

Gradín, Carlos

Working Paper

Trends in global inequality using a new integrated dataset

WIDER Working Paper, No. 2021/61

Provided in Cooperation with:

United Nations University (UNU), World Institute for Development Economics Research (WIDER)

Suggested Citation: Gradín, Carlos (2021) : Trends in global inequality using a new integrated dataset, WIDER Working Paper, No. 2021/61, ISBN 978-92-9256-999-0, The United Nations University World Institute for Development Economics Research (UNU-WIDER), Helsinki, <https://doi.org/10.35188/UNU-WIDER/2021/999-0>

This Version is available at:

<https://hdl.handle.net/10419/243387>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WIDER Working Paper 2021/61

**Trends in global inequality using a new
integrated dataset**

Carlos Gradín*

March 2021

Abstract: This paper presents preliminary evidence of the annual global income distribution since 1950 using a new integrated dataset that aggregates standardized country income distributions at the percentile level estimated from various sources in the World Income Inequality Database. I analyse the extent to which the main global inequality trends depend on specific distributive views, i.e. absolute or relative, or with more emphasis in specific parts of the distribution. The results show absolute inequality increasing almost continuously. Relative inequality has been clearly reduced in the long term but with most indices also showing a sharp decline since at least end of 1990s due to income differences between countries being narrowed. I also quantify the contribution of different countries and regions through changes in average income, in inequality or in population, showing that the main global trends are largely driven by the drastic changes experienced by China and India over these decades.

Key words: global inequality, income distribution, World Income Inequality Database

JEL classification: D31, D63, I31, O15

Supplementary material: The data appendix is available on the working paper's webpage—<https://www.wider.unu.edu/publication/trends-global-inequality-using-new-integrated-dataset>.

* UNU-WIDER, gradin@wider.unu.edu

This study has been prepared within the UNU-WIDER project [World Income Inequality Database \(WIID\)](#).

Copyright © UNU-WIDER 2021

UNU-WIDER employs a fair use policy for reasonable reproduction of UNU-WIDER copyrighted content—such as the reproduction of a table or a figure, and/or text not exceeding 400 words—with due acknowledgement of the original source, without requiring explicit permission from the copyright holder.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-999-0

<https://doi.org/10.35188/UNU-WIDER/2021/999-0>

Typescript prepared by Lesley Ellen.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

1 Introduction

The analysis of the global income distribution is the study of inequalities among all citizens in the world regardless of where they live. Since it started in the early 2000s, it has increasingly become a key topic of research with particular repercussions in the public and policy debates. The growing interest in assessing how income distribution changes over time among the world's population, and how this is shaped by countries' policies and by global megatrends like globalization or technological change, is not exempt from contentious discussions fuelled by a lack of appropriate data and different approaches taken or periods analysed. The lack of adequate data and a unified approach makes measuring global inequality quite complex. However, this does not stop some stylized facts emerging from the existing literature.

Taking a historical perspective, Bourguignon and Morrisson (2002) show a long-term trend of increasing inequality until 1980 followed by stagnation between 1980 and 1990 and posterior decline, according to survey data used by Bourguignon (2015). Anand and Segal (2008) also discuss in detail different measurement issues and review the earliest estimates of global inequality; for example, they provide ten series for the Gini index from eight different studies published between 2002 and 2006.¹ Although there is no general common trend, this evidence also points to an increase in inequality prior to 1970 or 1980, along with declines between 1990 and 2000. However, there is mixed evidence on what happened between 1980 and 1990 or 1995. The most recent increases in global inequality are also consistent with other findings, such as the analysis by Lakner and Milanovic (2016) for 1988–2008, later extended to 2013 in World Bank (2016), which takes a similar approach.

Most of the previous analyses were conducted using the Gini index as the main measure, although there has been increasing use of other indices with different distributive implications that can alter the conclusions about the trend in certain periods. This can be clearly seen in the stream of literature which moves the focus of inequality analysis to the concentration of income at the top of the distribution and combines survey data with information from tax administration and national accounts (World Inequality Lab 2018). It can also be seen in the approach in which the focus has shifted from a relative to an absolute view of inequality (as, for example, in Ravallion 2004, 2018, 2021 and Niño-Zarazúa et al. 2017).

This paper aims to contribute to this growing literature in different ways. First, I introduce a new integrated dataset on global inequality (percentile shares and summary measures) which provides annual information from 1950 to the present time. The dataset can easily be updated in successive years, enabling appropriate monitoring of past, current, and future inequality trends within and between countries in line with Sustainable Development Goal 10 (reducing inequalities within and between countries). This new dataset is a companion to the classical database from which it is obtained, i.e. the World Income Inequality Database (WIID). The WIID is the successor of the repository initially put together by Deininger and Squire (1996), which has already been widely used in different forms in this literature.²

¹ Anand and Segal (2008: Figure 1, page 62, and Table 1, page 63).

² Among the initial studies cited in Anand and Segal (2008: Table 3, page 75), these two datasets were used, for example, in Chotikapanich et al. (1997), Dowrick and Akmal (2005), Sala-i-Martin (2006), and Schultz (1998). They were used more recently in Davies and Shorrocks (2021), Jordá and Niño-Zarazúa (2019), Niño-Zarazúa et al. (2017), and Roope et al. (2018), among others.

This new companion global database, publicly available along with the main WIID compilation, presents estimates of the world's income distribution for nearly all current countries along with country aggregates by geographical region and income group. It uses rich within-country distributive information for each country based on household surveys. The income distribution in survey years is estimated at the percentile level from selected series of income shares (mostly by deciles, at least by quintiles) reported by various sources (such as the Luxembourg Income Study (LIS), PovcalNet, Eurostat, United Nations, and research studies, etc.) which best represent income distribution in each country over time. The original income distributions, which are heterogeneous across welfare concepts and other methods, have been adjusted in a simple and transparent way to allow more consistent comparisons across countries and over time. To maintain a balanced panel and avoid sample composition effects, missing country-year income distributions have been either interpolated (between closest survey years) or extrapolated (keeping the distribution constant before the first available survey or after the last one). The resulting dataset is a balanced panel of countries over time between 1950 and 2019. The actual survey-based information is rich enough to guarantee that there is a survey year falling within a bandwidth of five years from the target year for more than 50 per cent of the world population after 1950, reaching nearly 100 per cent in most of the 2000s. The database uses annual per capita income information from an integrated series of gross domestic product (GDP) estimates expressed in 2017 purchasing power parity (PPP) USD (based on the World Development Indicators and extended by the Maddison project and Penn World Tables whenever necessary).

Second, the paper uses this WIID Companion global dataset to provide a general overview of the long-term and short-term trends in global inequality, the contributions of its different components, and inequality between and within countries. It also quantifies the contribution of the main countries as well as country aggregates by region and income group to changes in total inequality or in each of its components. This overview takes a broad approach which incorporates analysis of the entire distribution and different summary inequality measures, income shares, and income share ratios. It attempts to fully describe the distributional changes that occur at different points of the distribution and establish the robustness of the inequality results to different inequality approaches, representing legitimate, and sometimes conflicting, normative views. This variety of approaches includes analysis of both absolute and relative inequality, as well as distributive sensitivities which put more emphasis on different parts of the distribution (i.e. the bottom, middle, or top).

The paper's results are highly consistent with previous evidence based on survey data. However, they enable a more detailed, systematic, and comprehensive analysis of the patterns in the global distributive trends and their between- and within-country components in terms of time and geographical coverage as well as the distributive approaches that can be used. The dataset connects the global and country income distributions, making the global trends highly consistent with the trends reported by country-level sources. This makes it easier to trace any distributive pattern found at the global level back to its origin at the country level. The annual estimates can be easily revised or updated as new or better information becomes available.

Absolute inequality, which requires larger dollar increases among poorer people over time for inequality to decline in a context of economic growth, has continuously increased since 1950, apart from short episodes around the main global economic recessions. This increase has affected both between- and within-country components, which reinforce each other. Lorenz dominance is the norm in this approach, indicating that the trend is unanimous and not affected by different distributive sensitivities.

A totally different storyline emerges, however, in terms of relative inequality, which requires higher relative growth of lower incomes for inequality to decline. The results show that after several

decades of increasing inequality, or with mixed evidence, that inequality has recently started to sharply decline. Strict Lorenz dominance occurs only in the long term, pointing to an overall decline but is rare when comparing ten-year periods, implying a more nuanced story if the focus is at the extremes of the distribution. However, the main trend is highly robust to the use of several popular inequality indices. The initial year in which global inequality starts to decline varies depending on the index used, and therefore on the different weights attached to changes affecting different parts of the distribution. The decline starts earlier (in 1976) if we factor in the substantial improvement of the bottom 40 per cent of the population (e.g. mean log deviation (MLD) or Palma index) or later (in 1998) if we consider the higher concentration of income at the top that occurred in the 1990s (e.g. coefficient of variation). If no particular focus is placed at either end of the distribution, then the starting point lies somewhere in between (i.e. 1991 with the Gini index). Therefore, the global trends in the 1980s and 1990s are more contentious, depending on distributive sensitivities or which years are being compared. What is less contentious is that all these indices agree with the sharp decline that followed afterwards. The story about relative inequality only changes when attaching extreme sensitivity to the very bottom of the distribution (e.g. GE(-1)), in which case, after sharply declining for several decades, inequality first stagnated and then increased after 2005.

To help disentangle these trends in the global income distribution, I quantify each country's contribution through different channels, i.e. inequality between countries and within countries (with constant population) and population growth (with constant income distributions). To do so, I use a Blinder–Oaxaca type of decomposition based on the statistical notion of the recentered influence function. The results highlight that the main trends in global inequality can be largely explained by the economic evolution of China and, to a lesser extent, India.

The initial decades are characterized by growing income differences between countries, with China and India being left behind. However, this is offset to a large extent by a declining trend in inequality in these countries and others. These trends later totally reversed, with within-country inequality starting to generally increase according to most relative measures, especially from the late 1980s and early 1990s, while between-country inequalities started to decline between the mid-1970s and 2000, depending on the aforementioned distributive sensitivities. Therefore, while most relative indices tend to agree on the within-country trends, which exhibit Lorenz dominance after 2000, it is how they assess the changes in between-country inequality that produces different conclusions. Although these trends are driven to a large extent by the contribution of the biggest emerging economies—China and India—to overall inequality and each of its components, the approach followed here also enables quantification of the particular contribution of every country and region. For example, I highlight the contribution of former socialist Eastern European countries during the transition to a market economy and the impact of the diverging inequality trends in various regions in recent years. Similarly, I show that the impact on global inequality of faster population growth in developing countries, which has recently been concentrated in the sub-Saharan African region, is also substantial, *ceteris paribus*.

The structure of the paper is as follows. The next section presents the new dataset, then Section 3 presents changes in per capita income. Section 4 discusses changes in the entire distribution, Section 5 analyses changes in inequality over time, Section 6 focuses on the between- and within-country components of inequality, and Section 7 focuses on the contribution of specific areas and countries. Section 8 concludes.

2 Data

The analysis of global inequality faces considerable data constraints due to the lack of information collected consistently over time and across countries. To address this, I have put together a new global inequality dataset which is based on a classical database for cross-country analysis of inequality—the WIID held by the United Nations University World Institute for Development Economics Research (UNU-WIDER). All the datasets are freely downloadable.³ The WIID was first launched in 2000, giving continuity to one of the first most successful initiatives, by Deininger and Squire (1996), for collecting cross-country information on inequality. The WIID has been updated several times, including an update by Deininger and Squire in 2004, and has been expanded to incorporate other sources. The most recent version is from March 2021. The WIID, which has over 20,000 data points, collects and stores information on income inequality for almost all countries in the world (196 countries or territories and four historical entities) over the longest possible period of time for which reliable data are available.⁴

The information is now mainly obtained from a variety of public sources, including international databases such as PovcalNet (World Bank’s Development Research Group), the LIS, Eurostat, the Socio-Economic Database for Latin America and the Caribbean (SEDLAC), the Organisation for Economic Co-operation and Development (OECD), United Nations agencies such as the UN International Children’s Emergency Fund (UNICEF) and the UN Economic Commission for Latin America (ECLAC), several national statistical authorities, and many independent research studies. Many of the historical sources in the WIID come from the original compilations by different authors and institutions in the 1970s and 1980s. The dataset is a unique combination of data from the most prominent current data providers and historical or independent sources, and it brings together this fragmented information in a systematic and organized way. However, we need to address some issues before using the WIID for the analysis of global inequality.

First, it is necessary to select the observations that will be used, because in many cases there may be more than one per country and year (for example, from different sources or referring to different measures of resources). Second, we need to deal with the heterogeneity in the welfare concepts measured, coverage, and sources. Although the most common welfare concept refers to some sort of income definition expressed in per capita terms, some observations refer to per capita consumption and others refer to income per household or per equivalent adult. Similarly, income can be gross or net (after taxes and social contributions have been deducted). Most observations refer to the national level and a few refer to urban areas or exclude specific parts of a country. Furthermore, the values reported by different sources can diverge in other methodological aspects, such as survey or treatment of non-responses, etc. Third, some observations only report the Gini index. Of those reporting income shares, many only have limited information (for example, they may report income shares by quintiles), while others report the full set of deciles and bottom and top 5 per cent. Finally, the information needs to be aggregated across countries to estimate global inequality, to address the fact that we will end up with a highly unbalanced panel with many missing observations for several countries and years.

In this section, I summarize how I addressed these issues in constructing the new global inequality database presented here. The entire process is discussed in more detail in a series of technical notes

³ See UNU-WIDER (2021a, 2021b).

⁴ For an earlier assessment of the WIID, see Jenkins (2015).

(Gradín 2021a, 2021b, 2021c) which include several country examples and the Stata codes used to construct the global database from the original WIID.⁵

Selection of inequality series

While the WIID only contains information from one series for some countries, for many others it displays information from several series that may overlap and refer to the same or different periods, measures of resources, or equivalence scales, among other things. This produces several possible estimates of inequality for each country and year, with incompatibilities among them that should be considered. For that reason, the first step was to produce a companion dataset (WIID Companion) with a careful selection of series or fragments of series that best represent the longest possible trend for each country with the highest possible internal consistency. To guarantee this internal consistency, the selection was made in terms of series (or fragments) rather than isolated year observations. It sought to avoid, as much as possible, the creation of spurious trends by mixing observations from different sources and using various methods over the same time span. I also prioritized those series that have richer information about the entire (net) income distributions (mainly at least deciles) and that are expressed in per capita terms at the national level. The main priority was to use LIS, but in countries and years for which LIS is not available, other sources with high international or regional comparability (e.g., Eurostat, ECLAC, SEDLAC), among other criteria (Gradín 2021a) were given maximum priority.

As a result of this selection process, I ended up with 2,142 country–year income distributions, representing the income distributions of 186 countries or territories for the period between 1947 and 2019. About two-thirds come from four main single contributors, namely PovcalNet (30 per cent), LIS (22 per cent), ECLAC (10 per cent), and Eurostat (9 per cent). The remaining observations are taken from various national statistical authorities (10 per cent) and other research studies (8 per cent), SEDLAC (3 per cent), and other international sources. From the original distributions, aggregated at the decile or quintile level,⁶ with or without bottom and top 5 per cent, I used the Shorrocks and Wan (2009) approach to estimate the entire synthetic distributions at the percentile level. This procedure first fits a log normal distribution and then guarantees that the synthetic distributions fit the same aggregated income shares as reported in the original ones. The synthetic distributions are accurate representations of the original underlying survey income distributions.

Integrating and standardizing income distributions

The selected series are, however, heterogenous across the methods used, particularly the welfare concept (measure of resources and equivalence scale).

One possible approach for addressing this heterogeneity involves limiting inequality comparisons to the most consistent cases. The greatest possible within-country consistency is obtained if the income distribution for all comparison years is obtained from the same source and welfare concept, and the greatest cross-country consistency is attained if the income distributions from different countries are harmonized to some extent (e.g. Eurostat and LIS). However, this can lead to a dramatic reduction in the number of comparison countries or years.

⁵ For a discussion of the main issues related to cross-country inequality databases, see Anand and Segal (2008), Atkinson and Brandolini (2001), or Ferreira et al. (2015).

⁶ Of which about 89 per cent originally referred to at least deciles (32 per cent also bottom and top 5 per cent), while the rest originally referred to at least quintiles.

A second approach for addressing this heterogeneity consists in standardizing the distributions so as to represent the distribution of the same welfare concept (e.g. net per capita income) in all countries and periods. This is also a common practice in the literature, although there is no single strategy for implementing it. This is the approach followed here. To achieve a higher degree of consistency, income distributions included in the WIID Companion underwent a two-phase adjustment process. As a result, each series always refers to the same welfare concept (household net income measured in per capita terms).

Some of the heterogeneity in the series is resolved in a first phase (integration) by taking advantage of the overlapping over time of the various series in each country originally using different methods. This is done by taking one series, usually the most recent, as a reference (e.g. LIS). This series is extended backwards (or forwards) using other series that overlap in at least one year. The series are shifted down or up using a common factor, so that they match the next series at the integration point (the year in which they overlap). Therefore, the i th percentile income share P_t^k for series k at year t is adjusted (AP_t^k) to match the next series, $k + 1$, at the overlapping year, $t_{k,k+1}$, using $AP_t^k = P_t^k + (AP_{t_{k,k+1}}^{k+1} - P_{t_{k,k+1}}^k)$. After adjustment, all the relevant information (resource, equivalence scale, population and geographic coverage) is updated to match the information from reference series. This is done to preserve the trend of the original series while matching the level of the following series.

After these adjustments, some country series refer to the target welfare concept, i.e. the distribution of net income per capita, but some refer to a heterogeneity of welfare concepts. In a second phase (standardization), the latter are converted to per capita net income using a simple regression approach that exploits the empirical relationship between percentile distributions for per capita net income and for other welfare concepts in the LIS sample in WIID. This sample comprises 3,826 country–year observations, of which 472 are for the target welfare concept (per capita net income) and the remaining refer to other welfare concepts (e.g., 584 for consumption per capita, 355 for gross income per household, etc.).

Although the LIS is known to have a good representation of high-income countries in Europe and North America, it contains a substantial number of countries from other regions and income levels that are then used as a reference for other similar countries. The LIS sample has 57 countries, of which 35 are high-income, 14 upper-middle-income, 6 lower-income, and 2 low-income countries.⁷ The regressions can generally be expressed as:

$$P_{i,t}^{net\ income,pc} = \beta_0 + \beta_g P_{i,t}^{r,s} + \beta_{r,s,g} r_{i,t} \#s_{i,t} \#group_i + u_i ,$$

where $P_{i,t}^{r,s}$ is the relative income (or income share) of any percentile, referring to any combination of resources and scale (respectively, r and s), where $\#$ stands for all possible interactions and $group_i$ indicates the country grouping that applies (i.e. country; region and income group; region; income group). There are 400 regression (100 percentiles and four country groupings).

⁷ Therefore, as well as having a good representation of European countries, the United States, Canada, and Australia, LIS includes Russia and other former USSR countries; several Latin American countries, particularly Chile and Mexico, but also Brazil, Colombia, Dominican Republic, Guatemala, Panama, Paraguay, Peru, and Uruguay; several Middle East and North African countries, including Egypt, Iraq, Israel, Jordan, Tunisia, and West Bank and Gaza; Cote d'Ivoire, Somalia, South Africa, along with Sudan in the sub-Saharan Africa region; and China, India Japan, Republic of Korea, Taiwan, and Vietnam in Asia. See Checchi et al. (2021) for a discussion of the challenges and solutions involved in incorporating middle- or low-income countries into LIS.

The corresponding adjusted values for welfare concepts other than per capita net income are then replaced by the prediction of the regression, involving a correction based on the average difference observed in the same or similar countries (in some cases, the same country in other years; in most cases, the same region and income group; and, in fewer cases, either the same region or the same income group).

That is, the conversion involves obtaining the final standardized values, $FP_{i,t}$, by replacing the adjusted income distribution values after phase 1, $AP_{i,t}^{r,s}$, by the prediction of this model for the country grouping that applies:⁸

$$FP_{i,t} = \hat{\beta}_0 + \hat{\beta}_{r,s,g} + \hat{\beta}_g AP_{i,t}^{r,s}.$$

For example, per capita consumption distributions in the Republic of Congo are adjusted to account for inequality being observed to be substantially higher in terms of income than in consumption in Côte d'Ivoire, which is the country that represents the same region and income group in the LIS sample used in the regressions.

The sum of adjustments made in both phases generally implies a higher level of inequality, which can be quantified by an average increase of 2.3 Gini points (3.6 among affected observations). These adjustments are especially large in developing regions such as sub-Saharan Africa or South Asia, because the series are originally expressed in per capita consumption terms. These large adjustments just reflect that there is enough evidence in these countries (or in others in the same regions and income groups) to show that inequality will be substantially higher if measured in per capita net income terms, as estimated by common household surveys.⁹ If these income distributions were kept in per capita consumption terms, within-country inequality would be severely under-estimated. Note that this approach preserves the trends observed in the original series, regardless of the welfare concept used, but adapts the level of inequality accordingly to gain constancy when comparing or aggregating across distributions.

Per capita income and population integrated series

To construct the global income distribution, it is necessary to estimate the average income of each country's fractiles, which in turn will determine where in the global distribution a given country income group falls. There are two main approaches: 1) using a macro aggregate, GDP, or gross national income (GNI); or 2) using the survey average income or consumption. As this database attempts to reconstruct an annual series of global income, it seems to be more appropriate to use a macro aggregate with much richer and consistent information.¹⁰ For that reason, as a proxy for average wellbeing across countries, I combined information primarily from the per capita GDP in

⁸ Whenever necessary, the final percentile incomes are re-ordered to guarantee monotonicity. Negative and zero values are assigned a value close to zero. However, the procedure does not include any smoothing of trends over time as it aims to recover what would come from household survey information, and sharp trends may often just reflect underlying structural changes. Avoiding dubious trends, however, was one of the criteria used in the data selection process.

⁹ Note that this empirical pattern exhibits great heterogeneity by geographical area, with adjustments being smaller in other areas, and therefore it is necessary to refine the conversion factors at least by region and income group when it is not possible to do so for the same country.

¹⁰ The use of mean income or consumption in surveys is more complex because it tends to be scarcer (it is available only for survey years, and not always), and tends to be expressed in different currencies and reference periods, needing an intensive process of standardization taking account of monetary reforms, inflation rates, etc., information that is not always possible to obtain.

2017 PPP USD series recently published by the World Bank (2021) as part of the World Development Indicators (WDI 2021), with complementary information from the Maddison Project Database (2020) and the Penn World Tables (PWT9.1) (Feenstra et al. 2015) in order to collect information for all countries for the period from 1950 to 2019. The different series were integrated by taking advantage of the overlapping of the different series for most countries (e.g. re-scaling values of the Maddison series or the PWT to match the WDI series in 1990 and integrating the relative trend before 1990 of these sources into the WDI series). For countries with no information in the WDI series, for adjustment, I used the ratio between the GDP for the country and that for the United States in 1990 in each source. Regarding country population, I mainly used the UNDESA (2019) series. In a few cases, the missing initial years are imputed based on the trend in the same income group and region.

Aggregation

Finally, I aggregated the income distribution for the whole world by combining the information for percentile income shares, per capita income, and population for currently existing countries.¹¹ Again, there are several possible approaches for tackling the lack of information about within-country income distributions for many countries and years. One involves limiting the comparisons using only survey-based information that falls within a specific bandwidth of those comparison years (e.g. two years before or after). However, this can create a largely unbalanced panel whose composition can change from year to year and affect the resulting trends, especially if we try to reconstruct the annual series for a long time span. If the sample is further restricted to include only countries with enough information in every comparison year so that the panel is balanced, the resulting sample will be a small and biased representation of the world, making it harder to discern global trends.

For those reasons, the missing information in the dataset used here is obtained for each year and for all countries by using linear interpolation of the percentile distributions in the closest years before and after the target year. For years before the first or after the last observation in the WIID Companion, the income distribution is held constant. In the few cases without any distributional information at all (e.g., Libya, Saudi Arabia, and several microstates), this is imputed using the population-weighted average income distribution in the same region and income group.¹²

Through the above procedure, I obtained a balanced panel of 209 countries and territories covering the period from 1950 to 2019, which prevents changes in the composition by countries spuriously determining the observed trends. Obviously, although all countries are included in all years, survey years are more common recently, particularly after 1990, implying that the quality of the within-country inequality trends is also higher in the most recent decades. In Figures 1 and 2, I report the various measures of available survey information that allow us to consider how representative are earliest estimates of trends in within-country inequality. Figure 1 reports the number of countries and the percentage of the world population using a survey that falls within five years (before or after) of the target year (e.g. number of countries and population share in 2000 with survey observations taking place between 1995 and 2005). Figure 2 displays the

¹¹ I thus completed all the necessary information for current existing countries in earlier periods with the trend based on the original entities (e.g., Czechoslovakia, Yugoslavia, USSR, Ethiopia, and Sudan). Because of the lack of distributive information for the German Democratic Republic in the WIID, covering the period before unification, here, I use the combined information for Germany for population and mean income, and the distributive information from the Federal Republic of Germany.

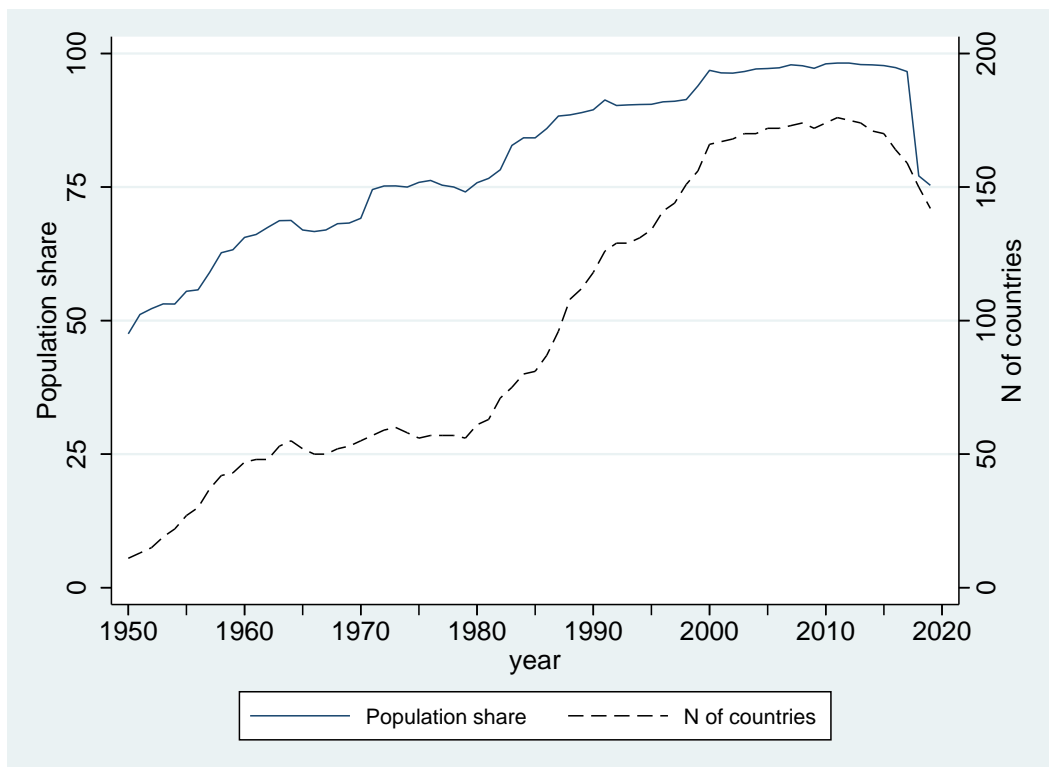
¹² I assigned the same mean income to all percentiles only in the case of the Democratic People's Republic of Korea (North Korea), so the country still contributes to inequality between countries.

population-weighted average gap between the target year and the closest survey year (e.g. the gap is two if the closest observation around 2000 is in either 1998 or 2002) falling within those bandwidths.

The initial number of countries with surveys within five years is relatively small around 1950, and we therefore need to keep in mind that the information for most countries in this initial decade comes from extrapolating the earliest available years and only changes whenever surveys are available for the first time in these countries. However, when countries are weighted by their population and, given that the most populous countries (including China, India, Mexico, and the United States) have survey information for before 1955, the world’s population covered with a survey within five years is about 50 per cent. This population share reaches 75 per cent around 1980 and continues to increase to nearly 100 per cent in most of the 2000s. It declines again in the most recent years due to a lack of updated information for some key countries. The population-weighted average distance in years between a survey year and the target year falling within the corresponding +/- five-year bandwidth mentioned above oscillates between 0.5 and 2 over most of the period analysed.

Overall, with its limitations, this is a good representation of the global trends in inequality. Note that in earlier years, when within-country inequality is highly extra/interpolated, as explained later in more detail, global inequality is strongly determined by the between-country component (differences in per capita income among countries), with generally more accurate information every year.

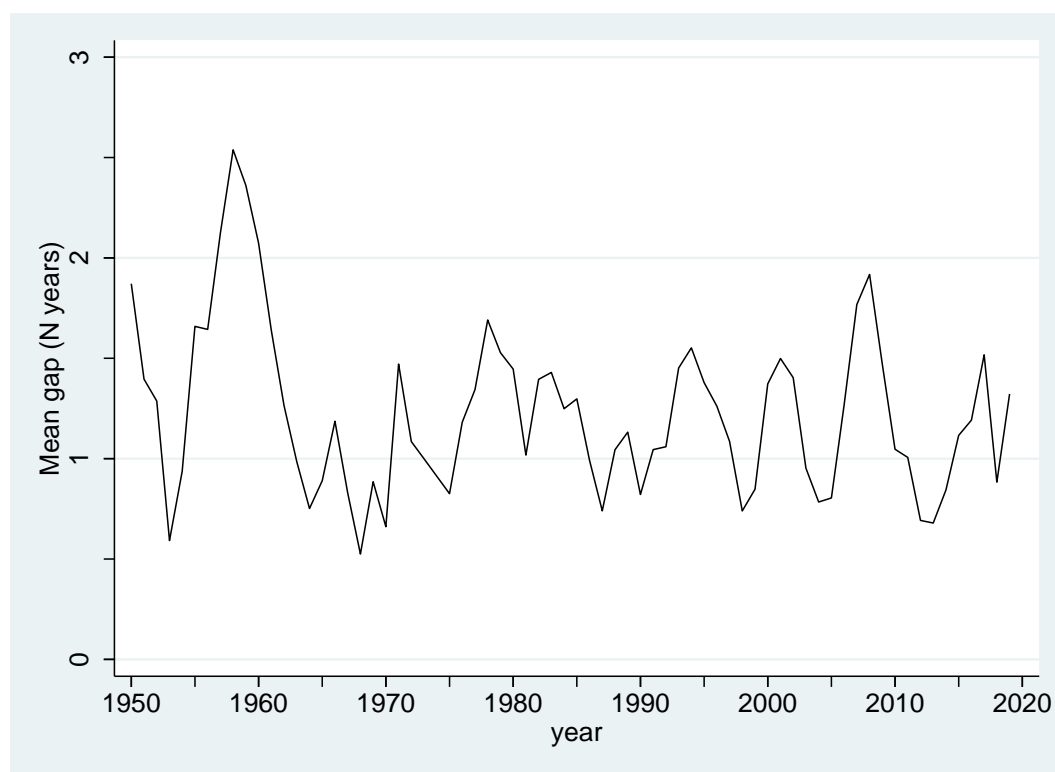
Figure 1: Number of countries and population share with income shares in each bin year using a +/-5-year bandwidth



Note: number of countries and share of the world’s population with information on income shares falling within a bandwidth of five years before and after the target year.

Source: author’s construction based on the WIID.

Figure 2: Population-weighted mean gap years between target year and closest survey year falling in a +/-5-year bandwidth



Note: population-weighted average gap in number of years between each target year and the closest survey income shares within a bandwidth of five years before or after the target year.

Source: author's construction based on the WIID.

3 The context: changes in per capita income

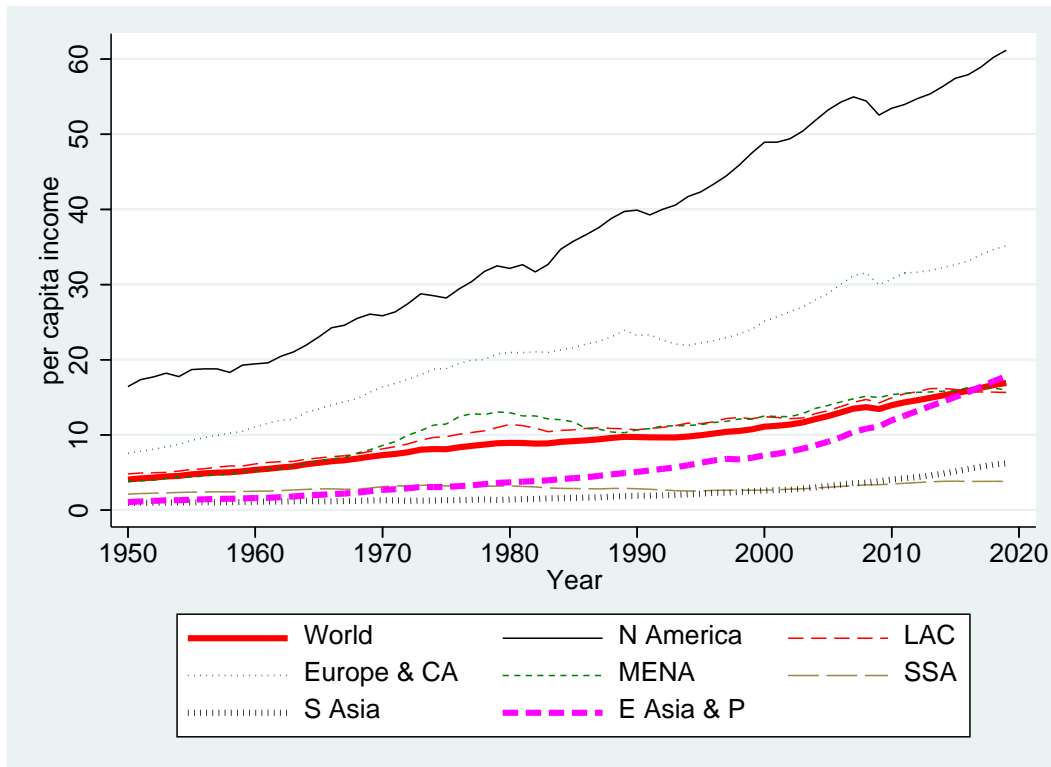
The period from 1950 to 2019 is characterized by strong and sustained global economic growth of about 2.1 per cent per capita annually, with the highest level of growth seen in the first 30 years (2.7 per cent on average in the 1950s, 3.3 per cent in the 1960s, and 2.3 per cent in the 1970s). The economy then slowed down to its lowest level of growth after the oil crises: the 1980s (0.9 per cent) and 1990s (1 per cent). Growth rates rose again to 2.3 per cent in the 2000s and 2010s. There were a few episodes in which global per capita income declined or the annual growth rate fell below 1 per cent, corresponding to the main global economic crises, such as in 1957–58, 1973–75, 1979–83, 1989–93, and 2008–09.

This generally strong economic growth trend was quite heterogenous. The trend was first much stronger in North America and Europe, with much more rapid growth recently being experienced in East Asia, particularly after the 1990s, followed by South Asia (Figure 3a). Growth was much weaker in sub-Saharan Africa over the entire period, with only an annual 0.9 per cent growth rate on average, as opposed to 4.2 per cent in East Asia and Pacific (6.3 per cent in China), 3.3 per cent in South Asia (3.6 per cent in India), and between 1.5 and 2.2 per cent in the other regions (1.9 per cent in the USA) (Figures 3a and 3b). Within the sub-Sahara African region, there were also stunning differences, with much higher growth rates, above 3 per cent, in Botswana, Cape Verde, Equatorial Guinea, Eswatini, Ethiopia, and Mauritius. Per capita average growth rates were negative in some countries, mainly in sub-Saharan Africa (e.g., Burundi, Central African Republic, Democratic Republic of the Congo, Republic of Congo, Liberia, and Madagascar), as well as in

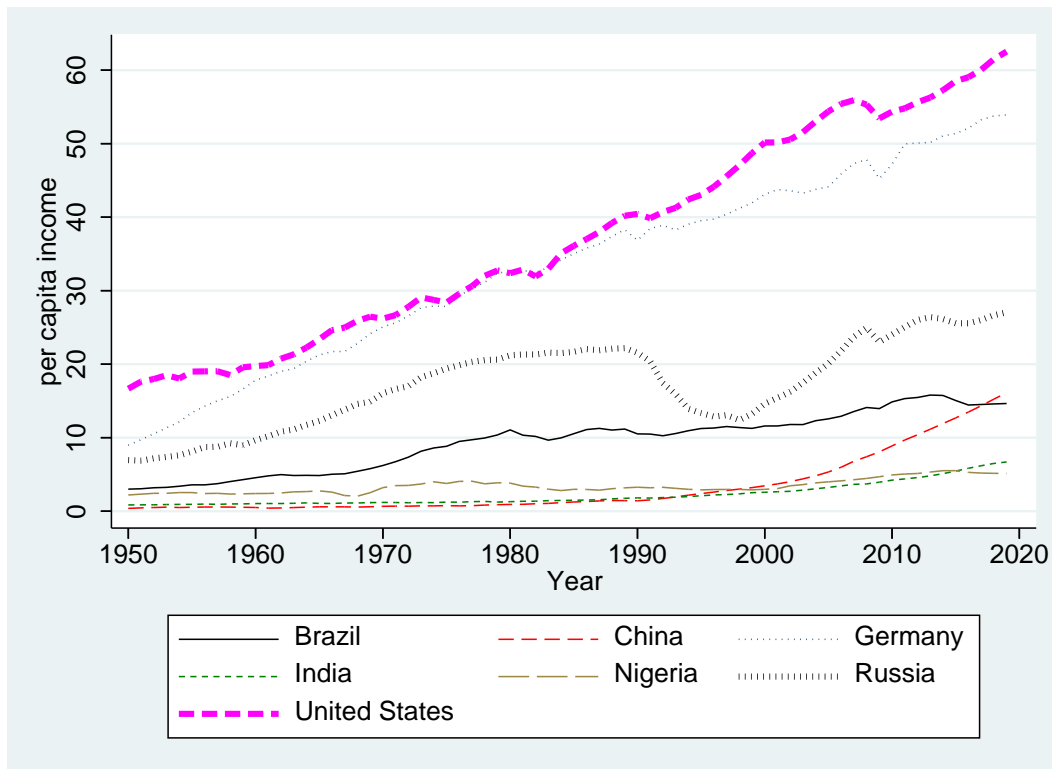
Haiti, Liberia, North Korea, Venezuela, Syria, and Yemen. The population of countries exhibiting negative growth rates over the period analysed was 145 million in 2019 (2 per cent of the world's population), and 512 million (nearly 7 per cent) if we include people in countries with an average growth rate below 1 per cent.

Figure 3: Per capita income

a. Overall and by geographical region



b. By selected countries



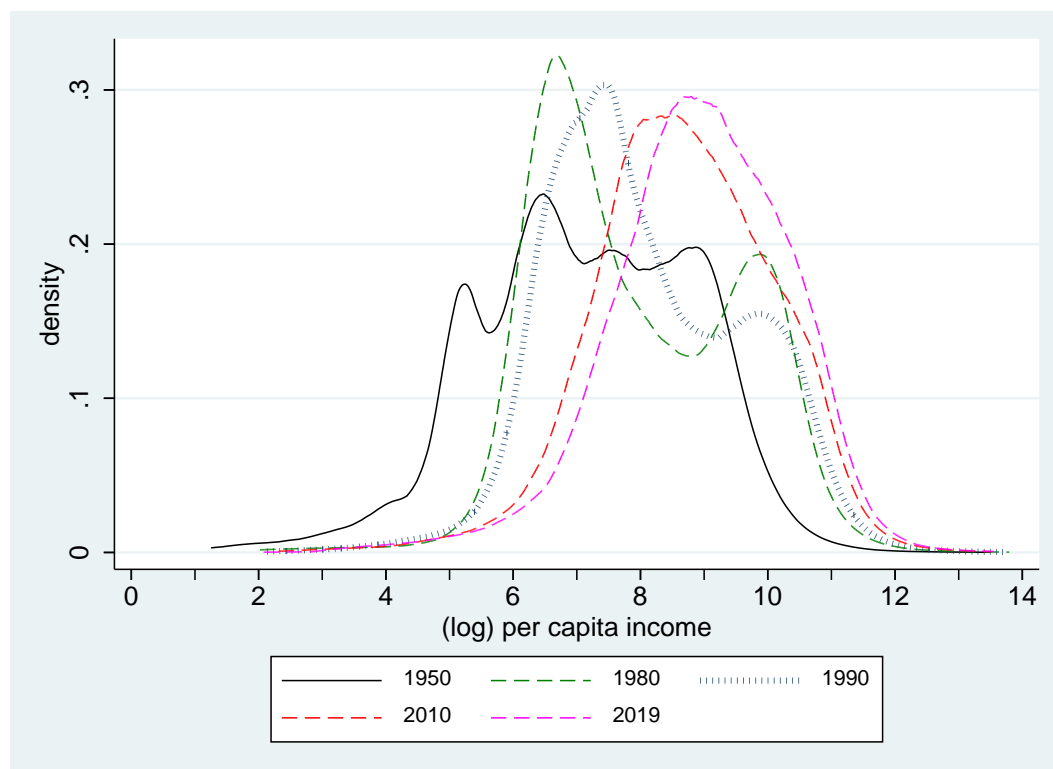
Note: integrated series constructed using WDI, Maddison and PWT GDP series. Amounts in thousands of 2017 PPP USD.

Source: author's construction based on the WIID.

4 Changes in the entire global income distribution

The global distribution of income among all citizens in the world has exhibited important changes over the last decades. The densities in Figure 4 highlight the huge shift of population mass from the very bottom of the (log-)income scale to higher levels. This process is the translation of the growth described above at the individual level once we account for existing high and changing within-country inequality. This process has clearly reduced the skewness of the global income distribution. These densities show an outstanding bimodality visible in the log-income scale which is accentuated in the first decades but starts to fade from 1990 onwards, and completely vanishes after 2000. This is a good indication of the huge structural changes that completely modified the shape of the distribution during those years.

Figure 4: Kernel density of the global income distribution



Note: per capita income 2017 PPP USD (log-scale).

Source: author's construction based on the WIID.

An alternative way of visualizing the changes in the individual income distribution is through the quantile curves (see data appendix¹³) and their accumulated change over time (growth incidence curve (GIC)). For each income percentile, the GICs can map either accumulated absolute income changes in 2017 PPP USD (see Figure 5) or relative changes (accumulated income growth rates, as in Figure 6). Growth rates are strong over most of the period analysed, but the global growth distributional pattern progressively shifted its shape from an initial U to an inverted U. The absolute GIC puts these relative gains in context, as they have very different implications depending on the initial incomes, which tend to be very small at the bottom and very large at the top.

Income growth in the first decades before 1980 shows a U-shaped pattern, with the strongest growth rates at the bottom and upper-middle levels of the income scale. Growth is weaker at the middle and very top, as reflected in the relative GICs. Growth for the first six deciles, however, is almost insignificant when represented in absolute terms as it starts at very low levels and tends to increase absolute distances among individuals. This income growth pattern substantially changed in the 1980s and 1990s with the collapse of communist regimes, the deceleration of growth in Japan and other economies, and the start of a trend of rising inequality in a large number of countries, with a decline in incomes of people between the 62nd and 83rd percentiles in the 1980s or between the 76th and 81st in the 1990s. This particular pattern of stagnation in the upper-middle part of the distribution and at the very bottom, combined with larger growth rates elsewhere (bottom, middle and top), is behind what has become known as the ‘elephant’ curve

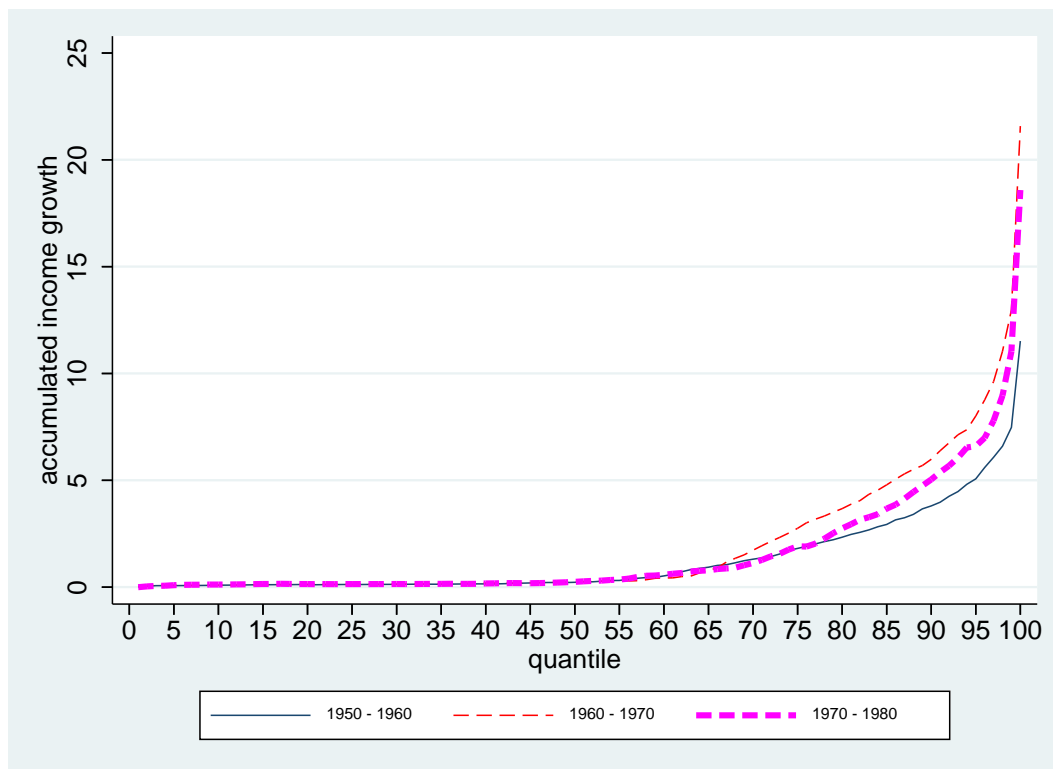
¹³ The data appendix is available here: <https://www.wider.unu.edu/publication/trends-global-inequality-using-new-integrated-dataset>.

(Lakner and Milanovic 2016). This pattern faded afterwards, leading to a clearer inverted U-shaped pattern in the 2000s where growth at the middle of the distribution becomes being stronger, reflecting the success of emerging economies like China, despite the evidence of growing inequality within countries, including in China, continuing to a large extent.

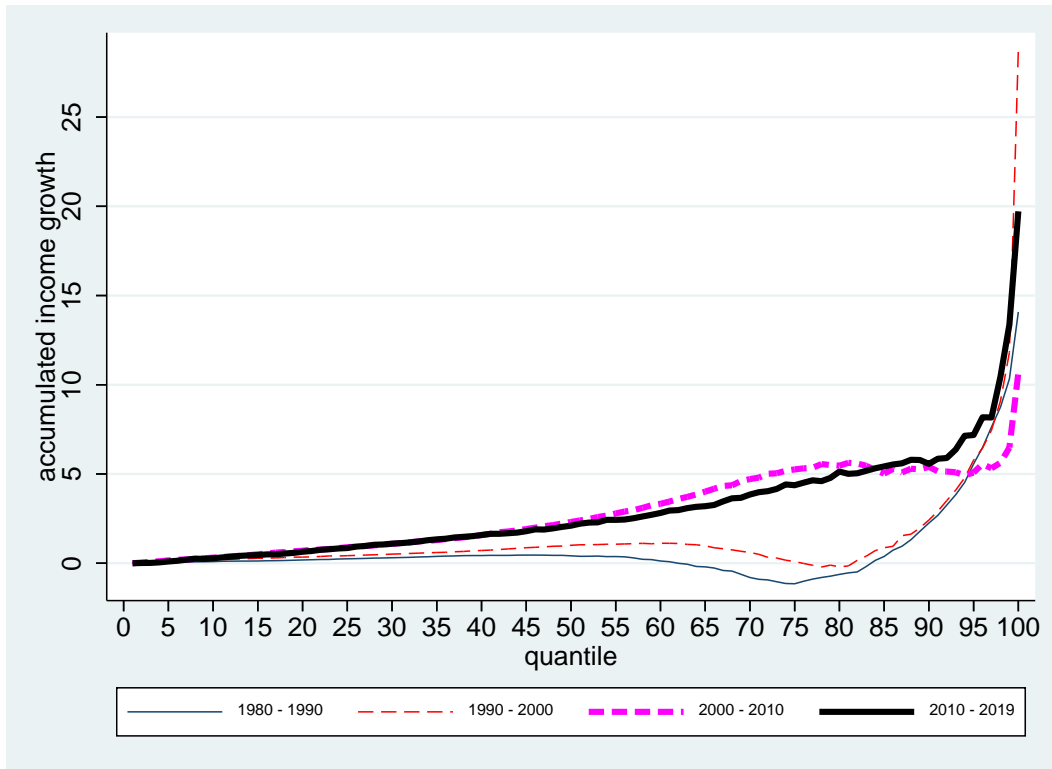
When considering changes over the long term (1950–2019 or 1990–2019), the serpent described by Ravallion (2018) stills dominates the absolute pattern, while the relative GIC reveals a strong pro-poor pattern, except at the bottom 10 per cent, for the entire 1950–2019 period and a more inverted-U pattern between 1990 and 2019 (with traces of the elephant curve due to weaker growth at the upper middle).

Figure 5: ‘Absolute’ growth incidence curves in the global income distribution (accumulated growth in 2017 PPP USD by sub-period)

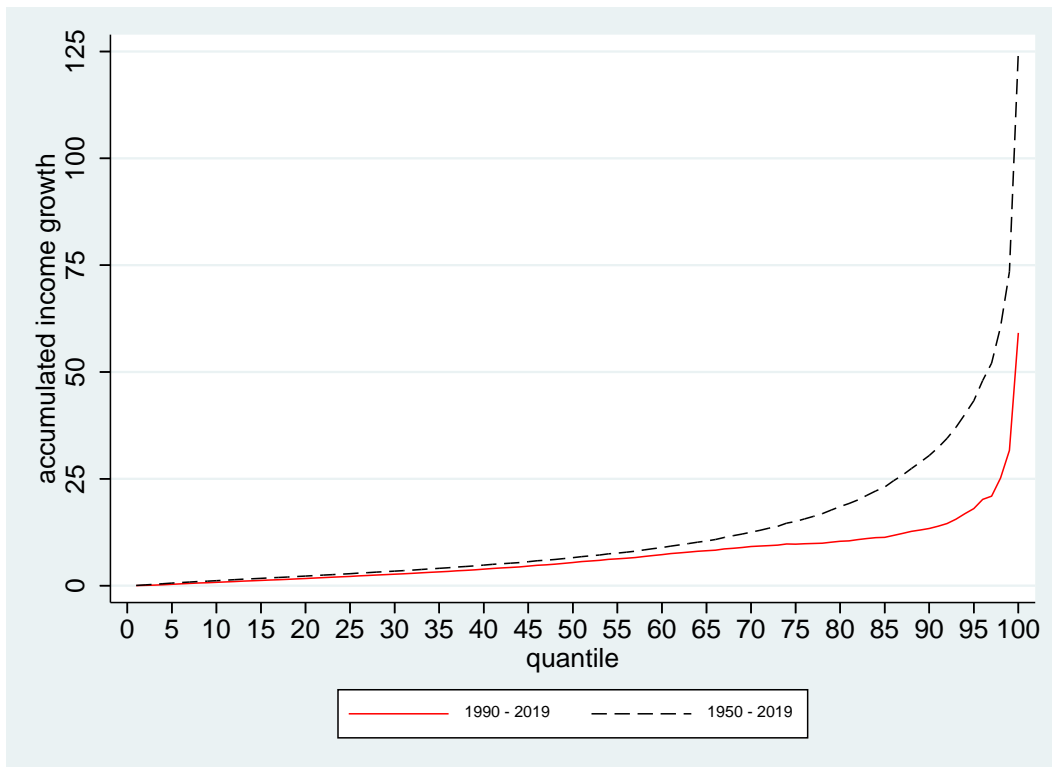
a. 1950–80



b. 1980–2019



c. 1950–2019 and 1990–19

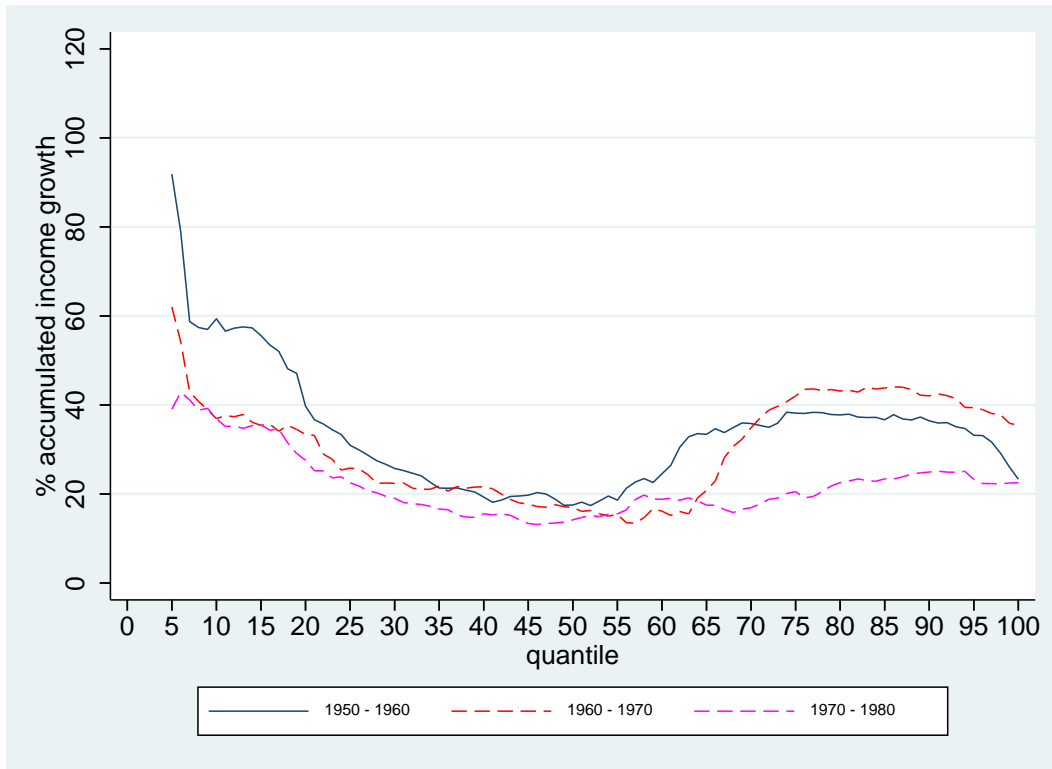


Note: amount in thousands.

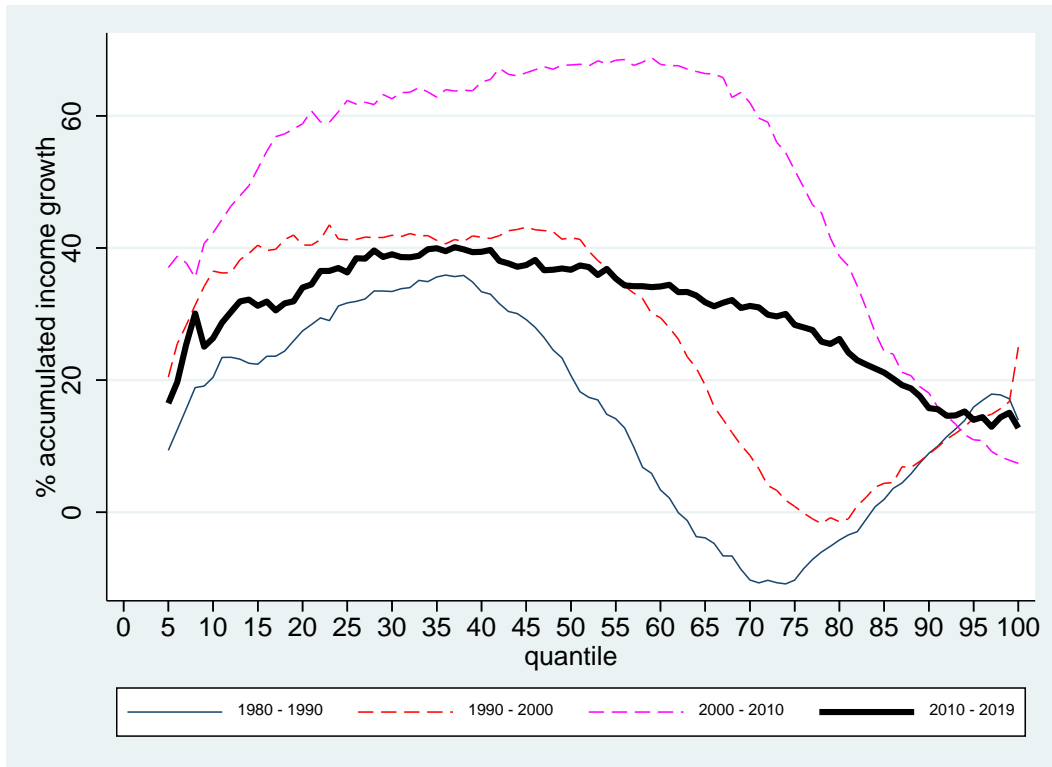
Source: author's construction based on the WIID.

Figure 6: 'Relative' growth incidence curves in the global income distribution (percentage accumulated growth rates, by sub-period)

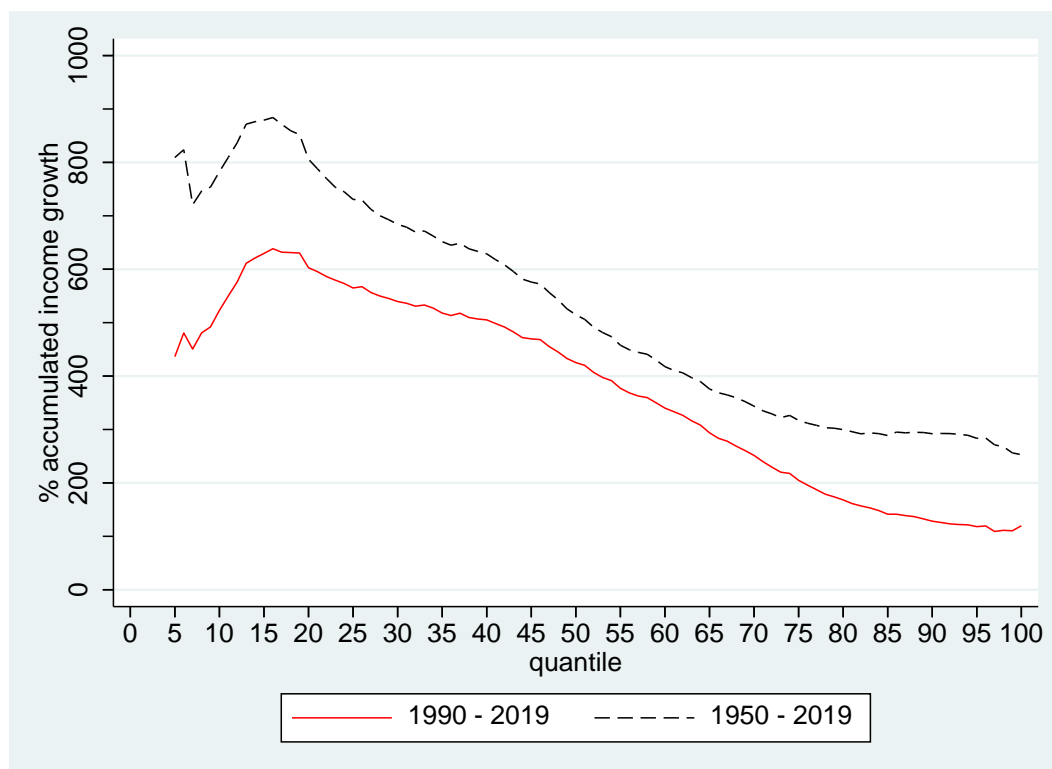
a. 1950–80



b. 1980–2019



c. 1950–2019



Source: author's construction based on the WIID.

5 Changes in global relative inequality

5.1 The approach

There are two main approaches for translating the changes in the income distribution over time into an assessment of whether inequality has increased, remained constant, or declined.

One approach involves comparing the corresponding Lorenz curves (Lorenz 1905) that map the cumulative shares of population and income for different income fractiles, producing an incomplete ordering of the distributions in terms of their inequality using a minimum set of value judgements (Atkinson 1970). These value judgements are given by only four principles: anonymity (or symmetry), population replication invariance, the Pigou–Dalton principle of transfers, and scale invariance. The first two are rather technical and necessary to make inequality comparisons using changes in distributions reporting the proportion of people with each income level and ignoring any other attributes of people (anonymity) and changes in the total number of people (population principle). The other two principles are more substantive and summarize the most common notion of inequality.

The transfers principle identifies what changes in incomes are considered to be unambiguously equalizing or disequalizing when total income is constant. It holds that a small progressive transfer from someone to someone else with relatively lower income reduces inequality, while the opposite change increases inequality. In both cases, the transfer should preserve the ranks of the people involved and total income. Scale invariance states how changes in total income should be distributed to keep inequality constant and is key in a context of economic growth such as the one analysed here. It introduces the notion of relative inequality because uniform income growth rates

along the income scale do not change inequality. The combination of these principles implies that one should pay attention to the slope of the income growth pattern of the different (anonymous) fractiles between two comparison years. If these are flat, inequality remains constant; if growth rates decline with incomes (pro-poor), inequality declines; and if they increase (pro-rich), inequality increases.

In terms of the Lorenz curve, one just needs to check whether the curve of one distribution is above the other one, with the former exhibiting lower inequality; that is, whenever the Lorenz curve in one year falls below the curve of another year, relative inequality is unambiguously higher in the former distribution if one agrees with the above-mentioned four principles. In other words, the same income shares across the population in the distribution with lower inequality can be obtained from those in the other through a sequence of progressive transfers (from anyone to someone relatively richer). The Lorenz curve allows for inequality assessments that will generate large consensus upon agreement on these principles but does not provide a measure of the intensity of the inequality change, and is unable to order distributions whenever the Lorenz curves cross at least once (for that reason, the order is incomplete). Crossing Lorenz curves reveal that one distribution can be obtained from the other through a combination of progressive and regressive transfers, which is likely what happens if the bottom and top of the income distribution have the largest (or the lowest) income growth rates.

Finally, the absolute Lorenz curve proposed by Moyes (1987), which maps the accumulated income differential between real income and the mean at each population share, plays the same role as the conventional Lorenz curve but for absolute inequality. The only difference is that the scale invariance principle is replaced by translation invariance. That is, uniform absolute increases in real incomes (2017 PPP USD) along the distribution leave inequality unchanged. It is important to note that in a context of economic growth, like the one we observe here in most of the period analysed, absolute inequality is much more demanding than relative inequality in terms of the amount of income growth that should be accrued by the poor compared with relative inequality. Uniform growth rates will increase absolute inequality (but leave relative inequality unchanged). However, the opposite is true in the context of a recession when uniform negative growth rates will decrease absolute inequality. In the case of average income stagnation, both inequality approaches converge to the same.

The second and complementary measurement approach is to use aggregate measures of inequality, which allow us to order all distributions and to quantify the intensity of changes but necessarily introduce additional value judgements that may generate less agreement. That is, completeness is achieved at the expense of consensus, particularly in relation to how to evaluate, for example, whether we give more relevance to the high growth of incomes at the bottom and middle of the distribution or to the also strong growth at the very top. Some may be more concerned with reducing relative poverty and will therefore give more relevance to the improvement of the poor, while others may be more concerned with the potential consequences of the higher accumulation of resources and power in very few affluent people. Based on the Pigou–Dalton principle of transfers, the first feature is equalizing and the second is disequalizing. If both occur at the same time, to assess the inequality trend, it is necessary to make explicit which one is given more relevance. The reality of the global income distribution described earlier is a bit more complicated because it also shows stagnation at the very bottom of the distribution in some periods and, therefore, anyone particularly concerned with the situation of the very poor might give more relevance to this fact than to the improvement of the rest of poor. Most traditional inequality measures such as the Gini index, the coefficient of variation (CV), the Theil (L-)index or the MLD or Theil M-index, are consistent with the use of the relative Lorenz curve, in that they do not contradict the corresponding ordering (when estimated at the same level of disaggregation) but may disagree in situations in which the Lorenz curves cross each other because they give different

weights to changes at different points of the distribution.¹⁴ The families of indices like the generalized entropy family (which embrace the Theil index, the MLD and the squared CV as particular cases), the Atkinson family, or the generalized Gini (much less used in empirical analyses) are also consistent with Lorenz orderings, and the parameter that differentiates the different members of the family explicitly reveals their degree of sensitivity to different parts of the distribution. Similarly, dominance in terms of the absolute Lorenz curve will imply unanimity among absolute inequality indices, including the absolute Gini or the standard deviation, regarding the direction of the inequality trend.

It has recently become quite popular to assess inequality trends based on partial measures reflecting the income share of specific parts of the distribution, like the top 1 or 10 per cent (e.g. World Inequality Lab 2018) or the share of the bottom 40 per cent (the concept of ‘shared prosperity’, World Bank 2016). These income shares can be combined to measure the ratio of total income held by each end of the distribution, as is the case with the Palma ratio between the top 10 and bottom 40 per cent (Cobham and Sumner 2014) or the S80S20 ratio (between the top and bottom 20 per cent). These ‘partial’ measures of inequality, however, are not consistent with the Lorenz criterion because they do not use information for the entire distribution but help to focus even more on sensitive parts of the distribution and tend to reach non-technical audiences more easily.¹⁵

In this section, I investigate what can be said about the trend in global inequality and to what extent the trend depends on potentially conflicting distributive views (therefore the inequality measure or approach used). To do so, I first compare the absolute and relative Lorenz curves between the key years. Then I estimate a battery of inequality measures that are consistent with Lorenz orderings, such as the Gini index, the generalized entropy family, and the Atkinson family in the relative case, and the absolute Gini measure and the standard deviation in the absolute case. I also estimate income shares for different world population groups as well as the Palma ratio.

5.2 Global Lorenz curves

I first address the analysis using the Lorenz curves for selected years. The relative Lorenz curves (Figure 7, summarized in Table 1) show that there is very limited evidence of strict Lorenz dominance over the period and, therefore, of unambiguous trends in inequality. This means that changes in the distribution between most years involve growth patterns with a mix of equalizing and disequalizing movements (as can be inferred from analysis of the GICs). Judgements about what happened with inequality will depend, at least to some extent, on the relative importance given to growth rates at specific parts of the global distribution.

When the analysis is conducted at the percentile level, for example, we find strong evidence of an unambiguous decline in inequality (Lorenz dominance) in the long term, i.e. between 1950 and 2019 (the former curve falling entirely below the latter). However, there is more ambiguity in the

¹⁴ As shown in Gradín (2020) using the statistical notion of recentered influence function (RIF), marginally increasing the proportion of the population with low or high income increases all well-known inequality measures but in very different proportions. The MLD is particularly more sensitive to the bottom, the Theil index to the top (but less than the CV), while the Gini index is less sensitive to both extremes than the other indices mentioned. Another way of looking at this is to compare how the change in the indices after a progressive transfer varies across the distribution if the distance between the donor and the receiver is fixed (Cowell 2011).

¹⁵ Some of these partial indices (top X per cent or ratios between top X per cent and bottom Y per cent) satisfy a weaker version of the transfer principle, meaning that they will not decline with a regressive transfer, but can remain constant if the donor and receiver are both either within or outside the relevant range (see, for example, discussion in Foster and Lustig 2019). These should not be confused with quantile ratios which compare the income of people occupying specific ranks, such as between the 90th and 10th percentiles, which is quite popular in labour economics.

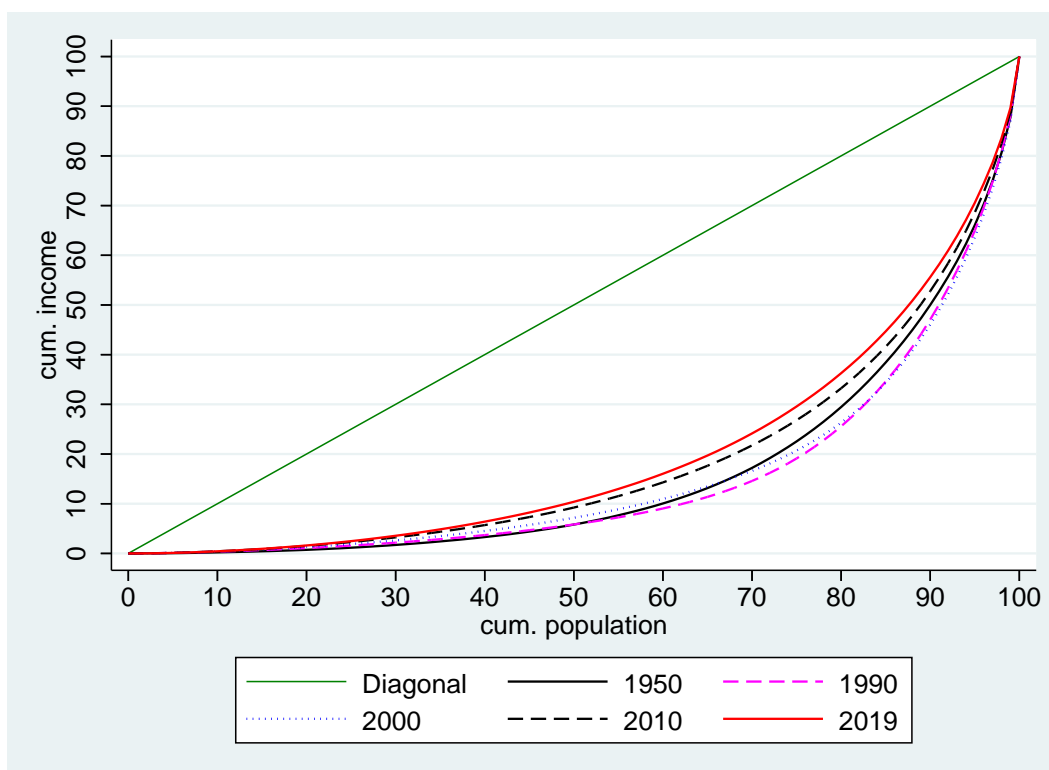
short term. The lack of dominance when each decade is compared with the next may imply multiple crossings. For example, the period between 1970 and 2000 involves two crossings at the bottom and one at the upper middle, pointing to higher inequality if this upper tail is given more relevance (e.g. percentile 78 when comparing 1980–2000, Figure 7b).

This lack of dominance when comparing most periods implies that the assessment of whether inequality declined or increased cannot generally be entirely made based only on the principles of anonymity, population invariance, scale invariance, and transfers. The combination of equalizing and disequalizing changes at the same time means that one needs to be more explicit about the importance attached to each of those changes in order to assess the inequality trend. However, the pattern in most crossings that occur at the extremes of the distribution generally mean that the trend can be assessed with a high degree of unanimity among most inequality indices (and their underlying normative criteria). For example, when comparing the most recent decades (e.g. 2000–19, Figure 7c) the only crossing occurs at the very bottom (percentile 7). This pattern points to inequality declining since 2000 unless we give a large weight to the relatively worse performance of the very bottom of the distribution.

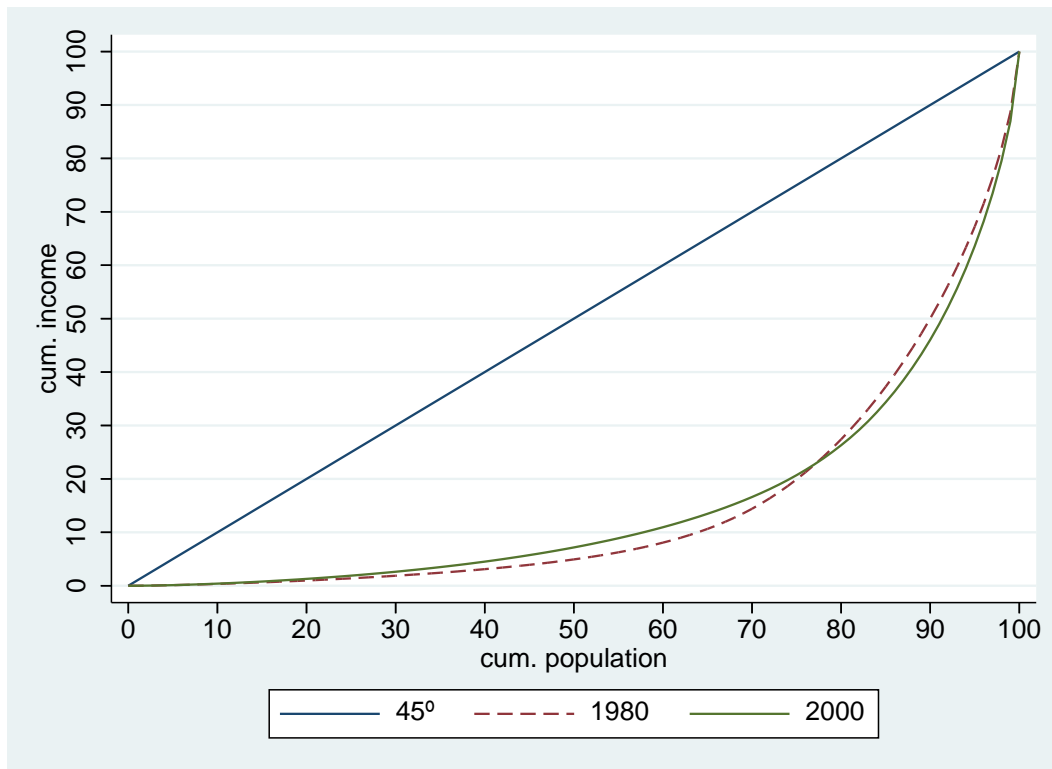
In terms of the absolute Lorenz curves (Figure 8), however, the results are quite robust, indicating an unambiguous increase of absolute inequality between the years being compared (the curves move away from the horizontal axis). Therefore, the trend does not depend on any distributive sensitivities and every inequality index which is consistent with these absolute Lorenz orderings will also point to an upward trend every decade and in the long term.

Figure 7: Global relative Lorenz curves, comparing different selected years

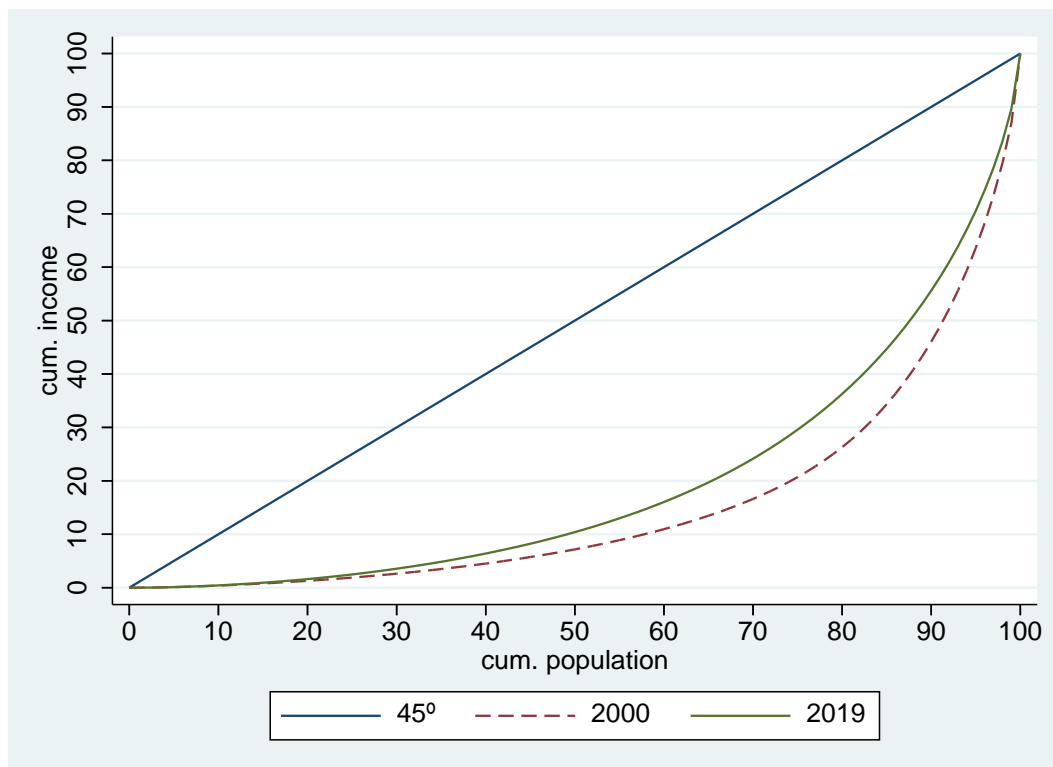
a. Multiple years, 1950–2019



b. 1980–2000



c. 2000–19



Note: relative Lorenz curves map the accumulated income share that corresponds to each accumulated share of the population (with percentiles ordered from poorest to richest).

Source: author's construction based on the WIID Companion.

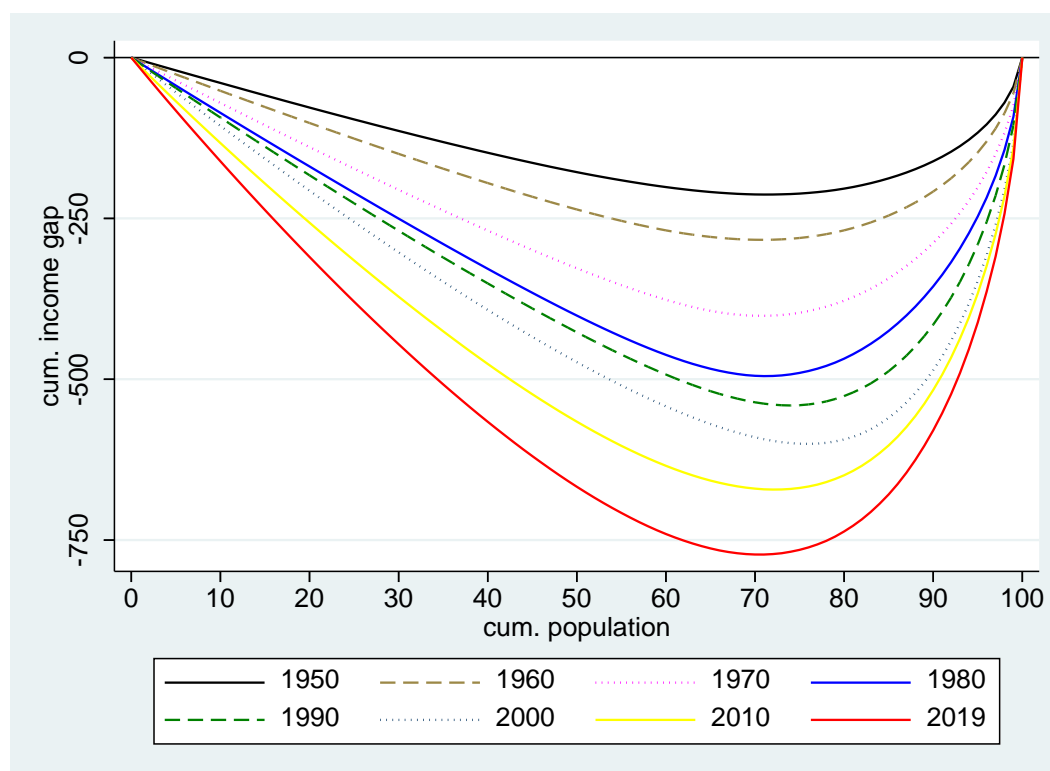
Table 1: Lorenz dominance (crossing percentile)

	1950	1960	1970	1980	1990	2000	2010
1960	-41 82						
1970	-34 89	2 -25 94					
1980	-36 91	3 -32 96	4 -43				
1990	-51 99	-56	-2 5 -69	-3 8 -72			
2000	-67	-70	-2 5 -76	3 6 -78	-3 4 -84		
2010	decline	4	5	6	5	6	
2019	decline	4	6	-2 7	6	7	-2 10

Note: decline = Dominance (unambiguous decline in inequality). Numbers indicate the percentile at which the most recent curve crosses the older one from below (positive) or from above (negative).

Source: author's construction based on the WIID.

Figure 8: Global absolute Lorenz curves, comparing different selected years



Note: absolute Lorenz curves map the accumulated income gap with respect to the mean (thousands, 2017 PPP USD) that corresponds to each accumulated share of the population (ordered from poorest to richest).

Source: author's construction based on the WIID.

5.3 Global aggregate measures of relative inequality

In contrast to the lack of Lorenz dominance, much clearer general trends emerge using main summary measures like the Gini index, the MLD, or the Theil index (Figure 9). In these cases,

there is an initial phase of certain stability (or a small increase) which is followed by a much sharper decline. The turning point year and the magnitude of changes vary with the inequality index used.

By construction, all inequality indices are sensitive to the extremes of the distribution, but to a different extent. For example, the Gini index, known to be less sensitive to the extremes of the distribution compared with other indices, starts with a level of almost 68 in 1950 or 1960, followed by an upward trend, reaching a peak of 70 in 1991. Later, it sharply falls to its lowest level of 61 in 2019. The MLD, with higher sensitivity to the bottom of the distribution compared with the Gini measure, initially shows more persistence, with some increase between the mid-1950s and mid-1970s (from 110 in 1953 to its peak of 119 in 1976), before exhibiting a much more sustained decline, reaching its lowest level of 80 in 2019. The Theil index, which is more sensitive to the top of the distribution than the other two indices, increases from 85 in 1958 to its maximum of 95 in 1994, before declining more strongly to 68 in 2019. If even more emphasis is put at the top (GE(2) or, equivalently, CV), inequality is rather stable until it sharply increases between 1983 and 1994, and then sharply declines thereafter: it starts with a level of 157 in 1950 before declining to 147 in 1983. It then exhibits a sharp increase to its maximum of 180 in 1998 before falling sharply again to 115 in 2019. The story, however, is different with indices that exhibit a much higher sensitivity to the bottom, e.g. GE(-1). Inequality continuously and strongly declines after 1950 until around 2005, increasing thereafter, especially after the financial crisis, ending at 340.

In summary and consistent with the analysis of Lorenz dominance, one common feature of all these indices is a decline in inequality over the entire period from 1950 or 1960 to the present. The initial decades generally exhibit more stability, with a tendency to increase inequality, except when we attach a large weight at the strong relative growth of bottom incomes, in which case inequality more clearly declines. In most recent decades, inequality clearly declines with most indices, i.e. with the MLD (from 1976), Gini index (from 1991), or Theil and GE(2) (from 1998). Inequality only increases with GE(-1) in recent years (from 2005), highlighting the income stagnation at the very bottom of the global distribution over this period. This story is somewhat corroborated by the Atkinson family of indices shown in the data appendix,¹⁶ with inequality declining from around 1990 with most of them (0.25, 0.50, or 0.75), earlier with higher sensitivity to the poor, i.e. 1976 with A(1). Meanwhile, A(2) exhibits a more pronounced continuous decline, but with an upward slope from 2005.

There is therefore a great level of consensus among indices that global relative inequality has strongly declined in recent decades, changing the previous trend. The main source of discrepancy about what happened with inequality over the last decades relies on putting a large weight at the extremes, particularly at the very bottom, in which case the story is totally reversed with inequality first falling deeply and then exhibiting a small increase more recently.

The reasons for this discrepancy between indices, which originate in crossing Lorenz curves, are well illustrated by looking at the two most common income shares, those for the top 10 and bottom 40 per cent (Figure 10a). The share of the global top 10 per cent tends to remain around 50 per cent of total income until around 1984, when it starts to rise sharply, reaching its maximum of 55 per cent of global income in 1994. Then, it declines to its current level of 45 per cent. At the same time, the income share of the bottom 40 per cent initially declines from 3.3 per cent in 1950 to 2.9 per cent in 1976, and then sharply increases to its current level of 6.3 per cent. That is, both income shares showed a similar trend until recently. They were initially stable and then increased in the 1980s and 1990s, with obviously opposite effects on inequality. There is a trade-off between

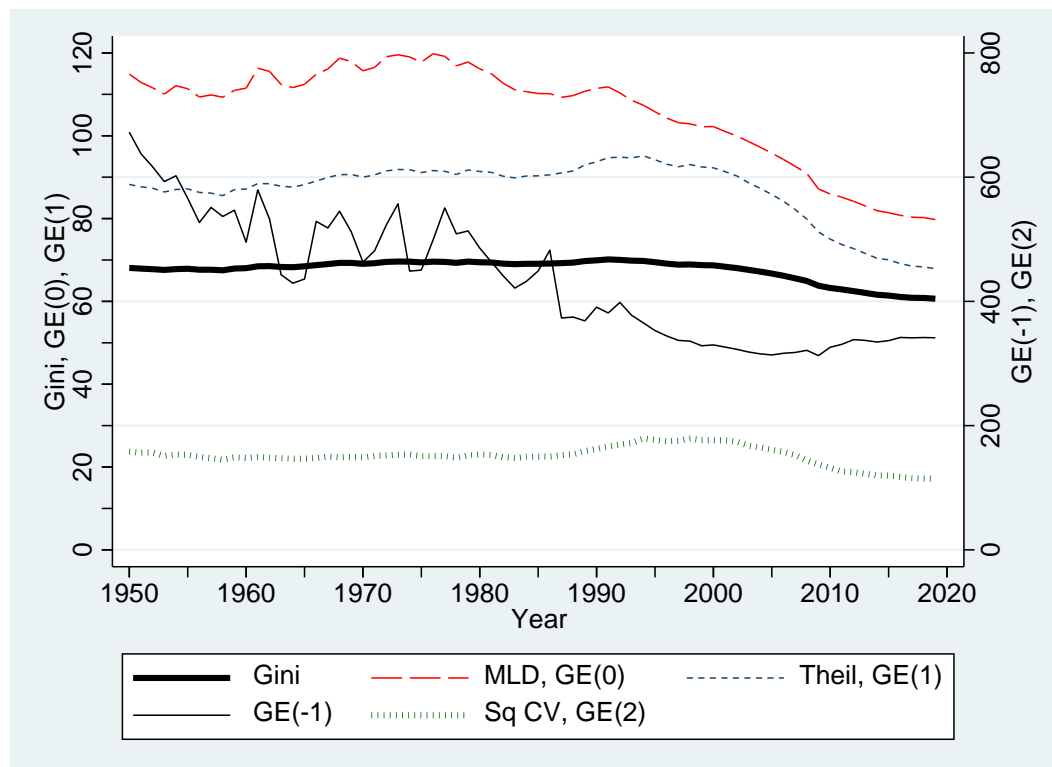
¹⁶ Note that $\varepsilon > 0$, $A(\varepsilon)$ is ordinally equivalent to $GE(\alpha)$, with $\alpha = 1 - \varepsilon$. Therefore, $A(1)$ is ordinally equivalent to $GE(0)$ or MLD, while $A(2)$ is ordinally equivalent to $GE(-1)$.

the disequalizing higher share of the global rich and the equalizing higher share of the global poor at the same time during these years. These gains obviously come at the expense of the share of the middle of the distribution. Whether one gives more weight to one or the other effect is a value judgement that may not generate consensus among people with different distributive sensitivities, who may conclude that inequality increased or declined. Indices that are more sensitive to the top, like the GE(2), will register the period with a sharp increasing share at the top as increasing inequality (even if the bottom 40 per cent also improves), while other indices will give more relevance to the improvement of the poor (despite the fast accumulation at the top).

Later, from around 2000, both income shares become more aligned in terms of their impact on inequality, reinforcing each other, with a fall in the top 10 per cent share and an increase in the bottom 40 per cent share, driving the strong decline in inequality reported by most indices over this period. The Palma index, which is the ratio between the incomes of these two groups, shows an increase in inequality from 14 in 1956 to its maximum of 17.5 in 1976, before continuously falling to its current level of 7.0. That is, this ratio shows a trend which is almost identical to the one we get using the MLD, implicitly giving greater weight to the improvement of the income share of the bottom 40 per cent of the world's population when the share of the top 10 per cent was also increasing.

The discrepancies shown by GE(-1), compared to other indices, originate in the relatively better performance of the very bottom during the initial decades and its stagnation in the last 30 years (Figure 10b, bottom 5 per cent compared with bottom 20 or 40 per cent).

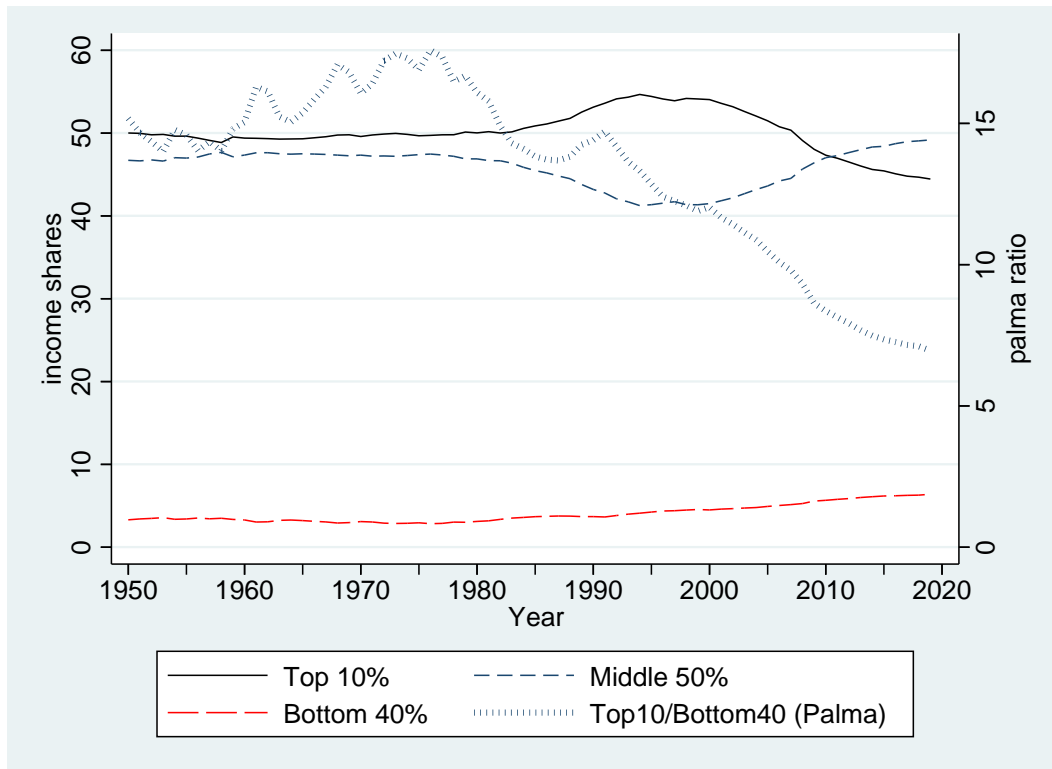
Figure 9: Relative measures of global income inequality, Gini and GE family



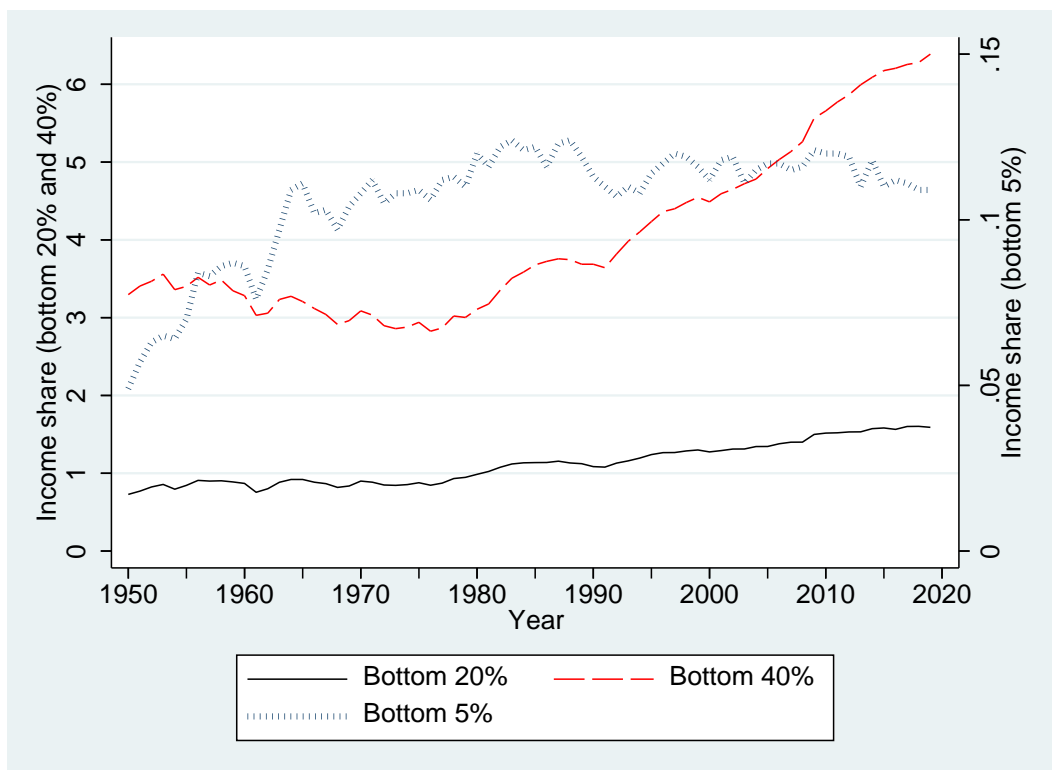
Source: author's construction based on the WIID.

Figure 10: Income shares

a. Income shares and Palma ratio



b. Bottom income shares

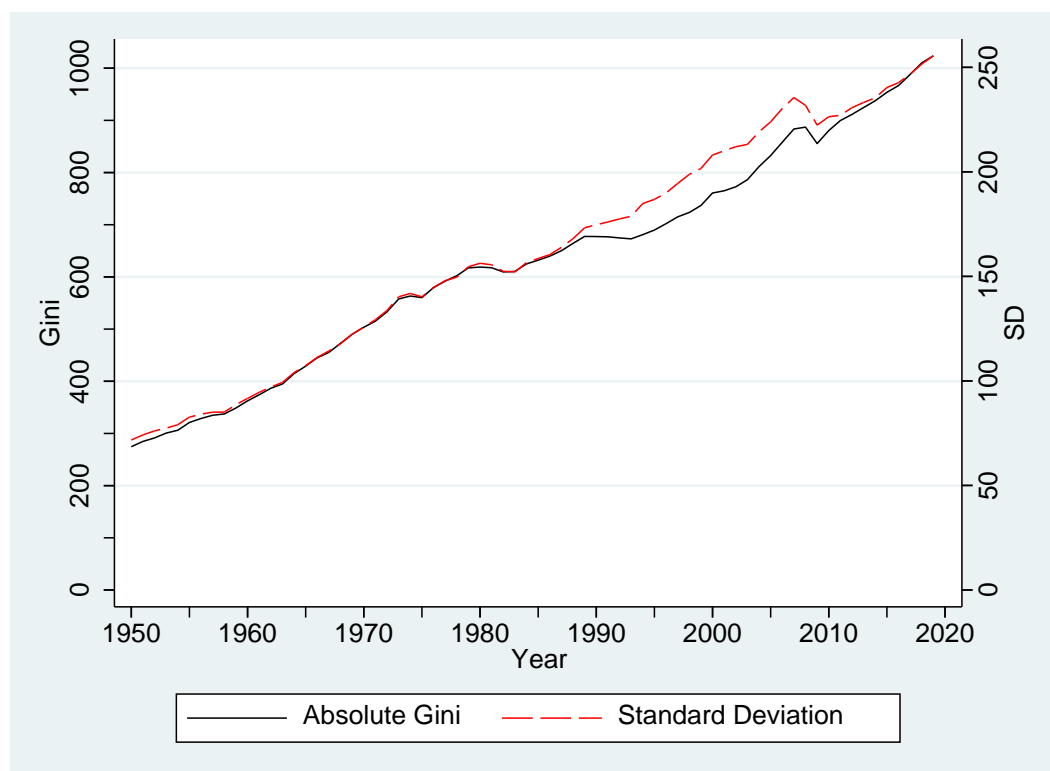


Source: author's construction based on the WIID.

5.4 Global aggregate measures of absolute inequality

From an absolute perspective, when the ruling principle to identify what happens to inequality when total income has changed shifts from scale invariance to translation invariance (i.e. uniform real income changes measured in PPPs do not affect inequality), it turns out that inequality continuously increases over most of the period analysed. This can already be inferred from the existence of (absolute) Lorenz dominance and therefore unambiguous increases in inequality in each decade, but the use of indices helps to highlight that the upward trend is almost continuous annually (Figure 11). The only exceptions are the short episodes of global recessions, such as in 1974–75, 1980–82, and particularly 2008–09, with both indices used here, the standard deviation, and the absolute Gini index (which multiplies the relative Gini by per capita income). The absolute Gini index also shows a decline during the 1990–93 recession. This is not surprising as the global distribution of income is characterized by strong sustained economic growth, a context in which it is unlikely that absolute distances between people are reduced, as reflected by the absolute GIC discussed above. This is true within countries but is even more the case when considering all the world's citizens given that initial income differences are even more striking.

Figure 11: Absolute measures of global income inequality



Note: Absolute Gini/1000; Standard deviation/100.

Source: author's construction based on the WIID.

6 Between-country versus within-country inequality

6.1 The approach

When investigating the drivers of the trends in global inequality, the first question that arises is whether the trend is the result of changes in inequality between or within countries. It is well known that, unlike what is usually observed in many countries, when it comes to describing global inequalities, the location component, i.e. inequality between countries, emerges as the main driver. This may sound like a paradox, as the very notion of global inequality is to break down country borders and consider the world as a unity itself. Even if ignoring borders makes sense from a normative point of view, the reality is that the country of residence has been and continues to be the main determinant of our position in the world's income distribution and cannot be ignored.

I estimate Lorenz curves and inequality measures in two counterfactual distributions. The approach is explained in more detail in the Appendix A. One is a counterfactual in which all inequality within countries has been removed, with all remaining inequality being attributed to 'between-country inequality'. This is done by giving each country percentile the mean income of the country, or the 'equally distributed equivalent' (EDE) income in the case of the Atkinson family.¹⁷

The other is a counterfactual in which all inequality between countries has been removed, with the remaining inequality being 'within-country inequality'. This is achieved by multiplying all incomes by the ratio between global and country mean incomes (or EDEs), in the case of relative inequality. The resulting level of within-country inequality is also the population-weighted average of country inequality in the case of the GE family, but not for Gini or the Atkinson family, in which case I also report the corresponding weighted average. In the case of absolute inequality, the within-country distribution is obtained by adding the differential between the global and country means instead (to keep inequality within each country constant).

It is well known that inequality indices have different decomposability properties. It is only in the case of the MLD that the between- and within-country inequality components as defined above add up to overall inequality. In the other cases, the reduction in inequality after equalizing incomes within and between countries can be seen as alternative measures of respectively inequality within and between countries. For that reason and to address the analysis in a meaningful way using any inequality measure, in line with Davies and Shorrocks (2021), I also use the Shapley approach (Chantreuil and Trannoy 2013; Shorrocks 2013) to estimate the share of overall inequality that is explained by each term. This is the average of each term obtained in the two possible sequences that can be followed to estimate them (i.e. first removing within-country inequality, or first removing between-country inequality).

6.2 Lorenz dominance in between-country and within-country distributions

In terms of relative Lorenz dominance, the between-country curves tend to cross multiple times. This is the ultimate reason for the lack of Lorenz dominance in the overall distribution and for the discrepancies among inequality measures, depending on the weight attached to each part of the distribution (Figure 12). The situation is close to dominance only in the last decades. For example, the curves between 2000 and 2019 only cross once at the very bottom (third percentile), indicating

¹⁷ That is, the level of income that, if equally distributed among the population, would give the same level of social welfare as the actual (unequal) distribution, using an isoelastic utilitarian social welfare function (Atkinson 1970).

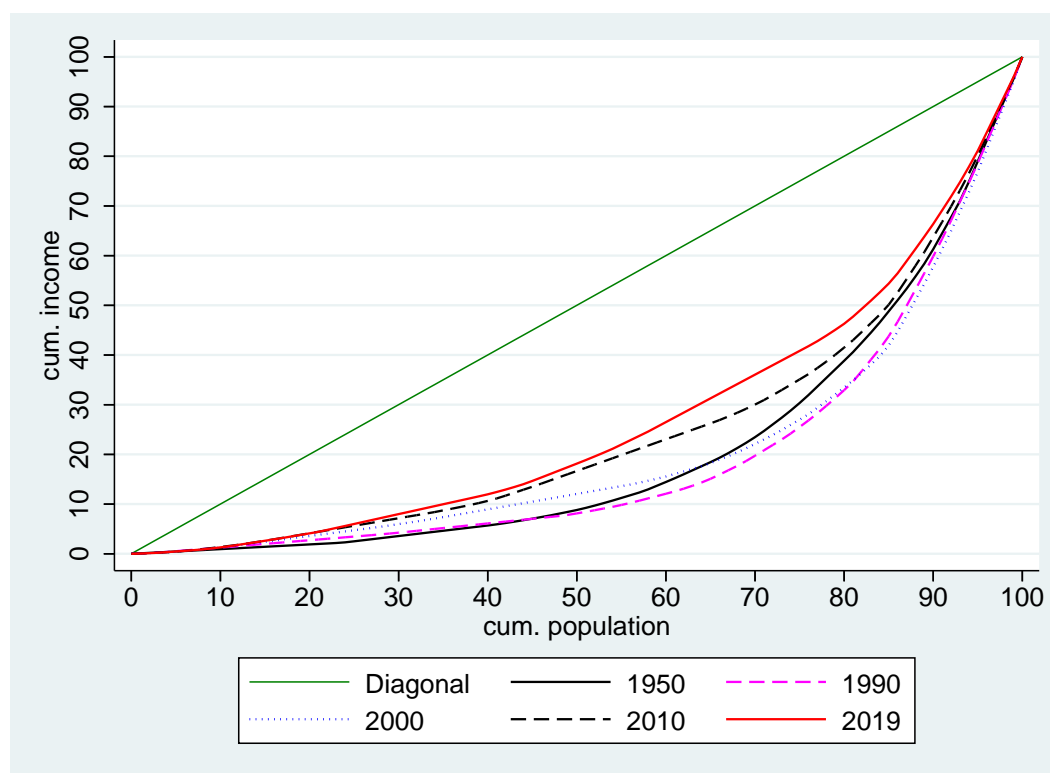
that there is almost unanimity among inequality measures, except for those extremely sensitive to the lack of improvement in incomes in the poorest countries.

Regarding the within-country distributions, the curves tend to be generally much closer to each other, indicating that the magnitude of changes is much smaller than observed between countries (Figure 13, Table 2). Inequality within countries unambiguously declines between 1950 and 1980, with dominance in the 1950s and 1960s, but not in the 1970s. Between 1980 and 2019, however, there is dominance, indicating an unambiguous increase in inequality within countries for that entire period. The situation is less clear by decade, with curves crossing either at the bottom, in the 1980s and the 1990s when most (disequalizing) drastic changes occur, or at the top, in the 2000s and 2010s when the curves are almost identical. This points to potential discrepancies between indices in the short term but to more general agreement again in the long-term trends.

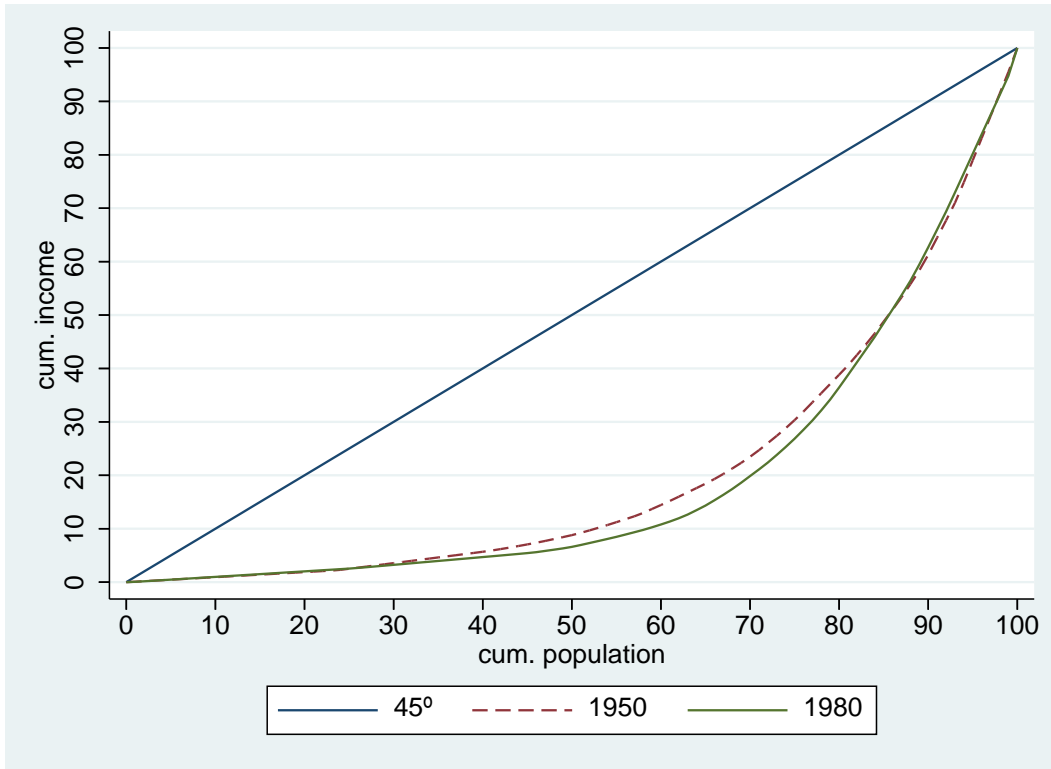
When compared together, the Lorenz curves of the distributions between and within countries cross at both extremes of the distribution every year (see data appendix).

In terms of absolute inequality, there is dominance every decade both within and between countries, indicating that inequality unambiguously increases over time pushed by both components.¹⁸

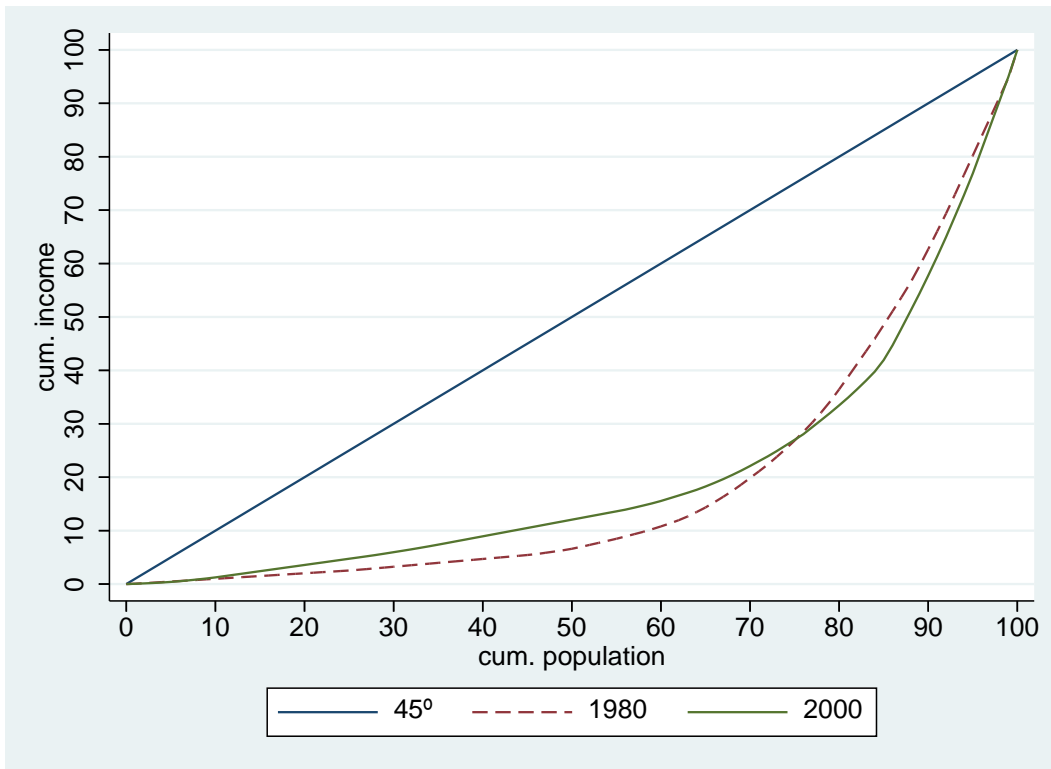
Figure 12: Lorenz curves, between-country distribution



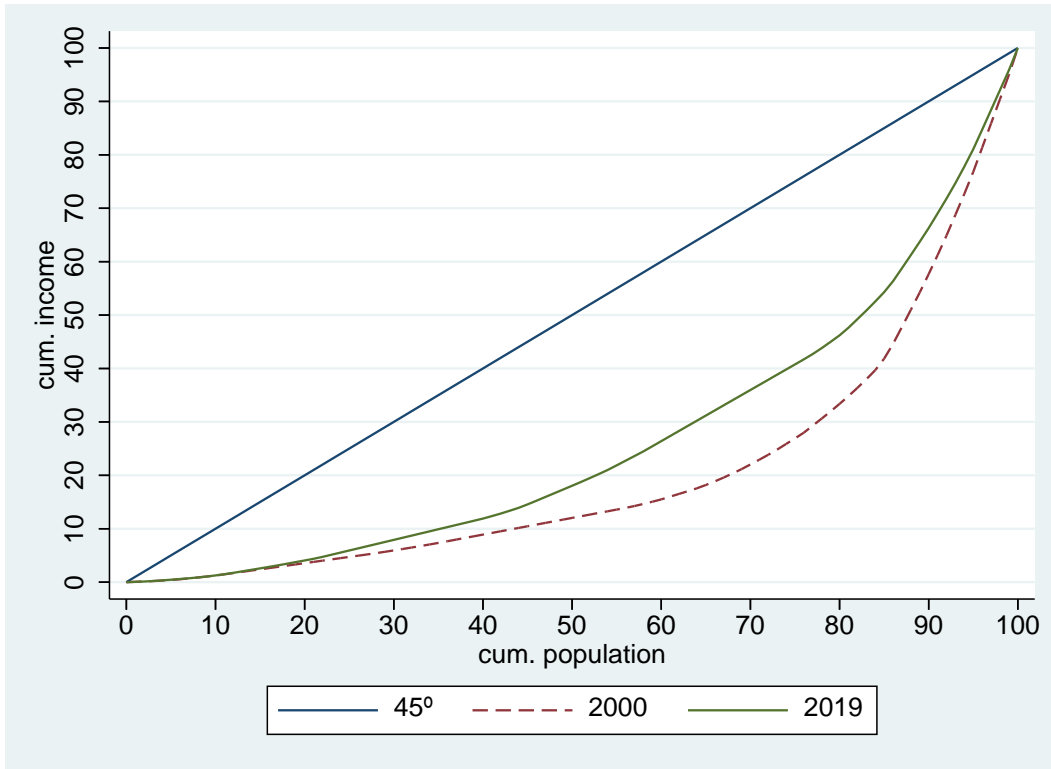
¹⁸ The only exception is a crossing at percentile 99 between the curves for 1980 and 1990.



Note: there is no dominance. The curves cross at various percentiles 4, 26, 87, and 98.



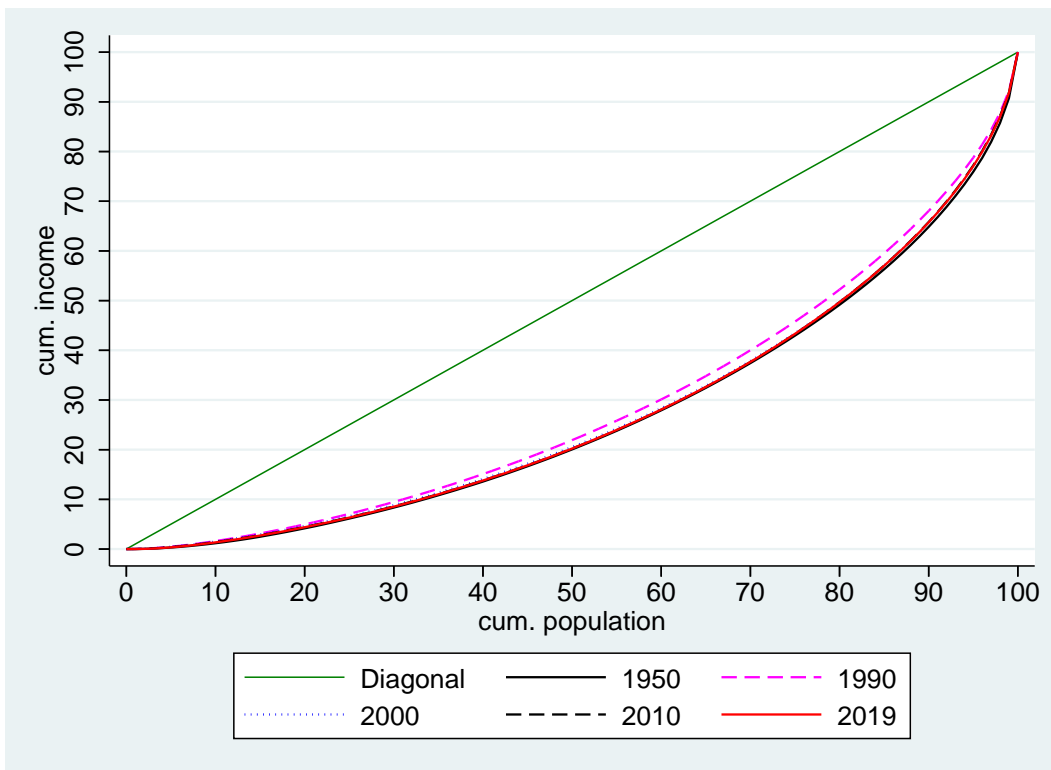
Note: there is no dominance. The curves cross at various percentiles 7, 76, and 99.

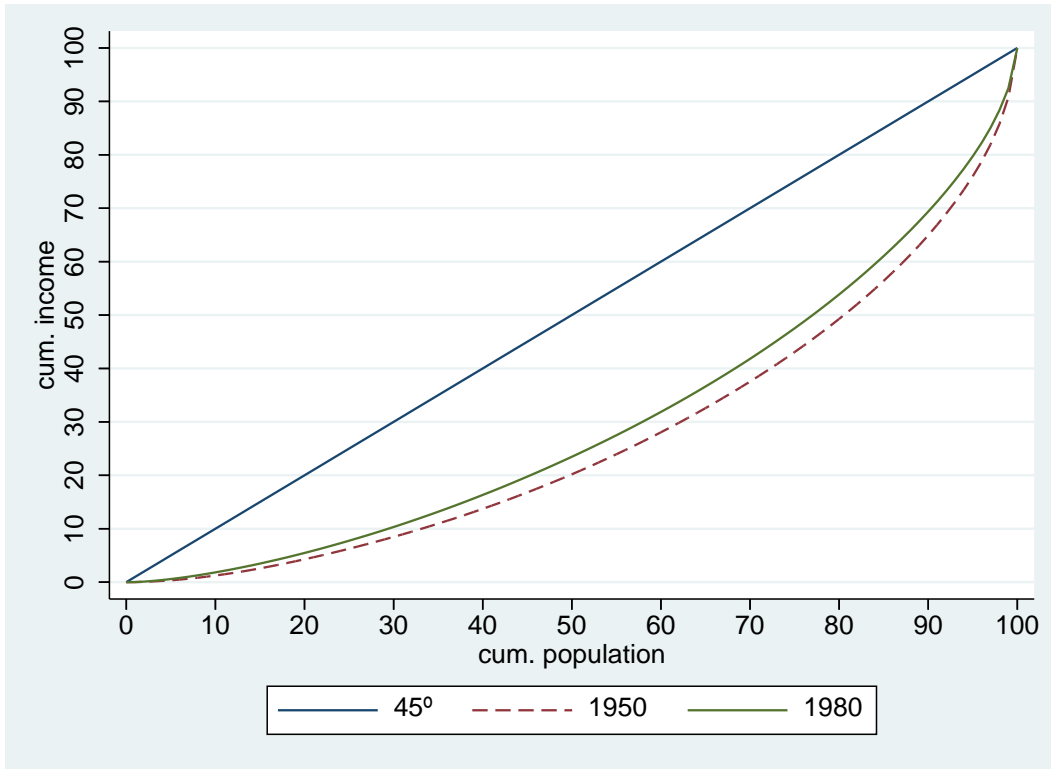


Note: there is no dominance. The curves cross at percentile 3.

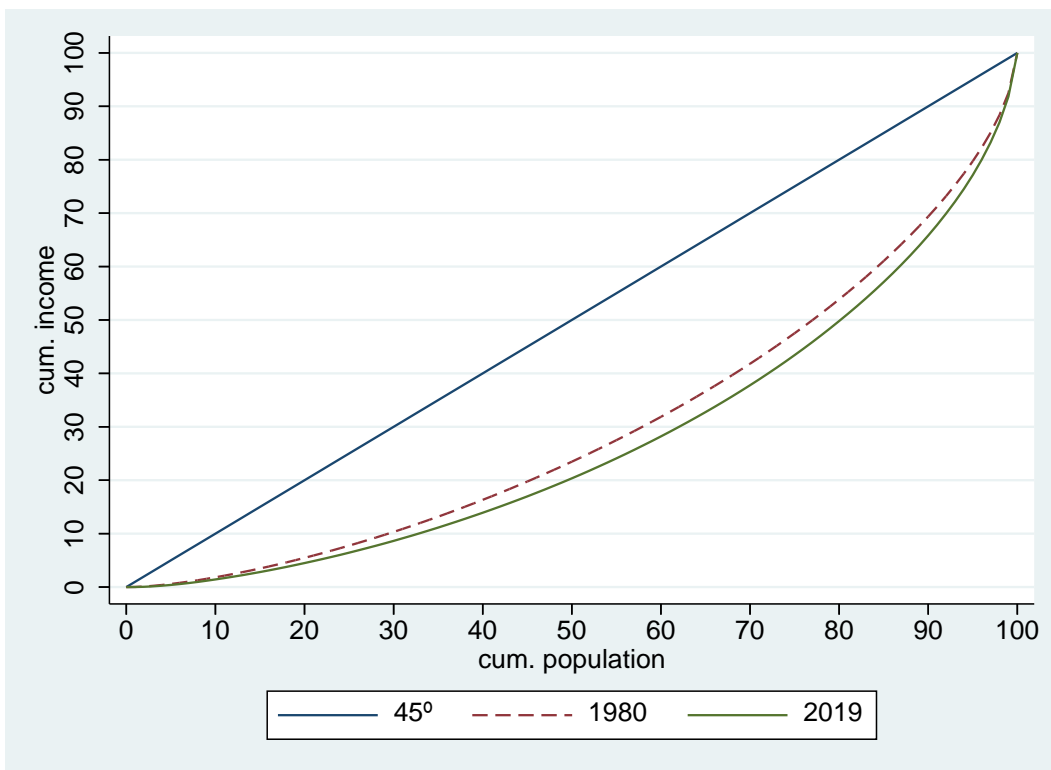
Source: author's construction using WIID Companion.

Figure 13: Lorenz curves, within-country distribution





Note: there is dominance. Inequality declined.



Note: there is dominance. Inequality increased.

Source: author's construction using WIID Companion.

Table 2: Lorenz dominance (crossing percentile), within-country distribution

	1950	1960	1970				1980	1990	2000	2010	
1960	decline										
1970	decline	decline									
1980	decline	decline	2	-6	14	-97					
1990	decline	-4 9	-3					-3			
2000	decline	-3 99	-3					-3	-2		
2010	decline	-2 98	increase				increase	increase	85		
2019	decline	99	increase				increase	increase	87	49	-82

Note: decline/increase = Dominance (unambiguous decline/increase in inequality). Numbers indicate the percentile at which the most recent curve crosses the older one from below (positive) or from above (negative).

Source: author's construction using WIID Companion.

6.3 Inequality measures in between-country and within-country distributions

The analysis using inequality indices allows us to quantify the importance of each component in overall inequality over time (Figure 14), highlighting a few stylized facts.

First, the between-country contribution tends to be larger than the within-country term with all indices until recently, when the within-country component becomes close to the between-country term (e.g., 1.8 Gini points in 2019 compared to 21 in 1980; 2.2 MLD points now compared with 55 in 1980), or even larger (e.g., Theil since 2018, GE(2), and Theil(-1) from 2009). The unanimity of the relevance of between- and within-country components across indices can be confirmed by the corresponding Shapley share of overall inequality, which is explained by inequality between countries, exhibiting an-inverse U over time (Figure 15). The maximum relevance of inequality between countries was achieved in the late 1970s and early 1980s in all cases.

Second, both components tend to move in opposite directions with the between-country term driving the general trend in global inequality (first increasing, later decreasing) while the within-country term partially offsets that trend (first decreases, later increases). This dealignment of both terms is more balanced during the first phase, resulting in greater overall stability, but is less so in the second phase in which the decline in inequality between countries is much stronger than the increase in inequality within countries.

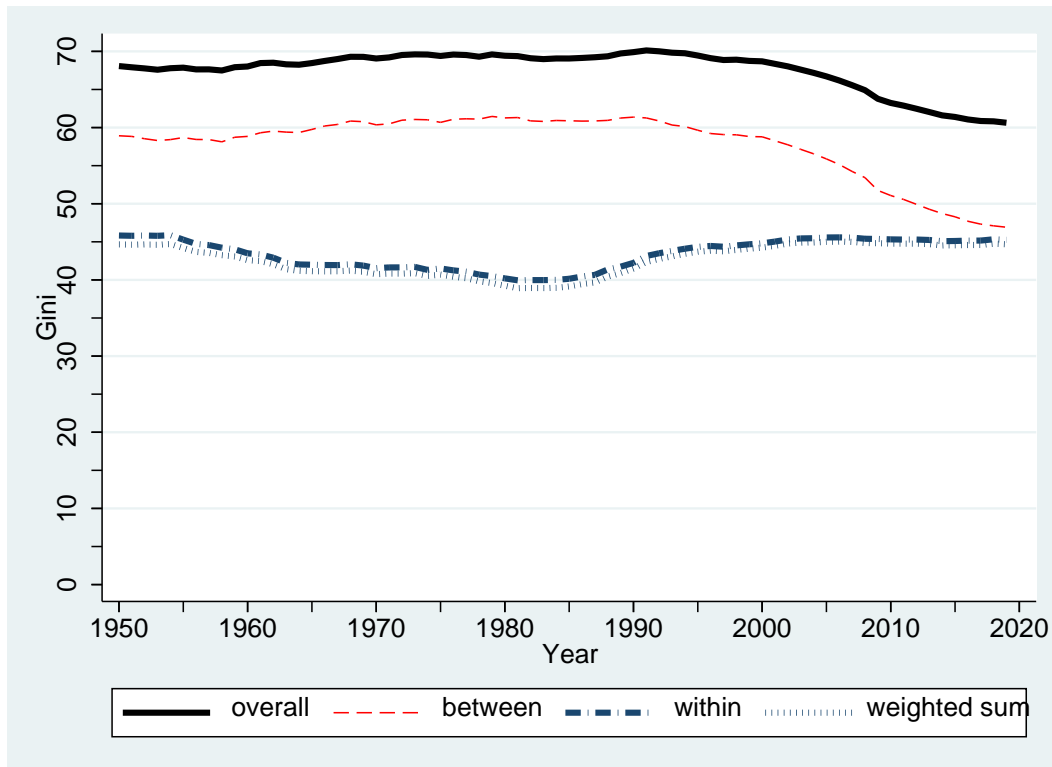
Third, as expected from the Lorenz analysis, these indices tend to agree more in pointing to an increase in inequality within countries from the mid-1980s, particularly before the mid-1990s. However, they disagree more in how they evaluate the trend in inequality between countries, especially related to when the decline starts (earlier with MLD, later with Theil index, in between with the Gini index). Therefore, we can conclude that the differences in how different relative measures evaluate global inequality are more related to the impact of changes in average incomes across countries than to changes within countries. The index with extreme sensitivity to the bottom of the distribution, GE(-1), also exhibits a decline in between-country inequality from the mid-1970s but is less steep than other indices so the increase in within-country inequality dominates the trend during the most recent years.

Fourth, all indices exhibit a deceleration in the decline of inequality between countries in the most recent years.

There is therefore no doubt that while declining within-country inequality helped to partially compensate for the increase in between-country inequality before 1990, the roles were reversed thereafter, with rising within-country inequality partially compensating for the strong decline in inequality between countries. The increasing within-country inequality during the last decades is the result of a heterogeneity of trends across regions and subperiods. Figure 16, for the Gini index (other indices in the data appendix) highlights the recent increases in population-weighted inequality predominant in East Asia and Pacific, South Asia, and North America, with declines in the Middle East and North Africa, sub-Saharan Africa (since 1992), Europe and Central Asia (since 1995), and Latin America and the Caribbean (since 1998).¹⁹

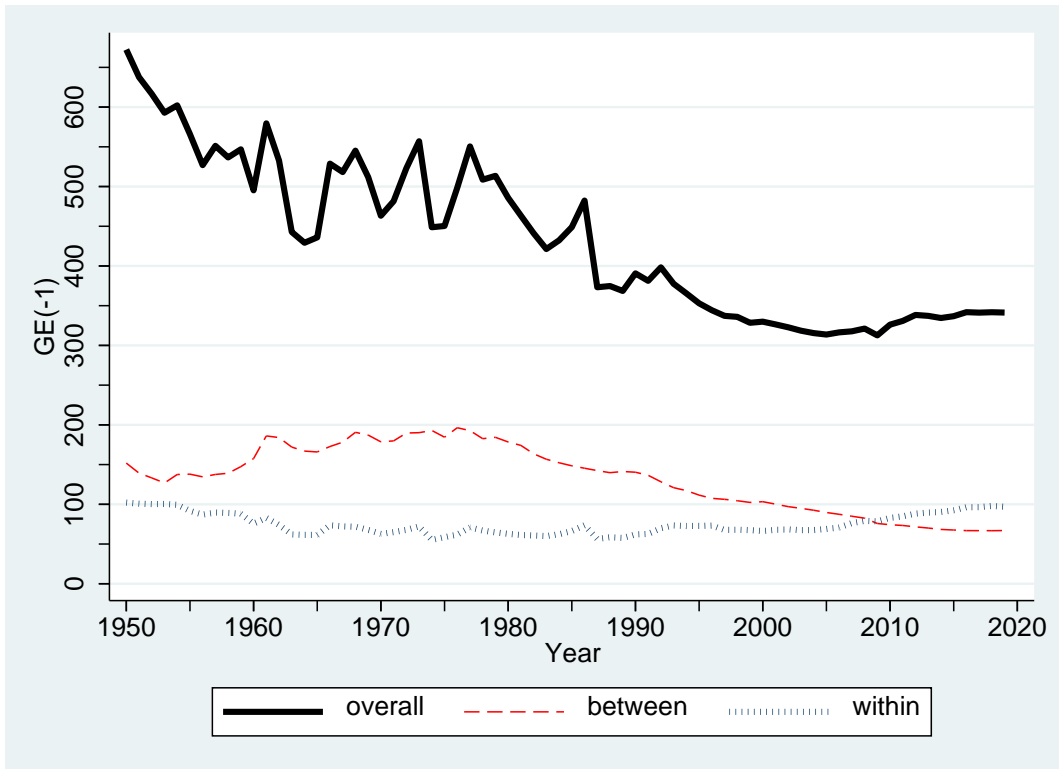
In the case of absolute inequality (Figure 14), both between- and within-country inequality components contributed to the sustained increase over time, even if, since the 2000s, absolute inequalities within countries seem to be more relevant to explain the upward trend.

Figure 14: Decomposition of overall global income inequality into between-country and within-country inequality
a. Gini

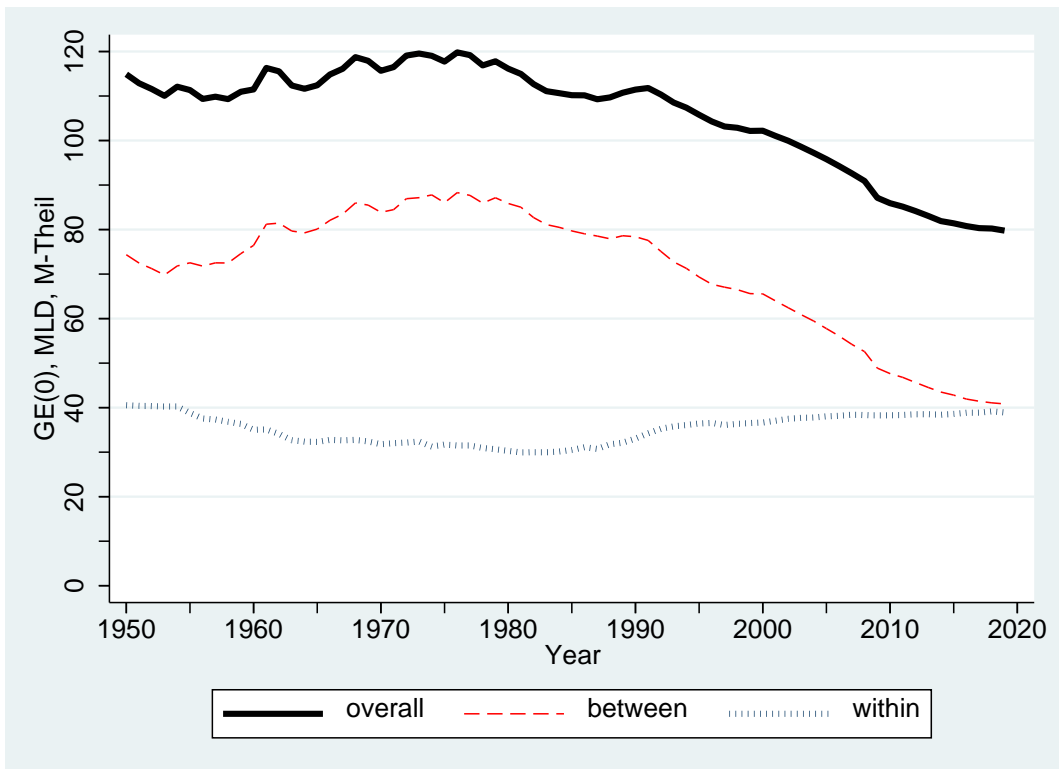


¹⁹ For an exhaustive analysis of inequality trends in five development countries (Brazil, China, India, Mexico, and South Africa), see Gradín et al. (2021a).

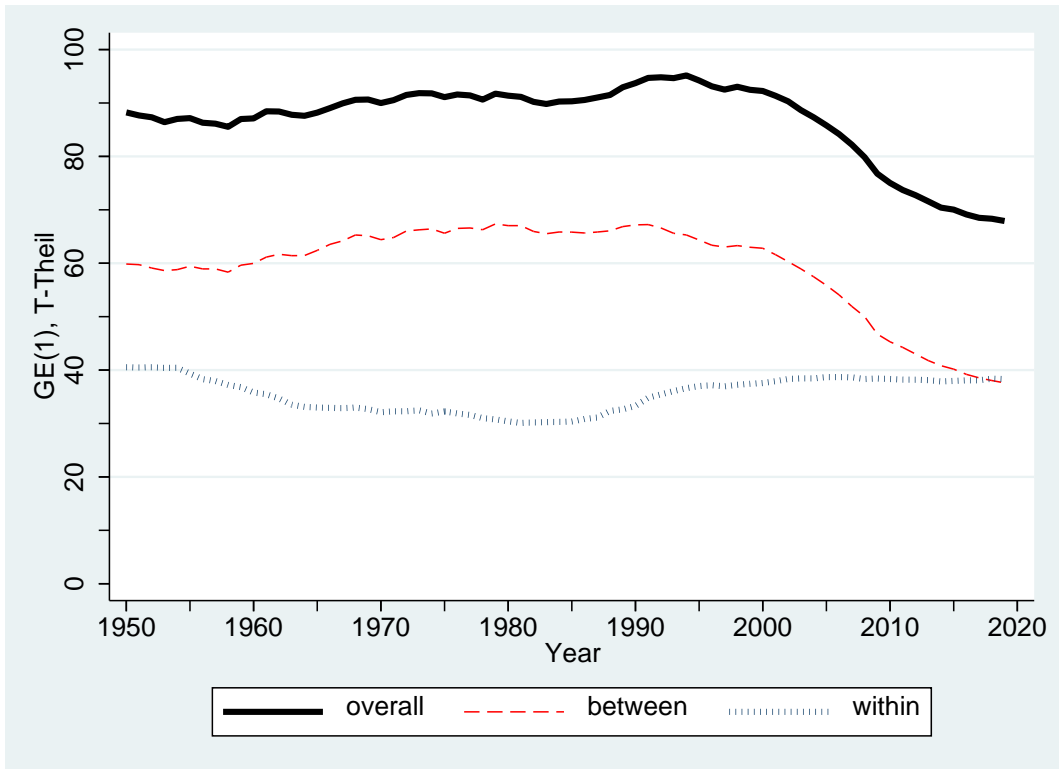
b. GE(-1)



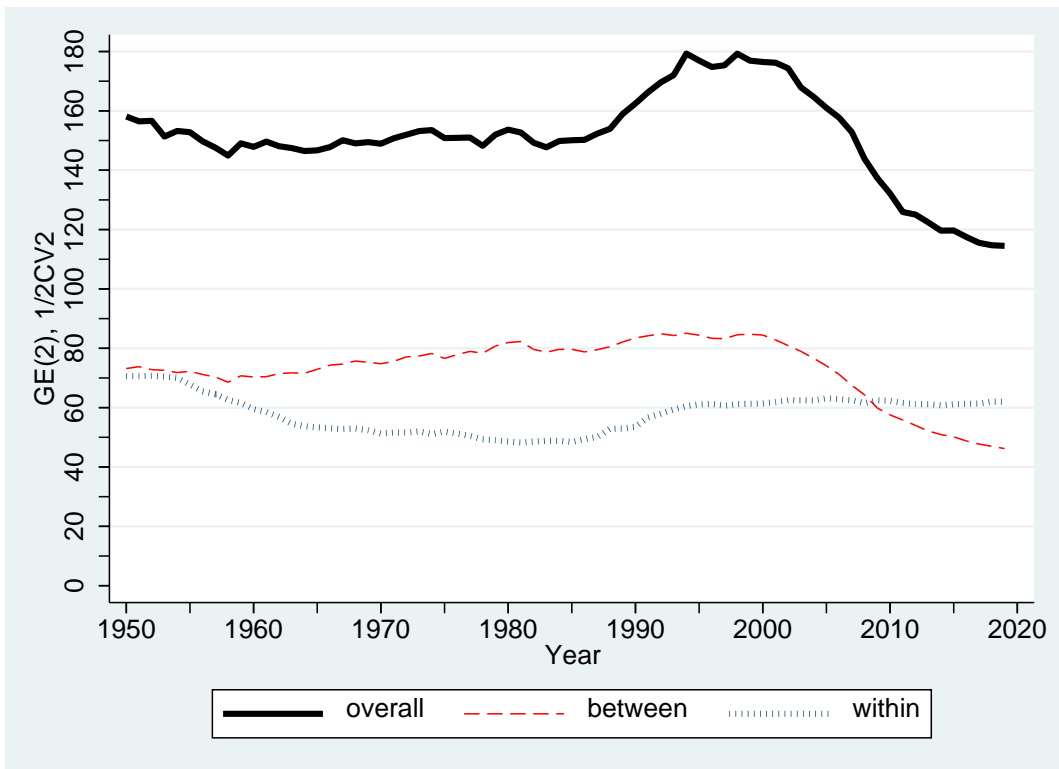
c. MLD, GE(0)



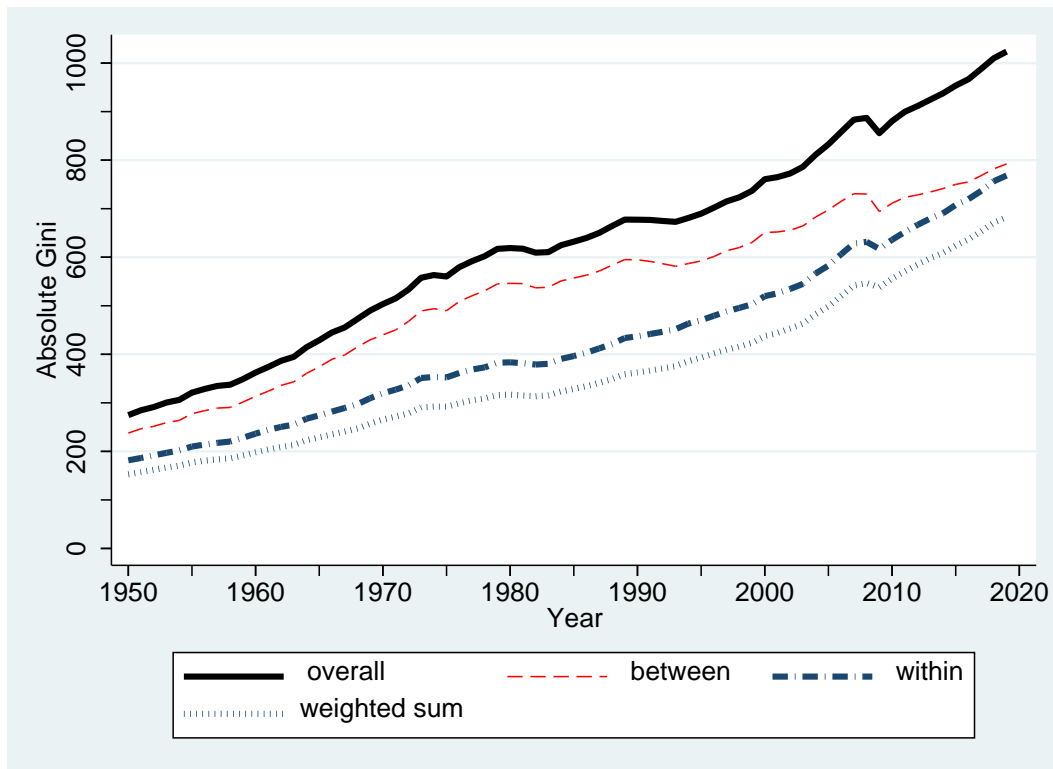
d. Theil, GE(1)



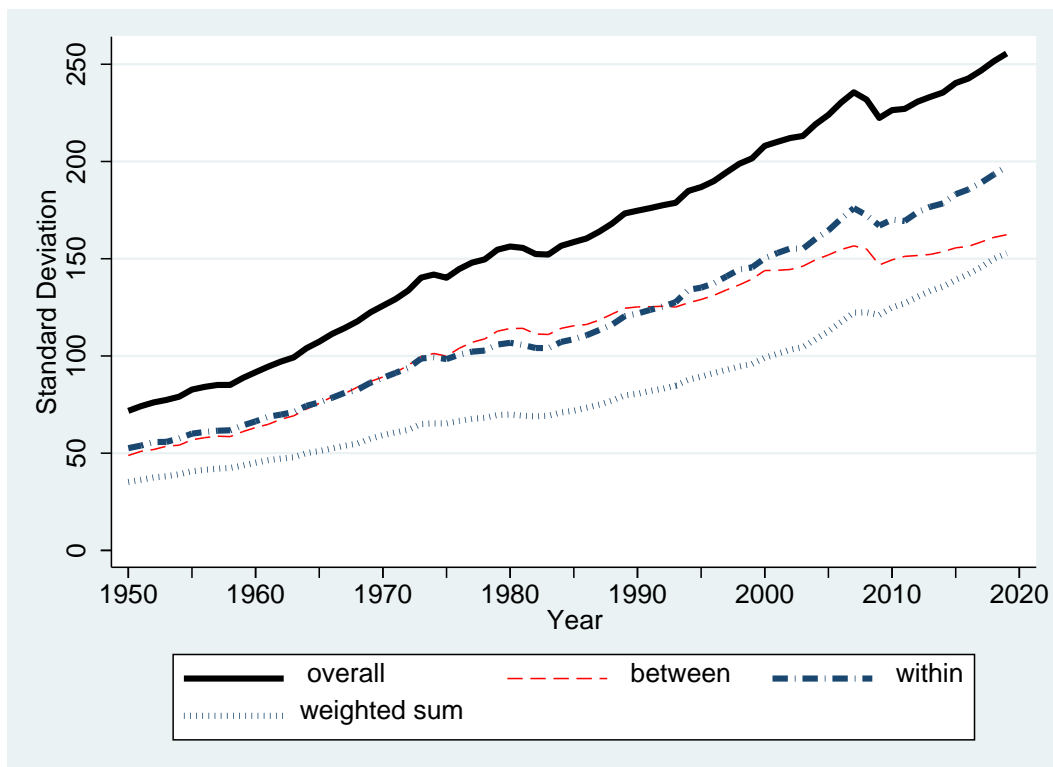
e. $\frac{1}{2}$ Squared CV, GE(2)



f. Absolute Gini

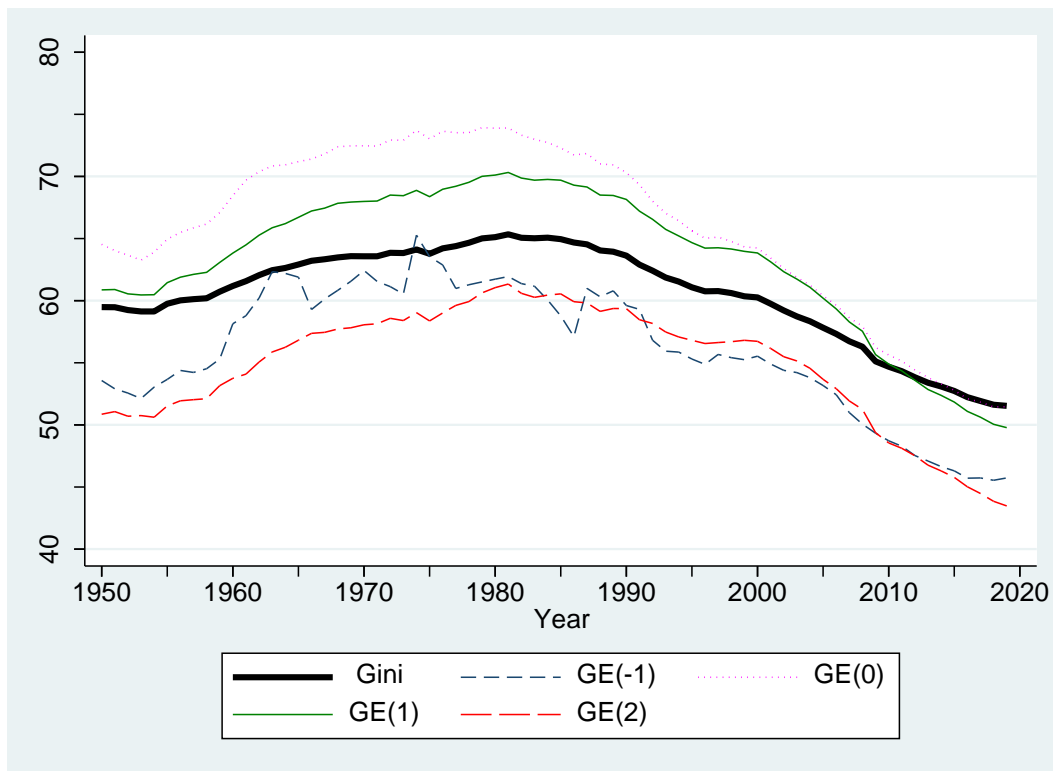


g. Standard deviation



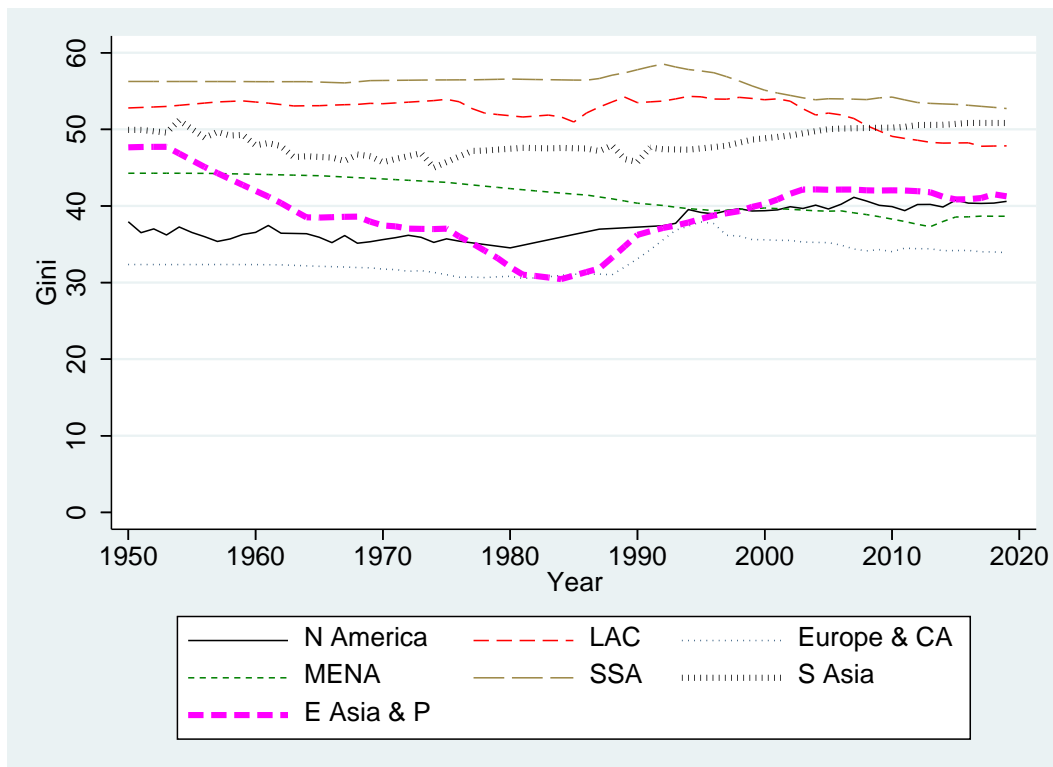
Source: author's construction based on the WIID.

Figure 15: Decomposition of overall global income inequality into between-country and within-country inequality: Shapley between-country contribution (%)



Source: author's construction based on the WIID.

Figure 16: Population-weighted average in inequality by region (Gini index)



Source: author's construction based on the WIID.

7 Main country contributors to the inequality trends

7.1 The approach

A second aspect of the global inequality trend that has been highlighted so far is the existence of big players, particularly China, in driving this trend. China's accounting for 19 per cent of the total world population in 2019 (down from 22 per cent in 1950) and its GDP per capita having multiplied by 42 over this entire period, along with the fact that the country has witnessed substantial changes in its income distribution (equalizing before the market reforms, mainly disequalizing afterwards), raises the question of the extent to which China alone is driving the global trends. A similar question can be asked of India, which has 18 per cent of the current world population and has recently shown sustained economic growth and structural reforms, or of other countries or regions. Similarly, the collapse of communist regimes in the 1990s followed by falling mean income and rising inequality can also explain the trend in inequality, at least during specific periods. However, demographic trends have also increased the share of the world's population living in developing countries, for example the doubling of the share of people living in sub-Saharan Africa (from 7 to 14 per cent), the geographical region which at the same time has become the poorest of all seven.

One common approach used in the literature to address China's contribution is to measure the trend in inequality with and without China in the sample, interpreting the change in inequality after adding China as its contribution to the trend. This 'marginal' approach to measuring the contribution of the country to total inequality can be misleading, though, and the values hard to interpret. If repeated with every country, the sum of the contributions of all countries will not equal the total observed level of inequality (the decomposition is inconsistent). One can also envisage a situation in which all countries have the same income distribution (same per capita income by percentile). Adding any country at the end would not change the global distribution of income and so the contribution of each country would be zero, leaving all global inequality unexplained.

Instead, here, I follow the approach in Gradín (2020), as explained in Appendix B, which estimates the contribution of any population group (country in this case) to inequality based on the sum of the contributions of people belonging to that group. These are estimated as the change in inequality after marginally increasing the population at each income level (given by the RIF of the corresponding inequality measure).²⁰ It is shown that, in the case of the MLD, this contribution is empirically equivalent to measuring the change in inequality after replacing the incomes of each group with the corresponding global mean using the Shapley approach (averaging across all possible sequencings of groups).²¹ This approach allows a more systematic analysis of the different contributions not only to overall inequality by any index but also to its within- and between-group components.

Thus, we can identify in a consistent and systematic way which countries more strongly contribute to the trend in inequality in each period, with their contributions always adding up to the total they

²⁰ Following the seminal contributions of Firpo et al. (2007, 2009).

²¹ The contribution of an income source to total inequality can be estimated as the change in inequality after either 1) removing the income source ('zero income' decomposition) or 2) equalizing the income source among all individuals in the population ('equalizing income' decomposition) (e.g. Sastre and Trannoy 2002). While removing a country from the sample is equivalent to the former in the context in which groups are seen as inequality sources, the approach followed here is equivalent to the latter.

intend to explain. A country's contribution to inequality generally increases, for example, when the incomes in the country move away from the global mean (above, below, or in both directions). This contribution can be channelled through the between-country or the within-country components. That is, on average, the entire country is moving away from the global mean (the country is getting richer or poorer) or is becoming internally more unequal, for instance.

Furthermore, it is worth noting that inequality changes can be driven by pure demographic trends due to some country populations growing faster than others even if relative per capita incomes remain constant (*ceteris paribus*, a country's contribution to global inequality and to its components will increase with its population size). Alternatively, a country's contribution to inequality can increase due to changes in the country's income distribution (with constant population); that is, the country becomes richer or poorer, or more or less unequal, keeping its population constant. To further disentangle these drivers, in each case using a Blinder–Oaxaca type of decomposition based on the RIF country contributions, I will identify whether these contributions to global, between-country, and within-country inequality are due to a demographic composition effect or a pure income distribution effect, and whether the distributional effect affects either the between- or the within-country income distributions—that is, the country that increases its contribution because the mean income moves to the extremes of the global distribution, or because it becomes internally more unequal. This is equivalent to undergoing a RIF regression decomposition and the details are explained in Appendix B.

7.2 Disentangling global distribution drivers

Figure 17 displays the contributions to the Gini index of a selection of countries, as well as by country region. Table 3 decomposes the change in those contributions between selected years into the distributive effects of inequality between and within countries, and the compositional effect of changes in population (which can affect between- and within-country inequality). This highlights the extent to which the main trends in global inequality are shaped by the economic and demographic trajectories of the most populous countries or regions. The results depend only to some extent on which inequality measure is used. I focus here on the case of the Gini index, but the data appendix gives results for other indices (also including disaggregation by country income group).

It becomes obvious that China's total contribution to inequality between countries dramatically increased from the mid-1950s and reached its maximum of almost 16 Gini points in 1977 (Figure 17a). This is about 36 per cent of total (Shapley) between-country inequality or 23 per cent of overall inequality that year. China's contribution sharply declined thereafter to barely 2 Gini points in 2019 (nearly zero in the case of the MLD and Theil). The deceleration of the impact of China on between-country inequality as it becomes richer is thus evident too, meaning that the main force that has pushed global inequality down in recent decades is about to end.²² At the same time, China's contribution to inequality within countries increased over the same period but to a much lesser extent (from 4.2 to 6.5).

²² Income per capita in China in 2019 is still slightly below the global average. When it goes above the mean, the impact on inequality is ambiguous as China growing faster makes the rest of the world, both rich and poor countries, relatively poorer (i.e. crossing Lorenz curves). Therefore, depending on the sensitivity to each end of the distribution, the impact may still be reducing inequality (for indices that put more emphasis at what happens at the top).

On the other hand, India reached its maximum contribution to between-country inequality in 1970 (8.5 Gini points), before reducing it to its current 5.6, which still gives room for future contributions to reduce global inequality as India catches up with the other countries.

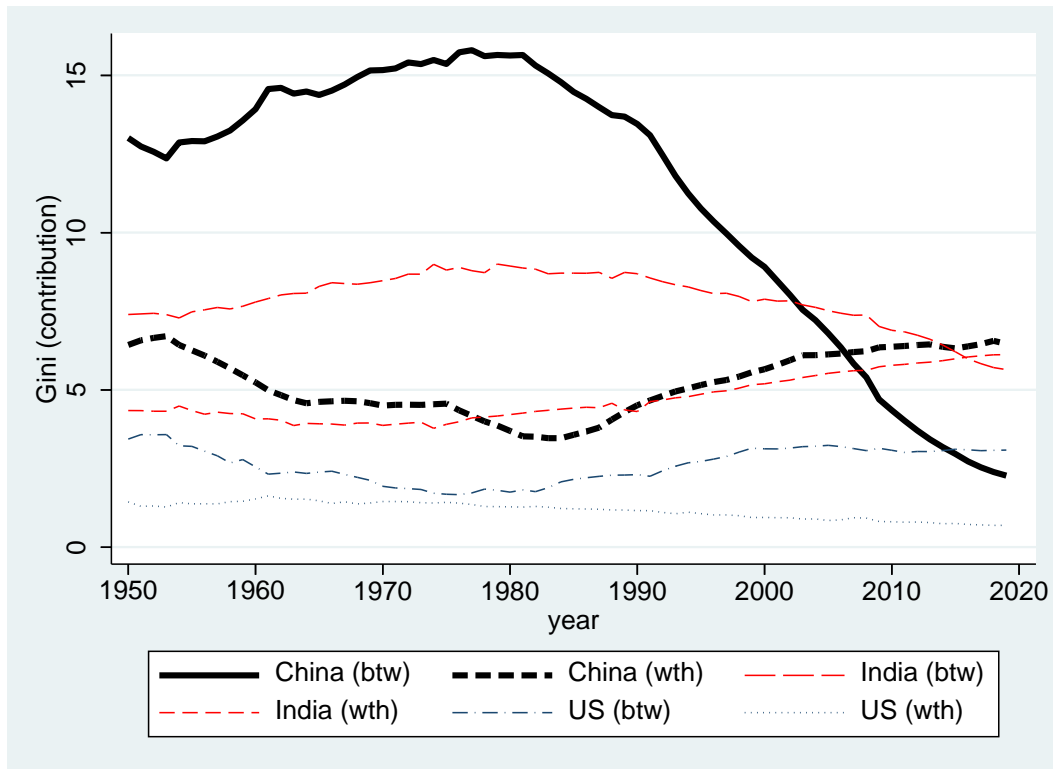
As a result, China and India being left behind initially contributed to increasing global inequality between countries: for example, estimated with constant population and in terms of the Gini index, 2.7 Gini points in the case of China and 1.4 in the case of India between 1950 and 1980. China also contributed to a much larger extent than India to reducing inequality within countries as measured by the Gini index over the same period (2.9 Gini points, versus only 0.4). The faster population growth in developing regions such as South Asia, sub-Saharan Africa, and East Asia, as compared with Europe, also explained another 2.3 Gini points of the increase in global inequality (total composition effect). As a result, the total contribution to the overall global Gini was close to zero in the case of China (as opposed to a higher level of inequality attributed to India's contribution of 1.4 Gini points).

In the most recent period, out of 8 Gini points of the total decline in global inequality between 2000 and 2019, China accounted for 5.8 Gini points (with constant population), driven by China accounting for more than a half of the reduction in inequality between countries (6.3 out of 11.5). Another Gini point was a composition effect due to the slower population growth in China over this period. On the other hand, China's contribution to increasing within-country inequality was 1.7, out of a total increase of 2. In the same period, India contributed to a reduction of 1.3 in global Gini (2.1 in inequality between countries). The contribution of the sub-Saharan African region prevented global inequality falling by 2.8 additional Gini points; this was entirely the result of the region's faster population growth rather than changes in the income distribution.

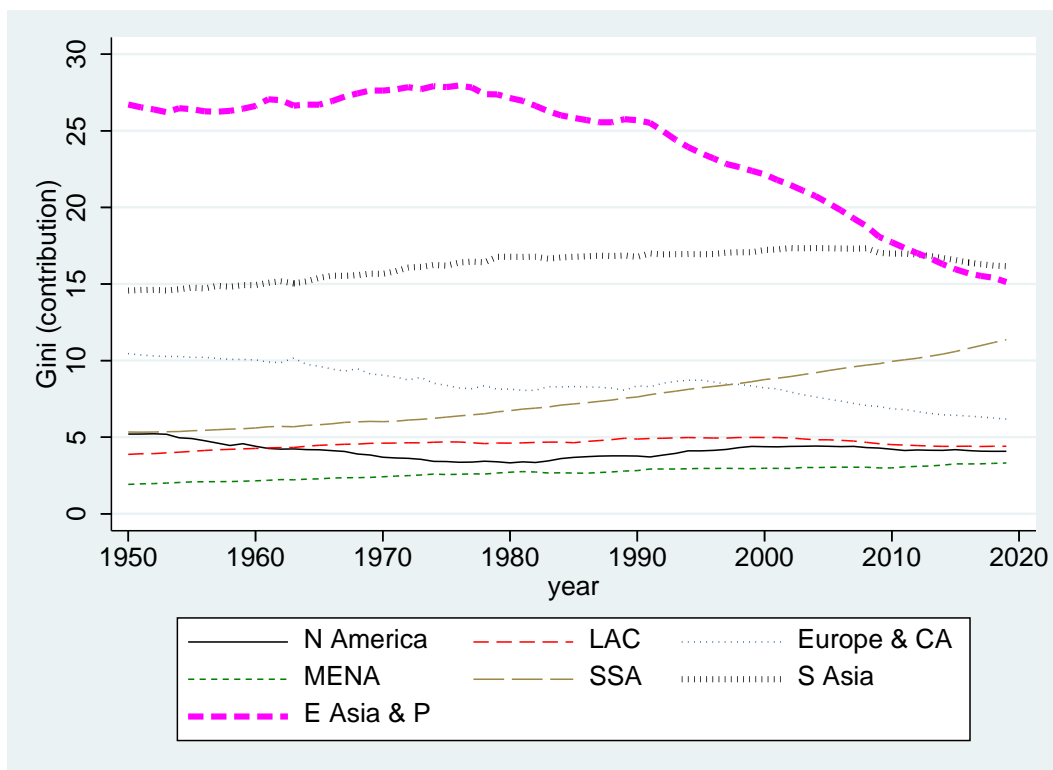
Finally, Figure 18 helps us to understand the impact of the collapse of communist regimes in Eastern Europe on the rise in global inequality between the late 1980s and mid-1990s with indices that are sensitive to the upper end of the distribution such as GE(2). The Figure shows that former socialist countries together contributed to most of the increase in between-country inequality (8 out of a total of 12 points). They also contributed to the increase in within-country inequality but to a lesser extent (5 out of a total of 14), indicating that the concentration of income at the top of the distribution in this period was not just explained by this process and was also driven by what happened in other countries.

Figure 17: Country contributions to inequality, Gini

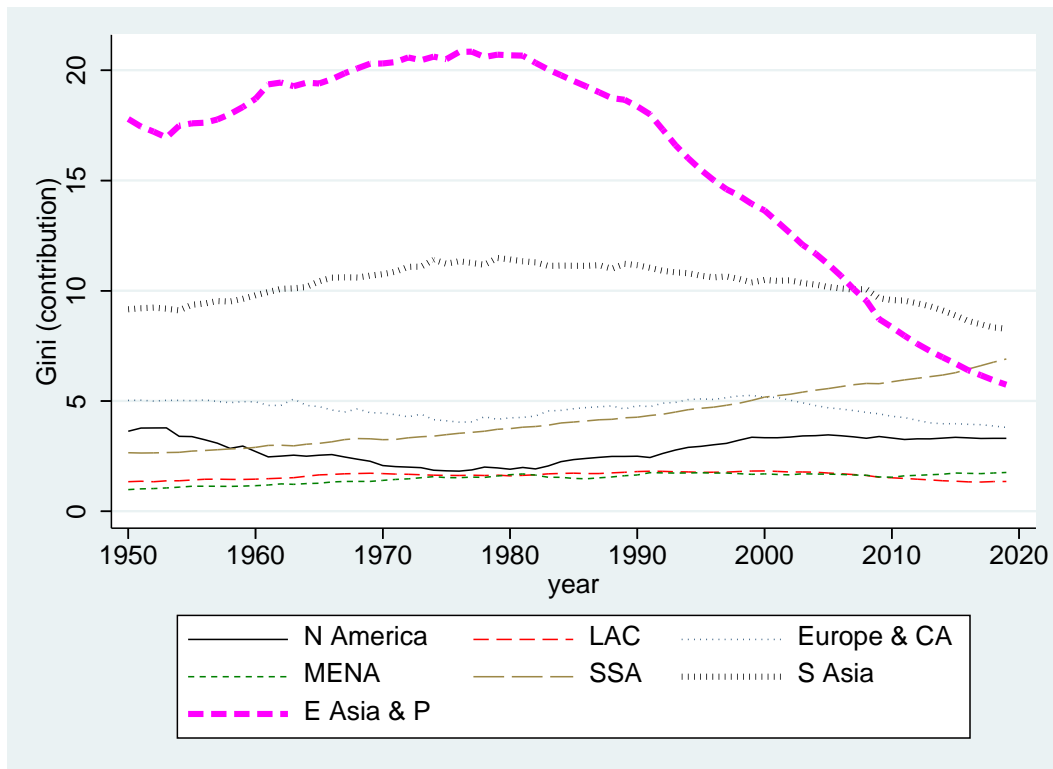
a. Selected countries



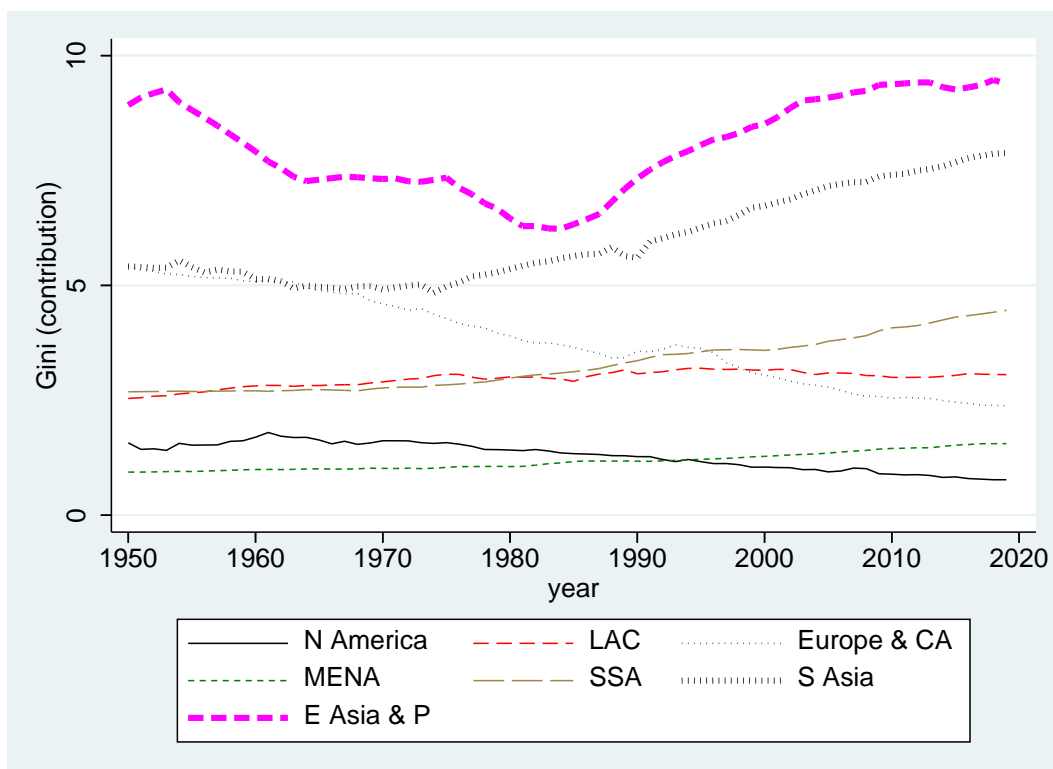
b. By region (overall)



c. By region (between-country)

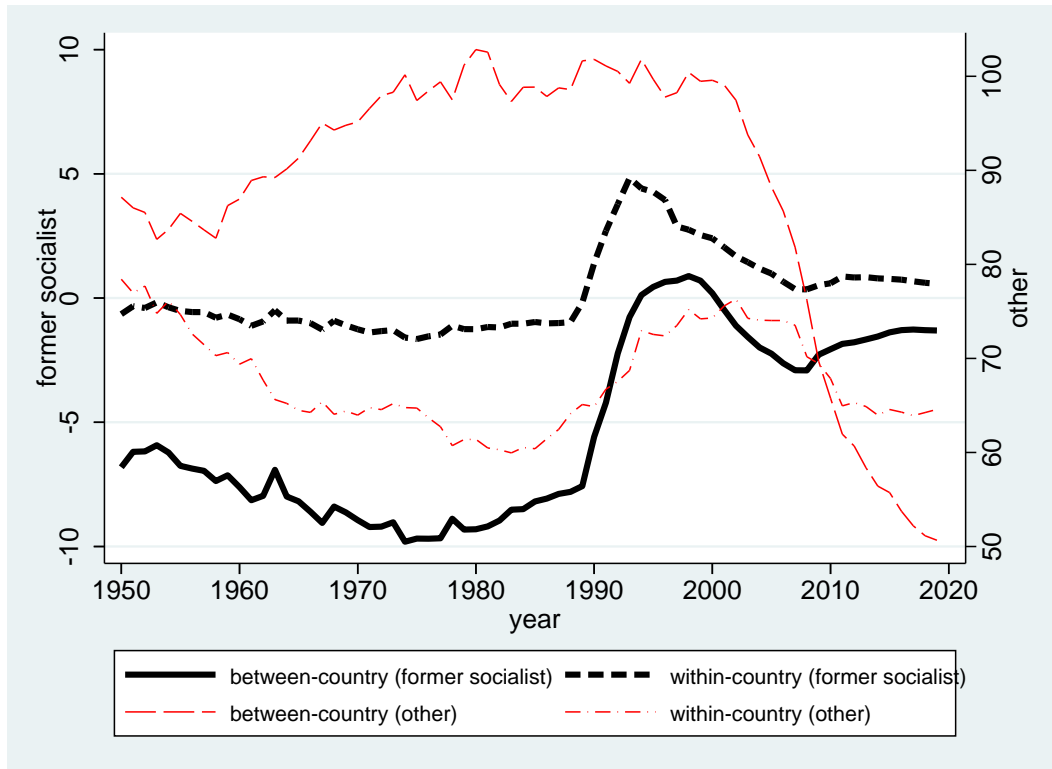


d. By region (within-country)



Source: author's construction based on the WIID.

Figure 18. Country contributions to inequality, former Soviet Union versus other countries, GE(2)



Source: author's construction based on the WIID.

Table 3: Contribution to global income inequality: changes over time by geographical region and selected countries, Gini index

	1950–80				1980–90				1990–2000				2000–19			
	T	B	W	C	T	B	W	C	T	B	W	C	T	B	W	C
<i>World</i>	1.35	2.82	-3.76	2.29	0.47	-1.05	0.99	0.53	-1.20	-3.42	1.70	0.51	-8.04	-11.47	1.97	1.46
<i>North America</i>	-1.87	-2.03	0.30	-0.15	0.44	0.43	0.06	-0.04	0.61	0.69	-0.13	0.05	-0.30	-0.06	-0.15	-0.09
United States	-1.84	-1.94	0.29	-0.19	0.44	0.41	0.07	-0.05	0.60	0.68	-0.13	0.05	-0.29	-0.06	-0.14	-0.09
<i>Latin America and the Caribbean</i>	0.73	0.08	-0.14	0.79	0.26	0.22	-0.02	0.07	0.11	-0.01	0.00	0.12	-0.57	-0.45	-0.06	-0.06
Brazil	0.20	-0.07	-0.04	0.32	0.11	0.06	0.01	0.04	-0.02	0.01	-0.07	0.03	-0.19	-0.15	0.04	-0.07
Mexico	0.18	-0.01	-0.04	0.23	0.06	0.04	0.00	0.03	0.04	0.02	-0.01	0.03	-0.11	-0.11	-0.02	0.02
<i>Europe and Central Asia</i>	-2.30	-0.54	-0.11	-1.66	0.22	0.50	0.14	-0.42	-0.14	0.83	-0.01	-0.96	-2.03	-1.10	-0.12	-0.81
Germany	-0.13	0.20	-0.02	-0.31	-0.07	-0.01	0.02	-0.08	0.02	0.11	-0.04	-0.06	-0.14	-0.03	0.01	-0.12
Russia	-0.29	-0.11	0.05	-0.23	0.10	0.06	0.16	-0.11	-0.16	0.00	0.03	-0.18	-0.41	-0.13	-0.11	-0.17
<i>Middle East and North Africa</i>	0.79	0.21	-0.11	0.69	0.11	-0.19	-0.04	0.33	0.15	-0.09	0.02	0.23	0.34	-0.21	0.09	0.46
<i>Sub-Saharan Africa</i>	1.39	0.72	-0.21	0.88	0.89	0.21	0.01	0.66	1.13	0.33	-0.21	1.01	2.60	0.07	-0.22	2.75
Nigeria	0.14	0.17	-0.10	0.07	0.16	0.05	0.03	0.08	0.16	0.08	-0.06	0.14	0.28	0.06	-0.16	0.39
<i>South Asia</i>	2.20	1.94	-0.51	0.77	-0.01	-0.45	-0.09	0.53	0.45	-1.14	0.64	0.95	-1.05	-2.32	0.64	0.62
Bangladesh	0.42	0.24	-0.03	0.21	0.10	-0.03	0.03	0.10	0.08	-0.05	0.03	0.10	-0.13	-0.20	0.06	0.01
India	1.41	1.44	-0.42	0.40	-0.15	-0.35	-0.13	0.33	0.07	-1.03	0.57	0.54	-1.30	-2.12	0.62	0.19
Pakistan	0.26	0.12	-0.03	0.17	0.11	-0.07	0.01	0.17	0.20	-0.04	0.04	0.20	0.35	0.06	-0.04	0.33
<i>East Asia and the Pacific</i>	0.41	2.43	-2.97	0.95	-1.45	-1.76	0.92	-0.61	-3.52	-4.03	1.39	-0.88	-7.02	-7.40	1.79	-1.42
China	-0.10	2.73	-2.89	0.06	-1.36	-1.59	0.85	-0.61	-3.38	-3.85	1.35	-0.88	-5.80	-6.28	1.66	-1.18
Indonesia	0.41	0.15	-0.05	0.30	-0.14	-0.12	-0.02	0.00	-0.13	-0.15	0.02	0.00	-0.25	-0.51	0.25	0.01
Japan	-0.51	-0.53	0.12	-0.10	0.21	0.23	0.05	-0.06	-0.01	0.08	0.00	-0.10	-0.38	-0.19	-0.02	-0.16

Note: changes in the inequality index between initial and final year. T=Total, B=Distributional effect between countries, W= Distributional effect within countries, C=Composition effect (total).

Source: author's construction based on the WIID.

8 Conclusions

Access to better data has improved our understanding of inequality trends globally between and within countries, especially during the last two decades. However, data on country income distributions based on household surveys are still sparse, and the information is dispersed and heterogenous. In this paper, I presented a new integrated dataset which enables more consistent comparisons of country and global income distributions obtained from the main international and country sources. This will facilitate a more systematic approach to these issues, including quantifying the contribution of its different components or monitoring future changes in the trends, in line with SDG 10. It should also facilitate analysis of different future or counterfactual scenarios, such as the distributive implications of major shocks.

This new database complements the WIID by simplifying the information selecting series that best describe the income distribution trend in each country for the longest possible period with the highest possible consistency. It makes the minimum necessary adjustments to the original survey data to integrate the information in a way that makes it more comparable across countries and over time, while maintaining the main data patterns that are already found in the original data based on household surveys. For that, the distribution will always refer to the same welfare concept, i.e. household net income per capita at the country level. These integrated series for country-level income distributions over time were aggregated to produce the global income distribution, where inequality is measured among the world population regardless of the place where they live. They also enable the study of between- and within-country components separately, disaggregating distributions by region and income group.

At the country level and globally, the new datasets enhance the information that used to be available in the WIID by providing the entire distribution of income at the percentile along with a variety of indicators of the inequality measures. This will facilitate more comprehensive and integral distributive analysis within and across countries or worldwide which can identify the degree of consensus about how to determine the type of distributional changes that take place. This will require admitting that there are different legitimate distributive sensitivities rather than imposing one specific approach and will give users the flexibility to choose their own, with the implicit or explicit value judgements that come with it.²³

Using this dataset, I have analysed the trends in the global income distribution using a comprehensive approach that embraces competing inequality views, including absolute and relative inequality evaluations of income changes, as well as different sensitivities to the performance of different parts of the distribution over time. While some people may pay attention to absolute income distances between people, others will focus on relative distances instead. Similarly, while some people may prioritize the relative or absolute improvement of the poor, others will legitimately be more concerned with the accumulation of income among the most affluent. Rather than imposing specific inequality views, the approach followed here allows us to investigate to what extent we can reach a consensus, regardless of our views on inequality, about what has happened to the global distribution of income. And when that consensus is not possible, it makes it possible to clarify where and how the discrepancy occurs.

²³ See, for instance, a recent discussion about different views on global inequality in Ravallion (2021) or Gradín et al. (2021b).

The preliminary results shown here indicate that it is only when income distances among people are evaluated in absolute terms that one can summarize the last seven decades using a single statement. Inequality unambiguously increased almost continuously between countries and within countries, and therefore globally. It is only deep recessions that seem to have temporarily reduced absolute income distances among people across the world.

Demanding higher income increases among the poor to consider that inequality was reduced in a context of strong global economic growth may seem too demanding or unfeasible for some people. Instead, whenever income distances are evaluated in relative terms, the story becomes more nuanced. The results using the Lorenz criterion unambiguously indicate a decline in relative inequality in the long-term (e.g. 1950–2019). This criterion does not help much in identifying the trend for shorter periods due to the lack of Lorenz dominance because we can observe simultaneously equalizing and disequalizing relative income changes at different parts of the distribution, and the magnitude and composition of these changes differ over time. But the lack of Lorenz dominance does not prevent a high level of agreement among most relative inequality measures. The preliminary results thus point to two well-distinguished phases.

The first decades are characterized by some overall stability, with a slightly upward trend, driven by the fact that the main developing regions, particularly China and India, were left behind in the post-war sustained economic growth that the world experienced, leading to increased inequality between countries. This upward inequality trend was aggravated by population growing faster in the developing world than in Europe. It was, however, largely compensated for, to a lesser extent and depending on the index, by lower inequality within countries, particularly China and India.

In the most recent decades, we observe a sharp decline in global inequality after the previous trends were totally reversed. This period is characterized by a large decline in inequality between countries, driven by stronger economic growth in emerging countries, especially in China and, to a lesser extent, India. The between-country trend has been clearly decelerating in the most recent years as China gets richer and can be expected to be reversed again if this process continues, as China's contribution to inequality is already close to zero with various indices. This decline in global inequality between countries is only partially compensated for by the disequalizing effects of faster population growth in sub-Saharan Africa, which has become the poorest region, and country inequality growing within countries in several areas but particularly in China and India.

The turning point in the global trend varies between the mid-1970s, if we pay more attention to the relatively good performance of the world's bottom 40 per cent, and the late 1990s, if we account for the higher concentration at the top 10 per cent of the income distribution which occurred in the 1980s and 1990s. The latter was due to the collapse of socialist regimes in Eastern Europe as well as increasing concentration at the top in several countries. With less sensitivity to either end of the distribution, the decline in global inequalities would have started in the early 1990s.

It is interesting to note that these discrepancies among different inequality sensitivities arise mainly from how the different indices evaluate the trend in inequality between countries rather than within countries. For the latter, the level of agreement with the direction is higher, even if the magnitude varies across indices.

One important point of discrepancy when assessing the global trends emerges if we pay much closer attention to the very bottom of the income distribution. In that case, inequality sharply declined in the first decades until around 2005, when it started to increase driven by stagnation in the incomes of the poorest 5 per cent of the world's population.

Note that the purpose of this database is to provide the type of distributive information that is usually represented in household surveys. I am well aware of the limitations of using survey-based information, particularly regarding the potential misestimation of the extremes of the distribution (incomes of the very poor and the very rich). There has been no attempt to make such corrections here and this will be explored in the future. I am also aware of the limitations of the approach followed here and the fact that results may depend to some extent on various methodological choices, such as the approach followed to tackle the lack of information in many countries and years or the use of per capita GDP to proxy each country's income, among other possible issues. Overall, I believe that, with its limitations, this is a good representation of the inequality trends, at least as they emerge from household surveys, given the paucity of information, even if the accuracy obviously improves in the most recent decades, particularly after 1980 or 1990. There is no way to overcome the fact that in earlier years, reliable information on within-country inequality is scarce and often of lower quality. Therefore, the global dataset needs to be highly imputed (e.g. extra/interpolated here), whether this is done implicitly or explicitly, and therefore trends may be misestimated.

I am also aware that people's wellbeing is not only determined by individual monetary incomes. An overall assessment of inequality, whether at the country level or globally, must factor in inequalities that affect the freedom of people in developing their capabilities along various other dimensions, especially health or education, which are not always perfectly correlated (and therefore well captured) by monetary income alone. However, disentangling the dynamics of income inequalities allows us to go one step further in understanding more general social inequalities.

References

- Anand, S., and P. Segal (2008). 'What Do We Know about Global Income Inequality?'. *Journal of Economic Literature*, 46(1): 57–94. <https://doi.org/10.1257/jel.46.1.57>
- Atkinson, A.B. (1970). 'On the Measurement of Inequality'. *Journal of Economic Theory*, 2(3): 244–63. [https://doi.org/10.1016/0022-0531\(70\)90039-6](https://doi.org/10.1016/0022-0531(70)90039-6)
- Atkinson, A.B., and A. Brandolini (2001). 'Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study'. *Journal of Economic Literature*, 39(3): 771–99. <https://doi.org/10.1257/jel.39.3.771>
- Bourguignon, F. (2015). *The Globalization of Inequality*. Princeton and Oxford: Princeton University Press.
- Bourguignon, F., and C. Morrisson (2002). 'Inequality Among World Citizens: 1820-1992'. *American Economic Review*, 92(4): 727–44. <https://doi.org/10.1257/00028280260344443>
- Chakravarty, S. (2009). *Inequality, Polarization and Poverty: Advances in Distributional Analysis*. Economic Studies in Inequality, Social Exclusion and Well-Being Series. New York, NY: Springer. <https://doi.org/10.1007/978-0-387-79253-8>
- Chantreuil, F., and A. Trannoy (2013). 'Inequality Decomposition Values: The Trade-off Between Marginality and Efficiency'. *Journal of Economic Inequality*, 11(1): 83–98. <https://doi.org/10.1007/s10888-011-9207-y>
- Checchi, D., A. Cupak, and T. Munzi (2021). 'Empirical Challenges Comparing Inequality Across Countries: The Case of Middle-income Countries from the LIS Database'. In C. Gradín, M. Leibbrandt, and F. Tarp (eds), *Inequality in the Developing World*. Oxford: Oxford University Press. <https://doi-org/10.1093/oso/9780198863960.001.0001>
- Chotikapanich, D., R. Valenzuela, and D.S. Prasada Rao (1997). 'Global and Regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data'. *Empirical Economics*, 22(4): 533–46. <https://doi.org/10.1007/BF01205778>

- Cobham, A., and A. Sumner (2014). 'Is It All About the 'Tails'? The Palma Measure of Income Inequality'. *Significance*, 11(1): 10–13. <https://doi.org/10.1111/j.1740-9713.2014.00718.x>
- Cowell, F. (2011). *Measuring Inequality*. 3rd Edition. London School of Economics Perspectives in Economic Analysis. Oxford: Oxford University Press.
- Davies, J., and A. Shorrocks (2021). 'Comparing Global Inequality of Income and Wealth'. In C. Gradín, M. Leibbrandt, and F. Tarp (eds), *Inequality in the Developing World*. Oxford: Oxford University Press. <https://doi-org/10.1093/oso/9780198863960.001.0001>
- Deininger, K., and L. Squire (1996). 'A New Data Set Measuring Income Inequality'. *World Bank Economic Review*, 10(3): 565–91. <https://doi.org/10.1093/wber/10.3.565>
- Dowrick, S., and M. Akmal (2005). 'Contradictory Trends in Global Income Inequality: A Tale of Two Biases'. *Review of Income and Wealth*, 51(2): 201–29. <https://doi.org/10.1111/j.1475-4991.2005.00152.x>
- Feenstra, R.C., R. Inklaar, and M.P. Timmer (2015). 'The Next Generation of the Penn World Table'. *American Economic Review*, 105(10): 3150–82. <http://dx.doi.org/10.1257/aer.20130954>
- Ferreira, F.H.G., N. Lustig, and D. Teles (2015). 'Appraising Cross-national Income Inequality Databases: An Introduction'. *Journal of Economic Inequality*, 13: 497–526. <https://doi.org/10.1007/s10888-015-9316-0>
- Firpo, S., N.M. Fortin, and T. Lemieux (2007). 'Decomposing Wage Distributions Using Recentered Influence Function Regressions'. Unpublished Manuscript. Vancouver: University of British Columbia.
- Firpo, S., N.M. Fortin, and T. Lemieux (2009). 'Unconditional Quantile Regressions', *Econometrica*, 77(3): 953–73. <https://doi.org/10.3982/ECTA6822>
- Foster, J., and N. Lustig (2019). 'Choosing an Inequality Index'. In *Human Development Report 2019: Beyond Income, Beyond Averages, Beyond Today: Inequalities in Human Development in the 21st Century*. New York, NY: UNDP.
- Gradín, C. (2020). 'Quantifying the Contribution of a Subpopulation to Inequality: An Application to Mozambique'. *Journal of Economic Inequality*, 18: 391–419. <https://doi.org/10.1007/s10888-020-09451-w>
- Gradín, C. (2021a). 'WIID Companion (March 2021): Data Selection'. UNU-WIDER Technical Note 2021/4. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/WTN/2021-4>
- Gradín, C. (2021b). 'WIID Companion (March 2021): Integrated and Standardized Series'. WIDER Technical Note 2021/5. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/WTN/2021-5>
- Gradín, C. (2021c). 'WIID Companion (March 2021): Global Income Distribution'. WIDER Technical Note 2021/6. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/WTN/2021-6>
- Gradín, C., M. Leibbrandt, and F. Tarp (2021a). 'Synthesis and Policy Implications'. In C. Gradín, M. Leibbrandt, and F. Tarp (eds), *Inequality in the Developing World*. Oxford: Oxford University Press. <https://doi-org/10.1093/oso/9780198863960.001.000>
- Gradín, C., M. Leibbrandt, and F. Tarp (2021b). *Inequality in the Developing World*. Oxford: Oxford University Press. <https://doi-org/10.1093/oso/9780198863960.001.000>
- Jenkins, S.P. (2015). 'World Income Inequality Databases an Assessment of WIID and SWIID'. *Journal of Economic Inequality*, 13: 629–71. <https://doi.org/10.1007/s10888-015-9305-3>
- Jordá, V., and M. Niño-Zarazúa (2019). 'Global Inequality: How Large is the Effect of Top Incomes?'. *World Development*, 123: 86–89. <https://doi.org/10.1016/j.worlddev.2019.06.017>
- Lakner, C., and B. Milanovic (2016). 'Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession'. *World Bank Economic Review*, 30(2): 203–32. <https://doi.org/10.1093/wber/lhv039>
- Lorenz, M.O. (1905). 'Methods of Measuring the Concentration of Wealth'. *Publications of the American Statistical Association*, 9(70): 209–19. <https://doi.org/10.2307/2276207>

- Maddison Project Database (2020). Version November 2020. Based on J. Bolt and J.L. van Zanden (2020), ‘Maddison Style Estimates of the Evolution of the World Economy: A New 2020 Update’. Maddison-Project Working Paper WP-15. Groningen: Groningen Growth and Development Centre.
- Moyes, P. (1987). ‘A New Concept of Lorenz Domination’. *Economics Letters*, 23(2): 203–07. [https://doi.org/10.1016/0165-1765\(87\)90040-1](https://doi.org/10.1016/0165-1765(87)90040-1).
- Niño-Zarazúa, M., L. Roope, and F. Tarp (2017). ‘Global Inequality: Relatively Lower, Absolutely Higher’. *Review of Income and Wealth*, 63(4): 661–84. <https://doi.org/10.1111/roiw.12240>
- Ravallion, M. (2004). ‘Competing Concepts of Inequality in the Globalization Debate’. In S. Collins and C. Graham (eds), *Brookings Trade Forum 2004*. Washington, DC: Brookings Institution.
- Ravallion, M. (2018). ‘Inequality and Globalization: A Review Essay’. *Journal of Economic Literature*, 56(2): 620–42. <https://doi.org/10.1257/jel.20171419>
- Ravallion, M. (2021). ‘What Might Explain Today’s Conflicting Narratives on Global Inequality?’. In C. Gradín, M. Leibbrandt, and F. Tarp (eds), *Inequality in the Developing World*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198863960.001.0001>
- Roope, L., M. Niño-Zarazúa, and F. Tarp (2018). ‘How Polarized is the Global Income Distribution?’. *Economics Letters*, 167: 86–89. <https://doi.org/10.1016/j.econlet.2018.03.013>
- Sala-i-Martin, X. (2006). ‘The World Distribution of Income: Falling Poverty and ... Convergence, Period’. *Quarterly Journal of Economics*, 121(2): 351–97. <https://doi.org/10.1162/qjec.2006.121.2.351>
- Sastre, M., and A. Trannoy (2002). ‘Shapley Inequality Decomposition by Factor Components: Some Methodological Issues’. *Journal of Economics*, 77(1): 51–89. <https://doi.org/10.1007/BF03052500>
- Schultz, T.P. (1998). ‘Inequality in the Distribution of Personal Income in the World: How It Is Changing and Why’. *Journal of Population Economics*, 11(3): 307–44. <https://doi.org/10.1007/s001480050072>
- Shorrocks, A.B. (2013). ‘Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value’. *Journal of Economic Inequality*, 11(1): 99–126. <https://doi.org/10.1007/s10888-011-9214-z>
- Shorrocks, A.B., and G. Wan (2009). ‘Ungrouping Income Distributions Synthesis Samples for Inequality and Poverty Analysis’. In K. Basu and R. Kanbur (eds), *Arguments for a Better World: Essays in Honor of Amartya Sen*. Volume I: Ethics, Welfare, and Measurement. <https://doi.org/10.1093/acprof:oso/9780199239115.003.0023>
- UNDESA (2019). ‘World Population Prospects 2019’. Online Edition rev. 1. New York, NY: United Nations, Department of Economic and Social Affairs, Population Division.
- UNU-WIDER (2021a). ‘World Income Inequality Database (WIID)’. Version 31 March 2021. Available at: <https://www.wider.unu.edu/database/world-income-inequality-database-wiid>
- UNU-WIDER (2021b). ‘World Income Inequality Database (WIID): User Guide and Data Sources’. Available at : https://www.wider.unu.edu/sites/default/files/WIID/PDF/WIID-User_Guide_31MARCH2021.pdf.
- World Bank (2016). *Poverty and Shared Prosperity 2016: Taking on Inequality*. Washington, DC: World Bank. <https://doi.org/10.1596/978-1-4648-0958-3>
- World Bank (2021). ‘World Development Indicators’. Available at: <https://databank.worldbank.org/source/world-development-indicators> (accessed 15 January 2021).
- World Inequality Lab (2018). *World Inequality Report*. Paris: World Inequality Lab, Paris School of Economics.

Appendix A. Between- and within-country components of inequality

Let $y = (y^1, \dots, y^K)$ denote the global income distribution made up of K countries, where $y^k = (y_1^k, \dots, y_{n^k}^k)$ indicates the country distribution of country k with population n^k , total population is then $n = \sum_{k=1}^K n^k$. Furthermore, $I(y)$ denotes any global inequality measure computed on incomes y .²⁴ \bar{y} denotes the global mean, while \bar{y}^k is the corresponding mean for country k .

Now, let us consider the distribution in two counterfactual situations.

The first counterfactual distribution is given by $y_b = (y_b^1, \dots, y_b^K)$, where in the distribution of each country k , $y_b^k = (\bar{y}^k, \dots, \bar{y}^k)$, the income of every person has been replaced by the country's mean income \bar{y}^k , while keeping inequality within each country unchanged. That is, this is the 'between-country global income distribution', in which all existing inequality within countries has been removed, i.e. $I(y_b^k) = 0$ for all countries.

A second counterfactual distribution is given by $y_w = (y_w^1, \dots, y_w^K)$, where in the distribution of each country, $y_w^k = y^k \frac{\bar{y}}{\bar{y}^k} = (y_1^k \frac{\bar{y}}{\bar{y}^k}, \dots, y_{n^k}^k \frac{\bar{y}}{\bar{y}^k})$, the income of every person (or percentile) has been rescaled by the same factor $\frac{\bar{y}}{\bar{y}^k}$ to have the global mean income \bar{y} , keeping relative inequality within each country unchanged. In the case of absolute inequality this is done by adding the differential instead, obtaining $y_{aw}^k = y^k + (\bar{y} - \bar{y}^k)$. This is the 'within-country global income distribution', in which all existing inequality between countries has been removed without affecting inequality in each country (all countries now have the same mean and therefore global inequality across countries using those means is zero).

In the case of the Atkinson family, as usual, to construct the two counterfactuals, I employ the concept of the 'equally distributed equivalent income' instead of the mean income (i.e. inequality-adjusted welfare: $e^k = \bar{y}^k(1 - A_\varepsilon(y^k))$), obtaining alternatively y_{Ab} and y_{Aw} .

Measures of inequality computed on y_b , $I(y_b)$, have been widely used as a true measure of between-country inequality. Alternatively, inequality between countries can be obtained as the reduction in inequality after equalizing average incomes across countries: $I(y) - I(y_w)$.

Similarly, $I(y_w)$ can be understood as the true measure of inequality within countries, while the reduction in inequality after equalizing within-country incomes can also be interpreted as a measure of within-country inequality: $I(y) - I(y_b)$.

In this paper, I use $I(y_b)$ and $I(y_w)$. The corresponding alternative measures, $I(y) - I(y_b)$ or $I(y) - I(y_w)$, can be easily inferred by comparing overall inequality and each component.

It is a known fact that the only inequality index in which inequality is the sum of the true between- and within-country inequality as defined above, is the MLD (GE_0): $GE_0(y) = GE_0(y_b) + GE_0(y_w)$. That is, this index is additively decomposable, and the magnitude of each term is the same obtained using both alternatives (path independence). Other indices have other well-known

²⁴ For a discussion of the underlying theory of inequality decompositions, see, for example, the discussion and related literature in Chakravarty (2009).

decomposability properties, but only this one guarantees that both terms are pure, in the sense that the within-country term is not contaminated with between-country inequalities and vice versa.

In the case of other members of the GE family, which verify additive decomposability, what is usually interpreted as the ‘within’ component is $GE_\alpha(y) - GE_\alpha(y_b) = \sum_{k=1}^K \frac{n^k}{n} \left(\frac{\bar{y}^k}{\bar{y}}\right)^\alpha GE_\alpha(y^k)$, which is a weighted sum of country inequality, with weights being a function of country means (except when $\alpha = 0$, i.e. MLD). These terms, therefore, are not true within-country in the sense that they reflect prevailing inequality across countries’ means too.²⁵ In the case of the Gini index, the decomposability is more complex as it also depends on the level of overlapping among country income distributions along the income space.

In the case of the Atkinson family, an index which is multiplicatively decomposable, and using the equally distributed equivalent income instead of the mean as the representative income of each country or globally, we get that:

$$1 - A_\varepsilon(y) = (1 - A_\varepsilon(y_{Ab})) \left(1 - \sum_{k=1}^K \frac{n^k \bar{y}^k}{n \bar{y}} A_\varepsilon(y^k)\right)$$

Note also that for all members of the GE family, the true within-country term (after the mean income has been equalized across countries) is just the population-weighted sum of country inequality: $GE_\alpha(y_w) = \sum_{k=1}^K \frac{n^k}{n} GE_\alpha(y^k)$.

To address this heterogeneity in decomposability properties, I also use an additional estimate based on the Shapley decomposition, in line with Davies and Shorrocks (2021).

The Shapley decomposition (Chantreuil and Trannoy 2013; Shorrocks 2013) is a simple method that allows us to obtain a consistent decomposition for all indices, with both terms adding up to overall inequality, regardless of their decomposability properties. It means, in this context, just computing the average between the two possible estimates for each component:

$$I(y) = S_b(y) + S_w(y);$$

$$\text{with } S_b(y) = \frac{1}{2}(I(y_b) + I(y) - I(y_w));$$

$$\text{and } S_w(y) = \frac{1}{2}(I(y_w) + I(y) - I(y_b)).^{26}$$

It is only in the case of the MLD (GE_0) that $S_b(y) = I(y_b)$ and $S_w(y) = I(y_w)$.

The importance of each component is then estimated as the percentage of total inequality:

$$s_b = 100S_b(y)/I(y) ; s_w = 100S_w(y)/I(y); s_b + s_w = 100.$$

²⁵ It also raises some normative issues as inequality in rich countries has a higher contribution to overall within-country inequality than inequality in poor countries.

²⁶ In the case of the income share of the $q\%$ of the population, $q(y)$, we need to account for the fact that inequality is given by $I(y) = |q(y) - q|$. The expression in the Shapley decomposition then becomes: $S_b(y) = \frac{1}{2}(|q(y_b) - q| + |q(y) - q| - |q(y_w) - q|)$.

Note that $s_b > s_w$ if and only if $I(y_b) > I(y_w)$.

Appendix B. Country contributions to inequality

To identify the individual contribution of a country to global inequality in a consistent way, with the sum of all contributions adding to the total level, I followed the approach in Gradín (2020), where any inequality measure is decomposed as the sum of group contributions:

$$I(y) = \sum_{k=1}^K \frac{n^k}{n} \overline{RIF}^k$$

Where $\overline{RIF}^k = \frac{1}{n^k} \sum_{i=1}^{n^k} RIF(y_i^k, I(y))$ is the mean value of the recentered influence function of global inequality index $I(y)$, estimated across country incomes, $y_i^k, i = 1, \dots, n^k$.

Furthermore, to separate the changes in the contribution of a country that is driven by demographic trends, after adding and subtracting inequality in a counterfactual distribution $\frac{n_0^k}{n_0} \overline{RIF}_1^k$ which keeps the initial population shares constant but uses the final average contribution, we can define the change in inequality between year 0 and 1 as:

$$\begin{aligned} I(y_1) - I(y_0) &= \sum_{k=1}^K \left(\frac{n_1^k}{n_1} \overline{RIF}_1^k - \frac{n_0^k}{n_0} \overline{RIF}_0^k \right) \\ &= \sum_{k=1}^K \left(\frac{n_1^k}{n_1} - \frac{n_0^k}{n_0} \right) \overline{RIF}_1^k - \sum_{k=1}^K \frac{n_0^k}{n_0} \left(\overline{RIF}_1^k - \overline{RIF}_0^k \right) \end{aligned}$$

Where $\frac{n_1^k}{n_1} \overline{RIF}_1^k - \frac{n_0^k}{n_0} \overline{RIF}_0^k$ is the total contribution of country k to the change in inequality, while $\left(\frac{n_1^k}{n_1} - \frac{n_0^k}{n_0} \right) \overline{RIF}_1^k$ is the contribution to the compositional effect (exclusively driven by changes in the country's population) and $\frac{n_0^k}{n_0} \left(\overline{RIF}_1^k - \overline{RIF}_0^k \right)$ is the contribution to the distributional effect (due to changes in the country's incomes).

Finally, combining this with the decomposition of any index into its Shapley between- and within-country components, the same is done separating the effects from each component:

$$\begin{aligned} I(y_1) - I(y_0) &= S_b(y_1) - S_b(y_0) + S_w(y_1) - S_w(y_0) = \\ &= \sum_{k=1}^K \left(\frac{n_1^k}{n_1} - \frac{n_0^k}{n_0} \right) \left(\overline{RIF}_{b1}^k + \overline{RIF}_{w1}^k \right) + \sum_{k=1}^K \frac{n_0^k}{n_0} \left(\overline{RIF}_{b1}^k - \overline{RIF}_{b0}^k \right) \\ &\quad + \sum_{k=1}^K \frac{n_0^k}{n_0} \left(\overline{RIF}_{w1}^k - \overline{RIF}_{w0}^k \right) \end{aligned}$$

Therefore, the change in total inequality is the sum of three terms. These terms indicate the sum of country contributions towards a compositional effect (i.e. changes in the distribution of population across countries over time, keeping income distributions constant within countries), and the corresponding contributions to the distributional effects between countries and within countries respectively (i.e. changes in global inequality between countries and within countries with constant country populations).