β matches the fraction of overall underreporting accounted for by the top 10% relatively poorly before the mid-1990s but matches it very well subsequently).

Figure 7 (7)



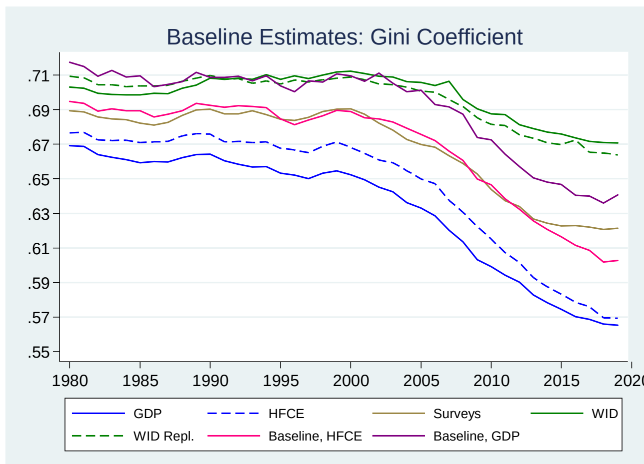
Note: The measures in panels a) and b) are the ratios of underreported income of the bottom 50% and top 10% to aggregate underreported income in each year, and the measures in panels c) and d) are the ratios of reported to true income of the bottom 50% and 10%. The "GDP" series combines national accounts GDP and PIP surveys' Lorenz curves. The "HFCE" series combines national accounts HFCE and PIP surveys' Lorenz curves. The "Surveys Alone" series uses only survey data from the PIP for both the mean and the Lorenz curve of each income distribution. The "WID" series uses country-level data from the WID. The "WID, Replication" series adjusts the PIP survey data by estimated values of β^{WID} in equation (5) to approximate WID data. The "Baseline GDP" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts GDP. The "Baseline HFCE" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts HFCE or disposable income.

Panels c) and d) of Figure 7 show the fraction of income captured by surveys for the bottom 50% and the top 10%. For survey income, this number is 1 by definition. The WID data (closely approximated by appropriate choices of β in our approach) implies that survey income captures roughly 40-50% of underlying income, the fraction rising slightly for the bottom 50% and falling somewhat for the top 10%, but largely remaining stable over time. Our baseline GDP series delivers similar results to the WID for underreporting by the top 10%, but an important difference for the bottom 50%, who, in our estimates, noticeably report a lower and lower share of their GDP in the household surveys starting in the early 2000s, diverging from the level of underreporting in the WID. In contrast, the measures that center surveys to HFCE (including the baseline HFCE series) suggest that about 80% of consumption was captured as survey income around 1980, with the share declining somewhat for the top 10% and more rapidly for the bottom 50% thereafter. Our results imply that as global growth increasingly reached the poor starting in the early 2000s, the poor began gradually reporting less and less of their income or consumption in the household surveys, while this effect was less pronounced for the rich.

3.2 World Poverty, Inequality and Welfare

Figure 8 presents estimates of the time series of the global Gini coefficient. It is identical to figure 2, except the estimates assuming $\beta = 2$ (the red series) are removed, and estimates in which Gini inequality in GDP and HFCE are calculated using regionally estimated betas (in purple and pink, respectively) are included. We see that all Gini inequality series, including the one constructed on the basis of the WID, are falling over time, with a particularly pronounced decline after the mid-2000s. Gini inequality for GDP and HFCE is lower and is falling faster than Gini inequality based on the WID, but is higher and falling slower than Gini inequality based on using national accounts aggregates to measure the mean of national income distributions and using household surveys to measure their dispersion. Gini inequality in GDP is close to Gini inequality in the WID until the mid-2000s and then begins falling more rapidly, consistent with our findings for China and India. Gini inequality in HFCE is close to global Gini inequality in the household surveys, although this is almost certainly an upper bound on actual Gini inequality in HFCE. We can roughly say that the increase in consumption inequality that we observe after adjusting household surveys for misreporting is equal to or less than the extent to which across-country inequality is greater in household survey income than it is in national accounts consumption.

Figure 8 (8)



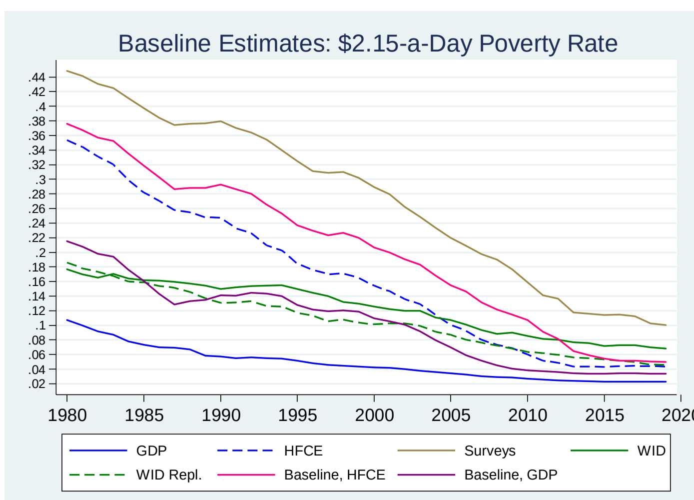
Note: The "GDP" series combines national accounts GDP and PIP surveys' Lorenz curves. The "HFCE" series combines national accounts HFCE and PIP surveys' Lorenz curves. The "Surveys" series uses only survey data from the PIP for both the mean and the Lorenz curve of each income distribution. The "WID" series uses country-level data from the WID. The "WID, Replication" series adjusts the PIP survey data by estimated values of β^{WID} in equation (5) to approximate WID data. The "Baseline GDP" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts GDP. The "Baseline HFCE" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts HFCE or disposable income.

Figure 9 presents estimates of global poverty rates at $2.15 dollars a day in 2017 PPP. This is the current equivalent of the old "$1-a-day poverty line", and is calculated as the median of the poverty lines of low-income countries (World Bank, 2023). The brown line, which presents estimates of the poverty rate using surveys alone, is very close to the official poverty rates reported by the World Bank PIP, with a slight difference during 2015-2019 owing to our use of the suppressed 2017 Indian NSS survey whereas World Bank PIP uses an alternative set of household surveys. Using the 2017 Indian NSS survey, global poverty using surveys alone falls from over 44% of the world population in 1980 to a little over 10%

of the world population in 2019 (compared to 8.5% in the official World Bank data).

What happens to this poverty estimate when we first, replace household survey income with national accounts consumption, alleviating the underestimation of income and consumption in the household surveys, and second, inflate household survey inequality measures to account for the effect of survey misreporting on income? Both of these changes are very conservative from the point of view of poverty measurement. While national accounts consumption is likely closer in concept to disposable income than is GDP, consumption still tends to be lower than disposable income, biasing our estimates of poverty upward. Moreover, our adjustment of household survey inequality is likely excessive if the target is the distribution of consumption or even of disposable income, as for many countries, especially in the developing world, we are essentially using the distribution of GDP, which is likely to be much more unequal (and similar in concept to pretax income). However, when we make both adjustments, we see that the resulting poverty estimate (in pink) suggests that poverty is lower than measured using household surveys alone. It falls from about 38% in 1980 (and 85% of the household survey estimate) to 5% in 2019. According to our results, global one-dollar-a-day consumption poverty is about half as high as we would think using surveys alone. We can quantify the impact of the survey adjustment: using national accounts consumption to center each survey income distribution but otherwise leaving these distributions unchanged would yield a poverty decline from 36% in 1980 to 4.3% in 2019 (dashed blue line).

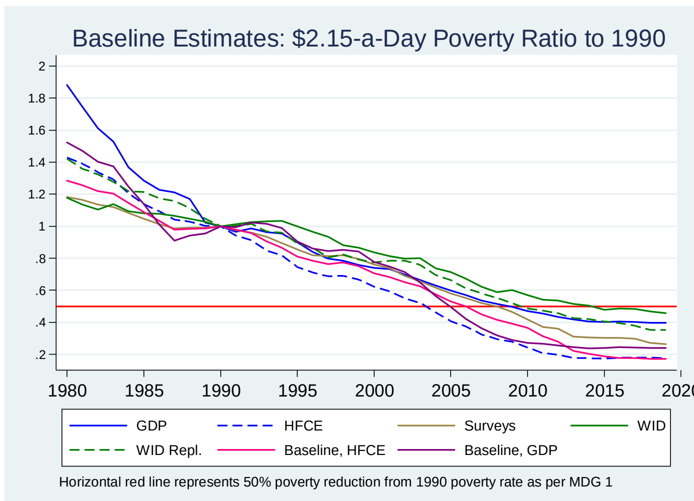
Figure 9 (9)



Note: See note to Figure (8).

If we use GDP per capita rather than HFCE per capita to center income distributions, \$2.15-a-day poverty becomes rare even in the 1980s and 1990s. Depending on whether we measure poverty using WID data (we note there are no official poverty measurements made by the WID), using GDP per capita and adjusted inequality estimates using regional data, or using GDP per capita and unadjusted inequality estimates, we obtain poverty rates ranging from 10% to 22% in 1980 and poverty rates ranging from 6.8% to 2.2% in 2019.

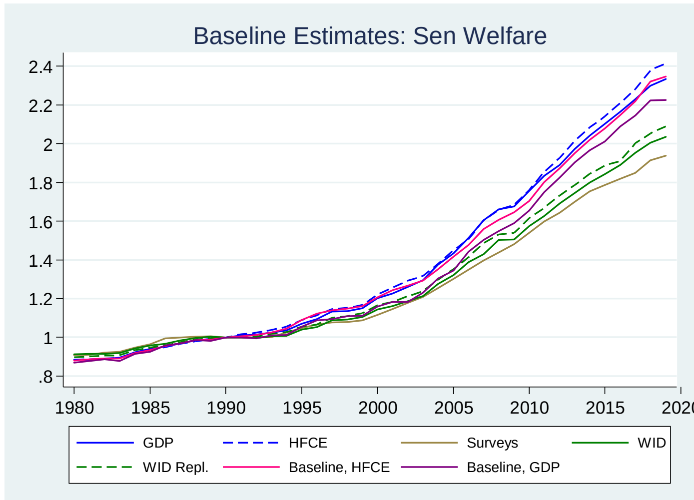
Figure 10 (10)



Note: See note to Figure (8).

To better see how our approach compares to the PIP in terms of the rate of global extreme poverty reduction, we present ratios of world \$2.15-a-day poverty to its level in 1990 in Figure 10. All estimates are consistent with sharply declining poverty over time and a halving of poverty between 1990 and 2015 as per the Millennium Development Goal. Estimates of the rate of poverty reduction using our HFCE per capita estimates (the pink line) show a more rapid rate of poverty reduction than do the estimates using surveys alone relative to 1990. Using GDP per capita to center each country distribution and either not adjusting the inequality measures or using the WID inequality estimates generates slower rates of poverty reduction, with global poverty continuing to remain at about 35-45% of its 1990 in 2019. However, using GDP per capita and adjusted inequality measures (the purple line) delivers a much more rapid rate of poverty reduction, with global poverty in 2019 being only 24% of the rate prevailing in 1990. Whether one wishes to use GDP per capita or HFCE per capita as the relevant income concept for poverty measurement, our results suggest that we should not revise upward global poverty estimates and their persistence, and likely we should revise them downward.

Figure 11 (11)



Note: See note to Figure (8).

It is useful to aggregate income and its distribution into a summary measure of welfare and consider its evolution over time. One such proposed measure is the Sen welfare index (Sen 1976), which is equal to the mean of the distribution multiplied by one minus the Gini coefficient, and thus increases with income and and falls with inequality. Figure 11 presents the evolution of the Sen index, normalized to 1 in 1990, for the different ways of measuring inequality we consider. All methods suggest that the Sen index has grown strikingly, increasing by at least 90% between 1990 and 2019. Using surveys alone generates the smallest increase, 94%, to a large extent because household survey income grows slower than either national accounts GDP or national accounts consumption. Estimates from the WID suggest a slightly higher welfare increase of 103%, suggesting that centering national income distributions around GDP per capita rather than household survey means has a larger positive impact on welfare than using the more unequal distributions in WID has a negative impact on welfare. Our replication of the WID using Lorenz curves in the family defined by equation 4 suggests a welfare increase of 109%. However, our results using adjusted inequality estimates suggest yet higher welfare growth: 123% for our estimate using GDP per capita and 135% for our estimate using HFCE per capita. It is notable that our approach for adjusting HFCE per capita, which is quite conservative as it frequently employs data on regional GDP, instead of regional disposable income or consumption, still finds that global welfare rose by more than 40 percent of its 1990 level than we would find if we used survey data alone. Using national accounts aggregates without adjusting inequality would result in welfare growth of 133% for GDP per capita and 141% for HFCE per capita. Thus, when we explicitly adjust survey inequality for misreporting, we still recover higher welfare growth estimates than either the WID or the World Bank. However, the survey adjustment is important to perform, as it decreases welfare growth by about 6-10 percent of the 1990 welfare level relative to using unadjusted surveys and national accounts data to center each country's distribution.

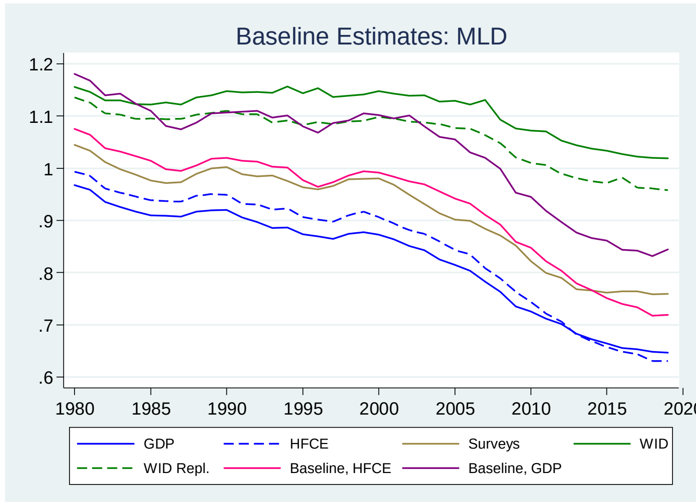
3.3 Decline in Within-Country Inequality

While it is well-known that global inequality has been declining overall, there are controversies over how much of this decline is the result of declines in inequality between countries as opposed to declines in inequality between countries. To explore this question, we shift from measuring inequality with the Gini coefficient and instead use the mean logarithmic deviation (MLD, alternatively known as the GE(0) index). The MLD has the unique property that it can be decomposed as the sum of the MLDs of individual countries' average incomes (the countries weighted by population) and of the population-weighted sum of individual countries' MLDs (Shorrocks 1984)[4]. Figure 12 reproduces Figure 8 using the MLD as the measure of inequality in place of the Gini and the trends are very similar. We should note that while our replication of the global MLD implied by the WID's country-level data is close to the original and follows the same trends it is somewhat worse than our replication of the corresponding global Gini coefficient.

[4]Specifically, it is the only inequality index satisfying

$$I_{total} = I_{between} + \sum_{c} \frac{pop_c}{pop_{world}} I_c$$

Figure 12 (12)

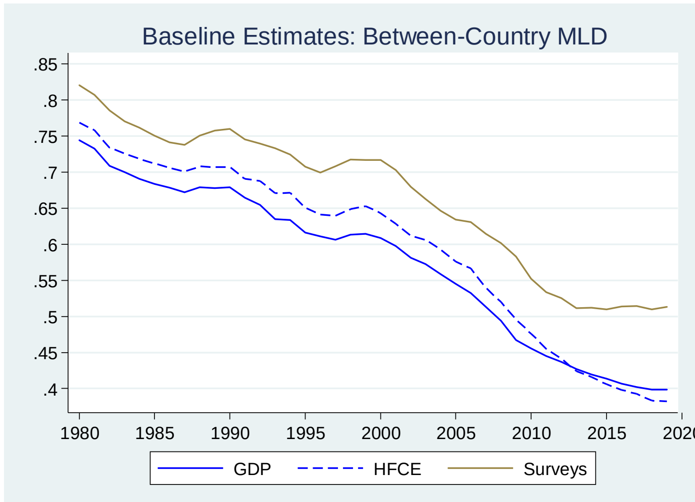


Note: See note to Figure (8).

It is well-accepted that declines in across-country inequality – e.g. China and India increasing their income relative to the U.S. and Europe – have been critical to the overall inequality decline. Indeed, as Figure 13 shows, across-country MLDs have declined by about 0.35 for national accounts GDP per capita and HFCE per capita (from a base of about 0.75) and by a slightly smaller number from a higher level for household survey income. These declines are similar in magnitude to the declines in the overall MLD recorded in Figure 12 – and considerably exceed the decline in the MLD for the WID (or its replication), which is about 0.15 (or 0.18 for the replication). According to the WID, within-country inequality must have risen by about 0.2 – making the rise in within-country inequality commensurate in importance with the fall in across-country inequality in telling the story of inequality over the last forty years. However, as we have seen in Section 2 and in the first part of this section, our region-based estimates of inequality deliver smaller and slower-growing (or declining) inequality estimates for some of the largest countries in the world – India and

China – than does the WID.

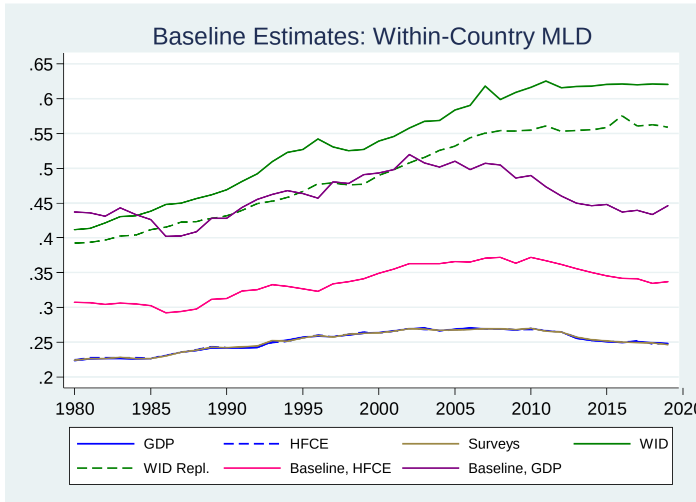
Figure 13 (13)



Note: See note to Figure (8).

Indeed, turning to within-country inequality in Figure 14 we see that our region-based inequality adjustment provides an alternative account of inequality to both the WID and the household surveys. According to the household surveys, within-country inequality has always been low, never exceeding an MLD of 0.27 since 1980, which is less than the decline in the across-country MLD over this time period. (Within-country MLDs are the same for all series that use unadjusted household surveys to measure inequality, whatever adjustment is made to the survey mean, so the GDP and HFCE within-country series are essentially identical to the within-country series if household surveys alone are used). Within-country inequality in the surveys flatlines in the 2000s and declines by about 0.02 during the 2010s. According to the WID, on the other hand, within-country inequality has been very high, starting out somewhat above 0.4 in 1980 and skyrocketing to over 0.6 by the early 2010s, thus delivering a 0.2 increase to overall MLD inequality and offsetting much of the decrease in across-country inequality (our replication finds a slightly smaller increase in within-country MLD inequality but the trends are very similar).

Figure 14 (14)



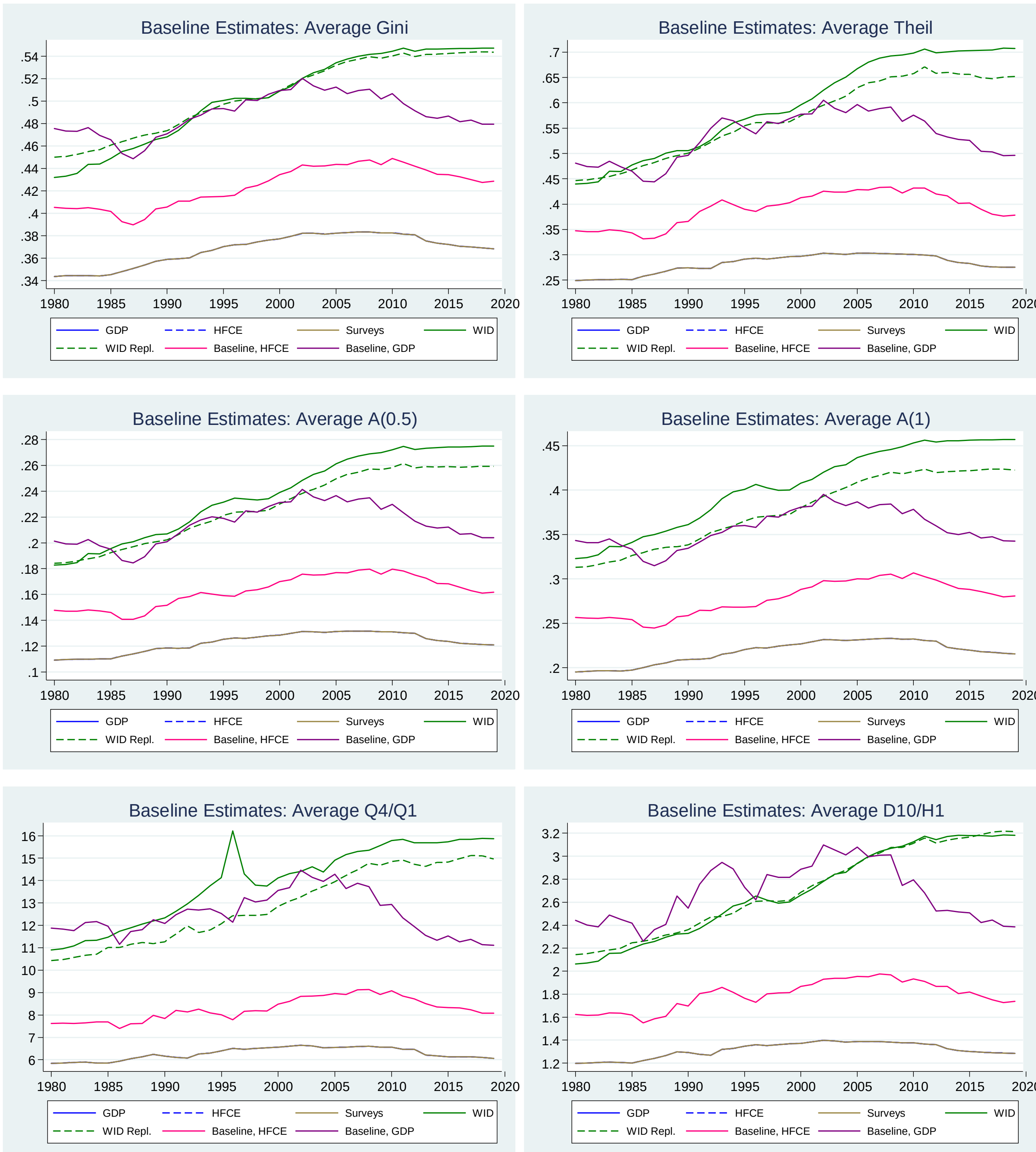
Note: See note to Figure (8).

The picture obtained by using our region-based inequality adjustment are different. Looking at within-country inequality in GDP (purple line) we see that it starts out as high as WID inequality in the 1980s, and grows together with WID inequality until the early 2000s to a level of about 0.5. However, while WID within-country inequality keeps growing through the 2000s and flatlines in the early 2010s, within-country inequality in GDP flatlines in the 2000s and declines outright by nearly 0.07 in the 2010s, back to its level in the early 1990s. Therefore, within-country inequality in GDP replicates the initial magnitude of within-country inequality in the WID (and comfortably exceeds within-country inequality in the household surveys) but within-country inequality in GDP does not rise for as long as does within-country inequality in the WID, and actually declines in the 2010s. The path of within-country inequality in HFCE (pink line) is qualitatively similar to that of within-country inequality in GDP, but at a considerably lower level and with shallower fluctuations, reaching a maximum of 0.37 in 2008 and declining to 0.34 by 2019, still somewhat higher than its level of 0.31 in 2019. Within-country inequality in HFCE is still considerably higher than within-country inequality in the household surveys (to which HFCE may be the closest income concept in the national accounts), but its trajectory over time is more similar to the surveys than to the WID.

Our region-based inequality estimates agree with the WID that within-country inequality is much higher than in the unadjusted household surveys, which is also reflected in the overall inequality estimates. However, they diverge in trend from the WID around the early 2000s, becoming flat while within-country inequality in the WID is still rising, and declining when within-country inequality in the WID becomes flat, much like the household surveys do. We note that all approaches, including the household surveys and the WID as well as our estimates, find that within-country inequality decelerates in the 2010s (and even the estimates using country-level data from the WID show that within-country inequality is down from its all-time peak in the early 2010s), so there is agreement that the growing trend in within-country inequality over the previous 30 years has stopped. However, there still is disagreement between the WID and the unadjusted household surveys over whether within-country inequality has declined or merely flatlined. Our methodology, which nests both the WID and the surveys in a common family of income distributions and uses external evidence to identify the right adjustment parameter, suggests resolving this debate in favor of the trends uncovered in the surveys.

It is worth noting that there are other ways of thinking about within-country inequality. While different inequality indexes can be differently decomposed into components that do and do not depend on within-country inequality, we focus on population-weighted average within-country inequality indexes. These indexes have the intuitive interpretation that if within-country inequality is relevant for welfare, it must generate some negative externality that affects only the residents of the country in question and not the rest of the world. This may be a political externality, like decreased political competition, or an economic externality, such as experiencing positive utility from not consuming less than one's compatriots. Now, if welfare is aggregative, then it should be the population-weighted average – rather than, say, the income-weighted average – of country-level measures. Therefore, population-weighted within-country inequality indexes should be welfare-relevant, even if they are not parts of exact decompositions of global inequality indexes into within and between. Figure 15 presents time paths of population-weighted within-country inequality measured with the Gini, the Theil index, the Atkinson indexes with inequality aversion equal to 0.5 and 1, and the ratio of the income share of the top 10% to the bottom 50%. We see that in all these cases, the time paths using unadjusted surveys, the WID and our within-country inequality measures for GDP and HFCE are similar to Figure 14.

Figure 15 (15)



Note: See note to Figure (8).

3.3.1 Inequality Declines for Several Large Countries

We conclude this section by asking how we measure inequality to have evolved in the world's largest countries and how they contribute to our finding that global within-country inequality has fallen. Table II presents our estimates of the change in the top 10% shares of the distributions of GDP and HFCE for the world's seven largest countries, accounting for roughly half of the world's population, together with similar estimates from the WID, our replication of the WID and from the World Bank PIP. Our replicated changes in top 10% shares in the WID parallel the WID closely, though not exactly. For each country, we compare the change in inequality between 1990 (when surveys generally become available and the point of departure for the Millennium Development Goals) and 2007 (roughly the beginning of the global financial crisis), and then between 2007 and 2019. As we show earlier in this section, our estimates suggest that within-country inequality (either in GDP or HFCE) rose until approximately the mid-2000s or the early 2010s, and subsequently declined, making 2007 a natural turning point to consider.

The first row of Table II presents estimates of the changes in the top 10% share for China. All sources agree that China experienced a rise in inequality during 1990-2007, and then a much smaller rise in inequality, or an outright decline in inequality, subsequently. Our estimates using the distribution of GDP suggest that the decline in Chinese inequality between 2007 and 2019 nearly reversed the initial rise from 1990 to 2007. The fact that we estimate smaller rises and larger declines in the Chinese top 10% share is consistent with misreporting in China becoming less progressive over time relative to what is implied in the WID.

In the second row of Table II, we present similar estimates for the change in the top 10% share for India. The income share of the top 10% in India skyrocketed according to the WID, rising by 14.5 percentage points between 1990 and 2007, and then by a further 8 percentage points between 2007 and 2019. The household surveys instead find a much more modest rise in inequality during the first period and an outright fall (though not as large as the initial rise) in the second period. Our estimates suggest a larger rise in the pre-2007 period than do the household surveys, but then also a larger fall that nearly cancels out the initial rise, suggesting little net change in the top 10% share. For the results of the WID to obtain, underreporting should have been becoming steadily more progressive in India over time, with the top 10% reporting a smaller and smaller share of their income relative to what the rest of the distribution was reporting. In contrast, our results suggest that underreporting in India remained at a constant level of progressivity since the 1990s.

The third row of Table II presents our estimates for the top 10% share in the U.S. We find a larger increase in the top 10% share of GDP than do the household surveys and even the

WID both before and after the global financial crisis. We also find a larger increase in the top 10% share of disposable income in the pre-crisis period though it is partially reversed in the post-crisis period. Part of our finding of a very high increase in U.S. inequality in GDP is likely because of the conceptual differences between GDP and pretax income (the income concept used to construct world inequality estimates in the WID). Our findings suggest that underreporting progressivity in U.S. household surveys is likely not smaller than implied by estimates in the WID, including the work of Piketty and Saez (2003).

For the next four largest countries (Indonesia, Pakistan, Brazil and Nigeria) we find inequality declines in the post-crisis period that outweigh inequality increases in the pre-crisis period, or inequality declines in both periods. This is generally consistent with the household surveys in the World Bank PIP, and, for some countries, also with the WID. These estimates suggest that underreporting progressivity generally did not increase as much for these countries as implied by the WID, with the top 10% in these countries reducing the fraction of income they report in household surveys by more than the rest of the distribution did. Specifically, we reach this conclusion because the elasticity of regional GDP per capita to the household survey mean income (or consumption) in these countries has generally not increased over time, and frequently decreased.

Considering the largest 7 countries as a whole, we see that their increases in the top 10% share during the pre-crisis period were partially (and, in fact, almost completely) reversed during the post-crisis period. This is consistent with the World Bank's household surveys, though less so with the WID, which records a rise in inequality in the post-crisis period at a much smaller rate than during the pre-crisis period. These results explain why our estimates suggest that global within-country inequality fell while the WID's estimates do not, but highlight that our estimates, along with those of the World Bank and the WID are consistent with a considerable deceleration in the growth of inequality within countries after the global financial crisis.

Table II (II)

Changes in Top 10% Share by Country

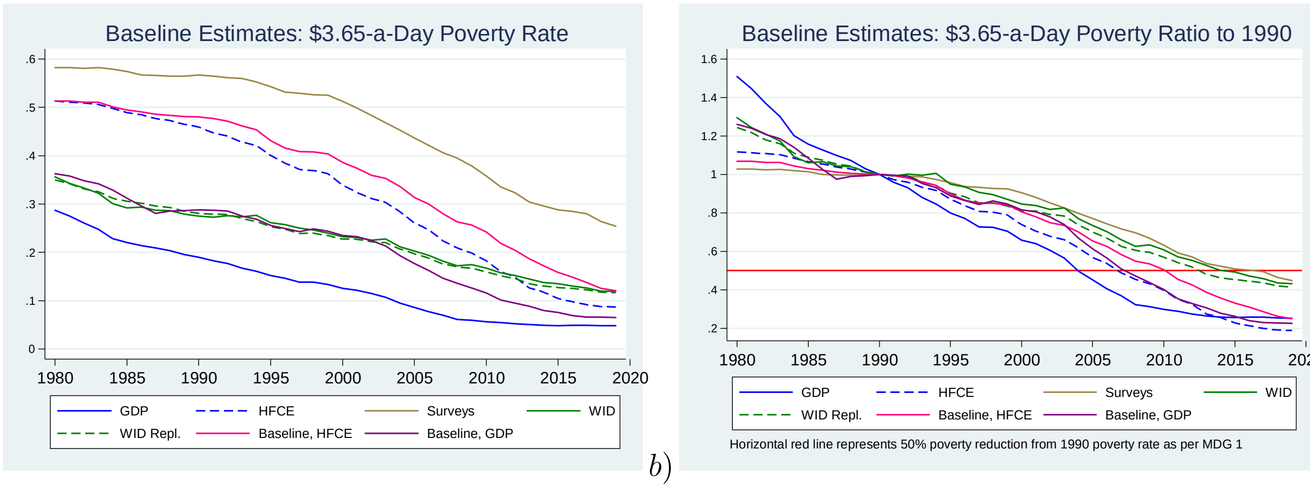
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Country	WID		WID Repl.		World Bank		Baseline GDP		Baseline HFCE	
	90-07	07-19	90-07	07-19	90-07	07-19	90-07	07-19	90-07	07-19
China	11.9%	-.9%	10.0%	.5%	5.6%	-1.9%	4.0%	-3.6%	7.4%	.3%
India	14.5%	8.0%	10.6%	5.2%	2.6%	-1.6%	4.4%	-4.1%	3.7%	-3.5%
USA	5.6%	1.3%	6.5%	.9%	2.7%	.3%	9.3%	7.9%	8.1%	-2.8%
Indonesia	3.7%	-1.8%	3.3%	-2.1%	2.5%	-.2%	7.2%	-11.0%	6.8%	-9.5%
Pakistan	.2%	2.3%	.5%	1.9%	-1.2%	-.2%	-1.2%	-.2%	-1.1%	-.2%
Brazil	-.2%	2.6%	.0%	2.8%	-4.9%	-1.2%	-2.3%	-4.8%	-2.6%	-2.3%
Nigeria	2.4%	-6.0%	1.9%	-4.7%	-2.4%	-1.4%	-.8%	-2.2%	-.8%	-2.2%
7 Largest Countries	10.0%	2.3%	8.0%	1.9%	2.8%	-1.4%	3.9%	-3.1%	4.7%	-2.2%

Note: Each column presents the absolute change in the top 10% share for the given series across the given time period. The "WID" series uses country-level data from the WID. The "WID, Replication" series adjusts the PIP survey data by estimated values of β^{WID} in equation (5) to approximate WID data. The "World Bank" series uses only survey data from the PIP for each Lorenz curve. The "Baseline GDP" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts GDP. The "Baseline HFCE" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts HFCE or disposable income.

3.4 Deeper and Faster Poverty Declines at Higher Poverty Rates

Earlier in this section, we have considered the behavior of the \$2.15-a-day poverty rate, showing that it is nearly 40% lower in 2019 when centering each national consumption distribution at national accounts HFCE and using our region-based approach to adjust inequality than when using unadjusted household survey data. However, the \$2.15-a-day poverty line represents a very specific definition of poverty. It is the median poverty line of all countries classified by the World Bank as low-income (with a GDP per capita of less than \$1,085; e.g. Afghanistan, Ethiopia and the DRC), and thus attempts to capture what fraction of the world population live below a material standard that can be considered as truly dire. We would have very different ideas of the world distribution of income for the same value of the \$2.15-a-day poverty rate if we were told that a large fraction of the nonpoor (according to that line) subsisted on income within a dollar a day of that rate, ever hovering on the edge of falling back into abject poverty than if we were told that only a small fraction of the nonpoor were in this precarious position. To the end of distinguishing these scenarios, the World Bank presents poverty rate estimates for two alternative lines. These are the \$3.65-a-day line, representing the median poverty line of the lower-middle-income countries (countries with a GDP per capita roughly between \$1,000 and \$4,000; e.g. India, Egypt, Bolivia), and the \$6.85-a-day line, representing the median poverty line of the upper-middle income countries (with a GDP per capita roughly between \$4,000 and \$13,000; e.g. China, South Africa, Mexico, Brazil, Turkey). In this subsection, we consider global poverty rates at these two higher lines. We find that our methodology suggests significant differences in the behavior of these poverty rates if the region-based inequality adjustment is used relative to both the approaches of the World Bank PIP and the WID.

Figure 16 (16)

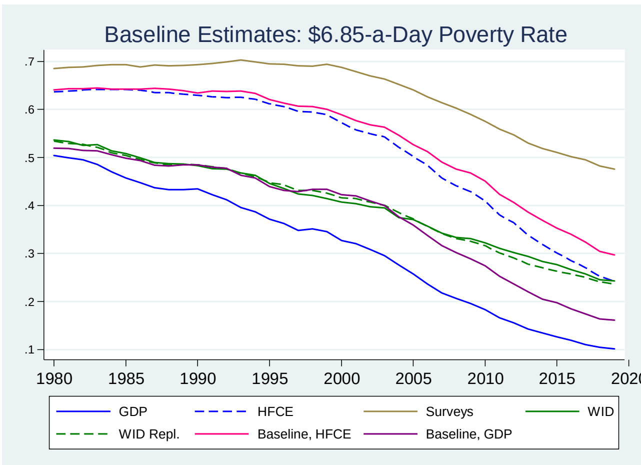


a) b)

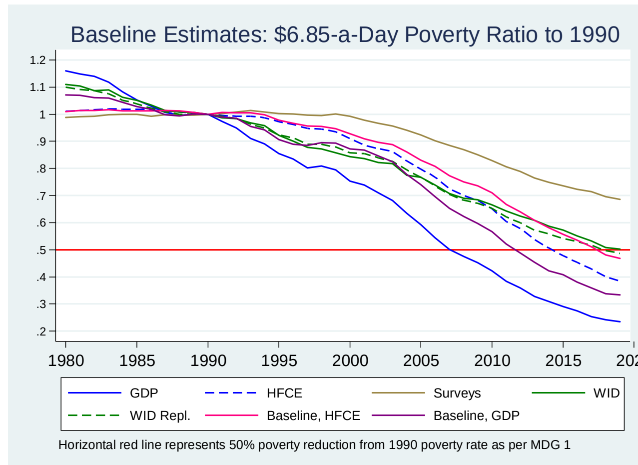
Note: See note to Figure (8).

Panel a) of Figure 16 presents estimates of $3.65-a-day poverty rates. We see that regardless of method, they have been declining over time. We also see that they are highest if surveys alone are used, and are much lower if we center distributions at national accounts consumption and adjust the surveys using our region-based procedure. Panel b) shows that while the $3.65-a-day poverty rate has been halved relative to 1990 only recently if surveys alone or the WID are used, our estimates using national accounts consumption suggest that it was halved in the early 2010s and that current poverty rates are about 25% of the ones prevailing in 1990. Estimates of $3.65-a-day poverty rates are lower if GDP is used to center the national distributions, whether adjusted or unadjusted. The unadjusted estimates using GDP or HFCE, as well as the misreporting-corrected estimates using GDP or HFCE, all show that by 2019, $3.65-a-day poverty declined to less than a quarter of its 1990 level.

Figure 17 (17)



a)



b)

Note: See note to Figure (8).

Finally, Figure 17 presents estimates for $6.85-a-day poverty. People above this poverty line would not be considered poor in countries such as China, Turkey or Brazil. Panel a) shows that if surveys alone are used (brown line), until the late 2010s, more than half the world population would have been under this poverty line. However, our HFCE-based estimates that adjust inequality using regional data (pink line) suggest that more than half the world population stopped being poor according to this generous definition of poverty by the mid-2000s, with less than a third of the global population being poor in 2019 according to this standard. It is worth noting that the large disparity in the surveys-only estimates and the HFCE-based estmates is relatively recent: in 1980, both measures would have agreed that $6.85-a-day poverty was the lot of 60-70% of the world's population. Using GDP rather than household surveys or HFCE to center income distributions, even with adjustments, further reduces this poverty rate, with only a quarter of the world's population being poor according to the $6.85-a-day threshold according to calculations using WID data, and less than a fifth being poor according to our methodology of correcting the household surveys using the regional adjustment. Panel b) suggests that global $6.85-a-day poverty has recently been halved from its 1990 level if our HFCE-based estimates or the WID-based estimates are used, and is just over 30% of its 1990 level if our GDP-based estimates are used, though it is far from being halved if we use surveys alone. While it is clear that adjusting the household surveys revises poverty estimates upward by substantial amounts – if we combined GDP with household survey distributions we would have obtained that only a tenth of the world's population is below the $6.85-a-day line – this adjustment suggests that inequality fell by much more during the 2010s than the WID suggests, and hence, that much more of the growth of that decade helped to lift people out of poverty by any definition.

It is worth noting that our HFCE-based estimates with the region-based inequality adjustments (the pink lines in all the figures) are conservative estimates of poverty at any poverty line for several reasons. First, they center national income distributions at final consumption expenditure, which likely is smaller than disposable income both on average and for the vast majority of individuals. Second, they frequently rely on the regional distribution of GDP, which tends to be much more unequally distributed than is disposable income or consumption, to implement the correction. Notwithstanding both conservative assumptions, these estimates deliver much lower poverty rates and much faster rates of poverty reduction than do the World Bank's official estimates in the PIP, offering a radically different picture of the role of poverty in the world. Poverty, even as it is understood in solidly middle-income countries rather than the extreme deprivation of people on the margins of subsistence, is rapidly becoming a relic of the past.

4 Accounting for Statistical Uncertainty in the Estimates

It is interesting and important to ask how robust our estimates of falling poverty, rising welfare, and falling within-country inequality are to statistical uncertainty in our estimating procedure. Statistical uncertainty enters our calculations in three places, 1) sampling error in the unit record data in the household surveys, 2) measurement error in log GDP or HFCE per capita at the region level, captured by the error term ε_r in Equation (3), and 3) the effects of these errors on the corrections and extrapolations described in Section 2.2. In this section we verify that none of these sources of variability affect our conclusions and provide bounds for the sensitivity of our approach to them.

We compute statistical uncertainty in the betas for each country as follows. For countries with no unit record data (the large Asian and African countries, as well as most Latin American countries), we estimate the standard error of the NLS estimate of β for each country and year using standard asymptotic formulas. For countries with unit record data (the US and other OECD countries), we estimate the standard error of the measurement-error corrected estimate β^{adj} (from Equation (8)) via the bootstrap, which we conduct by drawing a bootstrap sample with replacement for each region in each survey, and then drawing a bootstrap sample with replacement of the regions within each survey. This sampling scheme thus captures both sampling error in the unit record data (insofar as it contributes to our estimates of regional survey means and Lorenz curves) and measurement error in the dependent variable in Equation (3). We compute 100 estimates of β for each survey with unit record data (essentially each survey where we use LIS data) and compute the standard error of β as the standard deviation of these 100 estimates about their mean. (In robustness checks available on request, we have verified that using the mean of the 100 bootstrap iterations instead of the original nonlinear estimate β does not affect any of our world resuls).

Armed with a variance estimate of β for each country, we can compute uncertainty around statistics from the world distribution of income by sampling from each distribution of each estimate of β and following the interpolation and extrapolation steps of Section 2.2 to produce samples from the distribution of estimates of world poverty, inequality and welfare. To do this, we need to make assumptions about the cross-correlations of errors of β across countries and years. Two polar opposite assumptions are 1) that errors in β are independent across all surveys, and 2) that errors in β are perfectly correlated across all surveys within a country, but independent across surveys in different countries. Assumption 2 would be motivated by the idea that output in the same regions in a country may be consistently mismeasured from year to year. We provide variance estimates using both the Correlated and the Independent assumption, and expect the truth to lie somewhere in the middle. For each assumption, we construct 1000 different estimates of the world distribution and use them to compute uncertainty in our estimates.

Table III presents the results of our exercise. We supplement our baseline estimates and their uncertainty with corresponding figures from the WID, our replication of the WID using the parametric model in Equation (4) and from the World Bank's Poverty and Inequality Platform.

Table III (III)

Statistical Uncertainty in Main Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statistic	WID	WID Replication	World Bank PIP	Baseline GDP Correlated	Baseline GDP Independent	Baseline HFCE Correlated	Baseline HFCE Independent
1) P(Max Average Gini Attained Before 2000)	0	0	0	0	.062	0	.002
2) P(Max Average Gini Attained In 2000-12)	0	0	1	1	.926	.996	.949
3) P(Max Average Gini Attained After 2012)	1	1	0	0	.012	.004	.049
4) Inequality Decline between Max and 2019	-.000	-.000	-.014	-.040	-.040	-.020	-.020
(97.5% Upper Bound)				(-.020)	(-.012)	(-.012)	(-.002)
5) Welfare Ratio: 2019 to 1990, %	203	209	194	223	223	235	235
(2.5% Lower Bound)				(218)	(216)	(229)	(231)
6) $2.15\Day Poverty Rate Ratio, 2019 to 1990	.456	.351	.263	.238	.238	.170	.170
(97.5% Upper Bound)				(.328)	(.318)	(.230)	(.193)
7) $3.65\Day Poverty Rate Ratio, 2019 to 1990	.431	.415	.447	.226	.226	.250	.250
(97.5% Upper Bound)				(.302)	(.286)	(.317)	(.292)
8) $6.85\Day Poverty Rate Ratio, 2019 to 1990	.502	.486	.686	.333	.333	.468	.468
(97.5% Upper Bound)				(.413)	(.385)	(.521)	(.500)

Note: This table shows the results of a simulation exercise in which values of β were drawn from the sampling distributions of $\hat{\beta}_{adj}$ for countries with unit record data (LIS data) and of $\hat{\beta}^{NLS}$ for countries with regional data but without unit record data (SEDLAC and national statistical office data). Draws were either assumed to be independent across surveys (columns marked "Independent") or a common z-score was drawn from the standard normal distribution and used for all surveys belonging to a single country (columns marked "Correlated"). For each of 1000 iterations for each resampling method, given the values of $\hat{\beta}_{adj}$ for both the GDP-based and HFCE-based estimates of β, estimation proceeded as described in Section 2. Values of β^{WID} were assumed to be without error.

The "WID" series uses country-level data from the WID. The "WID, Replication" series adjusts the PIP survey data by estimated values of β^{WID} in equation (5) to approximate WID data. The "World Bank" series uses only survey data from the PIP for both the mean and the Lorenz curve of each income distribution. The "Baseline GDP" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts GDP. The "Baseline HFCE" series adjusts the PIP survey data by estimated values of $\hat{\beta}_{adj}$ in equation (8), with the Y variable being regional accounts HFCE or disposable income.

The maximum average Gini is defined for each replication as the maximum value of the population-weighted average of countries' Gini coefficients, a way to measure within-country inequality. Probabilities and 95% confidence intervals are defined based on the runsof 1000 replications for each series and each method of drawing replications.

We first consider our finding that within-country inequality has not been continuously rising into the 2010s, but peaked sometime during the aftermath of the global financial crisis and declined thereafter. In the first three rows of Table III we show what fraction of the 1000 replications of our baseline estimates feature this property by tabulating how many of them attain their maximum value of within-country inequality (the population-weighted average Gini) before 2000, between 2000 and 2012, and after 2012. We note that columns 1 and 2 show that within-country inequality attains its maximum after 2012 (in fact, in 2018) using inequality estimates from the WID, and in the same year using our approximation to the WID by employing the best-fitting value of β for every Lorenz curve in the WID. However, column 3 shows that within-country inequality in the World Bank's household surveys attains its maximum within the 2000-2012 period (in fact, in 2008)[5]. As we discussed in Section 3.4, our estimates that adjust household surveys for misreporting also suggest that within-country inequality peaked sometime during 2000-2012 both for the distribution of GDP and for the distribution of HFCE (columns 4-7). However, we now can describe how statistically certain we are in this finding. Of the 1000 replications in which all values of β are drawn independently across countries and years from our estimated sampling distributions, only 1.2% feature within-country inequality attaining its maximum after 2012 for GDP and 4.9% have within-country inequality attaining its maximum for HFCE (rows 1-3 of Table III). We can thus reject with 95% confidence the hypothesis that within-country inequality peaked after 2012 for both these series, consistent with the data from the World Bank PIP. If we assume instead that errors in β are independent across countries but perfectly correlated within countries, none of the 1000 replications feature within-country inequality peaking after 2012 for the distribution of GDP and only 4 have this behavior for HFCE. We thus conclude with confidence that our finding that within-country inequality has peaked by the early 2010s and has since then declined is not produced by pure chance.

Having concluded that within-country inequality has declined, how large can we argue the decline from peak inequality to have been? Both the WID and our replication of it suggest that it has been minimal (less than 0.0005 Gini points, row 4 and column 1). In contrast, the household surveys employed by the World Bank PIP suggest that the global within-country population-weighted Gini has declined by 1.4 Gini points. We find that inequality has declined by 4 and 2 Gini points between its peak and 2019 for the distribution of GDP and HFCE respectively. Of the 1000 iterations of our procedure that we run, 97.5% involve larger declines than 2 and 1.2 Gini points for GDP and HFCE respectively if we

[5] The World Bank does not construct this statistic and does not make the claim that global within-country inequality has peaked. However, the World Bank Poverty and Inequality Platform reports inequality statistics from the household surveys that allow us to reach this conclusion.

assume that sampling errors are perfectly correlated, and larger declines than 1.2 and 0.2 Gini points, respectively, if we assume that the errors are independent. Thus, we can reject the null hypothesis of only a minimal decline since within-country inequality has peaked, and for several sets of assumptions we can reject null hypotheses of declines in inequality of less than 1 Gini points at conventional levels of statistical significance.

We can similarly place lower or upper confidence bounds on our estimates of welfare growth and of poverty decline. Rows 5 through 8 present our estimates (with lower or upper confidence bounds) of the growth of Sen welfare, and the poverty ratio reductions using the three World Bank poverty lines discussed in Section 3.5, all relative to the value of welfare or the poverty rate in 1990. We see that welfare growth using our estimates for either GDP or HFCE comfortably exceeds welfare growth either in the WID or in the household surveys, even at 97.5% confidence (row 5). We also see that relative to their levels in 1990, global poverty rates have fallen for our distribution of HFCE by considerably more than they have fallen in the World Bank's household surveys, and that they have fallen for our distribution of GDP by well more than they have fallen using estimates from the WID (or our replication of them), all even if we take the 97.5% upper bound. Thus, this paper's main conclusions reflect underlying properties of the data rather than the noise in our estimation procedure.

5 Conclusion

In this paper, we attempt to tackle head-on the problem of adjusting household survey inequality measures to make them comparable to and consistent with using more reliable national accounts data to accurately measure economic growth (Pinkovskiy and Sala-i-Martin 2016). We propose a methodology that allows us to infer the degree of adjustment required by comparing the regional distributions of national accounts aggregates and of household survey means. Since household surveys are administered nationally, individuals with the same income should answer (and misreport their income) in the same way regardless of what region they live in, allowing us to estimate the structural parameters of the relationship between reported survey income and the individuals' underlying contributions to national accounts aggregates.

The main alternatives to our procedures are the estimates in the WID (Chancel et al. 2022) and the PIP (World Bank, 2023). The former also treat national accounts aggregates as the means of national income distributions, but adjust household survey distributions using tax data. The latter dispense with national accounts aggregates and rely on surveys alone for both the mean and the inequality of national accounts distributions.

We show that using our approach generates substantially different findings from both the

WID and the PIP, while corroborating their results in some dimensions. First, our methodology gives us a method to explicitly estimate, rather than assume, underreporting elasticities with respect to income in the household surveys. We find that underreporting progressivity has evolved smoothly over time, declining slightly around the world since 1990, whereas the estimates in the WID imply that it should have dramatically risen in the 2000s and 2010s after remaining essentially flat during the 1980s and 1990s. We also find that underreporting has been growing rapidly among the bottom 50% of the world income distribution, more so, in fact, than among the top 10%. One reason for this may be that as the global poor attain higher absolute incomes (but institutions incentivizing compliance with surveys don't change), their misreporting increases. Second, we confirm that the results of the previous literature – that global poverty and inequality have substantially declined since the 1980s, with both declines accelerating over time – even after correcting for survey underreporting, which is a first-order correction as it substantially affects the estimated level of global inequality and has the scope of completely revising our understanding of the trends (see figure 2). The level of inequality in our baseline measures is similar to the level of inequality in the WID and the PIP before 2000, but subsequently declines faster. Third, we show that not just between-country inequality but also within-country inequality in both GDP and HFCE have been falling since the global financial crisis, and therefore both types of inequality have contributed to falling global inequality. Finally, while poverty estimates are similar in the 1980s in the World Bank PIP and our HFCE-based measure, by the 2010s, the measures diverge quite substantially, with our approach showing considerably lower world poverty rates especially at high poverty lines.[6] In summary, our approach shows that when combining national accounts aggregates with suitably adjusted household survey inequality measures, we find that the world is a less poor and less unequal place than has been thought.

References

[1] Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. "Towards a system of distributional national accounts: Methods and global inequality estimates from WID. world." Economie & Statistique 517-518-519 (2020): 41.

[6] A concern in the literature on the world distribution of income is that because of China and India's size (nearly 40% of the world population), global results are unduly dependent on the quality of distributional statistics in these two countries. We have recreated all our estimates for the remaining 60% of the world that excludes China and India, and essentially all of our conclusions hold when considering this smaller group of countries. In particular, after adjusting for misreporting, both overall and within-country inequality remain on a declining trend after the late 2000s, and poverty reductions remain higher than in the WID or documented by the World Bank for the same group of countries.

[2] Bhalla, Surjit S. Imagine there's no country: Poverty, inequality, and growth in the era of globalization. Peterson Institute, 2002.

[3] Bourguignon, François, and Christian Morrisson. "Inequality among world citizens: 1820–1992." American economic review 92, no. 4 (2002): 727-744.

[4] Burkhauser, Richard V., Markus H. Hahn, and Roger Wilkins. "Measuring top incomes using tax record data: A cautionary tale from Australia." The Journal of Economic Inequality 13 (2015): 181-205.

[5] Chancel, Lucas, and Thomas Piketty. "Indian income inequality, 1922-2015: from british raj to billionaire raj?." Review of Income and Wealth 65 (2019): S33-S62.

[6] Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman, eds. World inequality report 2022. Harvard University Press, 2022.

[7] Chancel, Lucas, Denis Cogneau, Amory Gethin, Alix Myczkowski, and Anne-Sophie Robilliard. "Income inequality in Africa, 1990–2019: Measurement, patterns, determinants." World Development 163 (2023): 106162.

[8] Chen, Shaohua, and Martin Ravallion. "The developing world is poorer than we thought, but no less successful in the fight against poverty." The Quarterly Journal of Economics 125, no. 4 (2010): 1577-1625.

[9] Deaton, Angus. "Measuring poverty in a growing world (or measuring growth in a poor world)." Review of Economics and statistics 87, no. 1 (2005): 1-19.

[10] Gennaioli, Nicola, Rafael La Porta, Florencio Lopez De Silanes, and Andrei Shleifer. "Growth in regions." Journal of Economic growth 19 (2014): 259-309.

[11] Jha, Somesh. https://twitter.com/someshjha7/status/1678651674527948800. Accessed August 1, 2023.

[12] Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion. "Survey nonresponse and the distribution of income." The Journal of Economic Inequality 4 (2006): 33-55.

[13] Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org (multiple countries; microdata runs conducted most recently in September 2022- December 2023). Luxembourg: LIS.

[14] Meyer, Bruce D., Wallace KC Mok, and James X. Sullivan. "Household surveys in crisis." Journal of Economic Perspectives 29, no. 4 (2015): 199-226.

[15] Organization for Economic Cooperation and Development. 2023. OECD Statistics: Regions and Cities. Retrieved from https://stats.oecd.org/

[16] Piketty, Thomas. "Income inequality in France, 1901–1998." Journal of political economy 111, no. 5 (2003): 1004-1042.

[17] Piketty, Thomas, and Emmanuel Saez. "Income inequality in the United States, 1913–1998." The Quarterly journal of economics 118, no. 1 (2003): 1-41.

[18] Piketty, Thomas, Li Yang, and Gabriel Zucman. "Capital accumulation, private property, and rising inequality in China, 1978–2015." American Economic Review 109, no. 7 (2019): 2469-2496.

[19] Pinkovskiy, Maxim, and Xavier Sala-i-Martin. Parametric estimations of the world distribution of income. No. w15433. National Bureau of Economic Research, 2009.

[20] Pinkovskiy, Maxim, and Xavier Sala-i-Martin. "Lights, camera… income! Illuminating the national accounts-household surveys debate." The Quarterly Journal of Economics 131, no. 2 (2016): 579-631.

[21] Ravallion, Martin. "Poverty lines across the world." World bank policy research working paper 5284 (2010).

[22] Sala-i-Martin, Xavier. "The disturbing" rise" of global income inequality." (2002). NBER WP 8904.

[23] Sala-i-Martin, Xavier. "The world distribution of income (estimated from individual country distributions)." (2002). NBER WP 8933.

[24] Sala-i-Martin, Xavier. "The world distribution of income: falling poverty and… convergence, period." The Quarterly Journal of Economics 121, no. 2 (2006): 351-397.

[25] SEDLAC: Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank). Accessed September 2018.

[26] Sen, Amartya. "Real national income." The Review of Economic Studies 43, no. 1 (1976): 19-39.

[27] Shorrocks, Anthony F. "Inequality decomposition by population subgroups." Econometrica: Journal of the Econometric Society (1984): 1369-1385.

[28] Subramanian, Arvind. "India's GDP mis-estimation: Likelihood, magnitudes, mechanisms, and implications." CID Working Paper Series (2019).

[29] World Bank. (2023). Poverty and Inequality Platform. World Bank Group. www.pip.worldbank.org. Accessed August 1, 2023.

6 Appendix

6.1 Proof that Y Lorenz dominates $\hat{Y}$

Let $\ln y_i$ be the log unobserved true income of respondent i, and let ξ_i be their error in the reported to true income relationship, which by assumption is fully independent of $\ln y_i$. Then, as per equation (2), we define

$$\ln \hat{y}_i = \ln y_i + \xi_i = \alpha + \beta \ln x_i$$

Since we have the distribution of x_i, and once we estimate a value for β, the Lorenz curve of $\hat{Y}_i = \exp(\ln \hat{y}_i) = \exp(\alpha + \beta \ln x_i)$ is estimable by the formula (4). However, the Lorenz curve of Y_i is not estimable because of the presence of the term ξ_i.

We now show that $\hat{Y}_i$ is Lorenz dominated by Y_i. By a theorem of Atkinson (1970), Lorenz dominance is equivalent to second-order stochastic dominance for two variables Y and $\tilde{Y}$ such that $E(Y) = E\left(\tilde{Y}\right)$. Now, by a theorem of Rothschild and Stiglitz (1970) a variable Y second-order stochastically dominates another variable $\tilde{Y}$ if and only if $\tilde{Y}$ is a mean-preserving spread of Y: the variable $Z = \tilde{Y} - Y$ satisfies $E(Z|Y) = 0$. Define

$$\tilde{Y} = \frac{1}{E(\exp \xi_i)} \hat{Y} = \frac{\exp(\xi_i)}{E(\exp \xi_i)} Y$$

Then,

$$Z = \tilde{Y} - Y = Y \left(\frac{\exp(\xi_i)}{E(\exp \xi_i)} - 1 \right)$$

and

$$E(Z|Y) = YE\left(\left(\frac{\exp(\xi_i)}{E(\exp \xi_i)} - 1\right)|Y\right) = 0$$

the latter equality following by the independence of $\ln y_i$ and ξ_i

Therefore, $\tilde{Y}$ is a mean-preserving spread of Y and, defining $\tilde{L}$ and L to be the Lorenz curves of $\tilde{Y}$ and Y respectively, $\tilde{L} \leq L$. Now, $\tilde{Y}$ is just a rescaling of $\hat{Y}$, so $\hat{L} = \tilde{L} \leq L$. QED.