Poverty and Inequality Platform Methodology Handbook

Introduction

This site describes the methodology implemented by the World Bank to calculate global and regional poverty rates that are officially reported in the Poverty and Inequality Platform[1] (PIP). The next page summarizes the methodology in non-technical terms using five steps. The summary is followed by five chapters that dig into these steps in detail.

All decisions, assumptions, and protocols involved in these steps and in the chapters that follow are governed by the World Bank's Global Poverty Working Group, which is composed of staff from the Poverty and Equity Global Practice, the Development Data Group, and the Development Research Group.

Please cite the handbook as follows: World Bank. 2023. "Poverty and Inequality Platform Methodology Handbook." Edition 2023-09. Available at https://datanalytics.worldbank.org/PIP-Methodology/.

[1] https://pip.worldbank.org/home

Non-technical summary

This page summarizes the methodology in non-technical terms using five steps. The summary is followed by five chapters that dig into these steps in detail.

The five steps are illustrated in the figure below and summarized after the figure. First, household survey data are obtained from relevant sources. Second, the survey data are used to create an estimate of households' income or consumption, known as welfare aggregates. Third, the welfare aggregates are adjusted for differences in price levels across countries and over time to foster international comparability. Fourth, poverty and inequality are estimated for a particular country for a particular year. Fifth, the estimates of poverty are extrapolated or interpolated to a common year and the population-weighted poverty rate is calculated.

Step 1: Acquiring household survey data

Poverty rates are estimated from selected household surveys. In general, the surveys used ask a representative subset of households in a country about their consumption or income. These surveys are often the official surveys used by countries to monitor and report on poverty. Most household surveys from developing countries are obtained through collaboration with countries' National Statistical Offices, while most data for high-income countries are obtained from the EU Statistics on Income and Living Conditions (EU-SILC)[2] or from the Luxembourg Income Study (LIS) Database[3]. At times, household-level data cannot be obtained in which case, as a second best, aggregated data are used, such as income or consumption by population deciles.

[2] https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions

[3] https://www.lisdatacenter.org

Step 2: Constructing welfare aggregates

The Poverty and Inequality Platform (PIP) primarily utilizes a monetary measure of poverty. Monetary poverty is estimated from an aggregation of household's income or from the monetary value of their consumption. We refer to such aggregates jointly as welfare aggregates. Welfare aggregates are harmonized across countries and over time to maximize comparability, but country-specific decisions on issues such as whether income or consumption is used, the design of the questionnaire of the household survey, what components that are included in the welfare aggregate, and whether price differences within a country are accounted for, imply that full comparability is not feasible.

Step 3: Converting welfare aggregates

Welfare aggregates are often expressed in local currencies in the prices prevailing at the time of the collection of the data. To compare the consumption of an Indian household in 2011 to the consumption of a Nigerian household in 2018, the welfare aggregates need to be expressed in the same prices. To this end, consumer price indices are used to express all welfare aggregates in local 2017 prices, and purchasing power parities are used to account for price differences between countries. Once all welfare aggregates are expressed in the same units, a common poverty line is needed to estimate poverty. The international poverty line used in PIP is constructed from the poverty lines used by the poorest countries of the world.

Step 4: Calculating survey estimates of poverty and inequality

Armed with welfare distributions expressed in 2017 PPPs and an international poverty line, poverty and inequality can be calculated and compared across countries and over time. PIP contains a range of different monetary poverty measures, a multidimensional poverty measure, inequality measures, and other distributional statistics.

Step 5: Calculating global and regional poverty

Most countries do not conduct household surveys every year. Yet, to estimate regional and global poverty for a particular year, one needs an estimate of poverty for every country for the year in question. When a poverty estimate

is not available for a given year, the estimates are extrapolated or interpolated from other years. The extrapolations assume that everyone's income or consumption grows in accordance with per capita growth rates from national accounts between the time of the survey and the year in question. For countries without any household data at all, it is assumed that their poverty rate is equal to the population-weighted average poverty rate in their region. To ensure the quality of the regional and global numbers, coverage rules are used to determine whether a particular reference year has sufficient nearby survey data for global and regional numbers to be presented.

Chapter 1

Acquiring household survey data

This chapter describes the data sources used for measuring poverty and inequality as well as how the data are selected and obtained.

1.1 Selection criteria and source of data

For most countries, the Poverty and Inequality Platform (PIP) relies on household surveys carried out by countries' National Statistical Offices (NSO). Typically, official surveys used to report poverty for the country are used. This includes household budget surveys or household income and expenditure surveys. These surveys ask a representative sample households in the country detailed questions about their income and/or consumption and much more. In some occasions, the World Bank is a co-producer of these surveys.

Other surveys on consumption or income, such as surveys carried out by academic institutions and private sector companies are generally not included in PIP. The primary reason for this is that the estimates from PIP are the official vehicle for countries' reporting on the Sustainable Development Goal target 1.1.1[1] on extreme poverty, which ought to come from sources that are officially recognized by countries.

For high-income countries in the European Union (with the exception of Germany), the data are obtained from the EU Statistics on Income and Living Conditions (EU-SILC)[2]. EU-SILC data are partly obtained from surveys and

[1] https://sdgs.un.org/goals/goal1

[2] https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions

partly obtained from administrative records. For other high-income economies (Australia; Canada; Germany; Israel; Japan; South Korea; Taiwan, China; the United Kingdom; and the United States) data from the Luxembourg Income Study (LIS) Database[3] are used. These data rely on surveys carried out by NSOs or by academic institutions. Data from the LIS database are also used for high-income countries from the European Union for years predating EU-SILC, which naturally causes a break in the series.

The estimates for India from 2015-2021 rely on the Consumer Pyramids Household Survey conducted by the Centre for Monitoring Indian Economy, a private company. The survey data are adjusted to be representative of India and reflect the consumption basket of the National Sample Survey (see Sinha Roy and Van Der Weide (2022) and Aron et al. (2023)).

Occasionally, a survey fulfilling the above criteria is not incorporated into the platform because the World Bank does not have access to the data, because of concerns about the survey design or welfare aggregate, or because the auxiliary data needed (such as consumer price indices and purchasing power parities) are unavailable. Decisions to exclude a survey from the platform that has previously been available are discussed carefully in the Global Poverty Working Group within the World Bank and the reason for the exclusion of a data point is summarized in the "What's New" documents.

The input data used in PIP from these sources are most often unit record data of welfare and occasionally grouped data, which is converted to a full distribution. PIP cannot take as input data the resultant international poverty rate from a welfare distribution without the underlying distribution, such as a point estimate published on an NSO website. There are many reasons for this:

1. Extrapolating or interpolating country estimates to a common reference year, which is needed to calculate global poverty, requires a full distribution.

2. The lack of detailed information on the distribution would not allow for inequality metrics to be calculated, including the data needed for SDG 10.1.1 and 10.2.1, and would not allow users to access poverty rates as a function of the poverty line chosen.

3. Updates to poverty estimates in the face of revised PPPs and CPIs would not be possible to perform by the World Bank, imposing a higher burden on NSOs and potentially becoming an obstacle for timely publication of updated data.

4. Unit-record data allows for quality checking the data and ensuring that the choices used to create the welfare aggregate are as comparable as possible across countries.

[3] https://www.lisdatacenter.org

1.2 How data are obtained

In many low-income countries, the Poverty and Equity Global Practice of the World Bank is cooperating with the National Statistical Offices and supporting their efforts of conducting household surveys and measuring poverty. Data are obtained through these partnerships. This concerns most countries in East Asia & Pacific, the Middle East & North Africa, South Asia, and Sub-Saharan Africa.

For Latin America and the Caribbean, most of the data are obtained from and harmonized by the CEDLAS's Socio-Economic Database for Latin America and the Caribbean (SEDLAC), which is a partnership between the Center for Distributive, Labor and Social Studies (CEDLAS) at the National University of La Plata in Argentina and the World Bank's Poverty and Equity Global Practice for Latin America and the Caribbean.

For most high-income countries data are obtained through the Luxembourg Income Study or EU-SILC.

All the data go through validity checks by the Poverty and Equity Global Practice of the World Bank. In certain occasions, the Poverty and Equity Global Practice is also responsible for harmonizing and constructing the welfare aggregates. The harmonized welfare aggregates are passed on to PIP.

1.3 Representativeness

In the majority of cases, the surveys used are nationally representative of all households in a country. Often, the selection of households is based on the most recent census. Yet three factors prevent surveys from being representative of the entire population:

1. Some households are less likely to agree to participate in national surveys on income and consumption, or, if they participate, less likely to answer the questions used for calculating poverty. This is particularly a concern for wealthy households, which means that the distributions captured might have too thin a tail at the top.
2. Some individuals do not reside in households as defined in household surveys and hence are not captured by the above sampling frame. This particularly concerns homeless people and individuals residing in nursery homes, prisons, and displacement camps.
3. At times, conflicts and extreme weather events prevent enumerators from going to certain parts of a country. In these cases, some geographical areas may be excluded from the survey. The 2018-19 Nigeria Living Standards Survey, for example, is not representative of the Borno state, since parts of the state became inaccessible over the course of the survey.

For the purpose of calculating global or regional poverty, the poverty rates estimated in the surveys are assumed to be representative of the entire population despite of the three shortcomings above. This is largely because acquiring any alternative poverty estimate for the missing populations is challenging.

PIP also contains some surveys which explicitly only target the urban or rural part of the population. These surveys are not used for the purposes of calculating regional or global poverty, except for in Argentina and Suriname, which only have surveys representative of their urban population. For the urban population of Argentina and Suriname, these surveys are used for the regional and global poverty counts, while the rural population of Argentina and Suriname get the same treatment as countries for which we have no data.

1.4 Grouped vs. unit record data

In some cases, countries are unwilling to share unit-record data with the World Bank but they are willing to share grouped data. Grouped data are aggregated data, often to 5, 10, 20, or 100 quantiles of the distribution. Such data could reflect, for example, the mean income of each percentile or the income share of each percentile. For the purposes of estimating poverty and inequality, a Lorenz curve is estimated from these grouped data upon which inequality and poverty can be estimated (described here). Grouped data are used for Algeria, China, Guyana, Suriname, Turkmenistan, Trinidad & Tobago, Venezuela, and the United Arab Emirates and more countries for earlier years.

1.5 Multiple imputed micro data

In some countries, a full household survey with a comprehensive consumption measure cannot be collected due to concerns over the safety of enumerators and respondents, however, shorter interviews without a full welfare aggregate which minimize safety risks can be conducted. In these cases, a full welfare aggregate can be imputed using statistical models. These methods, described in greater detail in Arayavechkit et al. (2021) and Castaneda Aguilar et al. (2022), typically use multiple imputations, resulting in a dataset with many consumption values for each household. These multiple consumption values reflect the uncertainty that arises from the statistical modelling. Statistics of interest, such as poverty or inequality are computed separately on each of these consumption vectors and averaged to get to a final estimate. Currently such data are used for South Sudan 2016, Zimbabwe 2017, Nigeria 2010, 2012, and 2015, and India 2015-2021.

Chapter 2

Constructing welfare aggregates

Several different concepts of poverty exist, building on different notions of what a good life – or in the language of economists, welfare – entails. Poverty, in short, is the absence of welfare, and individuals are considered poor when their welfare levels fall short of a certain threshold. The Poverty and Inequality Platform (PIP) primarily utilizes a monetary measure of poverty. Monetary poverty is estimated from an aggregation of households' income or from the monetary value of their consumption. We refer to such aggregates jointly as welfare aggregates. This section explains how welfare aggregates are constructed. For a more detailed exposition of the construction of welfare aggregates, see Deaton and Zaidi (2002) and Mancini and Vecchi (2022).

2.1 Income or consumption

Monetary poverty estimates are based on either income or consumption aggregates. Consumption aggregates typically capture household expenditure on a set of items over a given period of time. These usually include purchased, own-produced, exchanged, and gifted food and non-food items (for example clothing, housing—including imputed rent—and the use value of durable consumer goods). Income aggregates capture the value of monetary inflow a household receives or earns over a given period of time. Household surveys usually provide information on labor income (salaries, own-business, and self-employment income), as well as non-labor income coming from pensions, subsidies, transfers, property income, scholarships, etc. Income aggregates in PIP aim to measure disposable income defined as the sum of labor and non-labor income (including transfers) less taxes and contributions. The exact definition and operationalization of income aggregates varies across different data sources.

Both income and consumption aggregates do not directly account for important non-monetary aspects of households' welfare such as access to basic services, education, health-care, and infrastructure, which are better captured by other indicators or by measures of multidimensional poverty.

Both income and consumption approaches to measuring monetary poverty have advantages and disadvantages. Countries typically choose the concept that can be more accurately measured and that is more relevant to their context, while balancing concerns about respondent burden in surveys.

The consumption approach is arguably more directly connected to economic welfare. Yet, consumption aggregates require a wide range of questions, detailed price data, and often post-fieldwork adjustments. The design of consumption questionnaires varies widely and, as shown by numerous experiments, can have significant effects on final poverty estimates (see section on questionnaire design below).

Income aggregates, on the other hand, often rely on no more than a handful of questions and can, at times, be verified from other sources. Yet they are difficult to obtain when a large fraction of the population works in the informal sector or is self-employed, which is frequently the case in poorer economies. When households produce their own food with limited market interactions, it is harder to measure income than consumption. Income aggregates also suffer from the disadvantage that incomes might be very low—even negative—in a given period, whereas consumption is smoothed to safeguard against such shocks. Subsistence requires a minimum level of consumption, which is strictly above zero (World Bank 2018). PIP's current practice is to drop observations with negative incomes, while zeros are included.

The differences between income and consumption matter for comparing trends and levels of poverty and inequality. Given that incomes can be very low or negative, poverty rates are typically higher when income is used rather than consumption. For a given poverty rate, poor households also tend to be further below the poverty line when income is used, as explained by the earlier point about very low incomes. Incomes are also more likely than consumption to be very high for a given year, which in conjunction with the very low values means that inequality often is higher when incomes are used rather than consumption.

2.2 Questionnaire design

Constructing a welfare aggregate, particularly a consumption aggregate, requires many specific questions to be asked in household surveys. There are many different ways to ask survey respondents about their consumption habits, and how one asks has a significant effect on how people respond. Consumption aggregates are sensitive to whether respondents recall consumption from memory or use consumption diaries, how many and how detailed consumption

items are listed, and the time period interviewers ask respondents to recall. De Weerdt, Gibson, and Beegle (2020) provide a recent review of the impact these choices can have on the final poverty estimates.

To give one concrete example, the recall period affects reported consumption through two main channels: memory decay and telescoping. A longer recall period is better at encompassing expenditure on infrequently purchased items, but it can lead to underreporting if respondents forget about past purchases. Despite lower average consumption, measured poverty might be lower under the longer recall period because it captures the purchases of low frequency items of households in the lower parts of the distribution. Short recall periods can mitigate underreporting but can lead to telescoping, where respondents mistakenly report the consumption that took place outside of the reference period.

India provides one example of the relevance of the recall period for poverty estimation. The official 2004–05 poverty rate for India based on a uniform recall period of 30 days was 27.5 percent. Using a mixed recall period of 7 and 30 days on food and tobacco and a 365-day recall period for non-food items, the corresponding figure was 21.8 percent (World Bank 2018).

Changes in questionnaire design imply that poverty estimates within countries become incomparable. Whenever such changes occur, they are marked in the comparability database.

2.3 Imputed rent

During the process of constructing a welfare aggregate, housing is arguably one of the most important and difficult components to include. From a conceptual point of view, housing could be understood as any other durable good: Its present value is not relevant for the analysis of current welfare because purchasing a house is such a large and relatively rare expenditure that it should not be included in the welfare aggregate (Deaton and Zaidi 2002, 35). The objective is instead to measure the value of occupying the dwelling for a month or another limited duration.

Whether the welfare aggregate is based on consumption or income, accounting for housing is important. When the welfare aggregate is based on consumption, household surveys usually record the value of rent paid by market tenants. Since rent is the market value of occupying a house for a given period of time, it is feasible and empirically viable to estimate the flow of housing services. Yet, the same information is not available for household owners, even though they also "consume" housing services. By not including the value of housing consumed by owners, it would look like two household with the very same consumption patterns have different welfare just because one pays rent and the other does not. In fact, the market tenant would look better off than the household owner because the consumption of the former is higher than the one of the latter.

When the welfare aggregate is based on income the logic is similar. In this case, the welfare aggregate should be composed of the remuneration of all the assets of the household including labor, capital, and durable goods such as houses (Gasparini and Sosa Escudero 2004). Families that own their housing receive an implicit value that is equivalent to the amount of money they would have to pay in the market if they had to rent a dwelling similar to the one they are currently living in. Hence, to properly account for housing in welfare aggregates, one needs to impute rent for owners.

Several methods exist to impute rent (see Balcázar et al. (2017) for a review). Due to the data needs of these approaches, imputed rent is not always computed, leading housing to be excluded from the welfare aggregate. Across the surveys used in PIP, it varies between country and within countries over time whether imputed rent is included. Due to the value of housing services relative to total welfare, this can matter a great deal for poverty comparability over time.

2.4 Within-survey spatial/temporal deflation

When constructing a welfare aggregate, one is met with the challenge that prices differ geographically within a country and change over the time of the fieldwork. This means that the same level of income can have a different value for different households in the survey, and that households with the same consumption patterns can have very different consumption aggregates if the prices they face differ. To make the welfare aggregates comparable across time and space, spatial and temporal price deflation are needed. Whether or not such deflation is carried out depends on the availability of price data and country-specific practices. Whenever a country switches from not spatially deflating to spatially deflating, this might lead to incomparable poverty estimates over time.

2.4.1 Spatial deflation

Suppose a household pays $1 for a kilo of rice in an urban center, while a rural household in the same country only pays $0.5 for a similar quality and amount of rice. If one were to assess poverty based on the value of the goods and services consumed without accounting for these price differences, everything else equal, one would conclude that the rural household is poorer than the urban household. Yet, from a welfare perspective, they are equally well off. To properly compare the welfare levels of the two households one needs to account for the differences in price levels that the two households face.

There are several ways to account for spatial price differences. One way is to compare the price of a representative basket of goods across locations and convert household consumption expenditure into the prices of a reference location, such as the capital city or the national average price level.

Currently most surveys from East Asia & Pacific, Latin America & the Caribbean, and Europe & Central Asia are spatially deflated, while it is less common in Sub-Saharan Africa, South Asia, and the Middle East & North Africa. China, India, and Indonesia have rural/urban purchasing power parities (PPPs) that are used to deflate welfare aggregates to account for rural/urban price differences within these countries.

2.4.2 Temporal deflation

Household surveys are often carried out over several months, leading to the possibility that prices evolve notably over this period. If unaccounted for, this means that two households at the same location that have the same consumption patterns but are interviewed at the beginning and the end of the fieldwork may have different consumption aggregates. Accounting for such temporal deflation implies choosing a reference period and deflating all consumption aggregates to that period. This can be a specific year, quarter or month. The exact time chosen for each survey is listed in the appendix table of the most recent What's New document. Temporal deflation within a survey is not carried out in all cases, often because the month of the fieldwork is not available.

Sometimes, a spatio-temporal price index is used to jointly adjust for both spatial and temporal price differences. The household welfare aggregate from the survey conducted in Ghana in 2016-17, for example, is evaluated at 2013 Greater Accra regional prices.

All surveys used by PIP are assigned to a particular year using two variables *year* and *datayear*. If all fieldwork for a survey took place during one calendar year, and relate to the consumption or income of that year, *year* and *datayear* are both equal to this calendar year. When household surveys span two calendar years, the *datayear* noted in the Poverty and Inequality Platform is not an integer. The Gambia, for example, has a datayear of 2015.31. This means that 69% of the fieldwork for this particular survey took place in 2015 while 31% of the fieldwork months took place in 2016. In these cases, the *year* information is the floor of *datayear*. The *datayear* is used for interpolations, extrapolations, and at times, for the inflation to convert the welfare vector to constant prices. Notice that in some cases neither the *datayear* nor the *year* align with the year of the survey. For EU-SILC surveys, for example, the income information asked in a survey in given year relates to income of the previous year. Here, the *datayear* and *year* will both be one year prior to the year the survey was carried out.

2.5 Equivalence scale

All welfare aggregates are converted into per capita terms by dividing the total household welfare with the number of household members. The use of a per

capita normalization is standard in the literature on developing countries. Behind the per capita conversion lies an assumption that all household members have equal consumption levels or derive equal welfare from the total household income. This is based on the assumption that there is little scope for economies of size in consumption for poor people.

This assumption can be questioned (Lanjouw and Ravallion 1995). If household members have different needs, which could be the case if children, for example, require less consumption than adults, then using per capita welfare might be misleading. The assumption would also be violated if household members with the same needs do not get the same share of household income and consumption. Evidence has shown that females often get less than an equal part of household income and consumption (see chapter 5 of World Bank (2018) and references therein). An alternative assumption often used by countries when measuring poverty, is to use an adult equivalence scale. Such a scale defines the consumption need of each household member—often based on age and the household size—relative to a benchmark. Total household welfare is then converted into welfare per adult equivalent household member.

Despite of its compelling features, the lack of agreement on which equivalence scale to use across countries and the more difficult interpretation of the final poverty rates have prevented the World Bank from using such numbers for global poverty monitoring thus far. Per capita estimates have the advantage that they have a clear counterpart in national accounts, which matters for extrapolating and interpolating poverty estimates to a common reference year.

Though equivalence scales aren't used, it is important to note that each household is weighted by the product of its sampling weight and the household size. The latter implies that the poverty rates show the share of the *population* living in poverty, not the share of *households* living in poverty.

2.6 Treatment of grouped data

A welfare aggregate is also needed to estimate poverty and inequality from grouped data. Grouped data are consumption expenditure or income organized in intervals or bins, such as deciles or percentiles. These bins are used to derive a continuous Lorenz curve, which plots the cumulative welfare share (on the y-axis) against the cumulative population share (on the x-axis). Together with information about mean welfare, the Lorenz curve can be used to construct a full distribution. Two approaches are used to derive a Lorenz function, the general quadratic (GQ) Lorenz function and the Beta Lorenz function (Datt 1998). Both functions are parameterized and estimated. The function that provides the best fit is selected for poverty and distributional statistics, conditional on passing normality and validity tests. The tests vary for poverty and distribu-

tional statistics (see here[1] for the code on the poverty tests and here[2] for the code on the distributional tests).

The GQ Lorenz function is estimated using the following specification:

$$L(1-L) = a(p^2 - L) + bL(p-1) + c(p-L),$$

where p is the cumulative proportion of the population, L is the cumulative proportion of consumption expenditure or income, and a, b, c are parameter estimates (Datt 1998). Poverty and inequality measures are based on the parameter estimates.

The Beta Lorenz function is estimated using the following specification:

$$L(p) = p - \theta p^{\gamma}(1-p)^{\delta},$$

where θ, γ, and δ are parameter estimates (Datt 1998).

2.7 Comparability of estimates

As countries frequently improve the questionnaire design of household surveys and the methodology for the construction of welfare aggregates, poverty estimates over time for a country are not always comparable. In order to guide users about when poverty can be compared over time within a country, PIP's line charts for countries only connect poverty estimates when these are comparable. In addition, PIP contains metadata on the comparability of poverty estimates within countries over time.

The metadata can be accessed through the PIP API[3] or the PIP Stata module[4]. The variables `comparable_spell` and `survey_comparability` contain the necessary information. The latter has the following logic: The oldest comparable series in each country starts with the value zero (0). When comparability is broken, the value changes to one (1) for the year of the break and it goes on until the comparability is broken again in a subsequent year. The process repeats until the most recent survey data point available.

Within a country, comparability of poverty estimates over time is assumed unless there is a known change to the survey instrument, survey methodology, measurement, or data structure. The assessment of comparability is country-dependent and relies on the accumulation of knowledge from past and current

[1] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_select_lorenz.R#L80-L132

[2] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_select_lorenz.R#L143-L175

[3] https://pip.worldbank.org/api

[4] https://github.com/worldbank/pip

World Bank staff, as well as close dialogue with national data producers with knowledge of survey design and methodology. More documentation of the comparability assessment is available in Atamanov et al. (2019).

Chapter 3

Converting welfare aggregates

Welfare aggregates from household surveys are often expressed in national currency units in prices around the time of the fieldwork. To use a welfare aggregate from a particular survey to estimate extreme poverty at the international poverty line, the welfare aggregates need to be converted to a unit comparable across time and across countries. To this end, first Consumer Price Indices (CPIs) are used to express the aggregates in the same prices within a country. Second, Purchasing Power Parities (PPPs) are used to express all welfare aggregates in the same currency by adjusting for price differences across countries.

3.1 Consumer Price Indices (CPIs)

Consumer price indices (CPIs) summarize the prices of a representative basket of goods and services consumed by households within an economy over a period of time. Inflation (deflation) occurs when there is a positive (negative) change in the CPI between two time periods. With inflation, the same amount of rupees is expected to buy more today than one year from today. CPIs are used to deflate nominal income or consumption expenditure of households so that the welfare of households can be evaluated and compared between two time periods at the same prices.

The primary source of CPI data for the Poverty and Inequality Platform is IMF's International Financial Statistics (IFS)[1] monthly CPI series. The simple average of the monthly CPI series for each calendar year is used as the annual CPI. When IFS data are missing, other sources of CPI data are obtained from IMFs World

[1] https://data.imf.org/?sk=4c514d48-b6ba-49ed-8ab9-52b0c1a0179b

Economic Outlook (WEO)[2] and National Statistical Offices (NSOs), among others. For more details on the different sources of CPI data used for global poverty measurement, see Figure 1 of Lakner et al. (2018) and the "What's New" technical notes accompanying PIP updates.

CPI series are rebased to the International Comparison Programme (ICP) reference year, currently 2017.

3.2 Purchasing Power Parities (PPPs)

Purchasing power parities (PPPs) are used in global poverty estimation to adjust for price differences across countries. PPPs are price indices published by the International Comparison Program (ICP[3]) that measure how much it costs to purchase a basket of goods and services in one country compared to how much it costs to purchase the same basket of goods and services in a reference country, typically the United States. PPP conversion factors are preferred to market exchange rates for the measurement of global poverty because the latter overestimate poverty in developing countries, where non-tradable services are relatively cheap (a phenomenon known as the Balassa-Samuelson-Penn effect).

The ICP updates PPP data periodically. For example, in the last three ICP rounds, PPPs were released for the 2005, 2011 and 2017 reference years. The 2017 PPPs are currently used to convert household welfare aggregates, expressed in local currency units in 2017 prices, into a common internationally comparable currency unit. The PPP conversion only affects the cross-country comparison of levels of welfare; the growth in the survey mean for a particular country over time is the same whether it is expressed in constant local currency or in USD PPP.

The PPP estimates used for measuring global poverty are the consumption PPPs from the ICP with a few exceptions. For eight countries, the PPPs used are derived as the geometric average of the PPPs published by the ICP and imputed PPPs derived from using a regression model the ICP team uses for countries that did not collect price data. This concerns Belize, Egypt, Guinea, Iraq, Nigeria, São Tomé and Príncipe, Sudan, and Trinidad and Tobago. These are countries where PPPs and CPIs move quite differently between 2011 and 2017 and additional evidence suggests that the official PPPs may be less appropriate for monitoring global poverty (Jolliffe et al. 2022). If official and imputed PPPs do not exist, PPPs are extrapolated from the revised 2011 PPPs using the ratio of domestic to US inflation between 2011 and 2017. This applies to six countries, namely Kiribati, Nauru, the Syrian Arab Republic, Tuvalu, the República Bolivariana de Venezuela, and the Republic of Yemen (Jolliffe et al. 2022). PPPs are also extrapolated for Timor-Leste even though imputed

[2] https://www.imf.org/en/Publications/SPROLLs/world-economic-outlook-databases#sort=%40imfdate%20descending

[3] https://www.worldbank.org/en/programs/icp

2017 PPPs exist. This is because of low credibility in the imputed PPP (Castaneda Aguilar et al. 2023).

Though PPPs are supposed to be nationally representative, to account for possible urban bias in ICP data collection, separate rural and urban PPPs are computed for China, India, and Indonesia using official national PPPs, the ratio of urban to rural poverty lines, and the urban share in ICP price data collection (Chen and Ravallion 2008, 2010; Jolliffe and Prydz 2015; Ferreira et al. 2016; Atamanov et al. 2020; Jolliffe et al. 2022).

3.3 Derivation of the international poverty line

Most countries have a national poverty line which summarizes the value of consumption or income per person or per adult equivalent needed to be non-poor. These national poverty lines are typically estimated by National Statistical Offices and reflect country-specific definitions of what it means to be poor. For low and middle-income countries, the lines usually reflect the cost of purchasing a bundle of food items necessary to obtain minimum daily calories to which a basic non-food component is added. For high-income countries the national poverty lines are often relative and are defined relative to the national mean or median income.

To compare poverty across countries one needs a common standard. Hence, national poverty lines, which differ from one country to the next, cannot be used. The international poverty line is an attempt to summarize the national poverty lines of the poorest countries.

Since 1990, the World Bank has derived international poverty lines from the national poverty lines of the poorest countries of the world (Ferreira et al. 2016). In 1990, this resulted in a the "dollar-a-day" poverty line. Whenever new rounds of PPPs have been released, the nominal value of the international poverty line has been updated. The real value of the international poverty line is usually unchanged.

The current international poverty line of \$2.15/day in 2017 PPPs has been derived as the median of harmonized national poverty lines of low-income countries (Jolliffe et al. 2022). This preserves the principle of using the national poverty lines of the poorest countries in the world as the basis of the international poverty line. In addition, it involves methodological improvements in the derivation of the international poverty line. For example, the new line is based on a larger sample of 28 low-income countries, compared to 15 countries from the previous lines. The previous lines were \$1.25 in 2005 PPPs (Ravallion, Chen, and Sangraula 2009) and \$1.90 in 2011 PPPs (Ferreira et al. 2016). Unlike the previous 15 national poverty lines, the new 28 national poverty lines have been harmonized across countries and are thus expressed in the same per-capita PPP

units. The $2.15 line is robust to several measurement assumptions and methods and keeps the real value of the old $1.90 line virtually unchanged (Jolliffe et al. 2022).

3.4 Derivation of other global poverty lines

In addition to the international poverty line, the World Bank uses two higher poverty lines to measure and monitor poverty in countries with a low incidence of extreme poverty. These higher lines, namely $3.65 and $6.85 in 2017 PPPs, are derived as the median values of national poverty lines of lower- and upper-middle income countries, respectively. These lines have been derived along with the international poverty line, using the same methodology (Jolliffe et al. 2022). All these poverty lines are absolute poverty lines.

The World Bank also uses a societal poverty line (SPL) that reflects a more relative concept of poverty. With 2017 PPPs, the societal poverty line is defined as $1.15 plus half the median level of consumption in a country, or the international poverty line if $1.15 plus half the median level of consumption is lower than the international poverty line (Jolliffe et al. 2022). This line increases as a country grows (and the median increases).

Chapter 4

Calculating survey estimates of poverty and inequality

4.1 Monetary poverty measures

Armed with welfare distributions expressed in 2017 PPPs and an international poverty line, poverty and inequality can be calculated and compared across countries and over time. The poverty and inequality measures used in the Poverty and Equity Platform (PIP) for micro data are mainly based on formula in Haughton and Khandker (2009), while the poverty and inequality measures estimated from grouped data are mainly based on formula in Datt (1998). In the formulas below we abstract from household weights for simplicity.

Poverty headcount ratio: The poverty headcount ratio (P_0) measures the proportion of the population that is counted as poor. When using micro data, the poverty headcount ratio is obtained using the following expression:

$$P_0 = \frac{1}{N}\sum_{i=1}^{N} I(y_i < z),$$

where N is the total population size, y_i is the welfare of individual i, z is the poverty line, and $I(.)$ is an indicator function that takes on the value 1 if the bracketed expression is true or 0 otherwise. You can find the source code where this calculation is carried out here[1]. For grouped data that rely on the GQ

[1] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_poverty_stats.R

Lorenz function, the poverty headcount rate is obtained using the following expression (source code[2]):

$$P_0 = -\frac{1}{2m}[n + r(b + 2z/\mu)((b + 2z/\mu)^2 - m)^{-1/2}],$$

where a, b, c are the parameter estimates of the General Quadratic Lorenz function, μ is mean welfare, z is the poverty line, $e = -(a + b + c + 1)$, $m = b^2 - 4a$, $n = 2be - 4c$, and $r = (n^2 - 4me^2)^{1/2}$.

If the Beta Lorenz function is used, the poverty headcount ratio is estimated using the expression below (source code[3]):

$$\theta {P_0}^{\gamma}(1 - P_0)^{\delta}\left[\frac{\gamma}{P_0} - \frac{\delta}{(1 - P_0)}\right] = 1 - \frac{z}{\mu}$$

where θ, δ, and γ are parameter estimates. See the section on treatment of grouped data for details on the GQ and Beta Lorenz functions.

Poverty gap index: The poverty gap index (P_1) is a measure that adds up the extent to which individuals on average fall below the poverty line (i.e. the depth of poverty), and expresses it as a percentage of the poverty line. The poverty gap index is obtained from micro data with the following expression (source code[4]):

$$P_1 = \frac{1}{N}\sum_{i=1}^{N}\frac{G_i}{z}$$

with $G_i = (z - y_i)\ddot{O}I(y_i < z)$, where G_i is defined as the poverty gap (i.e. poverty line (z) less welfare (y_i)) of poor individuals; the gap is considered to be zero for everyone else.

For grouped data, the poverty gap index is obtained using the following expression (Beta Lorenz source code[5], GQ source code[6]).

$$P_1 = P_0 - (\mu/z)L(P_0)$$

[2] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L546-L550

[3] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lb.R#L646-L670

[4] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_poverty_stats.R

[5] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lb.R#L806-L815

[6] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L568

with all variables defined as before. The expression holds for both GQ and Beta Lorenz functions.

Poverty severity index: The poverty severity index (P_2) is a measure of the weighted sum of poverty gaps (as a proportion of the poverty line), where the weights are the proportionate poverty gaps themselves. The poverty severity index is obtained from micro data with the following expression (source code[7]):

$$P_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{G_i}{z}\right)^2$$

with all variables defined as before. Also known as the squared poverty gap index, the poverty severity index accounts for inequality among the poor .

With the GQ Lorenz function, the poverty severity index is obtained from grouped data using the following expression (source code[8]):

$$P_2 = 2(P_1) - P_0 - \left(\frac{\mu}{z}\right)^2\left[aP_0 + bL(P_0) - \left(\frac{r}{16}\right)ln\left(\frac{1-P_0/s_1}{1-P_0/s_2}\right)\right],$$

where $s_1 = (r-n)/(2m)$, $s_2 = -(r+n)/(2m)$, and all other variables are defined as before.

With the the Beta Lorenz function, the poverty severity index is given as (source code[9]):

$$P_2 = (1-\mu/z)\left[2P_1 - (1-\mu/z)P_0\right] \qquad +\theta^2\left(\frac{\mu}{z}\right)^2\left[\gamma^2 B(P_0, 2\gamma-1, 2\delta+1) \qquad -2\gamma\delta B(P_0, 2\gamma,$$

with

$$B(k, r, s) = \int_0^k p^{r-1}(1-p)^{s-1}\, dp$$

where k, r, s are parameters and all other notations are defined as before for the Beta Lorenz function.

[7] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_poverty_stats.R

[8] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L572-L575

[9] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lb.R#L828-L847

The Watts index: The Watts index (W) is also an inequality-sensitive poverty measure that is given as (mirco data source code[10], GQ source code[11], Beta Lorenz source code[12]):

$$W = \frac{1}{N} \sum_{i=1}^{q} ln\left(\frac{z}{y_i}\right),$$

where the N individuals of the population are indexed in ascending order of welfare, and the sum is taken over q individuals whose welfare, y_i, falls below the poverty line z. This formula applies to both micro and grouped data. It also holds for both the GQ and Beta Lorenz functions, which are evaluated at the poverty line z and other percentiles y_i. A percentile is given as the product of the mean and the first derivative of the GQ or Beta Lorenz function at the respective rank in the distribution. The first derivative of the GQ Lorenz function is given as:

$$L^{'}(p) = -\frac{b}{2} - \frac{(2mp+n)(mp^2+np+e^2)^{-1/2}}{4}$$

with all variables and parameter estimates as previously defined. The first derivative of the Beta Lorenz function is given as:

$$L^{'}(p) = 1 - \theta p^{\gamma}(1-p)^{\delta}\left[\frac{\gamma}{p} - \frac{\delta}{1-p}\right]$$

with all variables and parameter estimates as previously defined.

An item for discussion regarding the Watts index and the Mean Log Deviation, defined below, is the treatment of zeros in the welfare distribution given that the logarithm of zero is not defined. As explained in section 2.1, PIP's current practice is to drop observations with negative welfare, while zeros are included when the welfare distribution is based on income. One approach is to transform the zeros by either maintaining the same mean of the distribution or by adjusting it. Another solution is to exclude the zeros from the sample, but that would create inconsistencies with other poverty and inequalities measures that use the whole sample. Currently PIP excludes zeros from the sample when calculating the Watts index and the Mean Log Deviation.

[10] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_poverty_stats.R

[11] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L402-L450

[12] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/gd_compute_pip_stats_lb.R#L328-L376

4.2 Multidimensional Poverty Measure

In many settings, important aspects of well-being, such as access to quality education or core services, are not captured by monetary measures of poverty. To address this issue, an established tradition of multidimensional poverty measures these nonmonetary dimensions and aggregates them into an index (Alkire et al. 2015). The United Nations Development Programme's Multidimensional Poverty Index (Global MPI)[13], produced in conjunction with the Oxford Poverty and Human Development Initiative[14], is a foremost example of such a multidimensional poverty measure. In 2018, the World Bank launched its own multidimensional poverty measure, building heavily on these prior efforts (World Bank 2018).

The dimensions, indicators, and weights of the World Bank's multidimensional poverty measure are given as follows:

Dimension	Indicator	Weight
Monetary poverty	Daily consumption or income is less than US$2.15 per person	1/3
Education	At least one school-age child up to the age of grade 8 is not enrolled in school	1/6
Education	No adult in the household (age of grade 9 or above) has completed primary education	1/6
Access to basic infrastructure	The household lacks access to limited-standard drinking water	1/9
Access to basic infrastructure	The household lacks access to limited-standard sanitation	1/9
Access to basic infrastructure	The household has no access to electricity	1/9

An individual is considered poor if she/he lives in a household that is deprived in at least 1/3 of the weighted indicators. PIP reports the share of individuals that are poor.

4.3 Inequality measures

Gini index: The Gini index is derived from the Lorenz curve, which plots the cumulative welfare share (on the y-axis) against the cumulative population share (on the x-axis). A 45-degree line is defined over the Lorenz curve as a line of perfect equality. The Gini index is the area between the 45-degree line and the Lorenz curve (multiplied by 100). Let A be the area between the 45-degree

[13] http://hdr.undp.org/en/2020-MPI
[14] https://ophi.org.uk/multidimensional-poverty-index/

line and the Lorenz curve, and let B be the area under the Lorenz curve. The Gini coefficient is given as: $\frac{A}{A+B} = 2A$ (since $A + B = 0.5$).

Formally, if x_i is a point on the x-axis, and y_i a point on the y-axis, then

$$Gini = 100 * (1 - \sum_{i=1}^{N}(x_i - x_{i-1})(y_i + y_{i-1})),$$

where there are N equal intervals on the x-axis (source code[15]). The Gini index ranges from 0 (perfect equality) to 100 (complete inequality) .

With the GQ Lorenz function, the Gini index is obtained from grouped data using the following expression (source code[16]):

$$Gini = 100 * \begin{cases} \frac{e}{2} - \frac{n(b+2)}{4m} + \frac{r^2}{8m\sqrt{-m}}\left[\sin^{-1}\frac{(2m+n)}{r} - \sin^{-1}\frac{n}{r}\right] & \text{if } m < 0 \\ \frac{e}{2} - \frac{n(b+2)}{4m} + \frac{r^2}{8m\sqrt{m}} ln\left[\text{abs}\left(\frac{2m+n+2\sqrt{m}(a+c-1)}{n-2e\sqrt{m}}\right)\right] & \text{if } m > 0 \end{cases}$$

with all parameters defined as before. With the Beta Lorenz function, the Gini index is given as (source code[17]):

$$Gini = 100 * 2\theta B(1 + \gamma, 1 + \delta)$$

Note that $B(1 + \gamma, 1 + \delta)$ is the Beta function $\int_0^1 p^\gamma (1 - p)^\delta \, dp$.

Mean log deviation: The mean log deviation (MLD) belongs to the family of generalized entropy inequality measures. It is given as:

$$MLD = \frac{1}{N}\sum_{i=1}^{N} ln\left(\frac{\mu}{y_i}\right) * 100,$$

where N is the total population size, μ is the mean welfare per person, and y_i is welfare of individual i. The MLD has a minimum value of 0 (perfect equality) and has no upper bound. This formula applies for both micro and grouped data (micro data source code[18], GQ source code[19], Beta Lorenz source code[20]). For

[15] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_gini.R
[16] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L254-L288
[17] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lb.R#L208-L221
[18] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_mld.R
[19] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L327-L353
[20] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lb.R#L253-L279

both the GQ and Beta Lorenz functions, the MLD is evaluated at the poverty line z and percentiles y_i. A percentile is given as the product of the mean and the first derivative of the GQ or Beta Lorenz function at the respective rank in the distribution (see section on the Watts Index for the derivation of the first derivatives).

In the same vein of the Watts Index, the MLD suffers from being undefined when there are zero values in the distribution. Currently PIP replaces zeros with ones when in order to use the same sample as with the other inequality measures.

Polarization: The polarization index (P), also known as the Wolfson polarization index, measures the extent to which the distribution of welfare is "spread out" and bi-modal. It is given as (micro data source code[21], GQ source code[22], Beta Lorenz source code[23]):

$$P = \frac{2(\mu^* - \mu^L)}{m},$$

where μ^* is the distribution-corrected mean (i.e. $\mu(1 - Gini/100)$), μ^L is the mean of the poorest half of the population, and m is the median. Like the Gini coefficient, the polarization index ranges from 0 (no polarization) to 1 (complete polarization). The polarization index is based on Wolfson (1994) and Ravallion and Chen (1997).

4.4 Other distributional statistics

Median:

The median is the amount of welfare per capita that divides the distribution into two equal halves. For micro data, the median is estimated as:

$$Med = \begin{cases} Y\left[\frac{n}{2}\right] & \text{if } n \text{ is even} \\ \left(Y\left[\frac{n-1}{2}\right] + Y\left[\frac{n+1}{2}\right]\right) & \text{if } n \text{ is odd} \end{cases}$$

where Y is the ordered list of welfare per capita in the sample, and n is the sample size.

For grouped data, the median is estimated as:

[21] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/md_compute_polarization.R

[22] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lq.R#L465-L475

[23] https://github.com/PIP-Technical-Team/wbpip/blob/0cf9aa7bc168400ecf78ae5455f4c4247b841e9d/R/gd_compute_pip_stats_lb.R#L424-L435

$$Med = \mu L^{'}(0.5)$$

where μ is the mean and $L^{'}(0.5)$ is the first derivative of the GQ or Beta Lorenz function at the 50th percentile. See section on the Watts Index for the derivation of the first derivatives.

Decile shares:

Imagine that we order all households according to their welfare per capita, such that the first decile contains the 10% of the population with the lowest welfare and so on. Decile shares measure the share of total welfare belonging to each decile. Formally, the decile share of decile d is given by:

$$decileshare_d = \frac{\sum_{i \in d} y_i}{\sum_{i=1}^{N} y_i}, d \in \{1, 2, .., 10\}$$

This formula holds for both micro and grouped data. For grouped data, ordered welfare levels, y_i, are the product of the mean and the first derivative of the GQ or Beta Lorenz function at the respective rank in the distribution.

Shared prosperity

Shared prosperity measures the extent to which welfare growth is inclusive by focusing on growth of the population at the bottom of the distribution. It is defined as the annualized growth rate of mean household per capita welfare of the bottom 40%. Promoting shared prosperity is one of the twin goals of the World Bank Group together with eradicating extreme poverty. The selection of the bottom 40% is a compromise between competing considerations; the bottom 40% still focuses on the bottom of the distribution but is not too small that it risks introducing measurement errors.

Formally, shared prosperity, sp, is calculated as

$$sp = \left(\frac{\mu_{b40,t_2}}{\mu_{b40,t_1}} \right)^{\frac{1}{t_2 - t_1}} - 1$$

,

where μ_{b40,t_2} and μ_{b40,t_1} are the mean welfare of the bottom 40% at time t_2 and t_1. Shared prosperity is anonymized, meaning that it does not track the same individuals in both time periods.

In order for shared prosperity to be calculated for a country the following three criteria must be met:

1. Two comparable household surveys must be available.

2. Among comparable surveys, one must be conducted within two years of the latest reference year (currently 2019) and the other within two years of five years prior to the reference year (currently 2014).

3. The period between the initial and end years should range between three and seven years. For example, a shared prosperity period of 2015–16 meets the second selection criterion but would not be allowed because it does not meet this third requirement.

In cases where multiple surveys fulfill these criteria, the most recent survey years are typically chosen. In countries in Europe and Central Asia using incomes as the welfare aggregate, households with negative incomes are included when estimating shared prosperity.

More information on shared prosperity can be found at the Global Database of Shared Prosperity[24].

Shared prosperity premium

The shared prosperity premium, *spp*, measures the difference in annualized growth of the bottom 40%, shared prosperity, *sp*, and the annualized growth of the mean:

$$spp = \left(\frac{\mu_{b40,t_2}}{\mu_{b40,t_1}}\right)^{\frac{1}{t_2-t_1}} - \left(\frac{\mu_{t_2}}{\mu_{t_1}}\right)^{\frac{1}{t_2-t_1}}.$$

Percentiles database

The percentiles database[25] reports 100 points ranked according to the consumption or income distributions for each survey data available in PIP.

Each percentile distribution reports 100 points per country per survey year ranked from the smallest income (percentile 1) to the largest (percentile 100). For each percentile, the database reports the following variables: The average welfare within the percentile group (avg_welfare); the value for the upper threshold of the percentile group (quantile); the share of the population in the percentile group (which might deviate slightly from 1% due to coarseness in the raw data) (pop_share); and the share of income held by each percentile group (welfare_share). In addition, the database reports the welfare measure (welfare_type) used in the survey data — income or consumption — and the region covered (reporting_level) - urban, rural, or national. The database can be merged with the PIP data using the country_code, reporting_level, year, and welfare_type variables.

For surveys with access to the microdata and for distributions based on binned data, the Stata quantiles function is used to allocate the survey population into

[24] https://www.worldbank.org/en/topic/poverty/brief/global-database-of-shared-prosperity

[25] https://datacatalog.worldbank.org/int/search/dataset/0063646

100 equally sized percentiles. The income and consumption averages, thresholds, and shares of each percentile is derived using the population and their income or consumption in each bin. The quantiles function to bin the distribution is used consistently throughout PIP when bins are necessary. For instance, the same approach is used to derive the 400 bins from LIS data.

Grouped data often have fewer than 100 observations, in which case the percentiles for grouped data are based on a parametric fit of the Lorenz curve. The percentiles are based on the Lorenz curve used for distributional statistics (like the decile shares, Gini coefficient, and median) in the PIP output.

Chapter 5

Calculating global and regional estimates

Few countries have survey estimates of poverty available every year. To estimate poverty at the regional and global level, the survey estimates need to be aligned to a reference year and aggregated. Such alignment and aggregation require assumptions about how to interpolate and extrapolate data as well as how to treat countries without any household survey data at all. This section explains how these calculations are made. Source code can be found here[1] and here[2].

5.1 Extrapolations

For countries that do not have welfare aggregates at or since a specific reference year, but which do have earlier welfare aggregates available, their most recent aggregate is extrapolated forward using growth rates from national accounts. This is done by first finding the growth in national accounts that occurred between the survey and the reference year and scaling the survey welfare distribution, $f(y_{survey})$, by this growth factor. This yields an estimate of the welfare distribution at the reference year, $f(y_{reference})$ as summarized below:

$$f(y_{reference}) = \frac{NA_{reference}}{NA_{survey}} \times f(y_{survey})$$

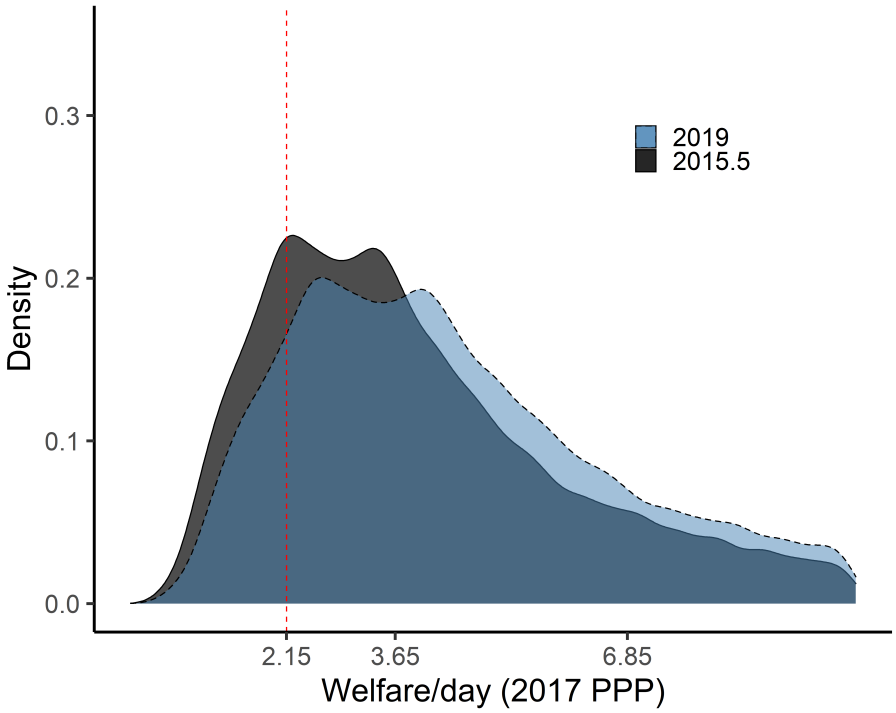
where NA_t is real GDP per capita or real HFCE per capita at time t, as explained in the national accounts data section. Poverty for the reference year

[1] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/predict_request_year_mean.R

[2] https://github.com/PIP-Technical-Team/wbpip/blob/master/R/fill_gaps.R

is then estimated using this extrapolated distribution. A similar approach is used to extrapolate backwards, when the earliest survey estimate available is more recent than the desired reference year. The extrapolation method assumes distribution-neutral growth, i.e. that everyone's welfare grows at the same rate. This implies that inequality is assumed to stay constant.

The below example charts the consumption aggregate from the 2015.5 Ethiopian survey expressed in 2017 PPP-adjusted dollars, and the 2018 Ethiopian distribution one would get from extrapolating this welfare vector to 2018 using the method described above. The 2015.5 daily consumption mean is 3.12 and the growth rate in national accounts between 2015.5 and 2018 is 14.5%. Note that the positive growth pushes the distribution to the right. This changes the share of extreme poor – those living below the $2.15 poverty line – from 30.8% in 2015.5 to 22.3% in 2018.



When household surveys span two calendar years, as the example above, the national accounts data used for extrapolating this survey forward is the weighted average of the two years, with the weights determined by the share of the fieldwork that took place in each year (see the end of the section 1.1 for details).

5.2 Interpolations

In cases where the reference year falls between two surveys, poverty is interpolated for the reference year using the nearest survey on each side of the reference year. One of two approaches are used, depending on the correspondence in growth between national accounts and survey data.

"Same direction" interpolation is used when growth in the survey mean between the two surveys is of the same sign as (1) the growth in national accounts from the first survey to the reference year, and (2) from the reference year to the second survey. Denoting $survey1$ the survey predating the reference year and $survey2$ the survey after the reference year, this means that same direction interpolation is used when

$$sign\left(\frac{\mu_{survey2}}{\mu_{survey1}} - 1\right) = sign\left(\frac{NA_{survey2}}{NA_{reference}} - 1\right) = sign\left(\frac{NA_{reference}}{NA_{survey1}} - 1\right)$$

"Diverging directions" interpolation is used when the equation above does not hold. The main difference between the two is that "same direction" interpolation tries to estimate how much of the growth in national accounts that had accrued at the reference year, and uses this to back out a predicted survey mean at the reference year. When the survey mean and national accounts grow in opposite directions this is not meaningful, and an alternative approach is used.

Interpolations are never done between consumption and income aggregates. Whenever both are available, specific rules are used to determine which aggregate to use.

5.2.1 "Same direction" interpolation

"Same direction" interpolation works in three steps.

First, the survey mean at the reference year, $\mu_{reference}$, is estimated using the following interpolation formula:

$$\mu_{reference} = (\mu_{survey2} - \mu_{survey1}) \times \frac{NA_{reference} - NA_{survey1}}{NA_{survey2} - NA_{survey1}} + \mu_{survey1}$$

Second, the welfare distributions at the two surveys are scaled to reflect this mean:

$$f(y_{reference,i}) = \frac{\mu_{reference}}{\mu_{survey,i}} \times f(y_{survey,i}),$$

where $i = 1, 2$. After this alignment, there will be two distributions with the same mean for the reference year but with different distributional shapes.

At the third step, poverty is estimated from each of these distributions, and the final poverty estimate at the reference year is the weighted average poverty rate from both distributions where each poverty estimate is weighted by the inverse of the relative distance between the survey year and the reference year:

$$P_{0,reference} = \frac{P_{0,reference,1} * (t_{survey,2} - t_{reference}) + P_{0,reference,2} * (t_{reference} - t_{survey,1})}{t_{survey,2} - t_{survey,1}}$$

For example, if a reference year falls two years after the first survey and one year before the second survey, the poverty estimate from the first survey is given a weight of 1/3 and the estimate from the second survey a weight of 2/3.

5.2.2 "Diverging directions" interpolation

If the growth rates in surveys and national accounts diverge, an approach similar to the extrapolation is applied to the two closest surveys; poverty is extrapolated forward by the national account growth rate using the early survey and backwards using the later survey. Poverty for the reference year is estimated using both distributions and the estimates are averaged using the formula above.

The mechanics of the extrapolation and interpolation are described in Ravallion (2003), Chen and Ravallion (2004), box 6.4 in Jolliffe and Prydz (2015), and in Appendix A of World Bank (2018).

5.3 National accounts data

All of the interpolation and extrapolation methods described above rely on national accounts data. Two national accounts variables are used: Household Final Consumption Expenditure (HFCE) and Gross Domestic Product (GDP) – both in real per capita terms. Due to its close conceptual relation with the welfare aggregate, when available, HFCE is used as the preferred national accounts data for extrapolation and interpolation methods. In Sub-Saharan Africa, however, HFCE estimates at constant prices are both sparse and the empirical correlation between growth in HFCE and growth in the survey mean is not as strong (Ravallion 2003). Hence, GDP is preferred for this region. The World Development Indicators[3] (WDI) database is the primary source for the HFCE data (we use the Households and NPISHs Final Consumption Expenditure per capita (constant 2015 US$)[4] series with the series code NE.CON.PRVT.PC.KD) and

[3] https://databank.worldbank.org/source/world-development-indicators
[4] https://data.worldbank.org/indicator/NE.CON.PRVT.PC.KD

the GDP data (where we use the GDP per capita (constant 2015 US$)[5] series with the series code NY.GDP.PCAP.KD).

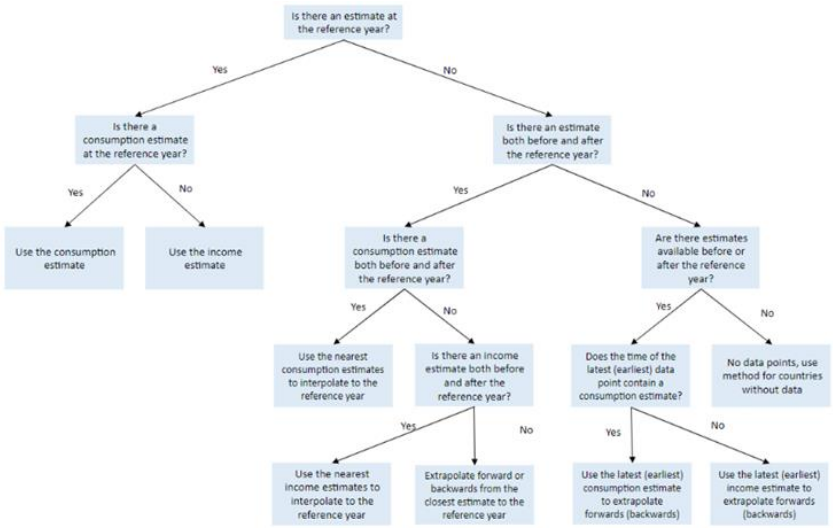
For a handful of countries, the data in WDI do not reflect calendar-year estimates but country-specific fiscal-year estimates. In those cases, the WDI data are converted to calendar-year estimates (see Castaneda Aguilar et al. (2022)). When WDI data are not available, we use data from the World Economic Outlook[6] and the Maddison Project Database[7]. For a detailed discussion on the sources of national accounts data see (Prydz et al. 2019).

It is important to note that the extrapolations and interpolations use only the growth rate from national accounts. In many countries, important differences in income levels between surveys and national accounts also exist (Jolliffe, Prydz, and Serajuddin 2019). Since the growth rate is the same whether PPPs or USD are used, it is immaterial that the constant USD series is used in the extrapolations and interpolations, while the survey welfare aggregate is expressed in PPPs.

5.4 Choosing between consumption and income estimates

The discussion on interpolation and extrapolation for simplicity assumed that only one welfare aggregate was available for a given country at a given year. Yet, a number of countries have poverty estimates available from both consumption and income aggregates. Due to its closer connection to welfare, whenever both income and consumption estimates are available for a given reference year, consumption estimates are preferred. Likewise, when both kinds of poverty estimates are available for the same years (but not for a particular reference year), interpolations and extrapolation are made using the consumption estimates. When consumption and income estimates are available at different years, the choice is a bit more complicated. The precise rules are outlined in the following decision tree:

[5] https://data.worldbank.org/indicator/NY.GDP.PCAP.KD
[6] https://www.imf.org/en/Publications/WEO
[7] https://www.rug.nl/ggdc/historicaldevelopment/maddison/?lang=en



5.5 Regional classification

In total, 218 economies are included in the global poverty estimates. The entire universe of economies considered corresponds to the universe of economies in the World Bank's Country and Lending Groups[8]. The regions employed in PIP, however, differ from the regional classifications used by the World Bank. Some economies, mostly high-income economies, are excluded from the geographical regions and are included as a separate group referred to as "other high income" (or "industrialized economies" or "rest of the world" in earlier publications). The list of economies included in each region is as follows:

- **East Asia and Pacific:** American Samoa; Cambodia; China; Fiji; Indonesia; Kiribati; Korea, Democratic People's Republic of; Lao People's Democratic Republic; Malaysia; Marshall Islands; Micronesia, Federated States of; Mongolia; Myanmar; Northern Mariana Islands; Palau; Papua New Guinea; Philippines; Samoa; Solomon Islands; Thailand; Timor-Leste; Tonga; Tuvalu; Vanuatu; Vietnam.
- **Europe and Central Asia:** Albania; Armenia; Azerbaijan; Belarus; Bosnia and Herzegovina; Bulgaria; Croatia; Czech Republic; Estonia; Georgia; Hungary; Kazakhstan; Kosovo; Kyrgyz Republic; Latvia; Lithuania; Moldova; Montenegro; North Macedonia; Poland; Romania; Russian Federation; Serbia; Slovak Republic; Slovenia; Tajikistan; Ukraine; Uzbekistan.

[8] https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

- **Latin America and the Caribbean:** Argentina; Barbados; Belize; Bolivia; Brazil; Chile; Colombia; Costa Rica; Cuba; Dominica; Dominican Republic; Ecuador; El Salvador; Grenada; Guatemala; Guyana; Haiti; Honduras; Jamaica; Mexico; Nicaragua; Panama; Paraguay; Peru; St. Kitts and Nevis; St. Lucia; St. Vincent and the Grenadines; Suriname; Trinidad and Tobago; Uruguay; Venezuela, Republica Bolivariana de.

- **Middle East and North Africa:** Algeria; Djibouti; Egypt, Arab Republic of; Iran, Islamic Republic of; Iraq; Jordan; Lebanon; Libya; Morocco; Oman; Syrian Arab Republic; Tunisia; West Bank and Gaza; Yemen, Republic of.

- **Other High Income (at times referred to as "Rest of the World"):** Andorra; Antigua and Barbuda; Aruba; Australia; Austria; Bahamas, the; Bahrain; Belgium; Bermuda; British Virgin Islands; Brunei Darussalam; Canada; Cayman Islands; Channel Islands; Curacao; Cyprus; Denmark; Faroe Islands; Finland; France; French Polynesia; Germany; Gibraltar; Greece; Greenland; Guam; Hong Kong SAR, China; Iceland; Ireland; Isle of Man; Israel; Italy; Japan; Korea, Republic of; Kuwait; Liechtenstein; Macao SAR, China; Malta; Monaco; Nauru; Netherlands; New Caledonia; New Zealand; Norway; Portugal; Puerto Rico; Qatar; San Marino; Saudi Arabia; Singapore; Sint Maarten; Spain; St. Martin; Sweden; Switzerland; Taiwan, China; Turks and Caicos Islands; United Arab Emirates; United Kingdom; United States; Virgin Islands, US.

- **South Asia:** Afghanistan; Bangladesh; Bhutan; India; Maldives; Nepal; Pakistan; Sri Lanka

- **Sub-Saharan Africa:** Angola; Benin; Botswana; Burkina Faso; Burundi; Cabo Verde; Cameroon; Central African Republic; Chad; Comoros; Congo, Democratic Republic of; Congo, Republic of; Cote d'Ivoire; Equatorial Guinea; Eritrea; Eswatini; Ethiopia; Gabon; Gambia, the; Ghana; Guinea; Guinea-Bissau; Kenya; Lesotho; Liberia; Madagascar; Malawi; Mali; Mauritania; Mauritius; Mozambique; Namibia; Niger; Nigeria; Rwanda; Sao Tome and Principe; Senegal; Seychelles; Sierra Leone; Somalia; South Africa; South Sudan; Sudan; Tanzania; Togo; Uganda; Zambia; Zimbabwe.

5.6 Population

To derive regional and global poverty rates, each country's poverty rate is weighted using its population share. Population data are taken from the World Bank's World Development Indicators[9] (WDI). For Kuwait and West Bank and

[9] http://datatopics.worldbank.org/world-development-indicators/

Gaza, population data is missing in the WDI for some years. In those cases, alternative sources are used as outlined in Arayavechkit et al. (2021).

The regional and global population may differ from the aggregates given in the World Development Indicators, because PIP uses different regional compositions of economies (See Regional classification section).

5.7 Treatment of countries without any poverty data

With the interpolation and extrapolation rules, all economies with at least one data point will have poverty estimates available for all years (conditional on national accounts data being present). Yet some economies have no household survey data that can be used to monitor poverty (or they lack credible PPPs or CPIs necessary to compute international poverty). In order to include these economies in the regional and global poverty estimates, it is assumed that they have a poverty headcount equal to their region's population-weighted average.

5.8 Coverage rule

Though the interpolation rules, extrapolation rule, and treatment of countries without any data assures that *some* poverty estimate is available for all regions and the world as a whole since 1981, the confidence one can have in these estimates depends on the how often and how long one has to extrapolate or interpolate the data, as well as the share of the population without any data at all.

To avoid presenting regional and global numbers based on outdated or no data, coverage rules are used to determine whether a particular reference year has sufficient data coverage. A country is considered to have sufficiently recent data if it has a survey-based poverty estimate at most three years from the reference year. Regional estimates are displayed for a given year if data cover 50 percent of the population in the region. For regions in which the surveys within three years either side of the reference year account for less than half of the regional population, the regional poverty estimate is not reported.

An additional coverage requirement is applied to govern when global poverty estimates are reported for a given reference year. This requirement addresses the goal of focusing the measurement of global poverty on economies where most of the poor live. Specifically, it tries to avoid a situation in which the global population threshold is met by having recent data in high-income countries, East Asia & Pacific, and Latin America & the Caribbean, which together account for a small share of the global extreme poor. Under this requirement, global poverty estimates are reported only if data are representative of at least 50

percent of the population in low-income and lower-middle-income countries, because most of the poor live in these groups of countries. The World Bank classification of economies according to income groups[10] in the reference year is used. This requirement is only applied to the global poverty estimate, not to regional estimates (for more details see Castaneda Aguilar et al. (2020)).

During the COVID-19 pandemic, countries experienced economic shocks and volatility at an unprecedented global scale, which means that extrapolating from a pre-COVID survey may work less well. For that reason, for 2020 and 2021, an annual coverage rule instead of a three-year window is used to decide on the regions for which to report post-2019 aggregates, such that only regions with sufficient data during or post-COVID have aggregates reported (Castaneda Aguilar et al. 2023). This is a conservative approach that is warranted by the extreme slowdown in surveys during COVID and the exceptional volatility in economic conditions.

Alkire, Sabina, José Manuel Roche, Paola Ballon, James Foster, Maria Emma Santos, and Suman Seth. 2015. *Multidimensional Poverty Measurement and Analysis*. Oxford University Press, USA. https://ophi.org.uk/multidimensional-poverty-measurement-and-analysis/.

Arayavechkit, Tanida, Aziz Atamanov, Karen Barreto Herrera, Nadia Belhaj Hassine Belghith, R. Andres Castaneda Aguilar, Tony Fujs, Reno Dewina, et al. 2021. "March 2021 PovcalNet Update: What's New." *Global Poverty Monitoring Technical Note* 15. https://openknowledge.worldbank.org/handle/10986/35340.

Aron, Danielle, R. Andres Castaneda Aguilar, Carolina Diaz-Bonilla, Maria Gabriela Farfan Betran, Elizabeth Mary Foster, Tony Fujs, Dean Jolliffe, et al. 2023. "September 2023 Update to the Poverty and Inequality Platform (PIP): What's New." *Global Poverty Monitoring Technical Note* 32. https://documents1.worldbank.org/en/publication/documents-reports/documentdetail/099624209142386941.

Atamanov, Aziz, R. Andres Castaneda Aguilar, Carolina Diaz-Bonilla, Dean Jolliffe, Christoph Lakner, Daniel Gerszon Mahler, Jose Montes, et al. 2019. "September 2019 PovcalNet Update: What's New." *Global Poverty Monitoring Technical Note* 10. https://openknowledge.worldbank.org/handle/10986/32478.

Atamanov, Aziz, Christoph Lakner, Daniel Gerszon Mahler, Samuel Kofi Tetteh Baah, and Judy Yang. 2020. "The Effect of New PPP Estimates on Global Poverty: A First Look" 12. https://openknowledge.worldbank.org/handle/10986/33816.

Balcázar, Carlos Felipe, Lidia Ceriani, Sergio Olivieri, and Marco Ranzani. 2017. "Rent-Imputation for Welfare Measurement: A Review of Methodologies and Empirical Findings." *Review of Income and Wealth* 63 (4): 881–98. https://onlinelibrary.wiley.com/doi/abs/10.1111/roiw.12312.

Castaneda Aguilar, R. Andres, Carolina Dias-Bonilla, Reno Dewina, Ifeanyi

[10] https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Edochie, Tony Fujs, Dean Jolliffe, Jonathan Lain, et al. 2022. "April 2022 Update to the Poverty and Inequality Platform (PIP)." *Global Poverty Monitoring Technical Note* 20.

Castaneda Aguilar, R. Andres, Carolina Diaz-Bonilla, Tony Fujs, Dean Jolliffe, Aphichoke Kotikula, Christoph Lakner, Gabriel Lara Ibarra, et al. 2023. "March 2023 Update to the Poverty and Inequality Platform (PIP)." https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099923403272329672.

Castaneda Aguilar, R. Andres, Tony Fujs, Dean Jolliffe, Christoph Lakner, Daniel Gerszon Mahler, Minh C. Nguyen, Marta Schoch, David L. Vargas Mogollon, Martha C. Viveros Mendoza, and Samuel Kofi Tetteh Baah. 2020. "September 2020 PovcalNet Update: What's New." *Global Poverty Monitoring Technical Note* 14. https://openknowledge.worldbank.org/handle/10986/34451.

Chen, Shaohua, and Martin Ravallion. 2004. "How Have the World's Poorest Fared Since the Early 1980s?" *The World Bank Research Observer* 19 (2): 141–69. https://doi.org/10.1093/wbro/lkh020.

———. 2008. "China Is Poorer Than We Thought, but No Less Successful in the Fight Against Poverty," Policy research working paper series, no. 4621. https://openknowledge.worldbank.org/handle/10986/6674.

———. 2010. "The Developing World Is Poorer Than We Thought, but No Less Successful in the Fight Against Poverty." *The Quarterly Journal of Economics* 125 (4): 1577–1625. https://doi.org/10.1162/qjec.2010.125.4.1577.

Datt, Gaurav. 1998. "Computational Tools for Poverty Measurement and Analysis." http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/125673/filename/125704.pdf.

De Weerdt, Joachim, John Gibson, and Kathleen Beegle. 2020. "What Can We Learn from Experimenting with Survey Methods?" *Annual Review of Resource Economics* 12 (1): 431–47. https://www.annualreviews.org/doi/10.1146/annurev-resource-103019-105958.

Deaton, Angus, and Salman Zaidi. 2002. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." *LSMS Working Paper*, no. 135. https://openknowledge.worldbank.org/handle/10986/14101.

Ferreira, Francisco HG, Shaohua Chen, Andrew Dabalen, Yuri Dikhanov, Nada Hamadeh, Dean Jolliffe, Ambar Narayan, Espen Beer Prydz, Ana Revenga, and Prem Sangraula. 2016. "A Global Count of the Extreme Poor in 2012: Data Issues, Methodology and Initial Results." *The Journal of Economic Inequality* 14 (2): 141–72. https://link.springer.com/article/10.1007/s10888-016-9326-6.

Gasparini, Leonardo, and Walter Sosa Escudero. 2004. "Implicit Rents from Own-Housing and Income Distribution: Econometric Estimates for Greater Buenos Aires." *Documentos de Trabajo Del CEDLAS* no. 14. http://hdl.handle.net/10915/3576.

Haughton, Jonathan, and Shahidur R Khandker. 2009. *Handbook on Poverty and Inequality*. World Bank Publications. https://openknowledge.worldbank.org/handle/10986/11985.

Jolliffe, Dean, Daniel Gerszon Mahler, Christoph Lakner, Aziz Atamanov, and Samuel Kofi Tetteh Baah [randomized order]. 2022. "Assessing the Impact of the 2017 PPPs on the International Poverty Line and Global Poverty," Policy research working paper series, no. 9941. https://openknowledge.worldbank.org/handle/10986/37061.

Jolliffe, Dean, and Espen Beer Prydz. 2015. *Global Poverty Goals and Prices: How Purchasing Power Parity Matters.* Policy Research Working Paper Series 7256. hhttps://openknowledge.worldbank.org/handle/10986/21988.

Jolliffe, Dean, Espen Beer Prydz, and Umar Serajuddin. 2019. "Mind the Gap: Disparities in Assessments of Living Standards Using National Accounts and Surveys." http://old.iariw.org/washington/Prydz-paper.pdf.

Lakner, Christoph, Daniel Gerszon Mahler, Minh C. Nguyen, Joao Pedro Azevedo, Shaohua Chen, Dean M. Jolliffe, Espen Beer Prydz, and Prem Sangraula. 2018. "Consumer Price Indices Used in Global Poverty Measurement." *Global Poverty Monitoring Technical Note* 4. http://documents.worldbank.org/curated/en/215371537208860890/Consumer-Price-Indices-Used-in-Global-Poverty-Measurement.

Lanjouw, Peter, and Martin Ravallion. 1995. "Poverty and Household Size." *The Economic Journal* 105 (433): 1415–34. https://doi.org/10.2307/2235108.

Mancini, Giulia, and Giovanni Vecchi. 2022. "On the Construction of a Consumption Aggregate for Inequality and Poverty Analysis." *World Bank Group, Washington, DC.* https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099225003092220001/p1694340e80f9a00a09b20042de5a9cd47e.

Prydz, Espen Beer, Dean Mitchell Jolliffe, Christoph Lakner, Daniel Gerszon Mahler, and Prem Sangraula. 2019. "National Accounts Data Used in Global Poverty Measurement." *Global Poverty Monitoring Technical Note.* Vol. 8. http://documents.worldbank.org/curated/en/664751553100573765/National-Accounts-Data-Used-in-Global-Poverty-Measurement.

Ravallion, Martin. 2003. "Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree?" *The Review of Economics and Statistics* 85 (3): 645–52. https://www.jstor.org/stable/3211703.

Ravallion, Martin, and Shaohua Chen. 1997. "What Can New Survey Data Tell Us about Recent Changes in Distribution and Poverty?" *The World Bank Economic Review* 11 (2): 357–82. https://www.jstor.org/stable/3990232.

Ravallion, Martin, Shaohua Chen, and Prem Sangraula. 2009. "Dollar a Day Revisited." *The World Bank Economic Review* 23 (2): 163–84. https://doi.org/10.1093/wber/lhp007.

Sinha Roy, Sutirtha, and Roy Van Der Weide. 2022. "Poverty in India Has Declined over the Last Decade but Not as Much as Previously Thought," Policy research working paper series, no. 9994. https://openknowledge.worldbank.org/handle/10986/37273.

Wolfson, Michael C. 1994. "When Inequalities Diverge." *The American Economic Review* 84 (2): 353–58. https://www.jstor.org/stable/2117858.

World Bank. 2018. *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/30418.