UNITED NATIONS DEVELOPMENT PROGRAMME

DEVELOPMENT FUTURES SERIES

FEBRUARY 2022

UNDP Global Policy Network Brief

The Inequality Gap: The Bottom 40 May Be Further Away Than We Thought

by Christian Oldiges and Shivani Nayyar[1]

This paper explores new data on income inequality by the World Inequality Database, which corrects underreporting of income in the top deciles of the income distribution. We find that within all low and middle income countries, the bottom 40 income shares are much lower than we previously thought, while the top 10 income shares are much higher. Important for Sustainable Development Goal 10.1, the bottom 40 income shares have been growing at a much slower pace than estimated earlier and often at a lower rate than the top 10 shares. Demonstrating the value of improved datasets, this paper calls upon practitioners to have these enhanced data and metrics in their methodological toolbox.

1 Introduction

The COVID-19 pandemic and the economic fallout from it have caused setbacks in terms of human development and the achievement of the Sustainable Development Goals (SDGs). At the peak of the crisis, roughly 1.5 billion children were out of school (UNESCO, 2020), and across countries, female labour force participation fell (UN Women, 2021). Decades of poverty reduction have been reversed, and approximately 100 million people have been pushed into extreme poverty (Mahler, Yonzan, Lakner, Castaneda Aquilar and Wu, 2020; World Bank, 2021), while global multidimensional poverty reduction is estimated to see a setback of almost ten years (Alkire, Nogales, Quinn and Suppa, 2021). At the same time, the impacts of the COVID-19 pandemic have not been even, as the effect on human lives is mediated by existing inequalities in human capabilities. For example, due to differences in education, training, skills and internet access (Hatayama, Viollaz and Winkler, 2020), as well as varying labour market structures and social protection systems, only a privileged minority has been able to socially distance and conduct their work and lives digitally. Globally, only 20 percent of jobs can be done at home, according to Dingel and Neiman (2020), 37 percent in the United States, while in low-income countries, this holds for only one in every 26 jobs (Garrote Sanchez, Gomez Parra, Ozden, Rijkers, Viollaz and Winkler, 2021).

Threatening livelihoods and well-being, there is wide recognition that COVID-19 has exacerbated pre-existing and systemic inequalities. The latter includes both income inequality[2] as well as inequalities in education and digital literacy, healthcare and living

standards including internet access (see Stantcheva (2022) for a comprehensive overview for pandemic-related increases in inequalities). Partly as a response to this, there have been calls for better measures and metrics for inequality (United Nations, 2021). Understanding the extent of inequality and tracking the trends is a prerequisite for comprehending the impact of shocks such as the COVID-19 pandemic and in preventing further deepening of inequalities.

This paper takes a step towards improved measurement of income inequality using data from the World Inequality Database (WID) as applied in the recently launched and UNDP-supported World Inequality Report 2022 (Chancel, Piketty, Saez, Zucman et al., 2021), compiled by the World Inequality Lab and available at https://wid.world/. Traditionally, measures of income inequality are based on income and household consumption surveys. These surveys usually significantly underestimate incomes of the rich, those at the top of the income distribution. The rich are often not part of the sample of households that are interviewed. Moreover, the survey results are based on self-reports of income for rich households, who when interviewed, have incentives to under-report their income for various reasons[3]. The major novelty of WID data is that they account for underestimation of incomes in top deciles and make adjustments for it, including incorporating tax-based information for the top part of the income distribution where available.

This paper makes three main contributions. One, we provide a proof of concept using this new source of data on income distributions for a development application. We call upon development practitioners to have these datasets and innovative measures in their toolkit, to gain an in-depth country level understanding of income distribution and inequality trends.

Second, we present interesting insights that are gained by using these data over a more standard, traditional data source, PovcalNet from the World Bank[4]. Our main results pertain to the top decile and the bottom 40 of the income distribution (to shed some light on SDG 10.1)[5]. We find that the income share of the top decile is higher than what appears in traditional data. This is the case across regions. This result is probably explained by the WID methodology and the adjustments with regards to the top income distribution.

We also see that over 2000-2021, the share of the top decile has grown faster than what the traditional data indicate. Conversely, we see that the share of the bottom 40 is lower than seen in PovcalNet, and over 2000-2021, it has grown more slowly than what was previously thought. Some of this may be a natural result of the underlying WID methodology and needs to be taken into account by researchers and policy-makers alike when choosing a data source.

Third, we produce growth incidence curves over 2000-2021 for those select countries that made tax data available. WID data show how growth rates at the tails of the income distribution are quite different once we move away from relying on just household surveys. Based on traditional data, it was believed that the income of the bottom 40 were growing at a relatively fast pace since 2000, a sign that SDG 10.1 was progressing along well. Our results based on WID data show that this is not the case and that the inequality gap has not been narrowing.

Finally, we discuss policy implications arising from the particular insights based on the analyses in the paper. The paper calls for better data to expand the number of countries with data that accurately reflect the incomes and wealth across the entire distribution. Going beyond this, an action agenda is proposed to tackle inequality from multiple angles with a better understanding of the factors that drive it.

2 New insights from the WID

In this section, we present new findings on inequality measures as derived from the WID. We show the new data can be used to measure progress towards SDG 10.1 by focusing on the bottom 40 percent and, for example, its relation to the top 10 percent of each country's income distribution. Throughout, we highlight how levels (Section 2.2) and trends (Section 2.3) of bottom 40 and top 10 income shares, as estimated by WID, differ from earlier estimates on inequality that rely solely on household surveys (e.g. PovcalNet).

2.1 WID methodology

Traditionally, income inequality estimates can be sourced from the World Bank's PovcalNet, the Organization of Economic Cooperation's and Development's Income Distribution Database, the UNU-WIDER World Income Inequality Database (WIID) and the Luxembourg Income Study Database (LIS). All of these sources rely almost exclusively on one source of information – household surveys that interview people about their consumption, income, wealth and other aspects of their lives.

WID data, on the other hand, are based on a combination of national accounts data, survey data and tax data when available. WID's annual estimates of the distribution of income and wealth rely on a Distributional National Accounts (DINA) methodology[6]. This method allows for the alignment of macroeconomic national accounts with information from micro household surveys. For a typical country, the following information on the adjustment of PovcalNet data is provided, This is available when the WID country level data are downloaded, as for the purpose of this paper: "Figures are obtained by correcting survey tabulations provided by the World Bank (PovcalNet) to account for conceptual discrepancies and the underrepresentation of top incomes. Surveys are available for the following years [example]: 1994, 1998, 2003, 2009, 2014. Income shares are interpolated linearly when surveys are available at the beginning and at the end of a given period. Inequality series are extrapolated backwards to 1990 and forwards to 2021 by keeping income shares constant when no data is available for these years."

In our sample of 115 low and middle income countries in 2021, estimates for 15 countries rely solely on regional imputations due to the lack of household surveys and tax data[7]. WID has assigned data quality scores ranging from 0 (least quality) to 4 (high quality), with the regional imputation based estimates assigned a 0. For 83 countries, estimates rely on 'adjusted surveys', while one is based on 'rescaled fiscal income' (South Africa, data quality score of 3). For 16 countries, tax data have been used for the adjustment and 11 of these country level estimates receive a data quality score of 4.

Convenient for researchers and practitioners, WID data are made available annually – relying on distribution neutral growth imputations for years between surveys – whereas traditional survey estimates are available for the years of the survey, usually at intervals of three to five years. For many countries with gaps in data collection and/or availability, as for example in India, where the main consumption survey conducted by the National Sample Organization has not been made available since 2011 (Drèze and Somanchi, 2021), the WID estimates provide additional and annual information until 2021[8].

While WID's annual estimates thus appear to have advantages over traditional income sources, users should, however, treat those that are based on interpolations or extrapolation with caution, as WID stresses, "These estimates, especially at the level of individual countries, can be fragile."[9] Thus, WID data are most reliable for policy advocacy when PovcalNet data (household surveys) are adjusted with tax data, at least for estimates of the higher end of the income distribution. At the moment, this is the case for a limited number of developing countries, which include, for example: Ivory Coast, India, Brazil, Chile, Colombia, Ecuador, Peru, and Uruguay. For these countries, the new insights can provide helpful guidance. Yet, even for countries without available tax data, the estimates based on adjusted PovcalNet data seem reliable and robust, as we show in our analysis.

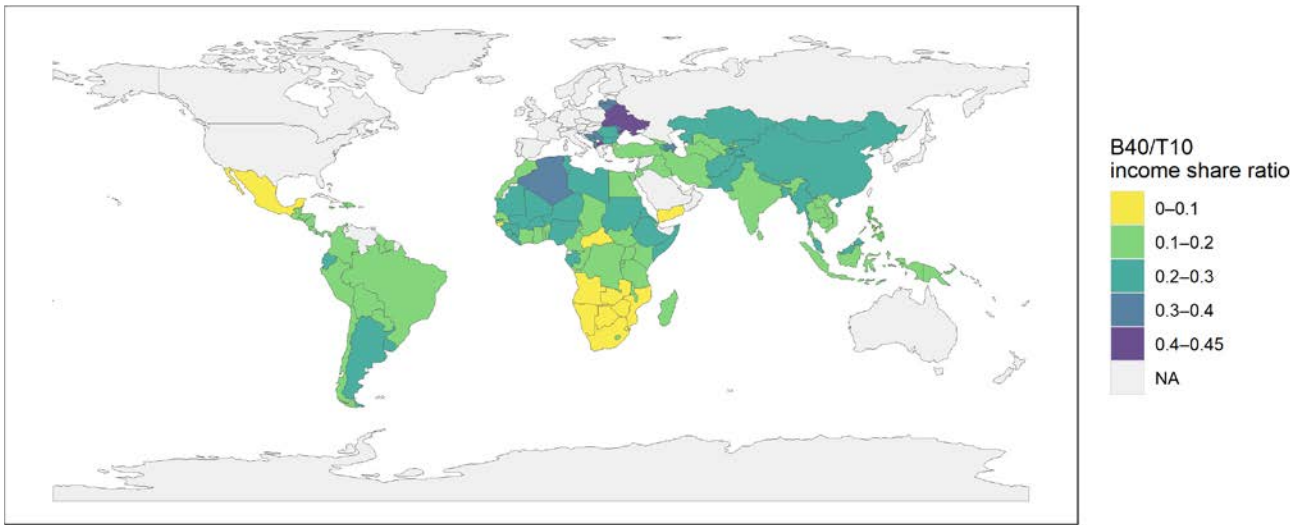
2.2 Levels and ratio of B40 and T10 income shares

The WID provide for annual snapshots of inequality allowing the end-user to define a measure of inequality. Keeping the SDG target of 10.1 in mind, we focus on the country-level income share of the bottom 40 percent (B40) and its relation to the top 10 percent (T10). We use the latest WID data available, pertaining to the year 2021, which in many instances is based on extrapolation of earlier data with the assumption of distribution-neutral growth. As shown in Figure 1, across all countries of interest, T10 shares are higher than B40 shares, since all B40/T10 ratios are far below one[10]. The lowest ratios of less than 0.1, implying that T10 shares are at least ten times higher than B40 shares, can be found across nearly all Southern African countries, and, for example, in the Central African Republic, Mexico, and Yemen. Higher ratios of between 0.2 and 0.25, which imply that T10 shares are 'only' four to five times larger, can be found in, for example, Argentina, across Western and Eastern Africa as well as in Central and Eastern Asia. Thus, overall and across all developing countries, T10 shares are at least three times larger than B40 shares.

Digging deeper and based on appendix Table A.1, we find that WID estimates for B40 income shares hover around only 12 percent in Europe and Central Asia, only half as much in Latin America and the Caribbean, around 9 percent in Middle East and North Africa, around 10 percent in South Asia, and even less in many of the Sub-Saharan African countries. On the other hand, across all world regions the T10 income share is largely between 35 and 60 percent and in some Sub-Saharan African nations it is even higher (e.g. Central African Republic, Mozambique, Namibia).

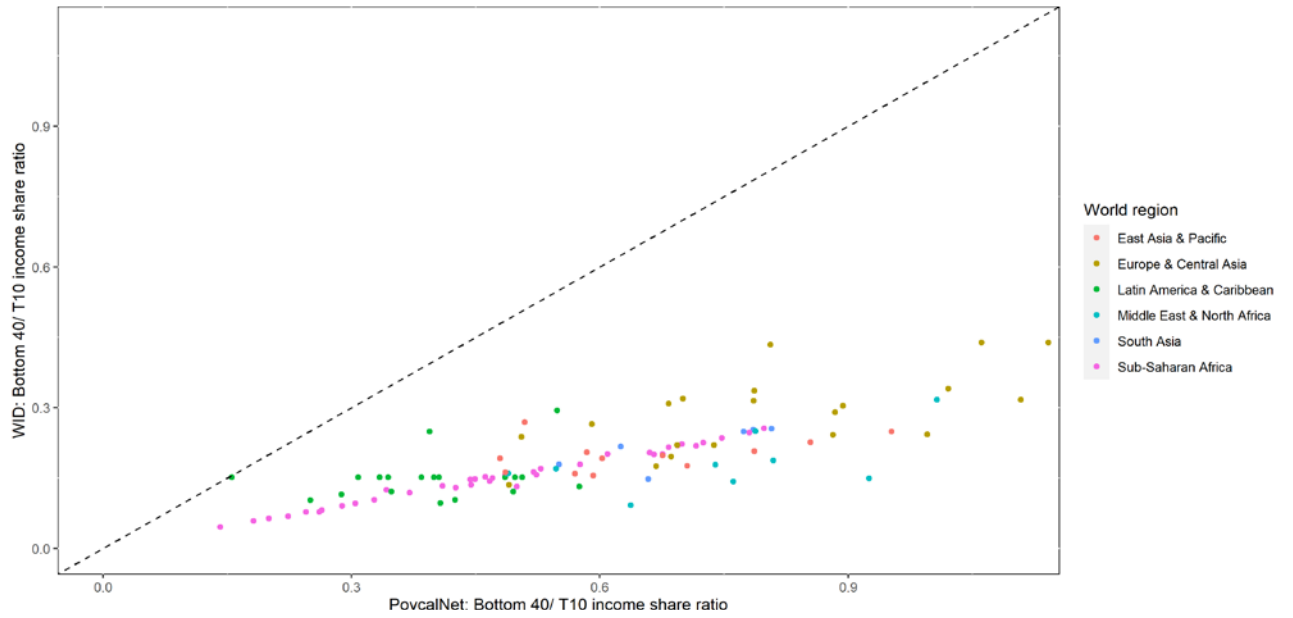
In the following, we examine how these insights differ from earlier estimates that rely solely on household surveys.

Figure 1: Ratio of B40 and T10 income shares in UNDP focus countries



Source: Authors' calculations based on latest WID data for 2021.

Figure 2: B40 and T10 income share ratio: WID and PovcalNet compared



Source: Authors' calculations with latest PovcalNet data and 2021 WID data.

Starting with Figure 2, we plot country-level B40/T10 ratios as derived with WID over B40/T10 ratios derived from PovcalNet, made publicly available by the World Bank and based on household surveys[11]. The dashed diagonal line indicates equal ratios, yet evidently, WID estimates are always lower than PovcalNet's (bar one exception). This implies that according to WID estimates inequality is higher than we thought. B40 income shares are lower than we had thought and T10 shares are much higher than presumed earlier.

Examining where the differences between PovcalNet and WID are largest (see Table 1), we notice that across world regions, T10 shares can be up to 80 percent larger (Middle East and North Africa) and at least 40 percent larger (Latin America and Caribbean). Average WID estimates of the T10 shares hover between 40 and 50 percent across world regions according to WID, whereas according to PovcalNet, the range is about 15 percentage points less (25 to 35 percent).

Table 1: Regional averages in B40 and T10

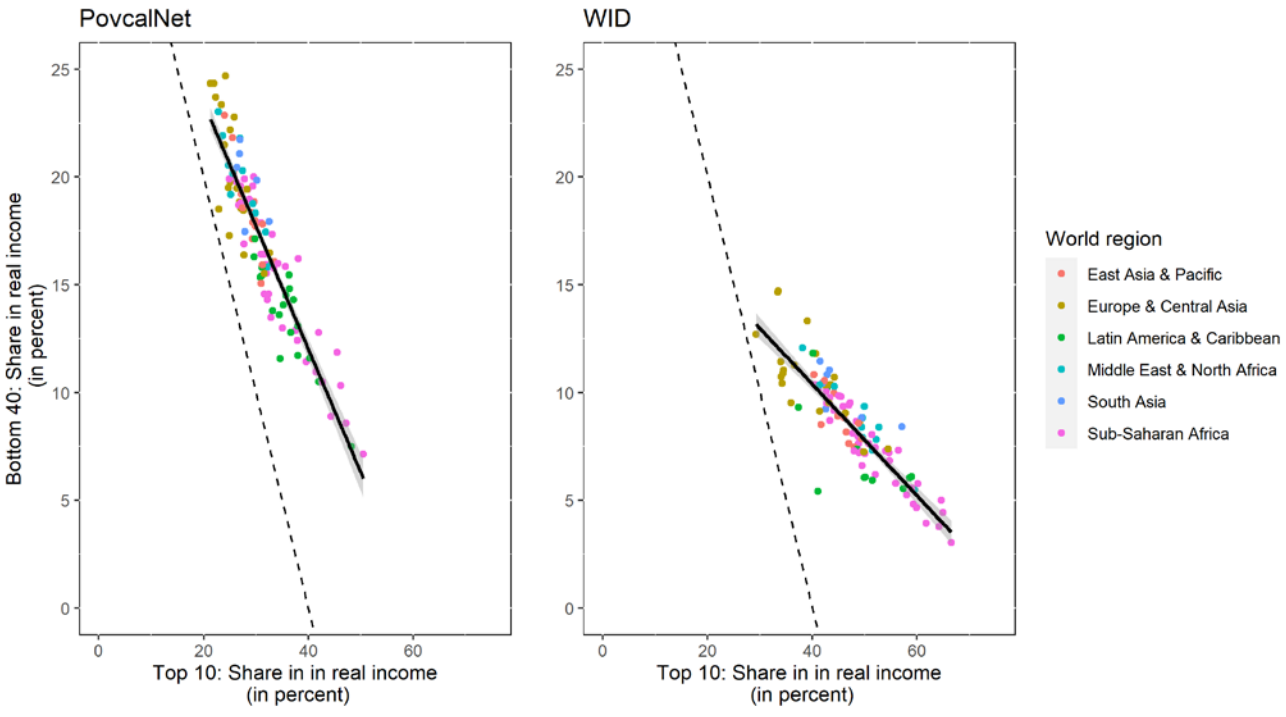
Region	B40			T10		
	WID	*PovcalNet*	*Rel. Diff.*	*WID*	*PovcalNet*	*Rel. Diff.*
East Asia & Pacific	8.9	18.4	-50.9	45.2	28.8	58.1
Europe & Central Asia	10.9	20.3	-46.1	39.2	25.7	52.5
Latin America & Caribbean	7.1	13.8	-46.7	49.1	35.6	39.1
Middle East & North Africa	8.3	19.4	-56.9	50.0	27.7	81.9
South Asia	9.8	19.8	-50.4	46.2	28.5	62.2
Sub-Saharan Africa	7.4	15.3	-52.2	51.5	34.2	52.0

Note: Authors' calculations with WID 2021 and latest PovcalNet.
Regional averages are simple averages of country estimates, and not population weighted.
Relative differences (Rel. Diff.) are in percentage of PovcalNet.

WID estimates of B40 income shares are always lower than PovcalNet's, by about 50 percent across all world regions (Table 1), implying that WID estimates are half as large as PovcalNet's. In most world regions, the average B40 share is close to 10 percent according to WID, whereas according to PovcalNet, this is around 20 percent. To check whether these results are caused by differences in data quality, we plot Figure 2 again, yet this time by the WID-level of data quality (see appendix Figure B.1). We find that neither the estimates of high quality nor those of lesser quality (reliant on regional imputations for lack of better data) seem to be outliers or potential drivers of any results.

WID data results in a higher T10 share, which is not surprising. As mentioned in Section 2.1, WID data adjusts for the underestimation of top incomes in underlying surveys. Higher T10 shares imply lower shares for some other part of the income distribution. This could be the middle or the bottom (or both). We do see lower B40 shares. It is important to keep in mind that this may be a result of the WID methodology and adjustment; however, we cannot be sure.

Figure 3: Income share of B40 and T10 percent for latest year available



Source: Authors' calculations with latest PovcalNet and 2021 WID data.

In Figure 3, the two key insights are emphasised visually. By plotting B40 income shares over T10 income shares separately with PovcalNet and WID, it is evident that WID-based T10 income shares are moved further to the right (and are thus higher) and B40 shares further down (thus lower)[12].

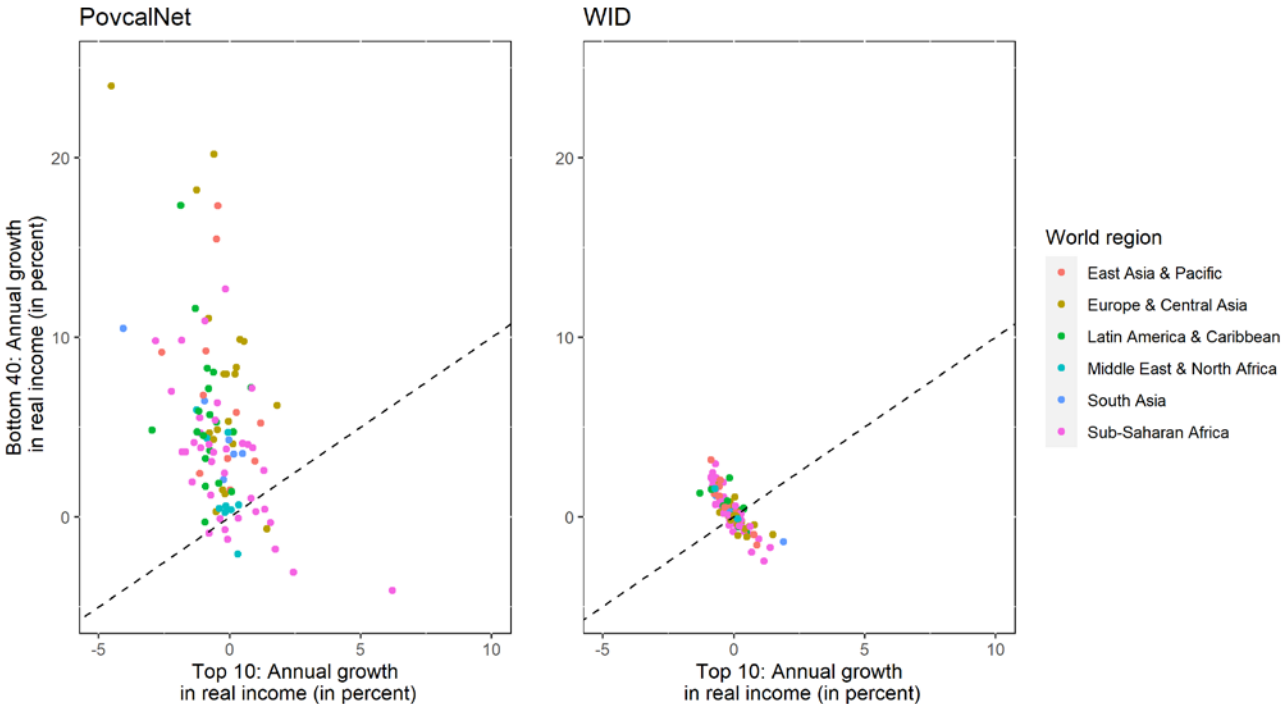
The changed nature of the B40 and T10 relationship carries additional implications. While the slope of the PovcalNet estimates is nearly parallel to the dashed line (-1), which indicates a one to one relationship, the slope as derived from WID is much flatter. With the relationship more tilted towards higher T10 shares and lower B40 shares, across-country differences are much larger than originally thought. Interpreting the slope across countries, a 10 percent reduction in the T10 income share would no longer translate into a 10 percent but just a 2.5 percent increase in the B40 share. Thus, increasing the B40 share via redistribution may be more challenging than we thought. Beyond the scope of this brief, it is an insight that deserves further research and attention.

2.3 SDG 10.1: Towards faster growth for B40?

Having examined the latest levels of inequality in terms of B40 income shares and T10 income shares, we now make use of the time series provided by WID to analyse trends in B40 as well as T10 income shares at the country level. While the B40/T10 ratio is widely applied (e.g. World Bank, 2018), growth rates of the two (B40 and T10) have rarely been studied together. With SDG 10.1 in mind, growth rates are of great interest, as faster growth for B40 shares is advocated for. As described by James Foster and Nora Lustig, the ratio of top 10 and bottom 40 makes for interesting comparisons, in particular from a communications point of view, despite not fulfilling certain axioms of inequality measurement (UNDP, 2019).

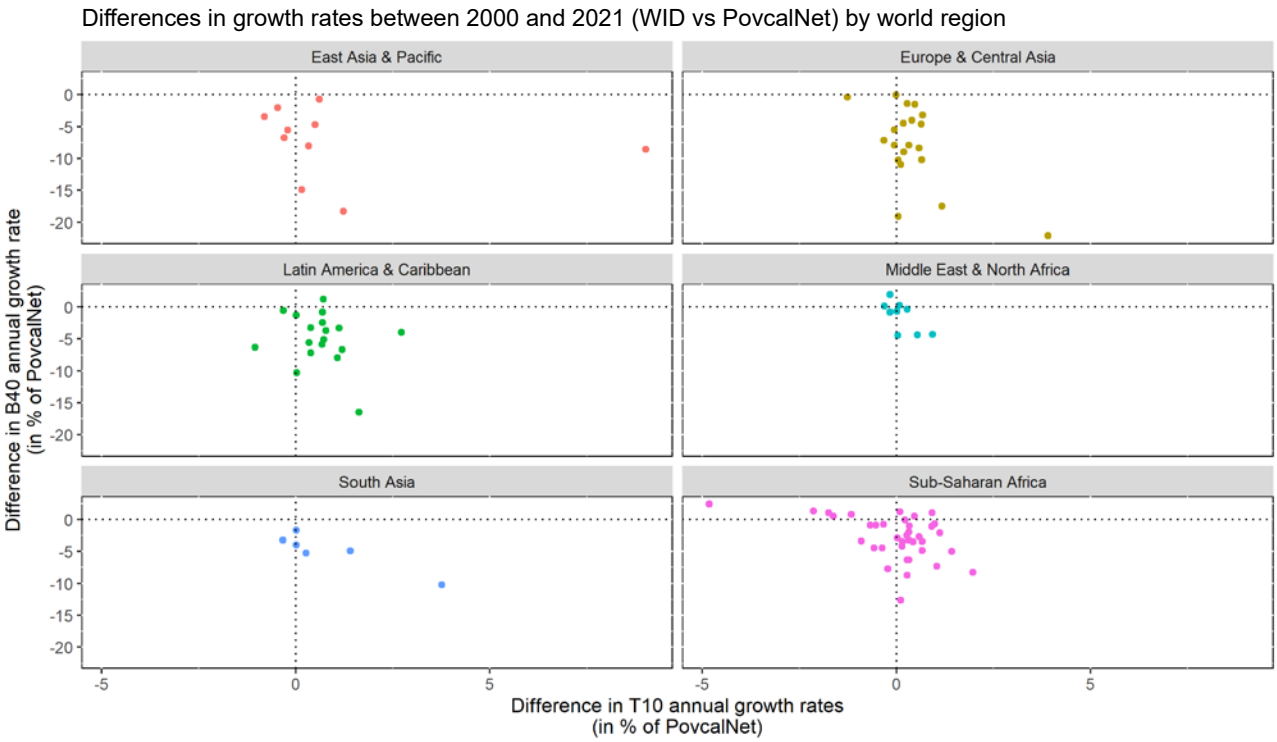
Plotting B40 annual growth rates over T10 annual growth rates yields the following insights. As shown in Figure 4, according to PovcalNet estimates (left panel), B40 growth rates are indeed much larger than T10 growth rates. Visually, these are shown above the dashed line, which indicates equal growth rates. On the basis of these data, it would appear that the world is making strong progress on SDG 10.1. WID estimates (right panel), on the other hand, suggest that B40 and T10 growth rates are much more similar, as country-level estimates tend to scatter more around the dashed line. Therefore, according to WID estimates, B40 income shares have been growing much more slowly than we thought. The outcome is, of course, the much lower B40 income share in 2021, as discussed in the previous section and the higher demands on achieving a higher income share for the B40 group. This shows that over 2000-2021, the bottom 40 have not made the kind of gains in terms of income share as PovcalNet data showed.

Figure 4: B40 growth rate and T10 growth rate between 2000 and 2021, PovcalNet and WID



Source: Authors' calculations with latest PovcalNet and 2021 WID data.

Figure 5: Difference in growth rates



Source: Authors' calculations with PovcalNet and 2021 WID data.

Examining from where the differences in growth rates stem, we plot the difference between WID and PovcalNet estimates as a percentage of the PovcalNet estimate for country-level B40 and T10 growth rates (Figure 5) by world region. Across world regions T10 growth rates are usually higher when WID data are applied, i.e. further to the right on the x-axis. In Latin America and the Caribbean, the majority of countries had higher T10 growth rates with WID, whereas nearly all B40 growth rates were lower according to WID. In Sub-Saharan Africa, almost all countries have lower or the same B40 growth rates, and about 60 percent have higher T10 growth rates. In South Asia, all countries have lower B40, and all but one have higher or the same T10 growth rates. In Europe and Central Asia, all countries have lower B40 growth rates, while for the majority of countries WID estimates of T10 growth rates are higher. Growth rates for countries in the Middle East and North Africa do not differ much between WID and PovcalNet.

Growth incidence curves at country-level

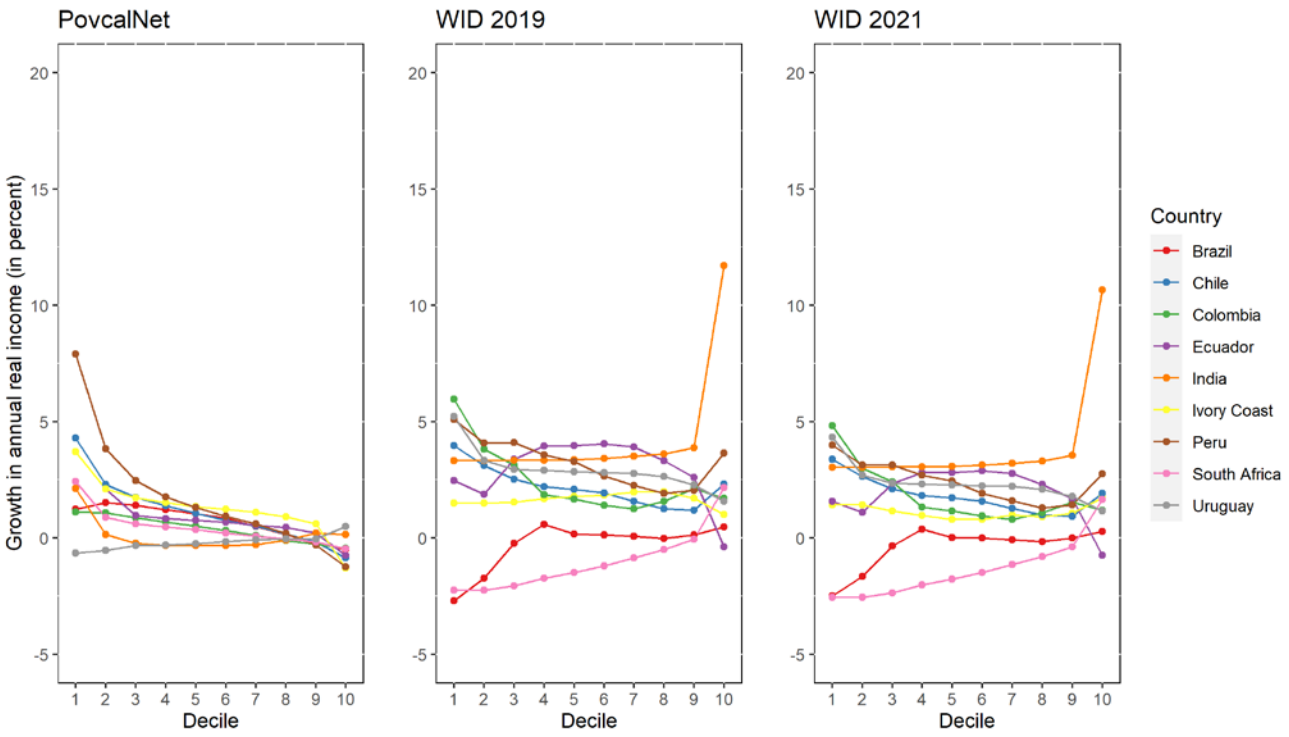
Another way to analyse inequality over time is to take into account growth rates for each part of the income distribution. One can do so via growth incidence curves (GICs) which plot growth rates over each percentile or decile, applied widely so to show income distributions over time – see for example the 'elephant curves' by Lakner and Milanovic (2016) for the entire world with data from PovcalNet as well as other sources. Due to imputations and extrapolations, WID allows for retrieval of income shares for every part of the income distribution for more countries than available at PovcalNet and at an annual basis up to 2021. In the following, we make use of the most complete time period, 2000 to 2021, and produce GICs for a limited number of countries. For the purpose of producing the most robust GICs, we use only estimates for countries with the two highest data quality scores (3 and 4).

In Figure 6, we plot GICs with PovcalNet (left panel) and WID (right panel) for several countries, for which both PovcalNet and WID data are available and WID rely on tax data, including several South American countries (Brazil, Chile, Colombia, Ecuador, Peru, and Uruguay), as well as India, Ivory Coast, and South Africa (rescaled fiscal income). Visibly, our earlier observations on just the B40 and T10 shares are also reflected in the GICs. For most estimates presented here, all of which rely on tax data, we see differences in the tails of the income distributions.

Whereas India accounts for a sharp upward tick in the last decile, a small downward slope is visible for Ivory Coast and Ecuador (WID). Interestingly, only for Brazil, the WID estimate of the GIC has the famous elephant shape with relatively high growth rates for the lower-middle and the top parts of the income distribution. It has to be noted here that PovcalNet and WID rely on slightly different years.

For example, WID's growth estimates for India rely on the time period 2000 to 2019/2021, whereas PovcalNet relies solely on the period 2004 to 2011, the years of the most recent household surveys. This may explain the major shift in levels for India. After all, India witnessed high growth rates in the period between 2011 and 2019.

Figure 6: Growth incidence curves for select countries for time period 2000 to 2019/2021



Source: Authors' calculations based on latest PovcalNet and WID data.

3 Discussion