2019

IARIW-World Bank

Special IARIW-World Bank Conference "New Approaches to Defining and Measuring Poverty in a Growing World" Washington, DC, November 7-8, 2019

Mind the Gap: Disparities in Assessments of Living Standards Using National Accounts and Surveys

Espen Prydz

Dean Jolliffe

Umar Serajuddin

Paper Prepared for the IARIW-World Bank Conference

Washington, DC, November 7-8, 2019

Session 4A: Issues of Poverty Data and the SNA

Time:13:30 – 15:30, September 8

Mind the Gap

Disparities in Assessments of Living Standards Using National Accounts and Surveys

October 15, 2019

Espen Beer Prydz[af], Dean Jolliffe[ab], Umar Serajuddin[a]

Abstract: Average per capita consumption and income from national accounts data can differ substantially from corresponding measures of consumption and income from household surveys. Using a new compilation of 1,664 household surveys matched to national accounts data, we find the gap between the data sources is larger and more robust than previously established. Survey means, on average, are 25 percent lower than corresponding means for household consumption from national accounts. The gap is largest in middle income countries, where also annualized growth rates for surveys are more than 1 percentage point lower than growth in national accounts, on average. We show how the gaps matter substantially for assessments of growth, poverty and inequality. We illustrate how typical survey measures, which do not account for these gaps, can exaggerate global poverty reduction and underestimate global inequality.

Keywords: National Accounts Systems, Household Income and Expenditure Surveys, Poverty, Inequality.
JEL Codes: I3, I32, E31, F01.

Acknowledgements: The authors wish to thank the UK Department of International Development for funding assistance through its Knowledge for Change Program (TF018077-KCP II - National Account vs Survey Based Welfare). The analysis in this paper would not be possible without data from national accounts and household surveys, and publicly available compilations by international agencies made available from the World Development Indicators database, PovcalNet and others. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the NPC, IZA, World Bank Group and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

[f] Corresponding author. Email addresses: eprydz@worldbank.org (E. B. Prydz), userajuddin@worldbank.org (U. Serajuddin), djolliffe@worldbank.org (D. Jolliffe)
[a] The World Bank, Development Economics Data Group (DECDG)
[b] Research Fellow with the Institute for the Study (IZA) of Labor in Bonn; Research Affiliate with the National Poverty Center (NPC) at the University of Michigan.

1. Introduction

Measures of per capita income and consumption are among the most frequently cited indicators of economic development. They are widely used in assessments of average living standards, economic growth, poverty and inequality, both within and between countries. Despite the prominence of these indicators, the two most common data sources for such measures – national accounts systems (NAS) and household surveys (HHS) – often give very different estimates of average living standards and growth. For example, in Pakistan in 2015, national accounts data suggest that average household consumption expenditure per capita was $9.3 per day at 2011 PPP, while the household survey suggests it was just a bit more than half of that, $4.9 per day at 2011 PPP.[1] In Botswana, the two recent household surveys suggest average consumption contracted at an annualized rate of -3.3 percent between 2010 and 2013, while the most closely aligned measure from the national accounts system, household final consumption expenditures (HFCE), indicated a robust expansion of household consumption at an annualized rate of 3.9 percent over the same period, as well did gross domestic product (GDP).

That national accounts data and survey data can lead to such diverging measures of the levels and rate of change of living standards is a recurring phenomenon across a wide range of countries and statistical systems. A frequently cited country-case is India, where there has been large discrepancy between measures of household consumption expenditure in national accounts and those in the National Sample Survey (NSS), fueling a vigorous debate about the evolution of poverty and its relationship to economic growth (see for example Deaton and Kozel, 2004; Subramanian and Jayaraj, 2015; Sundaram and Tendulkar, 2003). The issue of diverging estimates from national accounts and household surveys is not limited to poor countries. In the United States, per capita income from the two large national surveys, the Current Population Survey and the Consumer Expenditure Survey, are known to diverge from national accounts (see for example the recent assessment by Rothbaum, 2015). Nolan et al (2019) recently reviewed gaps between survey incomes and national accounts income in 27 OECD countries, finding discrepancies in annual growth rates of 0.32 percentage points in the United States and 0.55 percentage points in Germany.

The most complete reviews of the discrepancies between survey and national accounts data were conducted by Ravallion (2003), Karshenas (2001, 2003) and Deaton (2005), who all assessed the discrepancies globally with a sample of household surveys and national accounts data from 1980s and

[1] Survey data for this example comes from World Bank, Global Database of Shared Prosperity (GDSP) circa 2008 - 2013 (see www.worldbank.org/en/topic/poverty/brief/global-database-of-shared-prosperity) and national accounts data from World Development Indicators.

1990s. Since these global assessments, the availability of household survey data in poor countries has expanded considerably and many countries have revised both survey and national accounts data and methods. We also have better metadata on comparability and type of household surveys and income data, which can help us better understand the discrepancies.

The main objective of this paper is two-fold: First, using a larger compilation of NAS and HHS data, with broader geographic and temporal coverage stretching from 1979 to 2017, we revisit the NAS-HHS gap and the findings of Ravallion (2003), Karshenas (2003) and Deaton (2005) based on data primarily from the 1980s and 1990s. Second, we illustrate how the discrepancies can lead to divergent representations of how living standards, poverty and inequality, differ between countries and over time. We compile a dataset of 1,139 household survey means from 162 countries surveys matched to corresponding national accounts aggregates. The contribution is threefold. First, we assess the gap with a much larger sample of household surveys and national accounts data, over a larger time range than previous studies. Second, we overcome the heterogeneity often seen in household surveys by using a dataset of comparable spells. Third, we clearly illustrate the potential implications of the gaps in use of national accounts data in measuring global poverty and inequality.

Overall, our findings suggest that disparities between income and consumption measures from surveys and those from national accounts are much larger than found by both Ravallion (2003) and Deaton (2005). On average, across all countries, we find that per capita survey means are about 77 percent of the means of household consumption seen from national accounts, and per capita means income surveys are only about 45 percent of gross domestic product (GDP). This indicates considerably larger discrepancies than what was found in Deaton's and Ravallion's similar assessments. Ravallion found no statistically significant gap in means of consumption surveys and mans of household consumption from national accounts, while income surveys showed means that were 67 percent of national accounts. Deaton found that survey means were 88 percent of household consumption from national accounts and 57 percent of national income. We find that the size of the gaps varies systematically with economic development with the discrepancy being largest for middle income countries. In contrast to previous assessments of the gaps, we find relatively small differences between income and consumption surveys across all levels of development. We also find that the gaps for income and consumption surveys are similar, which contradicts Ravallion's (2003) suggestion that the gap is mainly due to underreporting of incomes in surveys. In contrast to Deaton (2005) and Karshenas (2003), we find that the gap is narrowing as countries get richer for both income and consumption surveys, possibly reflecting better integration of NAS and HHS data in high income countries due to improved efforts to align survey and national accounts in recent year. We also find that growth rates from national

accounts are higher than survey growth, particularly in middle income countries, in line with the economic gradient in the gaps in levels.

We illustrate how the implications of the observed gaps for measures of global poverty and inequality depend on assumptions about the reasons for the gaps and corresponding adjustments to measurement methods. On the one hand, if we assume that the gap is a result of surveys' underestimation of the full extent of household incomes and consumption and that this underestimation is proportionally uniform across the distribution, one could justify distribution neutral adjustments of distributions by using national accounts means for measuring poverty and inequality, in line with the methods used by Bhalla (2002), Sala-i-Martin (2006) and Pinkovskiy and Sala-i-Martin (2009; 2016). Such adjustments generally show that both global poverty and inequality is lower and falling faster than survey measures would suggests. On the other hand, if we assume that the gap is due to surveys disproportionally failing to capture the income consumption and income of the richest households would justify adjustments of top incomes proportional to the NAS-HHS gap, similar to what has been proposed by Lakner and Milanovic (2016) and Chandy and Siedel (2017a,b). Such adjustments have little effect on global poverty measures, but substantially revises upwards both global and national inequality, and significantly changes our understanding of the distributional nature of growth and 'shared prosperity'.[2] Notably, the relationship between observed levels of inequality and overall economic development changes significantly, strengthening evidence of a Kuznets curve where inequality first rises and then falls with economic development. Overall, we show that using NAS data to impute and extrapolate poverty measures give estimates that differ considerably from those directly derived from survey measures, given the heterogeneity of the gap across countries and time.

The remainder of the paper proceeds as follows. Section 2 briefly discusses conceptual differences between national accounts and surveys data, reviews existing literature assessing the gaps, and proposals adjustments in accord with the source of the gap for measuring poverty and inequality. Section 3 presents the data compiled for this paper. Section 4 presents our assessment of systematic difference between NAS and HHS, first in terms of levels and then in terms of growth rates. Section 5 discusses the implications of the discrepancies for the level and changes in global poverty, as well as for national and global inequality. Section 5, concludes.

[2] Shared prosperity is the World Bank's indicator of shared growth, measuring the income growth of the bottom 40 percent of the distribution in each country, and often compare this to the overall growth of the country.

2. National accounts and household surveys

That national accounts and survey measures of average household income and consumption can differ is well-established. Although the concepts they try to measure are broadly similar, there are also differences which would lead us to expect estimates to be different. Household surveys commonly measure consumption by using recall questionnaires administered to a sample of households designed to be nationally representative. Consumption expenditures are commonly estimated as the total market value of goods and services reported consumed by the household over the recall period. Survey methods and questionnaires vary widely across the world, and even within countries, which can have substantial impact on the average standard of living, poverty and inequality measures (see for example Jolliffe, 2001; Beegle et al. 2012; Jolliffe et al., 2014.). Estimates of household income are typically also derived from recall questions, asking about various components of revenue. As survey methods and questionnaire change in a country, past survey estimates are typically not revised, limiting the comparability of new estimates with previous data. Recent metadata on survey comparability suggests revisions in survey methods are very widespread, with two thirds of the 164 countries for which data is available having at least one break in the series, and a quarter of countries having three or more breaks.[3] The issue of limited comparability of surveys over time within countries is further exemplified by the fact that the World Bank is only able to produce comparable data over a recent five-year period is only available for 93 countries of the 164 countries in its database.[4] This highlights the heterogeneity of survey data, not only between countries but also within countries, which contribute to discrepancies both in growth rates and levels. Another limitation of household surveys is that they are conducted with uneven frequency and many countries lack surveys for at least five years or longer (Serajuddin et al, 2015).

National accounts systems aim to measure the value of total production and consumption of all goods and services among households, businesses, government and foreign countries (exports) within a country and period. The resulting national income measures are seldom derived from data collected in household surveys. The component of NAS which is most closely related to concepts captured in household surveys is Household Final Consumption Expenditures (HFCE). The definition of HFCE is broader than what is typically included in household surveys, encompassing things like spending of non-profit entities, such as religious groups, NGOs and foundations. Moreover, HFCE is not always directly

[3] The comparability metadata is available in the World Bank's GitHub Repository for PovcalNet: https://raw.githubusercontent.com/worldbank/povcalnet/master/metadata/povcalnet_metadata.csv
[4] Global Database of Shared Prosperity and Median Income/Consumption, circa 2011-2016
As of Sept 6, 2019. https://www.worldbank.org/en/topic/poverty/brief/global-database-of-shared-prosperity

measured but is often taken as a residual when other sectors of the economy are accounted for making it prone to errors.

While there are clearer internationals standards for compilation of national accounts data than for survey data, guided by the UN Statistical Division's System of National Accounts (SNA), there is also substantial heterogeneity in methods deployed across countries. As with household surveys, countries revise their NAS methods from time to time, which can lead to large revisions in national income and consumption. This has been particularly pronounced in Sub-Saharan Africa in the past decade, where several countries have revised their national accounts estimates to incorporate new methods and data sources in effort to better capture all economic sectors and includes emerging economic activities that were not previously captured. The implementation of the 2008 System of National Accounts has also had the effect of revising up estimates. For many economies these changes led to to upward adjustments of GDP. For example, Nigeria's GDP was adjusted upwards by 60 percent between 2013 and 2014. Similar revisions of more than 20 percent has been seen in Ghana, Kenya, Senegal and Zimbabwe (See Kouame, 2019). In contrast to household survey data though, common practice is to revise the entire national account series, ensuring as much comparability across time as possible. Arguably this ensures that national accounts data is more comparable over time and also available at higher frequency.

Previous studies have taken different approaches to what national accounts measures are more appropriate for comparison with income and consumption from surveys, and what national accounts data is most appropriate for measuring poverty and inequality. Anand and Segal (2008) argue that the better national accounts measure for poverty and inequality measures is HFCE, rather than GDP. Ravallion (2003) also suggest HFCE is the appropriate comparator to survey means and for poverty measures, while Deaton argues that for consumptions measures from surveys, HFCE is the right comparator, while for survey income means, GDP may be the more appropriate comparator. Pinkovskiy and Sala-i-Martin (2014, 2016) have primarily suggested using GDP for poverty measurement, rather than HFCE, mostly due to data availability. Ahluwalia et al. (1979) and Bourguignon and Morrison (2002) also use GDP measures in combination with distributional data. Nolan et al (2019) make an important distinction that GDP measures the economic output of countries, while household surveys capture income of resident households including income from other countries, while excluding income accruing abroad. Therefore, they argue, that the national accounts measure of gross national income (GNI), which incorporates net factor income from abroad, is conceptually closer to the geographical coverage of household surveys. While we believe HFCE is the more natural comparator to survey estimates of consumption, we also measure the gap between survey means and GDP, since there is a large literature that proposes measuring poverty and inequality using GDP and because of the greater availability and consistency of GDP data

potentially making it a more stable indicator of average national well-being and more readily available for imputation and extrapolation methods when survey data is not available. GNI per capital in constant terms is not readily available for a large set of countries over time to be used in our assessment.

The literature to on the survey-national accounts relationship have found that aggregate means are typically larger than corresponding measures in household surveys, but with varying magnitudes and level of robustness. Ravallion (2003) finds that gaps between household surveys and national accounts are mainly due to discrepancies for income surveys, and, in the case of consumption surveys gaps can be attributed to large data inconsistencies in transition economies in Eastern Europe but are negligible elsewhere. While identifying discrepancies in measures from the two data sources, Ravallion fails to reject that national accounts is an unbiased estimator of the level and growth rate of survey means of household consumption. He finds systematic differences in levels of income from surveys, but not for growth rates. Karshenas (2003) uses a slightly different dataset and methods to assess the gap and finds a small but statistically difference.[5] Deaton (2005), using a larger sample of more than 500 surveys, also finds a more robust gap in levels and argues that for understanding implications for global poverty it is also important to examine the gaps from a population weighted perspective.

Undoubtedly, both surveys and national accounts measure mean consumption and income with error and it is not obvious which is more suitable for measurement of global poverty and inequality. Several aspects of household surveys can lead consumption and income estimates to be underestimated. It is for example observed that the length of the recall period, often reduces the amount of consumption reported (Beegle et al, 2012). Another source of survey error is non-response. A particularly concerning phenomenon is the tendency of richer households being less likely to respond to surveys (Korinek et al, 2006). Second, surveys may not keep up with changing consumption patterns as economies and habits change. For example, food consumption away from home is often not appropriately captured as countries get richer (Farfan et al, 2017).

Similarly, national accounts data is measured with error and methodological changes and revisions can lead to large changes in estimates of national accounts aggregates and average living standards. However, in contrast to survey estimates, national accounts series are commonly include revision of past income estimates, such. Moreover, there are commonly established international standards for national accounts compilation which, arguably, could make it more comparable across space and time.

[5] Karshenas (2003) argues that using the log difference between the means is the preferable estimation method.

The claim that average living standards as measured by national accounts are more comparable across countries and consistent over time is supported by recent analysis which finds that national accounts means correlate much better with nighttime lights, than household survey means (Pinkovskiy and Sala-i-Martin, 2016). This finding has given additional weight to the arguments for considering the use of national accounts data in international poverty comparisons. However, even if means from national accounts may be more comparable, the implications for poverty and inequality measures critically depend on assumptions about the distribution of consumption and income within countries, which typically come from surveys. The mean may be very sensitive to whether consumption of the richest are captured, since they hold a large share of overall income.

The implication of the gap for estimates of global poverty and inequality depend on the understanding of the reason for the gap and corresponding methodological choices. Some analysts have chosen to measure poverty with national accounts means, arguing that these are more comparable and available and may capture a fuller extent of national consumption and income than surveys. Most prominent in this camp has been a series of studies by Pinkovskiy and Sala-i-Martin (Sala-i-Martin, 2006; Pinkovskiy and Sala-i-Martin, 2009; 2016) and a study by Bhalla (2002). This literature has generally measured poverty using national accounts means in combination with distributional parameters derived from survey data. Generally, because national accounts means are higher than survey means, and these methods have used the same poverty lines established for use with household surveys, they give much lower poverty estimates at the same poverty lines, than those directly measured from surveys. Implicitly, the methods used in this literature assumes a distribution-neutral error in household surveys, since they in keep the distribution of income from surveys, and in effect scale everyone's consumption by the NAS-HHS gap.

There are however good reasons to believe that the NAS-HHS gap is not a result of distribution neutral measurement errors. There is evidence to suggest that much of the gap may originate from the fact that richer households are less likely to fully participate in household surveys and NAS better capturing large transactions. Korniek et al (2006) show the probability of non-response in the US is increasing with income, which would lead to underestimation of the survey mean. For India, Banerjee and Piketty (2010) find that a substantial part of the NAS-HHS gap can be attributed to missing top incomes from surveys. From the perspective that the NAS-HSS gap originates from undercoverage and/or underreporting of income and consumption of the richest households, a natural methodological adjustment is not to simply scale survey means to national accounts means, but to rather adjust the top incomes by some proportion of the gap. Lakner and Milanovic (2016) and Chandy and Siedel (2017) provide proposals for assigning parts of the NAS-HHS gap to the top segment of the distributions, building on methods and literature

from rich countries which make similar adjustments to survey data in rich countries using tax data, which in turn have little effect on poverty but much larger effect on inequality. The methods are described further and applied in Section 5.

3. Data

Survey data

To assess the gap between surveys and national accounts, we compile a dataset of 1,780 household survey means for 163 countries from 1977 until 2015, together representing 93 percent of the world population in 2015. The vast majority of the surveys come from PovcalNet, the World Bank's database for monitoring of global poverty (see Ferreira et al, 2016 for details). The database contains income or consumption distributions from nationally representative household surveys typically carried out or supervised by national statistical offices or international agencies, used for national and international poverty monitoring. We obtain the per capita mean consumption or income denominated in 2011 PPP for each national survey. We also collect the deflators, currency conversion rates (in the case of currency devaluations and change of national currency) and PPPs used to convert the survey data to 2011 PPPs, so that we can recover the survey mean in current (nominal) local currency units.[6]

For most high income countries, the survey means available in PovcalNet are for income (rather than consumption), originating from the Luxembourg Income Study and the European Union Statistics on Income and Living Conditions (EU-SILC). To ensure better coverage of consumption surveys of high income countries in our sample, we supplement the data from other sources. For European countries, we derive consumption means from Eurostat's compilation of Household Budget Surveys (Eurostat, 2016). Eurostat publishes summary statistics from household budget surveys every approximately every five years since 1988, with the latest data available for 2010. The Eurostat compilation provides consumption data for all 28 EU Member States as for Montenegro, the Former Yugoslav of Republic of Macedonia, Turkey and Norway.[7] The United States is a large high income country for which survey-based consumption data is not available from international compilations. We therefore add survey means from the Consumption and Expenditure Survey.

[6] The ancillary data of deflators and exchange dates is available at http://iresearch.worldbank.org/PovcalNet/data.aspx.

[7] Eurostat provides survey means are given in per adult equivalent and in constant currency terms. We use available data about household structure, deflators and exchange rates to recover per capita means in current local currency units and also convert to 2011 PPPs, using the method applied in PovcalNet.

Consumption and income surveys are considered separate series for each country, so that we do not calculate growth rates from an income survey in one year to a consumption survey in another. We express all surveys values in annualized per capita basis.[8] In cases where household survey consumption data is only available at the household level or in per adult equivalent terms, we use adjust these to per capita terms. While per adult-equivalent and economies-of-scale adjustments may make sense for analysis of household welfare within a country, for comparisons with national accounts means which is expressed in per capita terms and cross-country comparisons over long periods of time, per capita measures are more suitable.

Overall, our survey compilation contains heterogeneous surveys which may not be directly comparable even within countries due to changes in survey methods and practice. Therefore, for the assessment of growth rates in particular, where comparability in the measures matter significantly, we complement our dataset from the main sources with the World Bank's 'Global Database of Shared Prosperity', where particular care has been taken to assure that consumption and income aggregates are comparable within 93 countries.[9] Since this database overlaps considerably with the PovcalNet database, we do not add these observations to our main dataset, but rather create a complementary sample based on more recent, harmonized and comparable data. Results for this dataset are presented separately as robustness checks and used for analysis of some aspects of the implications in Section 5.

National accounts data

In this study, we mainly focus on the component of national accounts that corresponds to households' expenditure, known as household final consumption expenditure (HFCE) under the 1993 System of National Accounts (SNA), since this is the concept closest to that commonly captured by surveys.[10] While HFCE is the variable that conceptually most closely corresponds to the measures from household surveys, we also compare to GDP, for two reasons. First, GDP is probably a more frequently cited indicator of development and may be measured with less noise than HFCE, which is often measured as a residual. Second, some have argued that for income surveys, in theory, is conceptually more closely aligned to

[8] The survey data is generally expressed in per capita terms, or in terms of per household, from the survey data. No additional population data is used in transforming the means to per capita terms.
[9] Global Database of Shared Prosperity, circa 2011 -2016, available at http://www.worldbank.org/en/topic/poverty/brief/global-database-of-shared-prosperity. Some estimates are based on the 2008-13 version and not yet updated.
[10] Previous literature, in particular Ravallion (2003) and Deaton (2005) has referred to HFCE as private consumption expenditure (PCE), which was used under previous systems of national accounts.

GDP, depending on how income is measured and what assumptions are made about how savings are accounted for in national accounts aggregates.

We extract national accounts from the World Development Indicators (WDI) database, for both HFCE and GDP, using the series expressed both in current local currency units and in constant PPP series. WDI's data is a compilation of World Bank and OECD national accounts datasets, obtained from official national sources. The per capita estimates are derived using the mid-year population estimates from the World Bank population series data. We obtain estimates both in local currency units and 2011 constant international dollars from the ICP.

Putting it all together

Since the data availability of national accounts data is incomplete, we match 1,666 of our survey means to a corresponding national accounts estimate for the same country and year.[11] Our overall sample is substantially larger and more recent than those included in past studies. Deaton (2005) presented results form 557 surveys from 127 countries between 1979 and 2000. Ravallion (2003) used a smaller sample of means for 90 countries, and a panel of 142 growth rates for 60 countries. We also have substantially better coverage of rich countries, than Ravallion, Karshenas and Deaton's reviews.The dataset also contains meta data about the origin of the survey estimate and national accounts data.

We take several steps to ensure that the data and our findings are not driven by adjustments for currency and prices. Our primary data is deflated to 2011 constant USD using purchasing power parities. To ensure that the process of converting the data to 2011 does not influence our results, we make some adjustments and checks using aggregates in local currency units. To assess differences in levels, we convert the amounts to current local currency units, using deflators and PPP conversion factors applied to the data. For growth rates, we do not make such adjustments, as we are interested in the difference in real growth rates between the data sources and using different deflators may be appropriate. However, we also conduct additional analysis where deflators between the data sources are set as equal, to understand if this drives the gap.

[11] In cases where the survey estimate covers more than one year, we match the survey year to a weighted national accounts aggregate with weights corresponding to the year of the survey data. (E.g. if a survey estimate is from a 12-month survey that is 9 months (75%) in 2011 and 3 months in 25 percent in 2012, we produce a weighted national accounts aggregate with corresponding weights.). This is consistent with the practice of dealing with surveys in PovcalNet that run over multiple years, as described in Chen and Ravallion (2010).

We generate growth spells with annualized growth rates by matching all observations for each country and type (income/consumption), which gives 6341 possible spells (including overlaps) and 1012 consecutive (non-overlapping) spells. In contrast, Ravallion (2003) generate 142 spells between successive household surveys for 60 countries in the 1980s and 1990s. Our practice of matching all observation with each other within each country-type panel maximizes our potential power in estimation, but we also conduct analysis of spells that are non-overlapping, and with limited length (e.g. spells of max 3 years, shorter than 5 years) etc.

We use three weighting schemes in our analysis. First, our preferred weighting is the one that weights each country equally. Where countries have varying number of observations of matched survey and national accounts data, each observation is weighted as the inverse of the number of observations for each country, so that the total weight given to each country, in each subsample, sums to one.[12] This ensures that our analysis does not assign more weight to countries because they have more household surveys available. Second, we also use population-weighted estimates for use in cases where we assess the gaps' overall effect on aggregate measures such as global changes in living standards or poverty – particularly relevant for the illustrative examples in Section 4. Last, we also present some unweighted results, where each survey-NAS observation is weighted equally, irrespective of country, with the drawback that this gives more weight to countries with more observations.

4. Results: Assessing the gaps

We first test for systematic differences in levels of survey and national accounts per capita means for income and consumption. We then assess systematic differences in growth rates.

Levels

To assess the systematic differences between national accounts and survey means, we use a simple measure for assessing the size and direction of the gaps. Following the approach of Ravallion (2003) and Deaton (2005), we first calculate a set of ratios of the survey to national accounts means. A ratio of 1 indicates that the measures are the same, a value of less than one suggests the survey measure is less than the national accounts measure. By subtracting one from the ratio, we get a measure of the proportional gap, *g*, between the HHS mean and the NAS mean. A negative (positive) gap suggests the survey mean is lower (higher) than the corresponding national accounts mean. For example, a ratio of 0.9, gives a gap

[12] Weights are re-calculated for each subsample. E.g. if a country has one consumption survey and one income survey, each survey is given 0.5 weight in the pooled sample of consumption and income surveys, while they each get a weight of 1 in the

measure of -0.1, suggesting that the survey mean is mean is 10 percent lower than the national accounts mean.

The distribution of the gap in our sample of matched survey and national account means is illustrated in Figure 2. Both the unweighted histogram and the country-weighted density function show that, on average, the gap is well below zero, suggesting that survey means on average are lower than national accounts means. The gap is clearly larger with regards to GDP than with HFCE, moreover consumption shows larger variation than income surveys. We estimate the average gap for various subsamples and calculate standard errors clustered at the country level. [13]

The gaps between national accounts and survey means are larger and more robust than previous studies have shown. Table 1 shows the value of the gaps for various sub-samples, by region and survey type. Overall, our estimates suggest that survey means are 23 percent lower than national account means for HFCE. There is no statistically significant difference in the gap observed for consumption surveys (25 percent) and income surveys (23 percent). The sample of more recent and comparable surveys from the Shared Prosperity database, suggests a gap of similar magnitude, with survey means being 24 percent (robust s.e. 2.3) lower than HFCE on average. Notably, these gaps are larger than the corresponding estimates from Ravallion (2003) and Deaton (2005). Deaton found a gap of 14 percent for consumption surveys and 9.6 percent for income surveys, while Ravallion estimated a gap of 17.4 percent for income surveys and 6.9 percent for consumption surveys, with the difference between income and consumption surveys being statistically significant. Karshenas (2003) finds a corresponding difference of only approximately 5 percent using a slightly different assessment method.[14] In addition to the point estimates, our larger sample compared with previous assessments also contribute to lower standard errors and establish the gaps more robustly.

We observe some distinct patterns in the gap by region and income group. Across all sub-samples by geographic and income grouping, the gap is negative and statistically significant. Across regions, Sub-Saharan Africa has the smallest average gap of 19 percent for HFCE and South Asia has the largest of 33 percent. There is a marked pattern of the ratio and overall level of economic development, with the gap being largest among middle income countries (27 percent for lower middle income countries and 33 percent for upper middle income countries). The gap is much smaller for low and high income countries, but still statistically significantly different from zero. The pattern across levels of development is shown

[13] Specifically, we estimate the mean gap for various sub-samples, by running an OLS regression of the gap on a constant: $g_i^n = \alpha + \varepsilon$ for various subsamples and weighting schemes, where α gives us the mean gap. We estimate this with clustered standard errors at the country level, allowing for errors to be correlated across observations within countries.

[14] Karshenas (2003) estimates the gap using mean difference in the logs of the measures, at -.05.

in Figure 3, indicating that, on average, the gap increases (gets more negative) as countries get richer but then is diminishing with income for high income countries. The line shows the results of locally weighted regressions of the latest ratio for each country showing a clear u-shape for both income and consumption surveys.

The pattern in the Figure 3 differs from the findings by Deaton (2005), Ravallion (2003) and Karshenas. Deaton and Karshenas both find a negative relationship between the gap and level of development over the entire range, with the gap increasing among the richest countries. In contrast, our sample shows that the gap is declining among high income countries. Moreover, we observe a very similar pattern for income surveys and consumption surveys, which Deaton and Ravallion did not. We believe this difference simply results from income surveys being more heavily concentrated in richer countries in the Deaton and Ravallion samples, where they observed a larger gap, while consumption surveys were concentrated in poorer countries, where the gap was smaller. This pattern was particularly strong in the case of Deaton who focused the analysis of rich countries on the UK and the US, two high income countries where the gap is particularly large also in our sample, but these appear somewhat unique cases among rich countries.

Another possible reason we observe a smaller and narrowing gap among richer countries than previous studies may be evolving practices and consolidation of data sources in these countries. Recently, it has been more common in rich countries to integrate administrative data, such as tax records, in survey data estimates. For example, the EU-SILC surveys, a major source for our income survey data for rich countries, explicitly allow countries to supplement income variables with administrative records, such as tax and payroll data (Jäntti et al 2013). This practice may better align estimates and contributed to a reduction in the gap between national accounts measures and survey measures. However, it is still interesting that consumption surveys, which, to our knowledge, do not typically rely on any such administrative data, show a similar narrowing of the gap among rich countries.

Last, in our assessment of differences in levels, we assess the gap between household surveys and GDP in Panel B of Table 1. Overall, surveys means are on average 49 percent lower than GDP across countries. For recent observations available in the shared prosperity database, we observe an even larger gap of 55 percent (s.e. 1.8). While this large discrepancy may not be surprising to national accounts and survey experts, it is worth noting that maybe the most commonly cited indicator of summarizing living standards in a country – GDP per capita – typically is twice the size of per capita income or consumption of that country. Interestingly, we do not see as strong of a narrowing of the gap among the richest countries for GDP, as with HFCE. Upper-middle income countries have the largest gap, on average but it is not much larger than that of high income countries in our sample.

Growth rates

We now turn to assessing how growth rates in national accounts differ from those of household surveys. To assess the relationship between growth rates in the per capita means of surveys and national accounts measures, we follow methods commonly deployed in evaluating forecasts of economic growth.[15] We derive a simple measure of discrepancy (or 'error' in the growth forecast literature) in growth rates, by subtracting the annualized growth rate in national accounts mean from the annualized growth rate in the surveys.[16] A negative value suggest the survey measure grow slower than national accounts measure.

The distribution of the gap is presented on Figure 4, showing a large variation in differences in growth rates between the two sources. To test the bias, or systematic differences, we estimate averages by subgroups of countries and datatypes.[17] If the gap is statistically significantly different from zero using robust standard errors clustered at country level, we conclude that there is a systematic difference.[18] We run this regression for various samples of spells and survey types and by region and income group. The main results are presented in Table 2. Overall, growth rates in surveys are, on average, lower than in national accounts. However, the difference is not statistically significant for the overall pooled sample. However, for the sample of low and middle income countries, the average gap is more than one percentage points (1.36 percent for LMICs and 1.17 for UMICs), and statistically different from 0 at the 95%. We also assess the gap in growth rates from the sample in the shared prosperity database, and find a negative point estimate of -1.75 percent, but insignificantly different from zero. Patterns are broadly similar for GDP.

The pattern of the larger gap in middle income countries, while in low and high income countries is consistent with the findings of levels. The gap in levels is clearly increasing across the middle income countries, suggesting that surveys must also grow slower than national accounts means over this segment. For high income countries, the point estimate suggests that growth rate of consumption in national accounts is higher than, which contributes to explaining the narrowing gap, but it is not large and statistically insignificant from zero. The pattern of gaps in growth rates across income groups, shown in

[15] The growth forecasting literature is concerned with assessing the precision and bias of forecasts in predicting actual growth rates. Similarly, we are interested in understanding the precision and bias of national accounts growth rates in predicting actual survey growth rates.

[17] We estimate subgroup averages by running a simple regression of the growth rate gap on a constant, analogous the preceding assessment of differences in levels

[18] This is consistent with methods used in the growth forecast evaluation literature, and the method proposed by Holden and Peel (1990) for testing bias in forecasts.

Figure 5, mirrors the pattern observed in Figure 3. The figure also illustrates the results for sub-sample by income and consumption surveys, showing small differences in the point estimates but larger confidence intervals for the income surveys. The 95 percent confidence intervals are based on robust standard errors clustered at the country level.

In addition to the direction of bias (or gaps), we are for many purposes interested in the precision (efficiency) of growth rates of national accounts as a predictor of survey growth rates, given that national accounts data is much more readily available and therefore frequently used to interpolate and extrapolated estimates of poverty to non-survey years. The mean absolute error and the root mean square errors presented in Table 2, gives an assessment of this the precision. Again, there is a clear economic gradient with the precision increasing with income, but it should be noted that high income countries typically experience lower growth rates and the gradient of relative errors may be smaller. The mean absolute error is large, more than 2 percentage points for all sub groups except for North America and High Income countries, highlighting the large average differences in growth rates observed between the data sources.

An alternative way to assess the degree of correspondence of the growth rates is to do a simple no-constant OLS regression of survey mean on national account mean, as done by Ravallion (2003). Table 3 presents such results. The resulting coefficient on the national accounts mean has typically been used in projecting distributions for World Bank poverty projections beyond the latest official reference year (see World Bank, 2015 for details). Note that regressions are weighted, giving each country a weight of 1, regardless of number of observations, not to give countries with many surveys overweight in the estimates. The results indicate a stronger correlation for richer countries than poorer ones. The Middle East and North Africa region stands out as a region with particularly poor correlation between national accounts and survey growth rates reflecting a . We find a much stronger relationship in Eastern Europe and Central Asia than Ravallion's (2003) assessment which mainly used data from the transition period.

The sample of comparable growth spells compiled in the most recent Shared Prosperity Database shows clearly how large the difference in annualized growth rates can be even for modest spells length. The gap between annualized real growth rate in survey mean and HFCE mean ranging from -3.9 percent to 10.7 percentage points, with a positive mean gap of 0.4 percent (std error 0.25) not statistically significantly different from zero at 95%-confidence level. Similarly, large gaps are found with GDP means in these spells. For the no-constant regression approach to assessing the gap, following Ravallion (2003), the regression coefficient is 0.72 for GDP and 0.63 for HFCE, with the national accounts growth explaining only 63 and 65 percent of variation in survey growth, respectively.

5. Implications of the gaps for measures poverty and inequality

The implications of the observed NAS-HHS gaps for measures of poverty and inequality depend on our assumptions about the origin of the gaps and consequently what adjustments are deemed appropriate. We analyze the implications under two scenarios for the origin of the gap and corresponding adjustments. The scenarios draw on two distinct approaches in the literature, corresponding to two different hypotheses for the origin of the observed gap. The adjustment can be broadly categorized into 'distribution neutral' and 'top income'-adjustments. We compare measures of poverty and inequality based on the adjusted data with data from surveys.

The first scenario effectively assumes that the NAS-HHS gap is the result of a 'distribution neutral' error in surveys, and that the appropriate adjustment for this error is to scale survey means to the national accounts means. This scenario is consistent with the literature which prefer using national accounts means in combination with survey measures of inequality, such as Bhalla (2002) and Pinkovskiy and Sala-i-Martin (2009, 2016). The approach taken by this literature assumes that surveys may not capture the full extent of consumption and that national accounts is a more complete and comparable measure of the true value of national household consumption. The mis-estimation of consumption surveys can result from various issues. First of all, there are elements of household consumption that often are not captured by household surveys, but that are included in national accounts, such as the value of public services and the value of owner-occupied housing. Second, there is a well-documented tendency for respondents of household surveys to underreport income and consumption, simply due things like recall decay and respondent fatigue. Third, respondents may also willingly underreport consumption and income if they are concerned, for concerns about privacy or tax liability. By using the observed distribution of consumption or income from surveys in combination with the national accounts mean, everyone's consumption or income is in effect scaled by a constant parameter equal to the NAS-HHS gap, an adjustment which is distribution neutral, keeping within-country inequality is unchanged.

Under the second scenario, it is assumed that the NAS-HHS gap is mainly a result of the richest households being inadequately captured in survey data. This scenario draws on the literature which suggests that surveys may do well at capturing incomes and consumption for most of the bottom of the distribution, but misses consumption by the top tail of the distribution. Lakner and Milanovic (2016) and Chandy and Seidel (2017a, 2017b) adjust the top of each survey distribution as a function of the NAS-HHS gap. Both fit a Pareto distribution to allocate the NAS-HHS gap to the richest household of the survey distribution. Their methods build on an approach suggested by Atkinson (2007) who uses a Pareto imputation to "elongate" the upper part of the distribution. Lakner and Milanovic add the full NAS-HHS gap to the top decile of the distribution in their data (though with an upper bound), while Chandy and

Seidel (2017b) add a top segment to the Lorenz curve from the survey distribution, with income corresponding to *half* the NAS-HHS gap, fitted with the pareto distribution from the top survey decile. In the latter method, the elongation of the distribution is thus a function of both the size of the gap and the observed inequality of the top survey decile.

In summary, our baseline scenario is the distributions from household surveys, as the standard data source for poverty and inequality measures, largely based on the data currently available in PovcalNet. Our 'distribution-neutral' adjustment scenario scales the survey means with national accounts means. Last, our 'top income' adjustment, adds half of the gap between HFCE and the survey mean to the top of the survey distribution following the method proposed by Chandy and Seidel (2017b). While all these distributions have errors, and we do not make a judgment of which better fits the true distribution of consumption. The scenarios simply help illustrate the implications for poverty and inequality measures, depending on how we understand the origin of the NAS-HHS gap and corresponding adjustments.

Poverty

Several researchers have suggested that global poverty can measured more comparably and with higher geographic coverage and temporal frequency, by directly using national accounts means, sometimes combined distributional measures from surveys. Bhalla (2002) uses distributions from surveys but means from the NAS, effectively scaling the survey mean to the national accounts mean of HFCE. Bourguignon and Morrisson (2002), Sala-i-Martin (2006), Sala-i-Martin and Pinkovskiy take a similar approach but use GDP means, rather than HFCE means. Ravallion and Chen (2010), in addition to surveys, propose a combination measure of national accounts and surveys. The approach of using national accounts data has been boosted by Pinkovskiy and Sala-i-Martin's (2016) finding that GDP per capita is more robustly correlated with levels and changes in nighttime lights and other indicators of standard of livings and non-monetary outcomes such as schooling and health status than surveys.

Since per capita income and consumption in national accounts are higher than in surveys, using national accounts data to measure poverty leads to much lower estimates of poverty, when holding the poverty line and the distribution the same. [19] Compared to our survey estimates, on average, country-level poverty estimates at the international line of $1.90 is 40 percent lower using HFCE, and 73 percent lower

[19] We simply scale the survey mean to the national accounts mean and apply the international poverty line of $1.90 per day.

using GDP.[20] Figure 6, Panel A, compares survey estimates of poverty at the $1.90 line to corresponding measures for national accounts. Only in 15 percent of observations is poverty higher using HFCE rather than surveys, and in 2 percent in the case of GDP. This illustrates the very different level of poverty resulting from national accounts estimates of poverty using the $1.90 line. Naturally, also when aggregated to global estimates, poverty by this measure is much lower as seen in Panel A of Figure 6. Our estimates based on National Accounts suggest that the World Bank's 3 percent poverty target for 2030 was reached by year 2013 in the case of GDP, and was at 4.5 percent when using HFCE.[21] But it is important to recall that with NAS measures being much higher than survey measures at any point in time, any starting point for these poverty trajectories in past reference years are also much lower.

Importantly, the $1.90 poverty line may not be the appropriate threshold for international poverty measurement when the distribution is adjusted to the level of national accounts means. Since the $1.90 line is itself estimated from national poverty lines originating from household surveys (see Ferreira et al, 2016; Jolliffe and Prydz, 2016 for details), the poverty line will likely be underestimated, and not suitable for use with living standards measured using national accounts. Applied to national accounts means, the international poverty line of $1.90/day no longer reflects the national assessment of poverty in the world's poorest countries due to the higher standard of living suggested by NAS data.[22] A rough and simple adjustment of the international poverty line for use with national accounts, acknowledging the systematic gap, would be to scale the poverty line by the average gaps. This would give a poverty line of $2.47 for use with HFCE and $3.73 for use with GDP.[23] Figure 6, Panel B illustrates the poverty estimate for each country using these poverty lines. Although poverty estimates with the adjusted lines are more closely align to survey estimates than in Panel A, there is still substantial variation, reflecting the variation in the gap across countries. This highlights that even if the poverty line is adjusted for use with national accounts data, the bias can be reduced, but the overall variation does not.

A different way of adjusting the IPL for poverty measurement based on national accounts means, is to set a country-specific IPL is based on the country-specific survey-national accounts gap observed in the data. Alternatively (and equivalently), one could use the national accounts series, scaled down by the gap, and use the $1.90 line. Such an approach could be justified from a perspective that national accounts

[20] Calculated for surveys where the household survey has more than 3% poverty. At low poverty estimates small changes can lead to very large changes in percentage terms.
[21] These measures are broadly consistent with Sala-i-Martin and Pinkovskiy (2016) who calculated measures up to 2010 using 2005 PPPs and the $1.25-line.
[22] Ravallion and Chen (2010) note that typical methods for setting national poverty lines will underestimate the poverty line if non-food spending is underestimated in surveys, and thus any correction for underestimation of non-food spending would also lead to higher poverty lines.
[23] These are simply conversions using the overall ratio observed in Table 1. For HFCE, the gap is -0.23, so adjusting the poverty line of 1.90, would give 1/(1-.23)*1.9=2.47. For GDP the gap is .49, resulting in 1/(1-.49)*1.9 = 3.73.

may be a more comparable measure of changes in living standards over time, but that surveys, on average, do a better job at measuring inequality and the level of poverty. Under such an approach, large changes to survey methodology that affect the mean (holding inequality constant), would have a much lower effect on the evolution of poverty. Such an adjustment can be done for a particular year, or by taking the average ratio for a country over a longer time period. We estimate such country specific IPLs for use with national account for 1990, ensuring that poverty estimates from national accounts and surveys are aligned in that year.[24] The results for global poverty are shown in Panel C of Figure 6. While the poverty estimates are equal (by design) in the reference year, there is considerable variation over time, and, in global measures the rate of decline is much larger, due to the faster growth rates of national accounts.

Across all the methods which involves substituting survey means with national accounts means distribution neutral, poverty is lower and falling faster than that of traditional survey measures. The more rapid decline seen in the poverty measures using national accounts points to a concern with current use of national accounts growth rates in extrapolating household survey estimates for years with missing surveys. Even if household survey means are used for poverty estimation for survey years, national accounts growth rates are used to interpolate such estimates to non-survey years and for nowcasts and projections of poverty in the future. Current methods use national accounts growth rates to align poverty estimates to non-survey years for global aggregation. Removing the bias implicit in this method, suggests a slower global decline in poverty than the Banks official poverty numbers. Because India, Indonesia and China, countries which historically have been the home of a large share of the global poor, typically have household surveys for most reference years that the World Bank report poverty for, the effect on global numbers is not that large.[25] However, for global extrapolations beyond the World Bank's latest reference year, for which there are very few surveys, this is a much larger concern. Such extrapolations commonly use an adjustment factor to adjust for the discrepancy in growth rates. But even if the systematic bias is removed and overall error is reduced by applying the adjustment factors, the precision of the method is still poor.

The various adjustments and implications discussed above assume a distribution neutral interpretation of the HHS-NAS gap. We now turn to the 'top income' adjustments which are more appropriate if the source of the gap is originating from top incomes being mis-measured in surveys due to biased response rates or underreporting by the richest households. Since this adjustment mainly effects the very top of the distribution, that in all countries is above the $1.90 threshold, this has little effect on

[24] The population weighted lines for 1990 are 4.98 for GDP and 2.45 HFCE.
[25] In recent years, no survey has been available for India.

poverty measures. However, since we add a segment to the survey distribution, poverty measures fall proportionally to the number of observations added to the top of the distribution. If poverty is 30 percent and we expand the number of distributions by 10 percent missing top incomes, the poverty rate would only fall to 27.2 percent. Since the method only adds a rather small segment the effect on aggregate poverty measures is small. However, other measures of the relationship between growth and poverty reduction will change.

Inequality and inclusive growth

While the implications of the top-adjustment have little effect on poverty measures, the implications for our understanding of levels and trends in inequality are larger. Figure 7 compares the Gini coefficient from the unadjusted distributions with inequality from the distributions with top-adjustments, using the method proposed by Chandy and Siedel (2014).[26] At the national level, observed Gini coefficients are on average 25% higher at the national level when using the top-income adjusted distributions, as shown in Panel A of Figure 7. Other measures of levels of inequality, such as the 90/10 ratio and the so-called Palma ratio would also be drastically affected by such adjustment. Measuring inequality globally using the top-adjusted distributions also result in a much higher level of global inequality, with the top income adjustment increasing the global Gini in 2013 from 62.6 to 67.3. Panel B of Figure 7 shows trends that are broadly similar, showing a robust decline in global inequality since about year 2000. The levels and patterns are broadly in line with those of Milanovic and Lakner (2016) who use a different form of adjustment and aggregation, providing estimates until 2008. Using the unadjusted distributions with national accounts means, global inequality appears lower and declining faster than with survey measured, driven by the tendency of national accounts measures to be larger and grow faster than surveys in middle income countries.

Generally, the larger NAS-HHS gap among middle income countries leads to relatively larger adjustments to inequality measures in these countries. These systematic differences along the developmental gradient of countries also lead to changing the relationship between economic development and levels of inequality. Evidence of the cross-sectional Kuznet curve – the hypothesis that income inequality first increases and then declines with development - has recently been questioned. Palma (2011) suggested that the "the 'upwards' side of the 'Inverted-U' between inequality and income per capita has evaporated" (Palma, 2011, p. 87). This is indeed true for the unadjusted Gini coefficients also in our sample. However, Gini coefficients from the top-income adjusted data, suggest that there is

[26] We simply adjust their code, which they made publicly available, with our data.

indeed an upwards sloping segment and firmly re-establish a Kuznet-like relationship between economic development and observe inequality, as seen in Figure 8. A quadric form regression of GDP explains more than double the variation of the adjusted Ginis than the unadjusted measures. However, we do not find a systematic patterns of within-country Kuznet relationship.

Measures of the degree to which economic growth is inclusive are also affected by assumption that at least some of the gap originates from missing top incomes in survey data. In measuring the degree to which growth is inclusive, the World Bank goals and the UN Sustainable Development Goals closely monitor growth in income and consumption of the bottom 40 percent in each country, relative to the growth for the overall population. For a recent set of comparable spells for 2008-13, 49 out of 83 countries (60 percent) reported a positive 'shared prosperity premium': income and consumption growth among the bottom 40 exceeded that of the mean of the overall population (World Bank, 2017; World Bank, 2016). Panel A of Figure 9 shows the growth in the bottom 40 percent on the vertical access, and in the mean on the horizontal axis, with the majority of observations about the 45-degree line. However, these estimates rely solely on survey data. Using the top-income adjusted distributions give a very different picture of the inclusiveness of growth over the periods. Among the 79 countries for which we are able to replicate shared prosperity measures for this period, only 35 countries (44 percent) had a positive shared prosperity premium. Panel B of Figure 9 illustrates the relationship and differences between growth in the bottom 40% and growth in the mean with the adjusted distributions. The population weighted average annualized shared prosperity premium for the period falls by 1.3 percentage points, from a positive 0.5 percent growth to -0.8 percent, inverting the positive trends in inequality measured by the official spells. This shows again how common measures of development are distorted if one assumes that the HHS-NAS gap originates from a lack of capture of consumption and income of the richest households.

6. Conclusion

This paper has compiled a large new dataset for assessing the correspondence between nation per capita monetary living standards measured in national accounts and household surveys. The data shows that the gap between the data sources are larger than measured in previous assessments. The gap does not seem to be an effect of income versus consumption surveys, as suggested by Deaton (2005) and Ravallion (2003), but rather aligned with the level of development, with the gap being largest for middle income countries, both in terms of levels and growth rates. The gaps, and corresponding implications for poverty and inequality measures, are largest for middle income countries which experience periods of rapid growth.

With more than 63% of the extreme global poor currently living in middle income countries, the implications for measurement of the evolution of global poverty and inequality aggregates are large.

We have argued that the gap, at least to some extent, originate from the lack of surveys to capture the full level of consumption and income for all households. Our paper has illustrated potential implications for common poverty and inequality measures, under different scenarios for the origin of the gap and corresponding appropriate adjustments. Under the distribution-neutral gap scenario, which would justify substituting survey means with national accounts mean, global poverty would be much lower using the $1.90 line, and the sustainable development goal of 'ending' extreme poverty would already be close to being met, or at least easily within reach. This is in line with findings of Sala-i-Martin and Pinkovskiy (2016) which uses this method to measure global poverty. However, we argue that the $1.90 line should be adjusted for use with national accounts. Using international lines adjusted for systematic differences levels of surveys and national accounts, poverty measures would be more in line with survey measures, but still cause relatively large changes to country level measures. Regardless of poverty line used with national accounts, rate of poverty reduction is greater than that measured in surveys because national accounts growth rates are typically higher. This latter observation, illustrates the perils of using growth rates from national accounts to extrapolate poverty, which is currently the practice in current global poverty measures and likely exaggerates the decline in poverty since 1990.

Under the scenario which assumes that the NAS-HHS gap is due to surveys not fully capturing consumption of incomes of the richest households in societies, and therefore justifying adjustments of top segment of the survey distributions, the implications for poverty measures are small, but typical inequality measures change drastically. Adjusting the survey data for missing top incomes with the observed gap in NAS-HHS, increased national and global inequality considerably. The global Gini using the adjusted data is much higher. Moreover, the hypothesis that as economies develop, inequality first increase and then decrease, also known as the Kuznet's curve, is much more strongly supported in our cross sectional sample of 'top income'-adjusted distributions.

Because of the large gaps between survey and national accounts data, and particularly the large variation in the gaps across countries and time, the prospect of filling gaps in poverty data with estimates of poverty imputed from national accounts growth rate or aggregates, is limited. The error (difference) of estimates based on national accounts data, in comparison to survey measures, are very large. As long as household surveys appear to be the preferred method of measuring poverty and inequality, national accounts data offer little hope for filling data gaps. Ultimately, more frequent, properly sampled household surveys, potentially with integration of tax records and administrative data as countries get richer, seems to be the best method for improving our understanding of poverty and inequality.

References

Ahluwalia, M.S., Carter, N.G. and Chenery, H.B., 1979. Growth and poverty in developing countries. *Journal of Development Economics*, 6(3), pp.299-341.

Anand, S. and Segal, P., 2008. What do we know about global income inequality?. *Journal of Economic Literature*, *46*(1), pp.57-94.

Atkinson, A.B., 2007. Measuring top incomes: methodological issues. *Atkinson, AB and T. Piketty: Top Incomes over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries*, pp.18-42.

Beegle, K., De Weerdt, J., Friedman, J. and Gibson, J., 2012. Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics*, *98*(1), pp.3-18.

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

Castaneda Aguilar, Raul Andres; Lakner, Christoph; Prydz, Espen Beer; Soler Lopez, Jorge Andres; Wu, Ruoxuan; Zhao, Qinghua. 2019. *Estimating Global Poverty in Stata: The Povcalnet Command.* Global Poverty Monitoring Technical Note; 9. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/836101568994246528/Estimating-Global-Poverty-in-Stata-The-Povcalnet-Command

Chandy, Laurence and Brina Seidel. 2017a. *How much do we really know about inequality within countries around the world? Adjusting Gini coefficients for missing top incomes*. Brookings Institution. https://www.brookings.edu/opinions/how-much-do-we-really-know-about-inequality-within-countries-around-the-world/

Chandy, Laurence and Brina Seidel. 2017b. *Allocating Income from Survey-National Accounts Gap to Top Incomes According to Pareto Distribution* https://www.brookings.edu/wp-content/uploads/2017/02/technical-appendix-02-13-17.pdf

Deaton, A., 2005. Measuring poverty in a growing world (or measuring growth in a poor world). *The Review of Economics and Statistics*, 87(1), pp.1-19.

Deaton, A. and Kozel, V., 2005. Data and dogma: the great Indian poverty debate. The World Bank Research Observer, 20(2), pp.177-199.

Deaton, A. and Dreze, J., 2002. Poverty and inequality in India: a re-examination. *Economic and political weekly*, pp.3729-3748. Vol. 37, No. 36 (Sep. 7-13, 2002), pp. 3729-3748

Dhongde, S. and Minoiu, C., 2013. Global poverty estimates: A sensitivity analysis. *World Development*, *44*, pp.1-13.

Eurostat. 2016. Household Budget Surveys Database. Accessed: April 2017.

Eurostat. 2015. Household characteristics by type of household.

Farfán, G., Genoni, M. E., & Vakis, R. (2017). You are what (and where) you eat: Capturing food away from home in welfare measures. *Food Policy*, *72*, 146-156.

Jäntti, M., Törmälehto, V. M., & Marlier, E. 2013. *The use of registers in the context of EU-SILC: challenges and opportunities: 2013 Edition*. Publications Office.

Jolliffe, D., 2001. Measuring absolute and relative poverty: the sensitivity of estimated household consumption to survey design. *Journal of Economic and Social Measurement*, *27*(1, 2), pp.1-23.

Jolliffe, D., Lanjouw, P., Chen, S., Kraay, A., Meyer, C., Negre, M., Prydz, E., Vakis, R. and Wethli, K., 2014. *A measured approach to ending poverty and boosting shared prosperity: concepts, data, and the twin goals.* Policy Research Report. Washington, DC: World Bank.

Karshenas, M., 2001, 'Measurement and Nature of Absolute Poverty in Least Developed Countries'. *Economic Working Paper Series* no 129. London: University of London, School of Oriental and African Studies.

Karshenas, Massoud, 2003, "Global Poverty: National Accounts Based versus Survey Based Estimates," Development and Change 34(4): 683-712.

Korinek, A., Mistiaen, J.A. and Ravallion, M., 2006. Survey nonresponse and the distribution of income. *The Journal of Economic Inequality*, *4*(1), pp.33-55.

Kouame , C. Y., Kalimili, B B N, Pirlea, Florina. 2019. *Many African economies are larger than previously estimated.* World Development Indicators. http://datatopics.worldbank.org/world-development-indicators/stories/many-economies-in-ssa-larger-than-previously-thought.html (Accessed 14 October)

Kuznets, S. 1955. Economic Growth and Income Inequality. *American Economic Review* 45 (March): 1–28.

Nolan, B., Roser, M., and Thewissen, S. (2019). GDP per capita versus median household income: What gives rise to the divergence over time and how does this vary across OECD countries?. *Review of Income and Wealth*, *65*(3), 465-494.

Pinkovskiy, M. and Sala-i-Martin, X., 2014. Africa is on time. *Journal of Economic Growth*, *19*(3), pp.311-338.

Pinkovskiy, M. and Sala-i-Martin, X., 2016. Lights, Camera… Income! Illuminating the national accounts-household surveys debate. *The Quarterly Journal of Economics*, *131*(2), pp.579-631.

Ravallion, M., 2003. Measuring aggregate welfare in developing countries: How well do national accounts and surveys agree?. *The Review of Economics and Statistics*, *85*(3), pp.645-652.

Rothbaum, J. 2015. Comparing Income Aggregates: How do the CPS and ACS Match the National Income and Product Accounts, 2007-2012, *U.S. Census Bureau SEHSD Working Paper,* No. 2015-01.

Sala-i-Martin, X., 2006. The world distribution of income: falling poverty and… convergence, period. *The Quarterly Journal of Economics*, *121*(2), pp.351-397.

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N. and Dabalen, A., 2015. Data deprivation: another deprivation to end. *Policy Research Working Paper Series*; no. No. 7252. Washington, D.C. : World Bank Group.

Sundaram, K. and Tendulkar, S.D., 2003. Poverty has declined in the 1990s: A resolution of comparability problems in NSS consumer expenditure data. *Economic and Political Weekly*, pp.327-337.

Subramanian, S. and Jayaraj, D., 2015. Growth and Inequality in the Distribution of India's Consumption Expenditure: 1983 to 2009-10. *Economic and Political Weekly*. Vol. 50, Issue No. 32, 08 Aug, 2015

Stiglitz, Joseph E. (1996). "Some Lessons From The East Asian Miracle".*The World Bank Research Observer*. 11 (2): 151–177. doi:10.1093/wbro/11.2.151.

World Bank Group. 2015. A Measured Approach to Ending Poverty and Boosting Shared Prosperity : Concepts, Data, and the Twin Goals. Policy Research Report;. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/20384

World Bank. 2016. *Global Database of Shared Prosperity*
http://www.worldbank.org/en/topic/poverty/brief/global-database-of-shared-prosperity

World Bank. 2017. *Poverty and Shared Prosperity 2016: Taking on Inequality*, Washington, DC: World Bank.

.

Tables

Table 1: Gaps in survey mean and national accounts means, across regions and income groups.

Panel A. Gap between survey mean and HFCE mean

	All				Consumption				Income			
Group	**N**	**Gap**	**S.E.**		**N**	**Gap**	**S.E.**		**N**	**Gap**	**S.E.**	
All	1139	-0.23	(0.02)	***	588	-0.25	(0.02)	***	551	-0.23	(0.02)	***
Low income	153	-0.17	(0.03)	***	139	-0.17	(0.04)	***	14	-0.12	(0.03)	**
Lower Middle Income	340	-0.27	(0.03)	***	203	-0.31	(0.03)	***	137	-0.19	(0.05)	***
Upper Middle Income	326	-0.33	(0.03)	***	151	-0.39	(0.03)	***	175	-0.29	(0.04)	***
High Income	359	-0.22	(0.02)	***	122	-0.23	(0.03)	***	237	-0.21	(0.02)	***
East Asia & Pacific	94	-0.25	(0.07)	***	82	-0.27	(0.08)	**	12	-0.22	(0.11)	
Europe & Central Asia	575	-0.26	(0.02)	***	324	-0.28	(0.03)	***	251	-0.22	(0.03)	***
Latin America & Caribbean	289	-0.26	(0.04)	***	22	-0.46	(0.06)	***	267	-0.24	(0.04)	***
Middle East & North Africa	43	-0.26	(0.09)	**	38	-0.25	(0.09)	**	5	-0.29	(0.00)	***
North America	48	-0.31	(0.16)		32	-0.46	(0.13)		16	-0.15	(0.12)	
South Asia	35	-0.33	(0.09)	**	35	-0.33	(0.09)	**	..	..	..	
Sub-Saharan Africa	91	-0.19	(0.05)	***	91	-0.19	(0.05)	***	..	..	..	

Note: Robust standard errors clustered at the country level (based on robust sandwich estimator). Statistical significance from zero, denoted by *p < = .1, **p < = .05, ***p < = .01.

Panel B. Gap between survey mean and GDP mean

	All				Consumption				Income			
Group	**N**	**Gap**	**S.E.**		**N**	**Gap**	**S.E.**		**N**	**Gap**	**S.E.**	
All	1210	-0.49	(0.02)	***	658	-0.49	(0.02)	***	552	-0.55	(0.02)	***
Low income	188	-0.39	(0.03)	***	175	-0.39	(0.03)	***	13	-0.41	(0.09)	**
Lower Middle Income	369	-0.52	(0.02)	***	232	-0.55	(0.03)	***	137	-0.50	(0.03)	***
Upper Middle Income	335	-0.60	(0.02)	***	161	-0.63	(0.02)	***	174	-0.59	(0.02)	***
High Income	363	-0.57	(0.01)	***	126	-0.58	(0.01)	***	237	-0.57	(0.01)	***
East Asia & Pacific	105	-0.49	(0.05)	***	93	-0.48	(0.05)	***	12	-0.57	(0.01)	***
Europe & Central Asia	595	-0.57	(0.02)	***	345	-0.57	(0.02)	***	250	-0.59	(0.01)	***
Latin America & Caribbean	297	-0.48	(0.03)	***	29	-0.58	(0.05)	***	268	-0.47	(0.04)	***
Middle East & North Africa	49	-0.51	(0.08)	***	43	-0.50	(0.08)	***	6	-0.58	(0.00)	***
North America	49	-0.56	(0.04)	***	33	-0.60	(0.04)	**	16	-0.50	(0.01)	***
South Asia	37	-0.56	(0.06)	***	37	-0.56	(0.06)	***	..	..	..	
Sub-Saharan Africa	120	-0.39	(0.04)	***	120	-0.39	(0.04)	***	..	..	..	

Note: Robust standard errors clustered at the country level (based on robust sandwich estimator). Statistical significance from zero, denoted by *p < = .1, **p < = .05, ***p < = .01.

Table 2: Growth in surveys vs national accounts – bias and errors

Panel A: Gap in growth rates HHS vs HFCE

	Spells	Gap (bias)			Precision	
Grouping	**N**	**α**	**(s.e.)**	**sig**	**RMSE**	**AE**
All	6341	-0.10	(0.21)		3.63	2.58
Low income	387	-0.18	(0.62)		4.89	3.73
Lower Middle Income	1853	-1.36	(0.44)	***	4.73	3.42
Upper Middle Income	2400	-1.17	(0.36)	***	3.89	2.88
High Income	1791	0.15	(0.21)		2.64	1.81
East Asia & Pacific	507	0.05	(0.58)		3.53	2.65
Europe & Central Asia	2920	-0.44	(0.26)	*	3.75	2.59
Latin America & Caribbean	2252	-0.50	(0.39)		3.17	2.33
Middle East & North Africa	88	-0.11	(0.47)		2.96	2.17
North America	395	-0.45	(0.68)		2.28	1.66
South Asia	93	0.53	(0.87)		2.87	2.22
Sub-Saharan Africa	170	0.71	(0.59)		4.14	3.01

Note: Robust standard errors clustered at the country level (based on robust sandwich estimator). Statistical significance from zero, denoted by *p < = .1, **p < = .05, ***p < = .01.

Panel B: Gap in growth rates HHS vs GDP

	Spells	Gap (bias)			Precision	
Grouping	**N**	**α**	**(s.e.)**	**sig**	**RMSE**	**AE**
All	6700	-0.27	(0.19)		3.63	2.62
Low income	486	-0.52	(0.5)		4.68	3.52
Lower Middle Income	2046	-1.25	(0.42)	***	5.03	3.63
Upper Middle Income	2483	-0.69	(0.3)	**	3.67	2.63
High Income	1787	-0.14	(0.26)		2.96	2.06
East Asia & Pacific	558	-0.24	(0.38)		2.82	2.04
Europe & Central Asia	3102	-0.38	(0.28)		3.99	2.89
Latin America & Caribbean	2307	0.00	(0.4)		3.37	2.37
Middle East & North Africa	113	-1.63	(0.55)	***	3.50	2.49
North America	408	-0.57	(0.28)	**	1.98	1.46
South Asia	110	-0.27	(0.65)		2.42	2.03
Sub-Saharan Africa	264	0.15	(0.5)		3.83	2.85

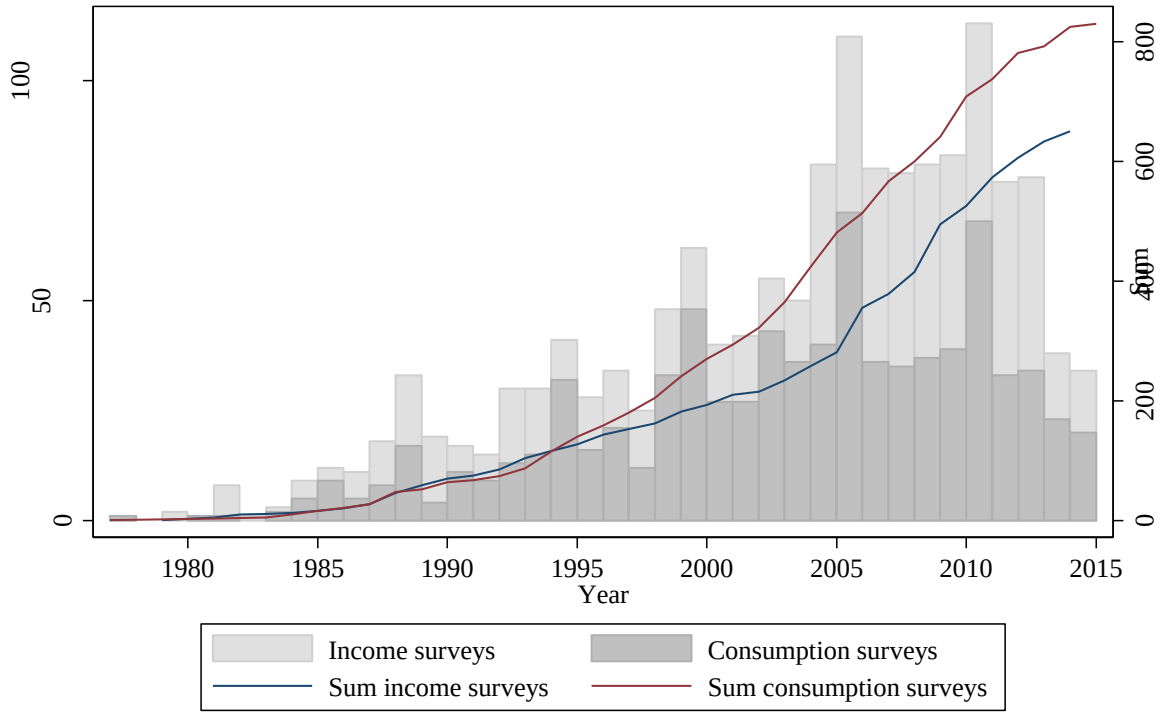
Note: Robust standard errors clustered at the country level (based on robust sandwich estimator). Statistical significance from zero, denoted by *p < = .1, **p < = .05, ***p < = .01.

Table 3: Growth in surveys vs national accounts – no-constant regression results

	HFCE				GDP			
	Coeff.	**Std. Err**	**N**	**R^2**	**Coeff.**	**Std. Err**	**N**	**R^2**
East Asia & Pacific	0.785	0.082	507	0.623	0.850	0.090	558	0.620
Europe & Central Asia	0.773	0.053	2920	0.509	0.796	0.077	3102	0.419
Latin America & Caribbean	0.829	0.101	2252	0.419	0.994	0.124	2307	0.438
Middle East & North Africa	0.544	0.100	88	0.278	0.365	0.191	113	0.106
North America	0.81	0.231	395	0.467	0.720	0.133	408	0.482
South Asia	0.912	0.163	93	0.569	0.897	0.191	110	0.717
Sub-Saharan Africa	0.708	0.180	170	0.245	0.748	0.137	264	0.246
Low income	0.602	0.107	387	0.283	0.651	0.130	486	0.235
Lower Middle Income	0.651	0.101	1853	0.370	0.648	0.117	2046	0.259
Upper Middle Income	0.702	0.054	2400	0.572	0.804	0.058	2483	0.523
High Income	0.892	0.052	1791	0.522	0.829	0.094	1787	0.405
All	0.758	0.043	6341	0.441	0.799	0.051	6700	0.392

Figures

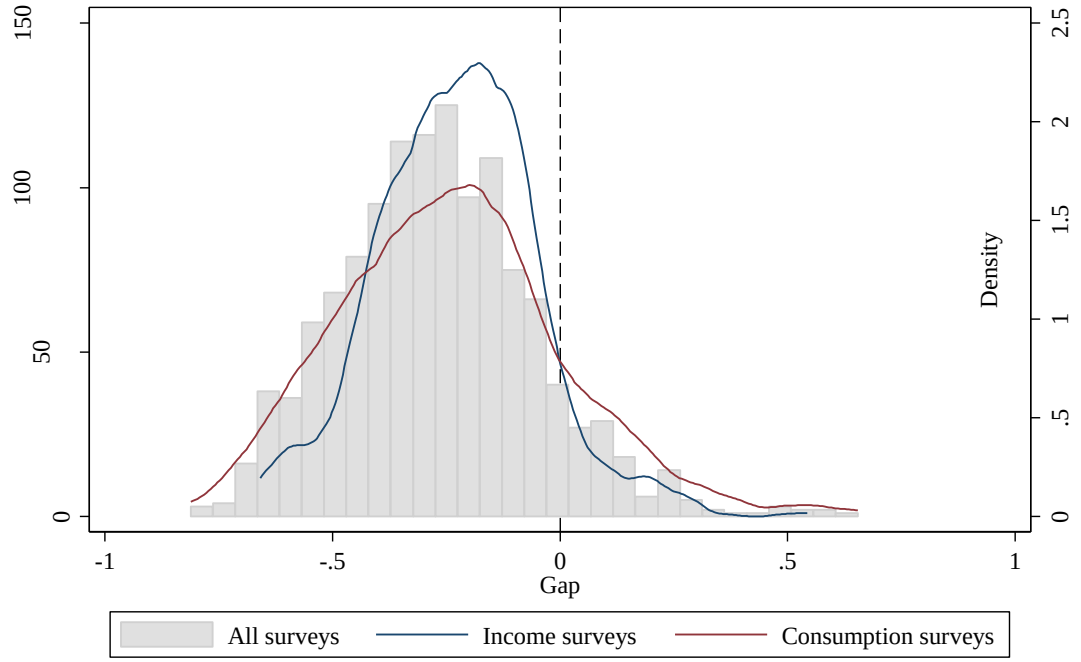
Figure 1: Available consumption and income surveys over time



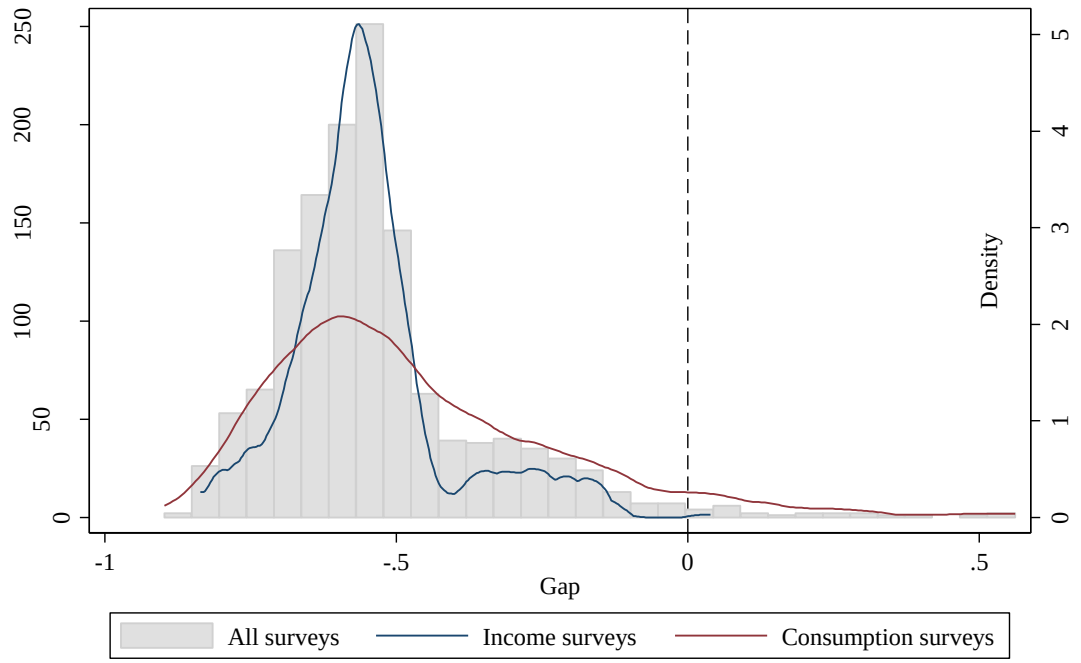
Note: The bars show the number of survey for each calendar year, separated by income and consumption surveys, measured on the left-hand axes. The lines show the running sum of surveys over time, measured on the right hand axes. Surveys are allocated to the year in which they started.

Figure 2: Distribution of gap of survey to national accounts consumption

Panel A: Survey mean vs HFCE mean

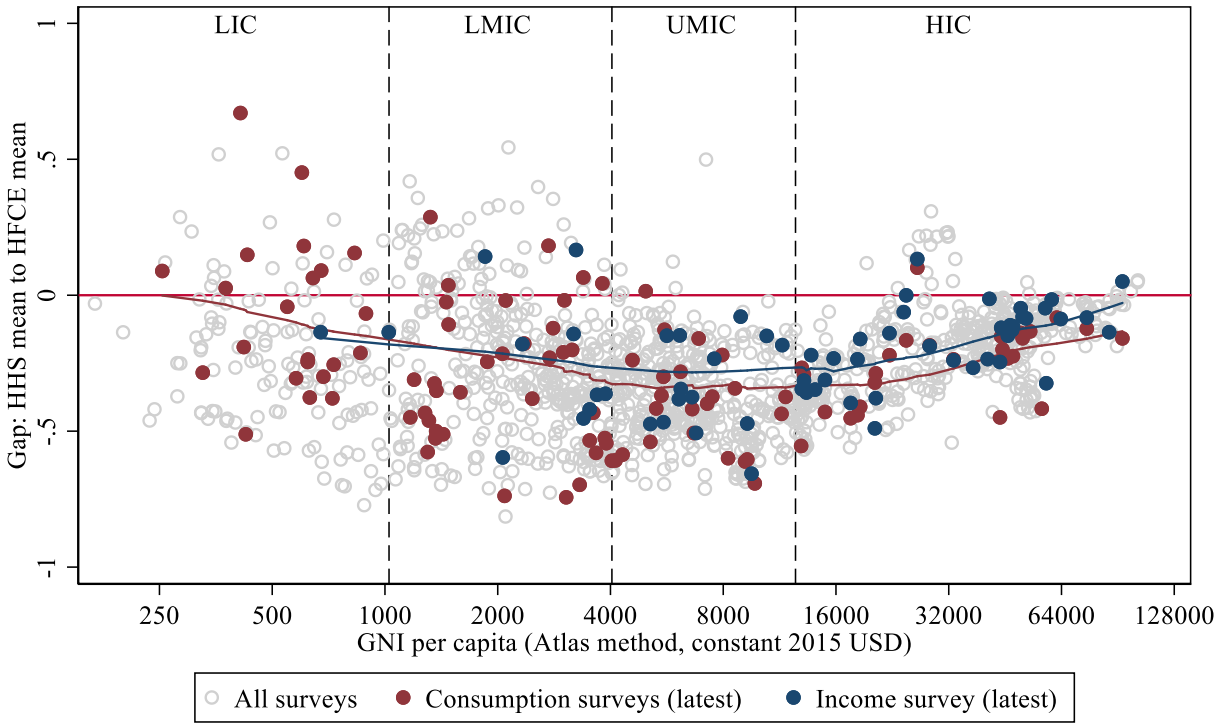


Panel B: Survey mean vs GDP mean



Note: The histograms are unweighted; the density functions are weighted so as to give each country equal weight.

Figure 3: Survey-national accounts gap and level of development



Note: The lowess lines (locally weighted scatterplot smoothing) are based on a non-parametric regression with bandwith of 0.8. The vertical bands demarcate the cutoffs for the World Bank's income classifications, expressed in 2015 USD, deflated using the Atlas method based on the SDR deflator.

Figure 4: Gap in growth rates in national accounts vs. household surveys

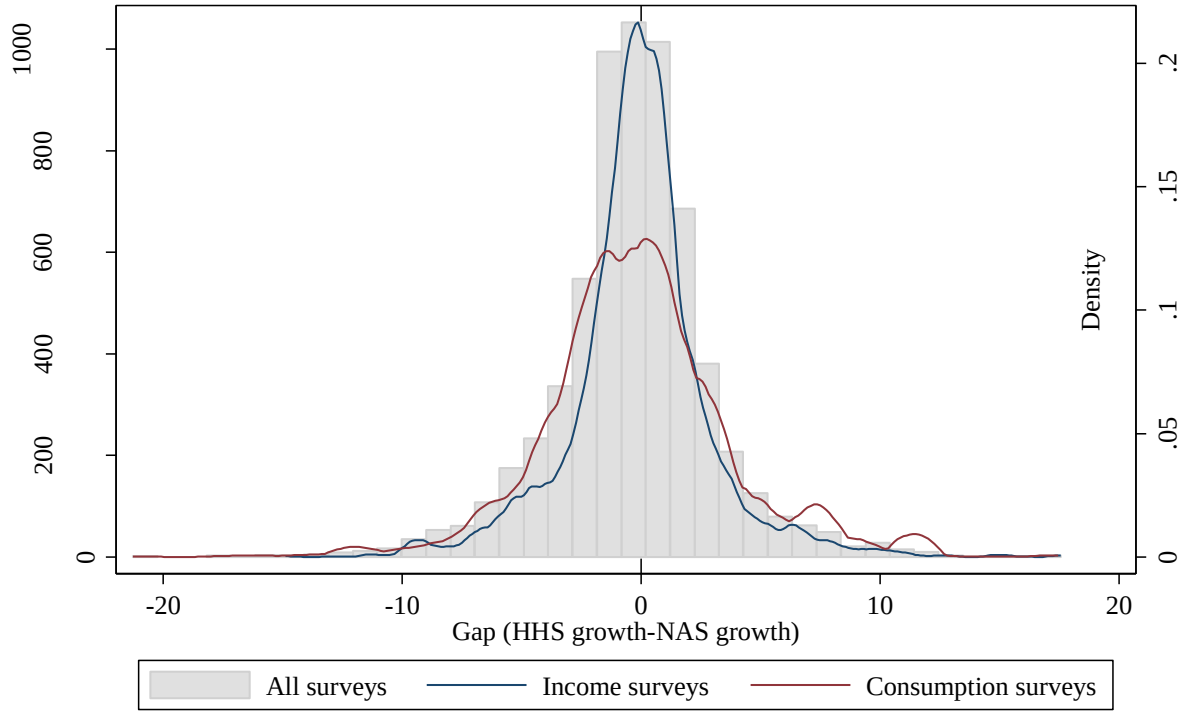
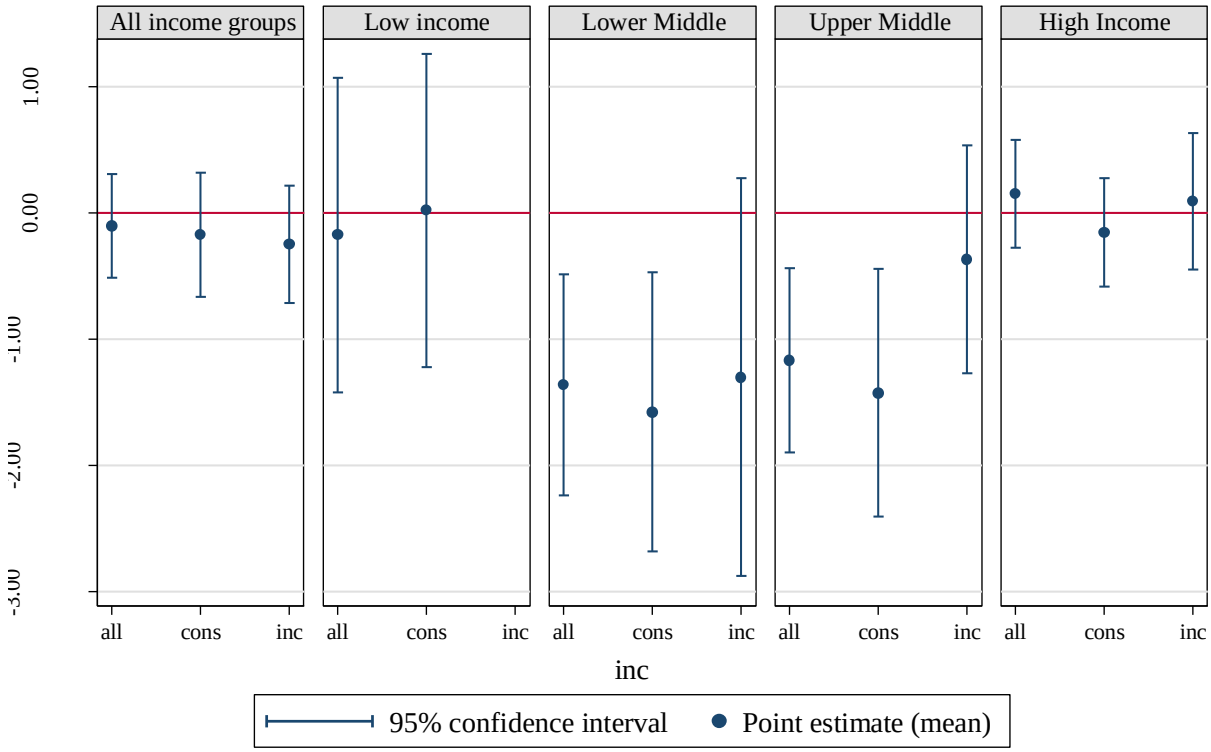


Figure 5: Difference in growth rates by income groups and by type of surveys.

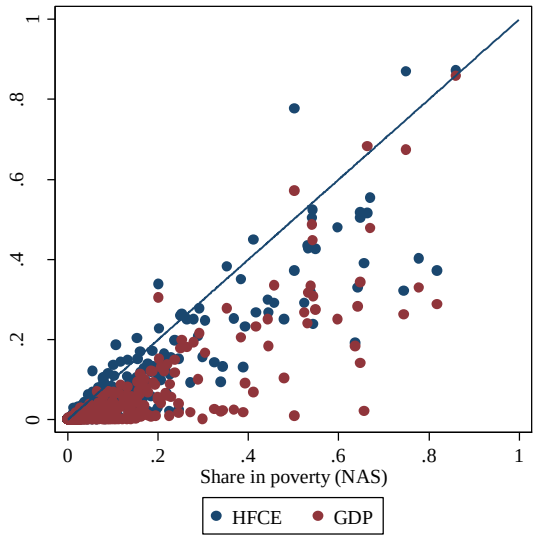


Note: The chart shows the result of the gap in growth rates, using the full country-weighted sample of growth spells. The markers show the point estimate (or the mean). Confidence intervals are based on robust standard errors, clustered at the country level.

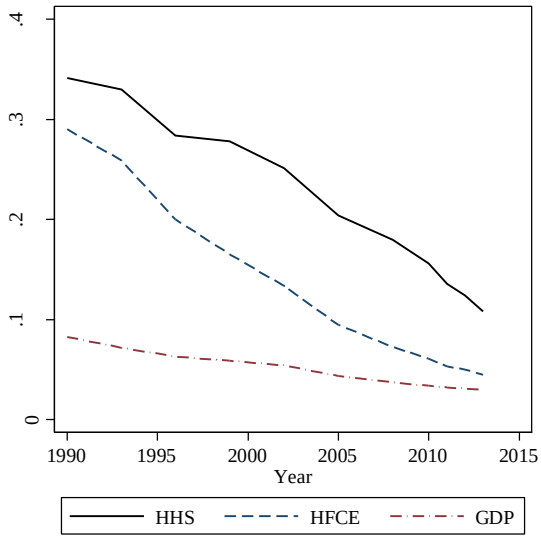
Figure 6: Poverty measures from National Accounts vs Household Surveys

A. $1.90 Poverty Line

National

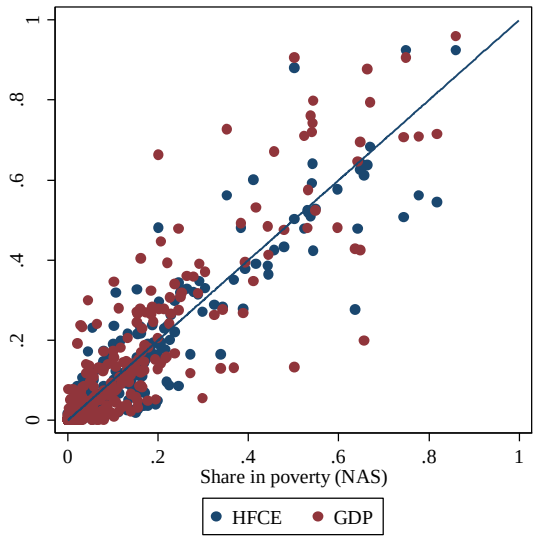


Global

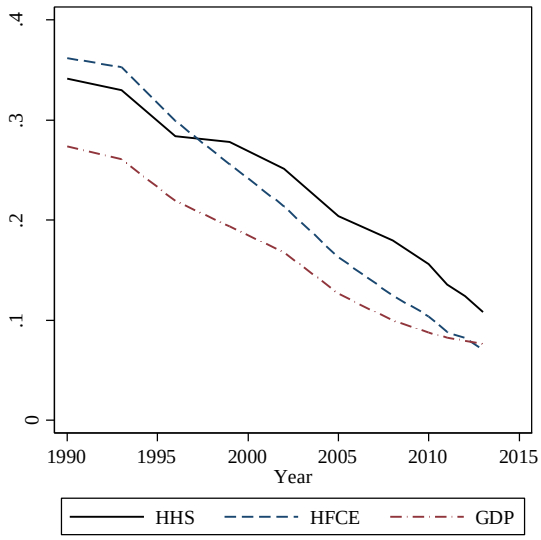


B. Uniformly adjusted Poverty Line (HFCE=$2.46; GDP=$3.73)

National

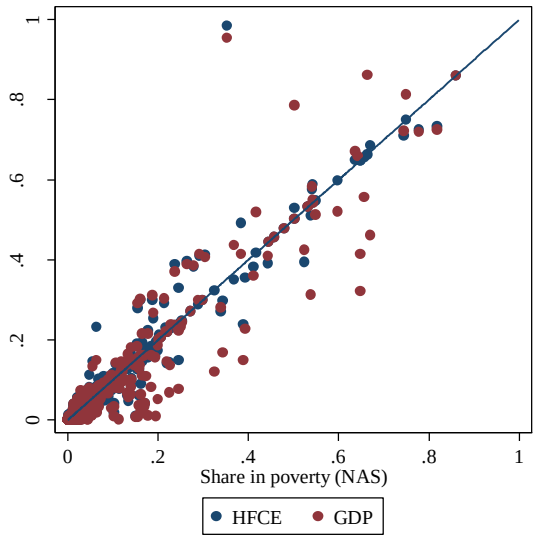


Global



C. Country-specific adjustment (1990)

National



Global

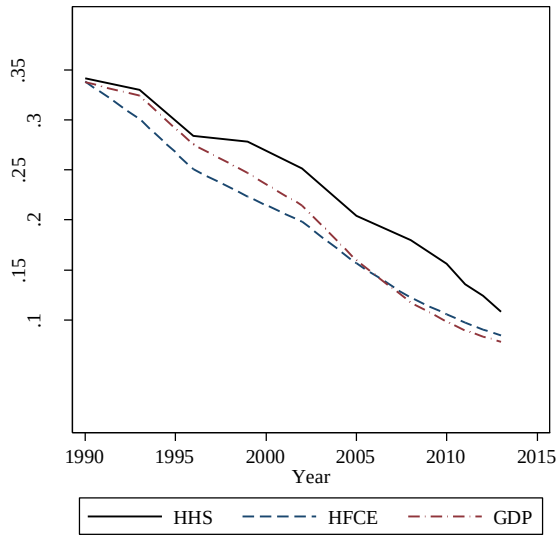
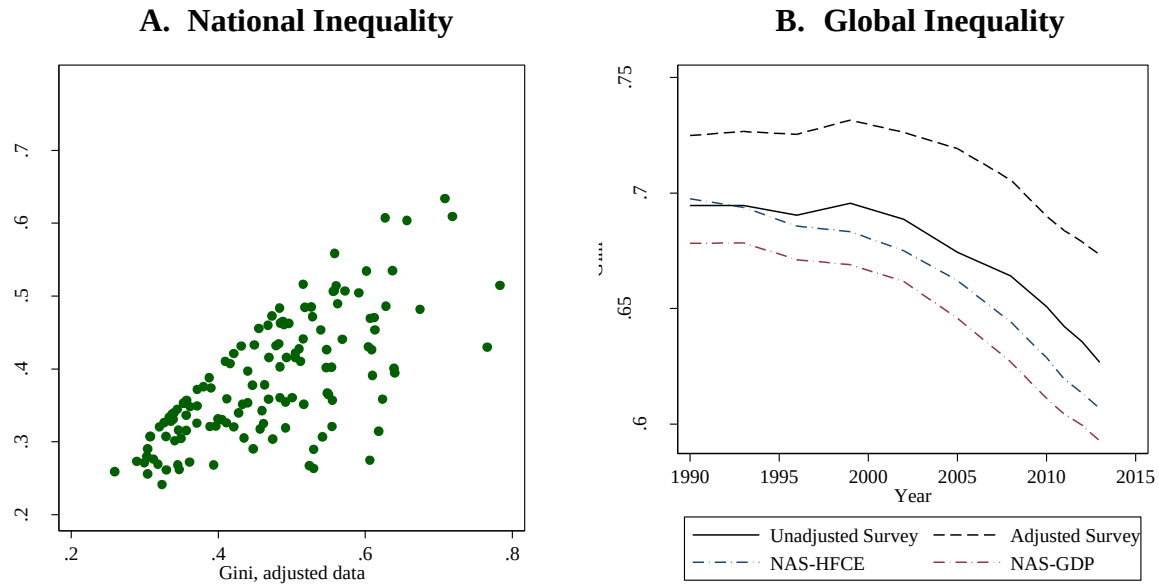


Figure 7: Inequality (Gini) Unadjusted and Adjusted Measures



Note: plots compare national and global inequality measures, based on surveys data and adjusted survey data. Left panel compares survey gini (on vertical axis) with gini from top-income adjusted data on horizontal axis for all surveys. Right panel shows global gini, for surveys, NAS-adjusted surveys and top-income adjusted surveys.

Figure 8: The 'Kuznets Curve': Inequality and Economic Development

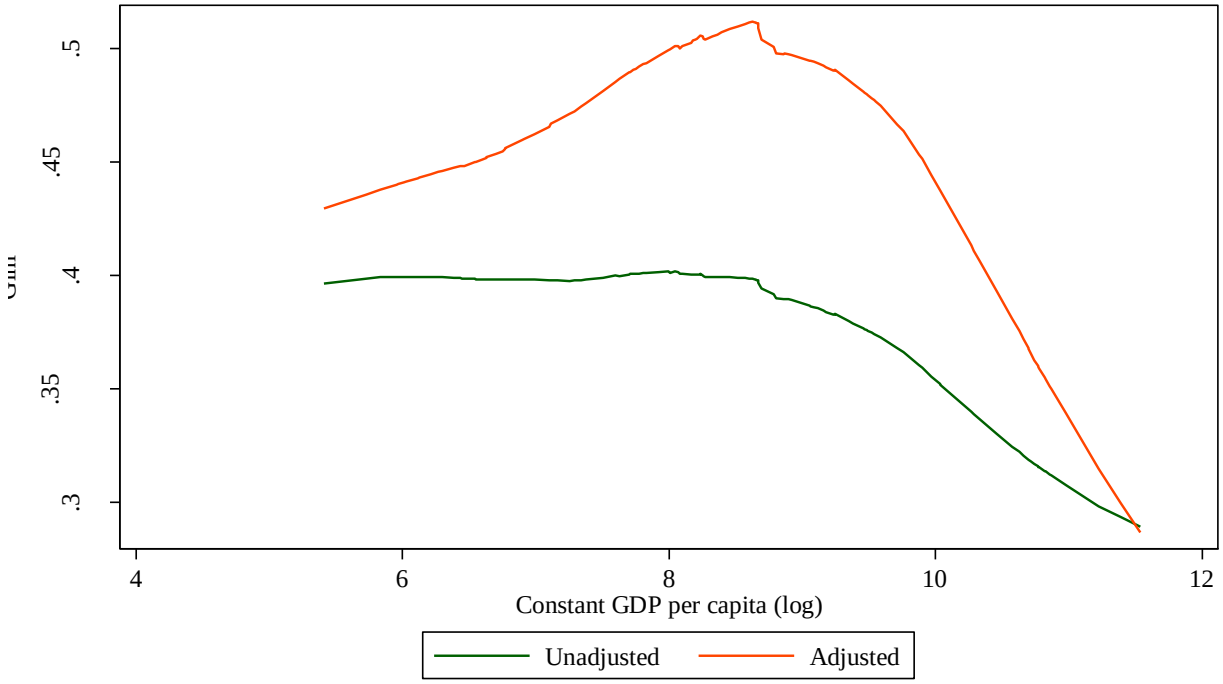
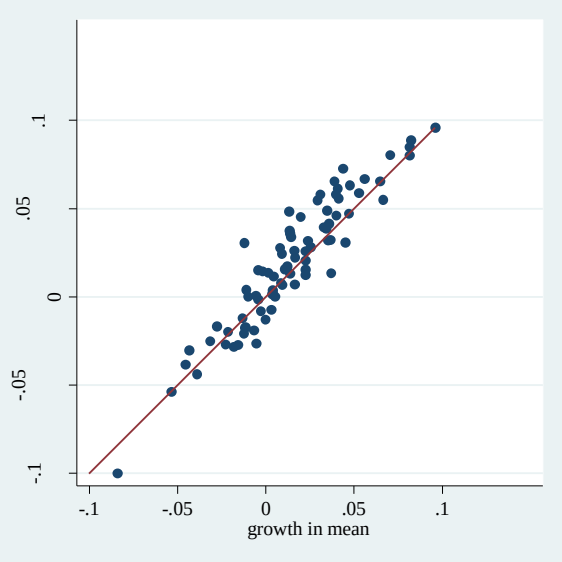


Figure 9: Shared prosperity (growth in bottom 40% vs growth in mean)

A. Survey estimates



B. Adjusted mean