Society for the Study of Economic Inequality

Working Paper Series

Inequality Bands: Seventy-five years of measuring income inequality in Latin America

Facundo Alvaredo

François Bourguignon

Francisco Ferreira

Nora Lustig

ECINEQ 2024 672

ECINEQ 2024 672
2024 August
www.ecineq.org

Inequality Bands: Seventy-five years of measuring income inequality in Latin America

Facundo Alvaredo
PSE / LSE / IIEP-UBa
François Bourguignon
Paris School of Economics and University of New York in Abu Dhabi
Francisco Ferreira
London School of Economics and IZA
Nora Lustig
Tulane University

Abstract

Drawing on a comprehensive compilation of quantile shares and inequality measuresfor 34 countries, including over 5,600 estimated Gini coefficients, we review the measurement ofincome inequality in Latin America and the Caribbean over the last seven decades. We find thatthere is quite a bit of uncertainty regarding inequality levels for the same country/yearcombinations. Differences in inequality levels estimated from household surveys alone are presentbut they derive from differences in the construction of the welfare indicator, the unit of analysis,or the treatment of the data. With harmonized household surveys, the discrepancies are quite small.The range, however, expands significantly when –to correct for undercoverage and underreportingespecially at the top of the distribution– inequality estimates come from some combination ofsurveys and administrative tax data. The range increases even further when survey-based incomeaggregates are scaled to achieve consistency not only with tax registries but with NationalAccounts. Since no single method to correct for underreporting at the top is fully convincing atpresent, we are left with (often wide) ranges, or bands, of inequality as our best summaries ofinequality levels. Reassuringly, however, the dynamic patterns are generally robust across thebands. Although the evidence roughly until the 1970s is too fragmentary and difficult to compare,clearer patterns emerge for the last fifty years. The main feature is a broad inverted U curve, withinequality rising in most countries prior to and often during the 1990s, and falling during the early21st century, at least until around

2015, when trends appear to diverge across countries. Thispattern is broadly robust but features considerable variation in timing and magnitude dependingon the country.

Keyword: income inequality, measurement, Latin America and the Caribbean

JEL Cassification: D31, D63, O54

Inequality Bands:
Seventy-five years of measuring income inequality in Latin America[*]

Facundo Alvaredo[†], François Bourguignon[‡], Francisco Ferreira[§], Nora Lustig[¶]

July 2024

Abstract: Drawing on a comprehensive compilation of quantile shares and inequality measures for 34 countries, including over 5,600 estimated Gini coefficients, we review the measurement of income inequality in Latin America and the Caribbean over the last seven decades. We find that there is quite a bit of uncertainty regarding inequality levels for the same country/year combinations. Differences in inequality levels estimated from household surveys alone are present but they derive from differences in the construction of the welfare indicator, the unit of analysis, or the treatment of the data. With harmonized household surveys, the discrepancies are quite small. The range, however, expands significantly when –to correct for undercoverage and underreporting especially at the top of the distribution– inequality estimates come from some combination of surveys and administrative tax data. The range increases even further when survey-based income aggregates are scaled to achieve consistency not only with tax registries but with National Accounts. Since no single method to correct for underreporting at the top is fully convincing at present, we are left with (often wide) ranges, or bands, of inequality as our best summaries of inequality levels. Reassuringly, however, the dynamic patterns are generally robust across the bands. Although the evidence roughly until the 1970s is too fragmentary and difficult to compare, clearer patterns emerge for the last fifty years. The main feature is a broad inverted U curve, with inequality rising in most countries prior to and often during the 1990s, and falling during the early 21st century, at least until around 2015, when trends appear to diverge across countries. This pattern is broadly robust but features considerable variation in timing and magnitude depending on the country.

Keywords: income inequality, measurement, Latin America and the Caribbean

JEL classifications : D31, D63, O54

---

[*] We are grateful to Leonardo Gasparini, Mauricio De Rosa, Ignacio Flores Beale, Álvaro Fuentes, Xavier Mancero, Marc Morgan, Christoph Lakner, Andrea Vigorito, and Nishant Yonzan for providing series and clarifications behind some of the figures; and to Leopoldo Tornarolli for kindly answering several inquiries concerning the recent evolution of inequality in Brazil. Gastón Nievas Offidani and Pietro Geuna provided excellent research assistance. We are also indebted for comments and discussion to Sergio Firpo; panel members of LACIR; participants in the LACIR workshops in Washington, DC (August, 2022) and Cartagena (March, 2023); and at the 28th LACEA Meeting in Bogotá (2023); and two anonymous referees. We acknowledge financial support from INET, EUR Grant ANR-17-EURE-0001 (FA) and LACIR.

[†] PSE-WIL, LSE-III, and IIEP-UBA.
[‡] Paris School of Economics and University of New York in Abu Dhabi.
[§] London School of Economics and IZA.
[¶] Tulane University.

1. Introduction

There are many 'windows' through which a country's income distribution can be observed and its inequality measured.[1] Useful information about people's incomes can be obtained from at least five sources: censuses; national account statistics; household surveys; tax records; and other administrative datasets on social security and wages.[2] Sometimes, researchers prefer to bypass income altogether and examine the distribution of consumption expenditures, which come primarily from household surveys, although not necessarily the same surveys that capture incomes. In Latin America, long considered one of the world's most unequal regions, all these different types of sources have been used to measure inequality. The region has a rich and long tradition in the development and production of inequality and poverty statistics. Studies attempting to estimate inequality and to describe the distribution of income can be traced back for more than seven decades.

This paper reviews the statistical evidence on income inequality in Latin America and the Caribbean (LAC) over the last seventy-five years or so.[3] We do not collect any data or produce new primary series ourselves. Instead, we examine a broad array of statistics and research findings produced in the region over time, as reported by statistical agencies, in academic studies, or in other repositories of primary and secondary sources. In so doing – as in any attempt to cover a large literature on income inequality in a large region – we must contend with a considerable heterogeneity of estimates and a range of comparability issues, which not only affect the extent to which a particular measure for a given country in a given year can be compared across time or space but also, more worryingly, generate disparate estimates for the *same* country and year.

In trying to make sense of this landscape, our primary objective is to summarize what can be said with confidence about the levels and dynamics of income inequality in Latin America, while recognizing the uncertainty that arises from conflicting information from different sources. We reach three main conclusions. First, the picture arising from household surveys alone is, broadly speaking, internally consistent, in the sense that the causes of most of the differences in estimates can be understood and derive from differences in the construction of the welfare indicator or the

[1] In an often-quoted passage, Zvi Griliches wrote: "The available economic statistics are our main window on economic behavior. In spite of the scratches and persistent fogging, we cannot stop peering through it and trying to understand what is happening." (1986, p.1509)

[2] Even this is not an exhaustive list. Income data is also present in registries, private company databases, rich lists, and so on.

[3] Our remit focuses narrowly on incomes, with occasional forays into consumption expenditures and earnings. Inequalities in wealth, land, education, health, opportunities, and political power are the subject of companion papers in the LACIR project.

unit of analysis.[4] There is some residual variation among data sources that present inequality numbers for purportedly the same distribution, but the ranges are relatively narrow and, again, can usually be understood.

Second, the range of estimates is considerably augmented when information is drawn from tax or social security records, and even more so when attempts are made to reconcile income totals from the 'microeconomic' sources with National Accounts aggregates. While some reasons for the disparities can be surmised – relating, for example, to differences in the coverage of richer or poorer households; in the reporting of capital incomes; or in the treatment of earnings retained by firms – it is much harder to zoom in on a single preferred estimate. All of the various sources have strengths and weaknesses; furthermore, attempts to combine them inherently require assumptions and methodological decisions that are – to a greater or lesser extent – arbitrary. This gives rise to a situation of genuine uncertainty, where we become fairly confident that inequality is underestimated when based on household surveys alone, but are less confident that the estimates combining various sources under those multiple assumptions are accurate. This uncertainty, in turn, leads us to deal with ranges, or bands, of inequality estimates, which we argue represent the true (and uncertain) state of knowledge about inequality levels in Latin America.

Third, the uncertainty about levels does not, in the main, affect the broad dynamic pattern of how inequality has evolved over time. Here, there is remarkable consistency, with most countries in the region having followed an inverted U curve between the 1970s and the 2010s. The levels, the magnitude of the changes, and the exact timing of the peaks and troughs vary across both countries and sources, as one would expect. But the overall picture of rising inequality in the first half of the period, and declining inequality from the mid-1990s or early 2000s to the mid-2010s, is consistent and robust to different data constellations – with few exceptions.

The paper proceeds as follows. Section 2 briefly discusses the four main sources of heterogeneity in the measurement of income inequality: different data; different concepts; different treatments of the same data; and different summary statistics. It also briefly describes the (meta) dataset we have assembled. Section 3 zooms in on estimates derived from household surveys and presents a picture of the levels and dynamics of inequality in eighteen countries in the region, as described by most of the series available since the 1940s. The picture for sixteen additional countries, including many in the Caribbean, for which evidence is sparser, is given in the appendix. We discuss the existing dispersion of estimates arising from different analysts working on the same underlying data but with different concepts or assumptions and show that, while levels can vary

[4] Inequality, as used here, is a measure of dispersion in a distribution: a distribution of something, amongst a group of units. We use the term 'welfare indicator' to refer to the variable that is distributed – income, consumption, etc. – and 'unit of analysis' to refer to the units among whom the variable is distributed – individuals, adults, households, etc.

and while the direction of year-on-year changes can differ, the broad dynamic patterns are consistent.

Section 4 introduces information from administrative sources (tax and social security) and national accounts, briefly discusses the recent developments in methods to analyze them, and examines the implications for our understanding of inequality levels and dynamics in the region. We note that incorporating these additional sources gives rise to an additional dispersion in estimated inequality levels, and that the new estimates also suffer from differences in methods, in the treatment of data, in tax codes, and in data quality.

Section 5 discusses the likely economic mechanisms driving the main robust finding, namely the inverted U in inequality dynamics which is broadly common to all countries in the region for which the available statistical evidence allows a suitable time coverage. It also introduces some of the important differences among countries in the region, the broad common pattern notwithstanding. These final remarks are necessarily brief and non-exhaustive, and certainly do not attempt to preempt the detailed discussion of inequality drivers in the 27 chapters that follow this one in the LACIR project. Section 6 concludes.

2. Why is there disagreement on how much inequality there is?

When surveying the literature and data compilations on income inequality in LAC, one is faced with the fact that there are often a few different estimates for a particular country in any given year.[5] Any attempt to understand what is really going on with inequality in the region must confront this fact head on and seek to understand it. Are all but one of these estimates wrong? Or perhaps, are they all wrong? Is there one, unequivocally right way to measure income inequality? And if not, what do we really know about the phenomenon?[6]

There are multiple sources for this variation in inequality statistics which, for our purposes, can be helpfully classified into four broad groups or varieties. The first group (*Variety 1*) is the existence of different data sources that give rise to different distributions for any particular country and time. In developing countries, household surveys and tax records seldom have the same population coverage. Household surveys typically fail to capture top incomes properly, whereas tax registers will almost by definition miss out on workers that do not have to declare incomes to the tax system, or those who work outside the formal sector entirely. Item and unit non-responses, as well as

[5] Although it is not part of our remit, the same remark applies to most other parts of the world, particularly developing ones.

[6] We are certainly not the first to raise these questions. For an earlier discussion of the sensitivity of inequality measures in Latin America to how data is treated, see Székely and Hilgert (2007).

misreporting, can create serious biases for some household surveys, just as tax evasion and avoidance strategies can generate serious measurement errors for tax data. Some statistical offices try to address nonresponse before releasing the surveys (not necessarily in a correct and transparent manner) and others do not.[7] Informality limits the representativeness and accuracy of social security and other administrative wage databases in the region. When incomes from (any of) these sources are aggregated to a national total, that total is almost always at odds with income totals computed through the National Accounts System (SNA) and reconciling them must deal with possible errors in all data sources – including the SNA itself – as well as conceptual differences in how some incomes are treated.

The second set of sources for the variation in inequality statistics (*Variety 2*) is that, even for one particular dataset, statistical offices and researchers may choose to work with different welfare indicators altogether. Common choices include household incomes, individual earnings, or consumption expenditures. Incomes may include estimates of the value of production for own consumption and other forms of income in kind, or just monetary income. Household incomes or consumption may be used as a total, or in per capita terms, or using some alternative equivalence scale that seeks to compensate for differences in needs within the household, or for the existence of economies of scale in consumption.[8] They may also be reported in gross terms (i.e., before taxes and social security contributions), or net of taxes and transfers. There may also be differences in the units of analysis: household income per capita, for example, may be distributed among households or individuals – and this will affect overall inequality.

Something as simple as mixing income-based with consumption-based inequality measures can introduce significant misconceptions, some of which may permeate to the public discourse and have long-lasting effects. Consider, for example, the long-held view that Latin America and the Caribbean is unambiguously the world's most unequal region, with Gini coefficients averaging approximately 0.50, a level higher than in any other continent.[9] This view, like Figure 1 below which depicts average inequality levels for various regions of the world between 1990 and 2020, draws on inequality statistics from multiple countries, some of which are based on income per capita, and some on consumption expenditure per capita. However, when the cross-country

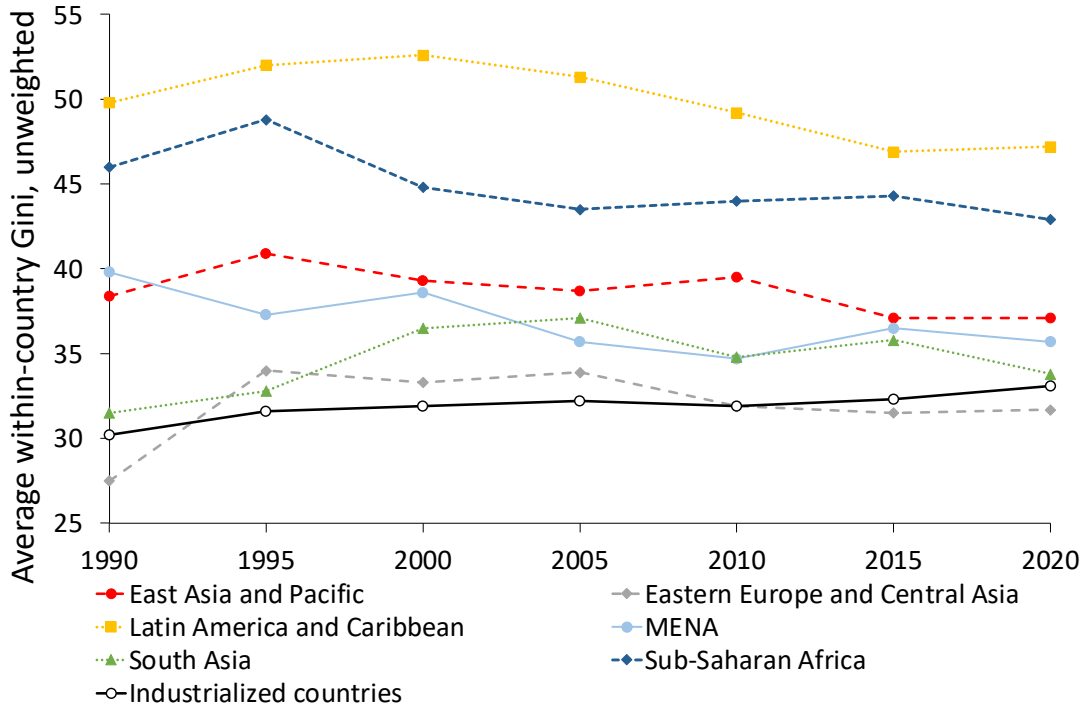
[7] We focus here on differences across data sources for the same country and year. Naturally, comparability challenges are even greater across surveys from different countries. Questions on income from self-employment, for example, are phrased differently and are asked with different reference periods in different Latin American countries. In a region where incomes from self-employment are volatile and informal employment is substantial, questionnaire differences such as these can have real impact on cross-country comparisons. We are grateful to an anonymous referee for this observation.

[8] See Coulter, Cowell and Jenkins (1992) for a classic survey of equivalence scales.

[9] Drawing on a relatively long tradition of nationally representative household surveys, researchers felt generally confident of their assessment of inequality, including global rankings. Deininger and Squire (1996), for instance, stated that their data set confirmed the "familiar fact that inequality in Latin America is considerably higher than in the rest of the world" (p. 566).

comparison is made on a more consistent basis, using consumption as the common variable, Sub-Saharan Africa comes out as the geographic area with the highest mean (and median) Gini coefficient, but also with the highest dispersion (possibly due to measurement errors), as shown in Figure 2.[10]

Figure 1. Levels and dynamics of income/consumption inequality in the world 1990-2020

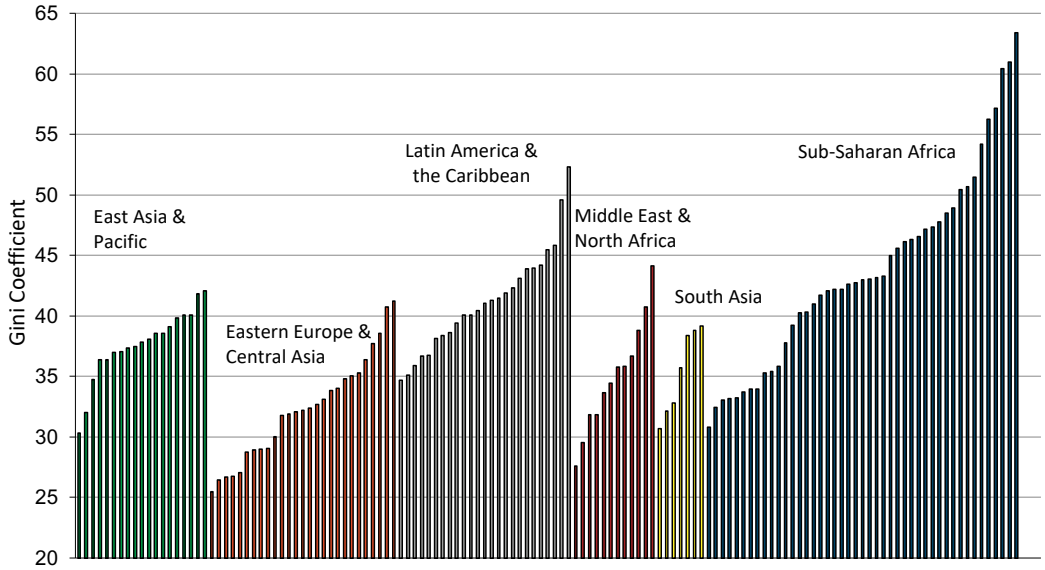


Source: Data for 1990-2015 are from Ferreira, Lakner and Silwal (unpublished), drawing primarily on PovcalNet-PIP and kindly updated by Nishant Yonzan. Data for 2020 are based on Mahler et al. (2022).
Notes: Each point represents an unweighted average of coefficients available in a -2/+2-year window around the reported year, with the closest year used when multiple observations are available. The panel of countries is unbalanced, with sample sizes ranging from 73 in 1990 to 139 in 2010. The series for Latin American and the Caribbean, Eastern Europe, and Industrialized countries are based mostly on Gini coefficients of household per capita income. The series for East Asia and Pacific, Central Asia, MENA, South Asia and Sub-Saharan Asia are based mostly on Gini coefficients of household per capita consumption. Additional details are available from the authors on request.

[10] But of course, as usual, data availability is behind many such choices. The reason the World Bank and others often mix income and consumption statistics is that information on both variables are not available for many countries. In constructing Figure 2 (as well as its subsequent updates), Alvaredo and Gasparini (2015) had to rely on an assumption: they considered a subset of Latin American countries for which household surveys do include reasonably good data for both consumption and income and estimated the average ratio of the consumption/income Ginis (0.861). They then applied this ratio to scale down the income Gini coefficients for all the other countries for which data on consumption was not available. A similar procedure was used for Eastern Europe and Central Asia (coefficient 0.931). The sample underlying Figure 2 is slightly different from that for Figure 1, but they largely overlap and come from the same source.

The third set of sources for the variation in inequality statistics (*Variety 3*) is that, even when the same data source, welfare indicator, and unit of analysis are used, the data may still be treated differently by different analysts, often in perfectly justifiable ways. For example, even when attempting to construct a distribution of per capita household incomes by individuals, two different analysts might correct for missing observations (unit nonresponse) through different kinds of imputations; might trim the distribution at the tails or not; might deflate for differences in the cost of living across space in different ways, might consider that wage incomes have been reported in net terms, but self-employment incomes in gross, and so on.

Figure 2. Inequality in the developing world around 2020: Gini coefficients of the distribution of household consumption per capita by region



Source: Update of Alvaredo and Gasparini (2015), Gasparini et al. (2018) and Bracco et al. (2021) based on PovcalNet/PIP.

Notes: Each vertical bar represents a country in the given geographical region. PovcalNet/PIP includes information for the distribution of per capita consumption expenditures, except in almost all Latin American and a few Caribbean and Eastern European countries, for which income inequality statistics are reported. In the figure, the income Gini coefficients in LAC are adjusted to reflect the gap between income and consumption inequality estimates, as described in the main text.

The fourth and final source of variation (*Variety 4*) is, in some sense, the easiest to detect. This is the well-known fact that, once a distribution of a well-defined welfare indicator by a unit of analysis is obtained, one can summarize its dispersion by means of different inequality indices.

Different inequality measures are sensitive to different parts of the distribution, and may, quite correctly, rank a pair of distributions in opposite ways. The literature presents measures that satisfy a set of desirable properties (axioms), and on establishing the conditions under which all indices within certain classes will yield consistent rankings (e.g. Lorenz dominance).[11]

Important though it is, we do not concern ourselves much with Variety 4 in this paper. We assume the readers are familiar with the issues and we report a very small number of indicators, namely the Gini coefficient and the shares of different segments in the distribution, such as the bottom 50%, the next 40%, the top 10% and the top 1%. These choices are driven primarily by the availability of information: these measures – the Gini coefficient in particular – are by far the most commonly calculated and reported in the literature on LAC (and beyond). In any case, the Gini coefficient is an attractive measure of inequality for several reasons. It satisfies the four key axioms of anonymity, Pigou-Dalton, scale invariance and population replication.[12] It is the only summary index that can be represented graphically using a conventional Lorenz curve. It also has an intuitive interpretation: a Gini coefficient of G percent means that, if we take any two individuals randomly from the distribution, the expected (income) difference between them is 2G percent of the mean. For instance, a rise in the Gini coefficient from 50 to 70 percent implies that the expected difference goes up from 100 to 140 percent of the mean. As any inequality measure, the Gini offers a specific representation of inequality. Hence, it is useful to complement it with judiciously chosen income shares.

In this paper, we are primarily concerned with the variation in inequality estimates in LAC that arises from Varieties 1, 2, and 3: differences in data sources; differences in concepts (welfare indicators and units of analysis); and differences in the treatment of the data. We would like to understand the extent to which these variations affect the confidence with which individual inequality estimates should be treated, and what can be said robustly about inequality levels and trends in the region. To this end, we have assembled a meta dataset of summary statistics on inequality in Latin America and the Caribbean, which builds primarily on the compilation in the World Income Inequality Database (WIID), curated by UNU-WIDER, an aggregator of inequality indices for as many countries and years as possible. In this dataset, UNU-WIDER researchers have assembled information from regional data producers, from independent studies, and from data harmonizers, such as the Economic Commission for Latin America (ECLAC, or CEPAL in the Spanish/Portuguese acronym), the Universidad de La Plata and World Bank Socio-Economic

[11] See, for example, Atkinson (1970) and Cowell (2000). Also, see Duclos and Araar (2006).
[12] It does not satisfy a fifth, namely strict sub-group decomposability. See Cowell (1980) on the axiomatic approach to constructing inequality indices.

Database for Latin America and the Caribbean (SEDLAC), and the Luxembourg Income Study (LIS). These are discussed in greater detail in the next section.

We complement the WIID with various additional observations coming from individual studies by independent scholars, so that we present in this paper the most comprehensive set of series and observations that we could find – although we presume that other studies could still be found and added. Considering the information from Gini coefficients and quantile income shares alone, there are 64,322 observations covering 34 LAC countries over the 1948-2021 period (out of which 5,630 are Gini coefficients). The earliest estimates are Gini coefficients for Guatemala in 1948, Mexico in 1950, and Argentina in 1953. The most recent observations are for 12 countries in 2021 (Argentina, Bolivia, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Panama, Paraguay, Peru and Uruguay). The complete set of estimates of Gini coefficients are presented in Figure 3 and appendix Figure A1. Figure 3 contains the series for the eighteen countries with the highest number of observations, namely Argentina (urban only), Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, and Venezuela. Appendix Figure A1 depicts similar plots, albeit considerably less dense, for an additional sixteen countries, namely the Bahamas, Barbados, Belize, Cuba, Dominica, Grenada, Guyana, Haiti, Jamaica, Puerto Rico (a US protectorate), Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago, and the Turks and Caicos Islands (a British overseas territory).

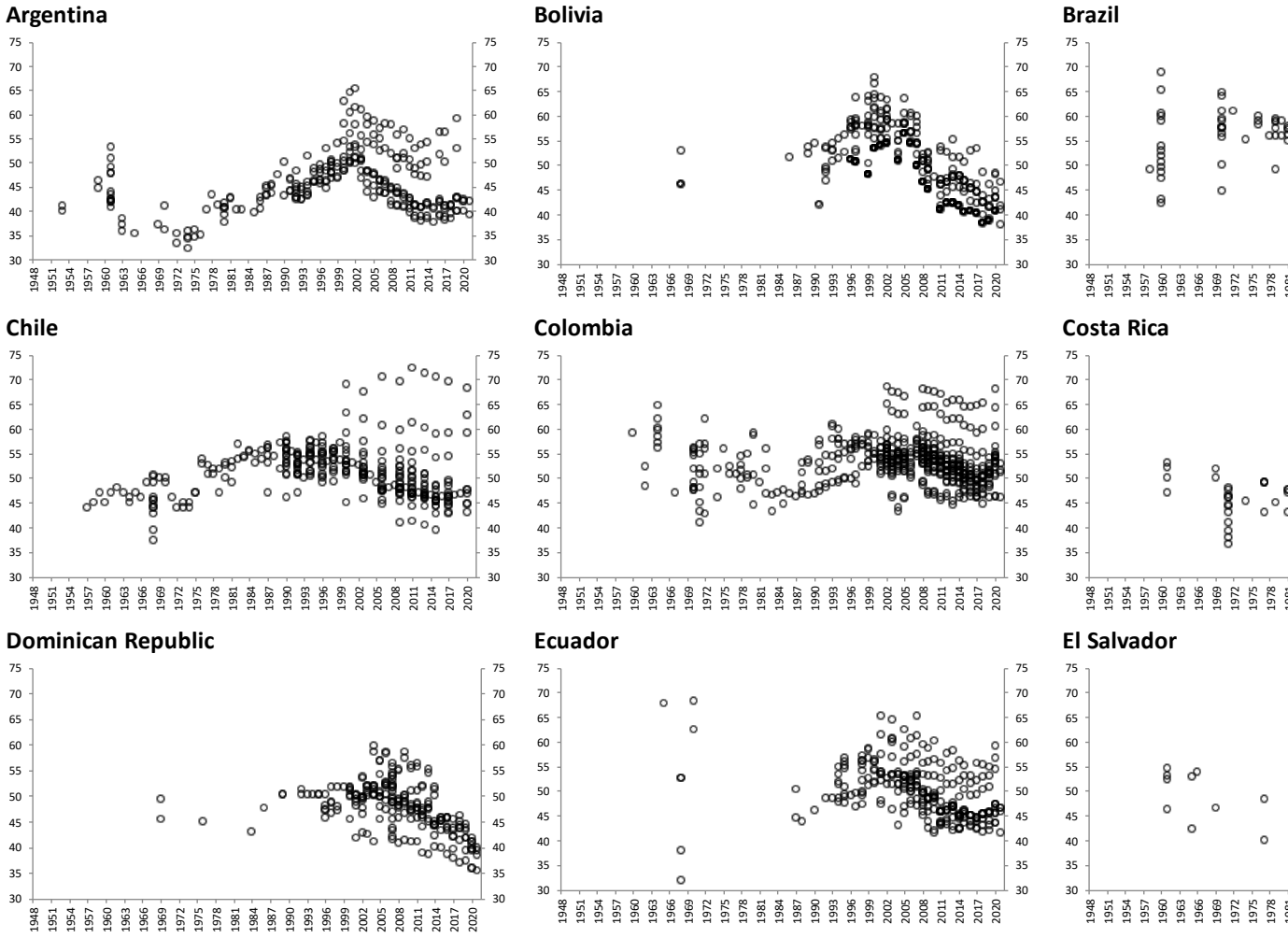
The indices in Figures 3 and A1 comprise the full range of estimates, including those based only on household surveys; those based only on censuses; those based on tax data; and those combining surveys, tax data, and national accounts. The full range of variation described under Varieties 1, 2, and 3 of sources of heterogeneity is present in these figures. The resulting range of inequality estimates is considerable. It is relatively narrow for some countries and periods, such as (urban) Argentina in the 1970s, Costa Rica in the 1990s, or perhaps Panama in the 2000s. For other countries and periods, the range is rather staggering: a reader asking what the Gini coefficient was in Brazil in 1960 can find answers ranging from 0.43 to 0.70. And this is not a feature exclusive to the earlier part of the time series: The range for Chile in 2010-11 is also from 0.40 to 0.70, and for Mexico in 2017 it is even greater.

The levels of dispersion in Figure 3 suggest that the sources of non-comparability discussed above are far from trivial. An understanding of the levels and dynamics of inequality in Latin America cannot ignore or abstract from them. In the next two sections, we try to "unpack" some of the heterogeneity. Section 3 focuses on series that are primarily based on household surveys, and Section 4 considers the more recent estimates that incorporate additional sources.

One clarification is in order. On top of the sources of variation across inequality estimates described above, there is also the margin of error inherent to any point estimate. This error is a

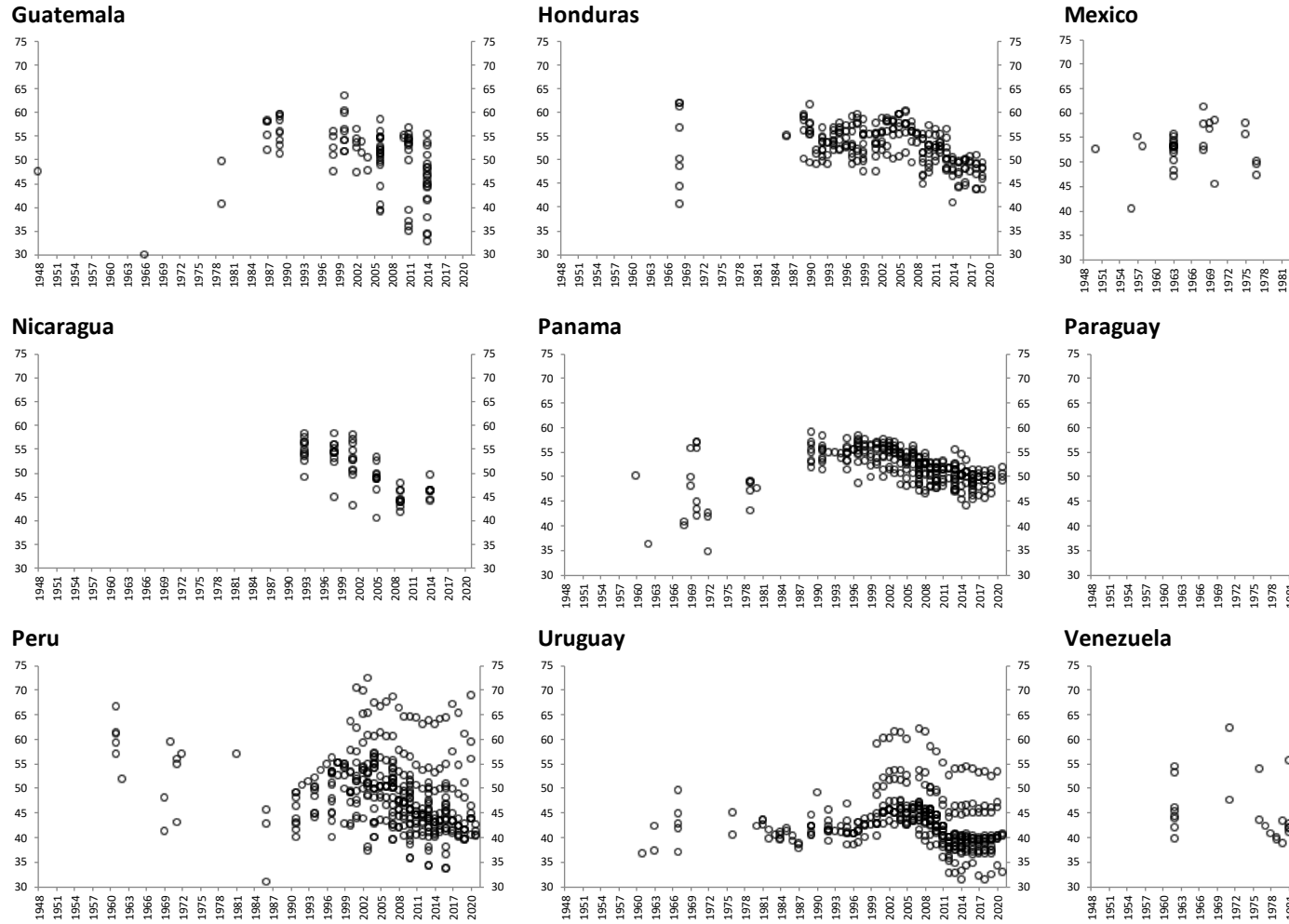
combination of sampling and non-sampling factors, the latter being potentially very important when disparate datasets are combined. These include errors arising from unit non-compliance (in surveys), from tax evasion (in administrative data), from the production of National Accounts, etc., which may all compound, calling for a systematic 'total error' approach in the measurement of inequality, as suggested in Atkinson (2019) and World Bank (2017) for the case of global poverty. The inequality bands discussed in Section 3 and Section 4 abstract from (at least part of) those errors, and focus on point estimates.

Figure 3. Gini coefficients in Latin America and the Caribbean, 1948-2021



Note: The figure shows all the Gini coefficients from WIID, from De Rosa, Flores, and Morgan (2022), and from ad
studies collected by the authors. The plots for the remaining countries in Latin America and the Caribbean can be foun

Figure 3. Gini coefficients in Latin America and the Caribbean, 1948-2021 (continue



Note: The figure shows all the Gini coefficients from WIID, from De Rosa, Flores, and Morgan (2022), and from ad
studies collected by the authors. The plots for the remaining countries in Latin America and the Caribbean can be foun

3. Income inequality measures from household survey data

We try to begin our review of inequality statistics in LAC by focusing on those that draw only on household surveys. However, that is easier said than done, particularly if one wants to go back to the 1950s and 1960s, because many of the earliest studies on income distribution in the region combined information from the first household surveys with other sources, including censuses and the (then incipient) national accounts computations. The development of economic statistics is a lengthy historical process that involves the nature of prevailing paradigms, the construction of a body of conventions, and the limits of available data. The production of such statistics engages governments, central banks, official statistics offices, research institutions, as well as independent scholars, at different stages of the process. And in the early years of the System of National Accounts, national accountants were also experts in distributional issues, as the inter-linkages between the estimation of national income and its distribution were clearly recognized.

Latin America participated fully in this tradition. Most studies of income distribution in the 1940s, 1950s, and 1960s systematically attempted to present their results in the context of the estimation of national income, or of input-output matrices, or as the result of the information provided by population censuses.[13] In Argentina, for example, in parallel with the establishment of the system of national accounts, CONADE-CEPAL (1965) set out to estimate the personal and the functional distribution of income in great detail for the years 1953, 1959 and 1961, making use of surveys, population and industrial censuses, income tax registries, and social security records, and attempted a reconciliation with national income.[14] In Brazil, the seventh population census of 1960 was the first to enquire about incomes. Samples were exploited by Hoffman (1971), Fishlow (1972), and Langoni (1973) to estimate the distribution of income at the national level for the labour force (the População Economicamente Ativa, or PEA). In Mexico, the 1950 census included questions on income; additionally, surveys were introduced rather early in the country; the first one in 1956 (and subsequently in 1958 and 1963). Both sources, together with the national accounts and input-output matrices, were exploited by Navarrete (1960) and later critically reviewed by Altimir (1982) to produce an estimate of income inequality for 1950.[15]

[13] In several Caribbean countries (in particular, those former colonies or territories of the British Empire), it is possible to estimate the shares accruing to top income groups over several decades of the first half of the 20th century based on tax data. While these countries have today sparser and more problematic information on the distribution of incomes than the rest of the region, they had the best existing raw data at that time, following the introduction of the income tax. There is thus scope for expanding the time coverage of the picture presented in Figure 3.

[14] Ironically, later and to this date, Argentina's surveys cover urban areas only. It is the only country in Latin America whose inequality estimates do not include rural areas (about one third of the population).

[15] A comprehensive description of inequality series and studies covering the years 1940-1985 can be found in seven reports published by CEPAL/ECLAC on Argentina, Brazil, Chile, Colombia, Mexico, Peru and Venezuela (1986a,b, 1987a,b, 1988a,b, 1989).

All these early studies contain a wealth of information on inequality in the region in the 1940s, 50s and 60s. Unfortunately, they suffer from a number of the sources of non-comparability discussed in Section 2 – of Varieties 1, 2 and 3 – which hamper their usefulness for cross-country comparison of inequality levels, or for long-run dynamic studies. The unit of analysis, the definitions of income, the set of imputations, the spatial coverage, and the experts' judgements vary over time and across countries.

For better or worse, in the decades that followed, beginning in the late 1960s and 1970s, the link between the production of SNA aggregates and the research on distributional data weakened, and the dominant literature in the two fields (national accounts and distribution studies) went their separate ways, globally and in Latin America. At this point, the applied analysis of the distribution of income began to rely almost exclusively on household surveys, which were first fielded at different times in different countries in the region.

The first attempts were not always led by governments, through their statistical institutes or other official bodies. In some cases, they were isolated and exploratory inquiries. The process to make surveys systematic and have them at regular intervals took more time. Many if not most of these surveys covered urban areas only, or even just the capital city – at least at first. For instance, in Chile, the first representative surveys started in 1956 with the *Encuestas de Ocupación y Desocupación*, a constant-methodology data source for the Gran Santiago area, conducted by the Universidad de Chile until 1979, which is the source of the inequality series produced by Heskia (1980). These were followed in 1963/64 by the *Encuesta de Presupuestos Familiares*, conducted by the Universidad Católica.

In Colombia, household surveys representative at the national level are rather recent; they were developed progressively, increasing geographical coverage in particular between 1996 and 2001. Yet, labor force/household surveys began in 1970 and were carried out by the *Departamento Administrativo Nacional de Estadística* (DANE), with some irregularity and varying coverage. It is only after 1975 that quarterly surveys became available for the urban sector (7 cities) and a consistent annual series could be constructed, although covering only part of the urban sector and mostly labor incomes. Following several surveys conducted in the previous decades, Argentina launched the official *Programa de Encuesta Permanente de Hogares,* EPH, in 1972. Brazil began the development of a systematic survey system in 1967, leading to the *Pesquisa Nacional por Amostra de Domicílios*, PNAD. Both the EPH and the PNAD survive, with changes, to this day.

Viewed from a current perspective, when a great deal of work is once again seeking to combine those different sources of information, this decoupling between surveys, administrative registries, and the SNA may appear regrettable. Nonetheless, that path had at least one positive consequence:

the increasingly systematic collection of surveys since the 1970s, as well as the possibility of exploiting microdata sets, expanded the time coverage and improved cross-country comparability. This has allowed scholars, official statistical offices, and research institutions to produce more comparable inequality series, both over time and across countries.

Thus, we have reasonably consistent – albeit imperfect – inequality series starting in the 1970s, 1980s or the 1990s, depending on the country, and a patchwork of shorter time series or individual observations, many times produced by independent researchers, for the years going back to the 1940s. A few international institutions and research centers, both in Latin America and beyond, have made efforts to compile these series in public-use databases, and we build on their work in our own compilation.

Four of these databases produce their own estimates from available microdata collected through household surveys. These are: (a) the *Socioeconomic Database for Latin America and the Caribbean* (SEDLAC), produced by the Center for Distributive, Labor and Social Studies (CEDLAS) at the University of La Plata; (b) *PovcalNet*, now replaced by the *Poverty and Inequality Platform* (PIP) of the World Bank; (c) the series produced by the UN Economic Commission for Latin America and the Caribbean (*CEPAL* or ECLAC); and (d) the *Luxembourg Income Study* (LIS), which has been progressively incorporating Latin American countries (so far Brazil, Chile, Colombia, Guatemala, Mexico, Panama, Peru, Paraguay and Uruguay) to its database.[16] There are also secondary-source datasets, which compile summary statistics reported elsewhere, rather than generating their own from the microdata. These include, among others, the *All the Ginis Dataset* constructed by Branko Milanovic at CUNY, and the *World Income Inequality Database* (WIID) maintained by UNU-WIDER.[17]

Microdata-based datasets are often preferred because their series tend to be internally consistent. But our purpose here is precisely to take stock of the full range of available estimates, so as to assess the variability across the field. To this end, we greatly benefit from the comprehensive collection of the existing historical series that has been assembled by UNU-WIDER in WIID. We use it as the main building block for the analysis in this section, for two main reasons. First, its coverage goes back further in time than any other database for LAC. Second, it includes the core of the inequality indicators produced by the leading primary, microdata-based datasets mentioned above. As such, the WIID is a useful information aggregator that combines the merits of the longer

[16] There is a close relationship between SEDLAC and PIP, in that the World Bank is a partner of CEDLAS in the project, and draws on SEDLAC to build its own estimates for Latin America. That said, there is additional 'treatment' at the Bank, which leads to there being some differences between the two series, as we will see below.
[17] The OECD also produces inequality statistics, following its own guidelines, for its three LAC members (Costa Rica, Mexico and Chile), and one non-member (Brazil). They are given in OECD.Stat. These observations are also included in Figure 3 above, and in Figure 4 below.

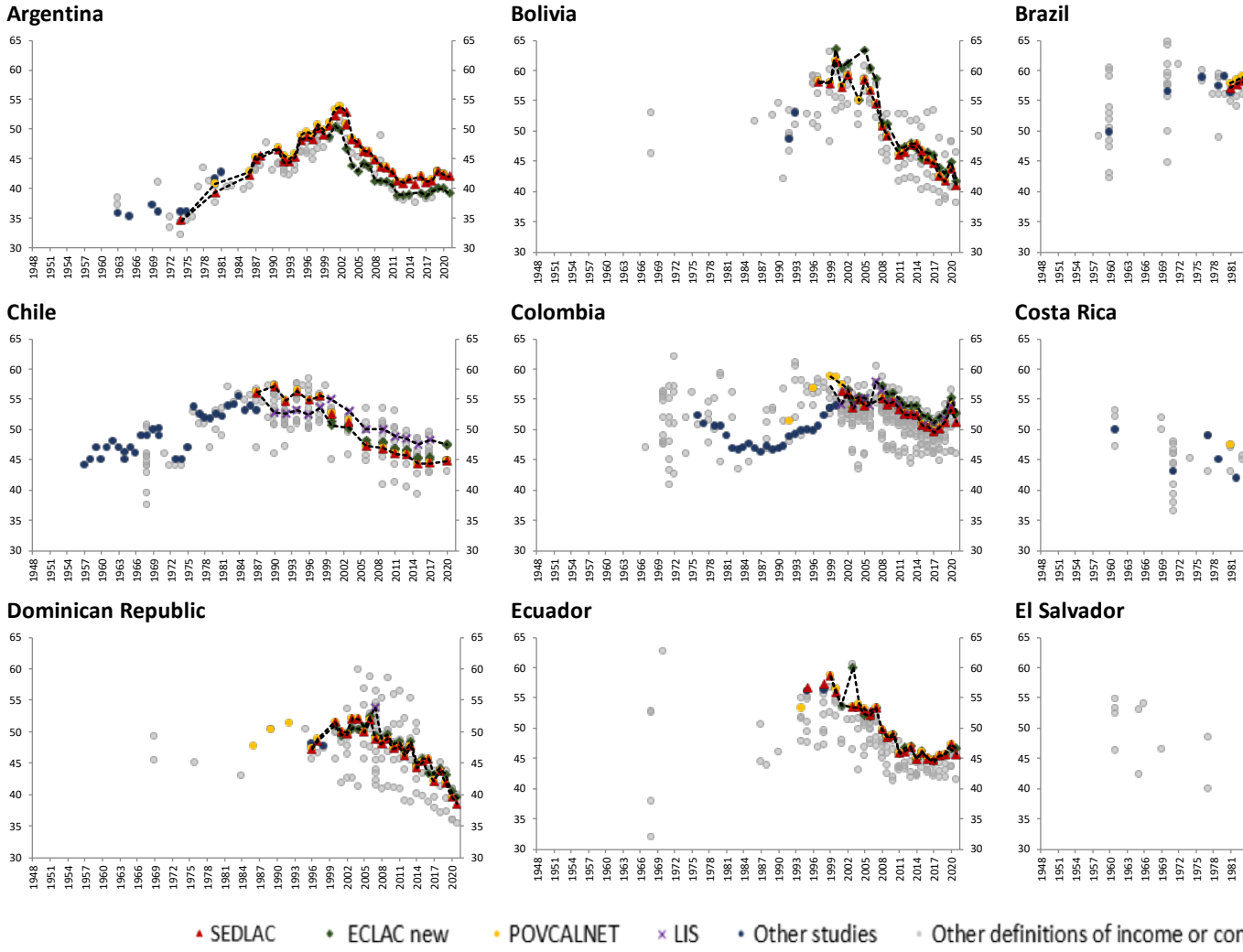
time spans allowed by secondary source datasets, with the efforts towards time and cross-country comparability afforded by the series built on the original microdata sources for the more recent period.[18]

As noted in Section 2, we complement WIID with observations from additional independent studies. Restricting the database from Figure 3 to Gini coefficients arising only from household surveys after 1970 (but also including the aforementioned early studies that use data from censuses prior to that year) gives rise to the sub-dataset presented in Figure 4, which covers the same eighteen countries shown in Figure 3. As before, an analogous figure for the other sixteen countries is shown in the appendix (Figure A2). Given the sparsity of information on this last set of countries, which unfortunately includes all the Caribbean countries in our data, we focus the remainder of our discussion on Figure 4.

Because the observations in Figure 4 are all Gini coefficients and, at least in the latter period, exclusively from household surveys (and censuses, on the grounds that the process of collection of incomes is related to that of surveys), the figure eliminates the fourth variety of comparability problems (different indices), and some of the first variety (different data sources).[19] Most of the methodological variation in Figure 4 therefore comes from Varieties 2 and 3: differences in welfare indicator or unit of analysis, and different treatment of the data. We use a color scheme to highlight some of these differences and to structure the discussion.

[18] Yet there are no perfect data or dataset. For a thorough review of WIID, see Jenkins (2015). The same paper explains why we do not include the Standardized World Income Inequality Database – yet another alternative – in our analysis.
[19] Analogous series for the income shares of the top 10% and 1% are available from the authors on request.

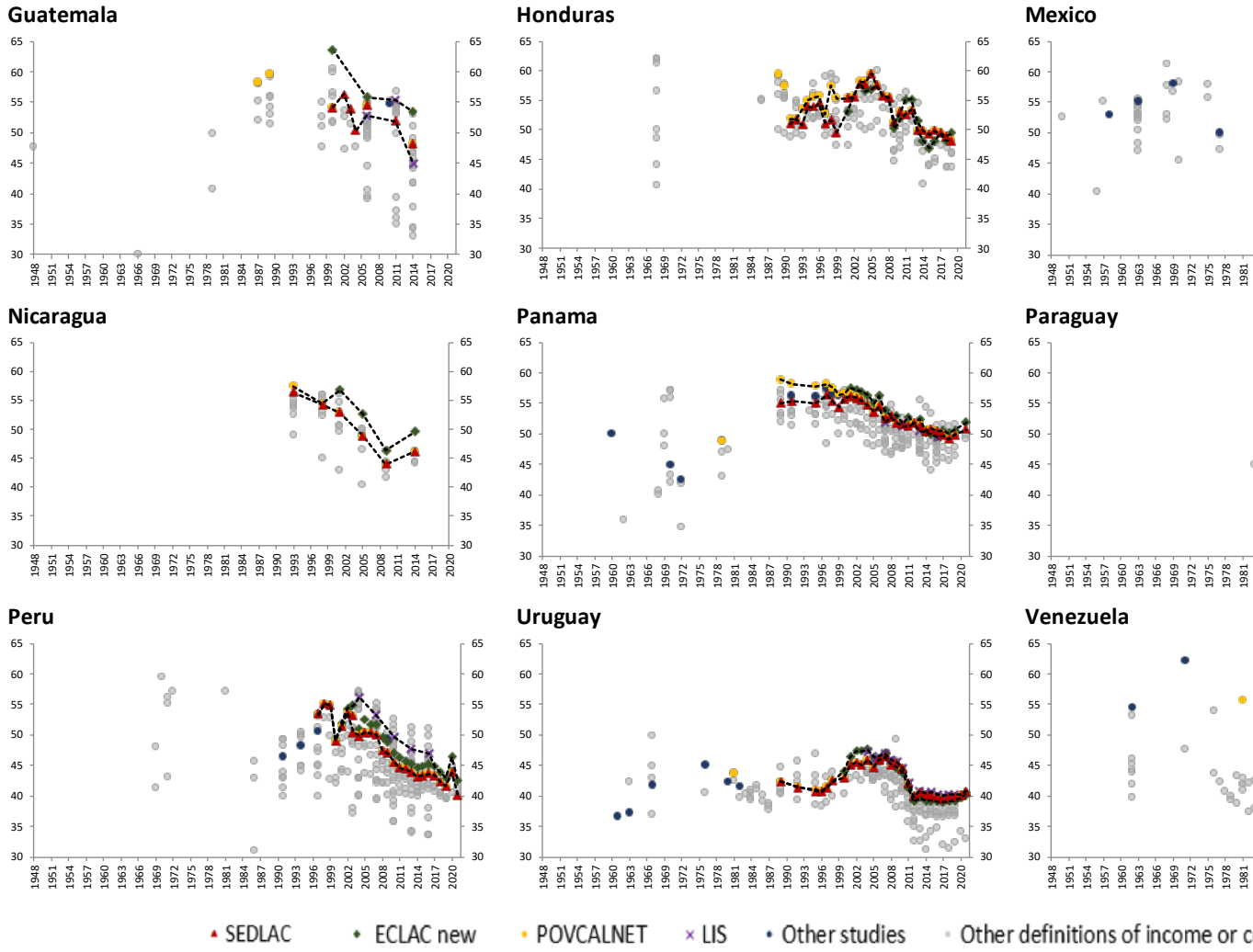
Figure 4. Predominantly HHS-based Gini coefficients in Latin America and the Caribb



Source: WIID complemented by the authors.

Notes: Band 2 is shown. The plots for the remaining countries in Latin America and the Caribbean, which have fe fragmented series, can be found in Appendix Figure A2.

Figure 4. Predominantly HHS-based Gini coefficients in Latin America and the Caribbean (c



Source: WIID complemented by the authors.

Notes: Band 2 is shown. The plots for the remaining countries in Latin America and the Caribbean, which have fe fragmented series, can be found in Appendix Figure A2.

The series highlighted in red (SEDLAC), green (ECLAC new), yellow (PovcalNet-PIP), and purple (LIS) come from the four microdata-based compilations mentioned above.[20] All these series use as welfare indicator household per capita income, but there is considerable uncertainty as to whether that income is pre- or post-taxes. Sometimes, the uncertainty originates from vagueness in the questionnaires or in the enumerator instructions, which means that respondents are not asked clearly and specifically to report either net or gross incomes. This is a surprisingly common and serious issue in the region, which requires attention. The geographical coverage is always the country, except for Colombia before 2001, Argentina and Uruguay, where it is urban areas (Uruguay started nationally representative series in 2009).[21] The series marked in blue are observations from independent studies, but with an income definition equal or close to household per capita income. This is typically the case for studies covering the 1950s, 1960s and 1970s.

All points in gray refer to Gini coefficients calculated using welfare indicators *other than* household per capita income. They refer to series based on equivalized income, total household income, income of the head of household, gross income, market income, consumption, or earnings. They also include national, rural, and urban figures; and sometimes refer to a partial coverage of the national population. They may also use different units of analysis: households, individuals, adults only, etc.

In this paper we do not discuss the merits and the limitations of each of the series, nor do we focus on appraising the quality of the data.[22] Instead, we take the broad span of the (meta) data as our subject of study, and document three main facts that jump out of Figure 4. First, the dispersion in level estimates across the full set of points, on any given year, is considerably narrower than in Figure 3: narrowing the scope to include primarily household survey-based measures removes a large part of the variation. Second, though, some variation remains, even when most of these summary statistics are based on the same raw surveys. This remaining dispersion should be telling of how important definitions are, and of how much the analyst's methodological choices matter when looking at the levels or year-to-year fluctuations in inequality.[23]

[20] In WIID there are two series for ECLAC: an older one which includes a (relatively coarse) adjustment for missing capital incomes based on the gap with National Accounts and which has been discontinued, and a newer one that is solely based on surveys. Both were included in Figure 3 but only the newer one is included in Figure 4.

[21] Gini coefficients at the national (urban and rural) level for Argentina, Colombia, and Uruguay, when available, are plotted in grey in Figure 4. The Gini coefficient for urban areas allows for the longest historical series in these countries. More precisely, in Argentina household surveys referred only to the Greater Buenos Aires between 1974 and 1991; the number of urban areas covered has been increasing since then.

[22] Interested readers are referred to the special issue of the *Journal of Economic Inequality* dedicated to the appraisal of inequality databases (Ferreira, Lustig and Teles, 2015).

[23] At the same time, there have also been attempts to produce long-term, comparable series based on those many existing studies shown in Figure 3 and Figure 4, including within WIID (Gradín, 2021). The levels and trends shown

Third, despite this large variability in *levels* (even across comparable definitions), Figure 4 also suggests that household survey estimates of inequality tell a consistent story in terms of *trends*, or medium-term dynamics. When considering the big picture over this long period (that is, abstracting from annual changes), the wide array of estimates can be read and interpreted in terms of *inequality bands*: intervals within which available estimates lie every year. Indeed, given our color-coded classification of the observations in Figure 4, we define two different bands. Band 1, the widest, is simply the annual range of estimates across the entire dataset (including blue and gray points). Band 2 gives our preferred, narrower interpretation of household survey-based inequality indices, namely the range of estimates across the four most comparable microdata driven series (SEDLAC, new CEPAL, PovcalNet-PIP and LIS). While Band 1 can be seen for the full timespan covered in each country, Band 2 begins later, depending on when the microdata series commences. It can go as far back as the early 1970s for Argentina, or the early 1980s for Brazil and Mexico, or as late at the early 2000s for Guatemala.

For ease of visualization, only Band 2 is plotted in Figure 4, as dotted black lines. Band 1 is the easiest to see on its own: it is simply the range of the data, year on year. Table 1 presents (the decadal averages of) the relevant information for both definitions of "inequality bands': the maximum Gini, the minimum Gini, and the difference between them. Figure 5 shows the average width of the two bands across the entire period, for each of the eighteen countries in Figure 4. It shows that when summary measures of inequality use different concepts for the welfare indicator or unit of analysis (that is, when the gray points are included), the heterogeneity is still considerable, ranging from around three Gini points (in Argentina) to more than ten in Peru, Mexico, and Guatemala.

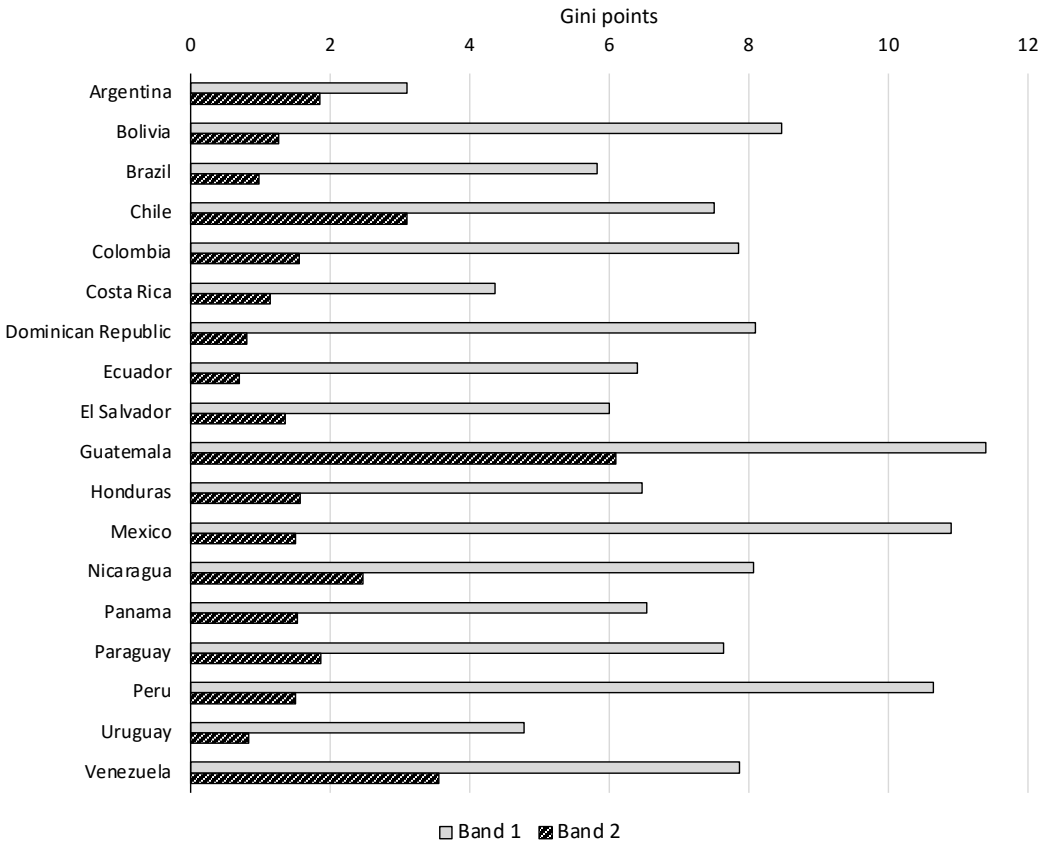
The band width is considerably reduced when looking only at the harmonized series in Band 2. These colored series depict Gini coefficients for the same welfare concepts and from the same datasets, so that only differences in treatment of the data (Variety 3) are left. The average range (over time) for this set of estimates, shown in Figure 5, is typically below two Gini points – a much more acceptable degree of uncertainty to have about inequality levels. Nonetheless, it is three points or above in three countries, namely Chile, Guatemala, and Venezuela.

in those attempts – which lie beyond our remit – should be read with caution, in light of the discussions by Atkinson and Brandolini (2001), Ferreira, Lustig and Teles (2015), and Jenkins (2015).

Table 1. Gini coefficient across comparable survey-based series and inequality bands 1980s-2020s

		Gini Coefficient		Gini Coefficient			Band 1	Band 2
		All survey-based series (Figure 4)		SEDLAC / LIS / ECLAC new / PovcalNet-PIP			Annual range of estimates across all survey-based series	Annual range of estimates across SEDLAC / LIS / ECLAC new / PovcalNet-PIP
	Years	Max	Min	Max	Min	Mean	decennial average Gini points	decennial average Gini points
Argentina	1980s	47.6	37.6	45.5	39.3	43.0	2.1	0.9
	1990s	50.7	42.3	50.7	44.4	47.6	3.2	0.7
	2000s	53.8	41.0	53.8	41.3	47.6	4.2	3.5
	2010s	43.6	37.7	43.6	38.8	41.1	2.9	2.3
Bolivia	1980s	52.5	51.6					
	1990s	63.0	42.0	58.2	57.9	58.0	8.0	0.2
	2000s	66.5	45.0	63.5	49.2	56.8	7.0	2.5
	2010s	53.4	38.2	48.0	41.6	45.6	10.4	1.1
Brazil	1980s	63.3	52.9	63.3	54.9	58.8	3.6	0.7
	1990s	60.6	53.2	60.5	58.1	59.5	3.7	0.5
	2000s	59.5	45.5	58.4	50.9	55.6	9.1	1.3
	2010s	58.2	46.0	54.0	49.0	52.5	6.9	1.4
Chile	1980s	57.3	47.0	56.2	56.1	56.2	3.3	
	1990s	58.5	46.0	57.2	52.5	54.8	7.8	2.8
	2000s	55.0	41.0	55.0	46.9	50.0	9.6	3.3
	2010s	53.2	39.4	48.9	44.3	46.1	9.3	3.2
Colombia	1980s	59.1	43.4				7.1	
	1990s	61.0	46.9	58.7	51.5	55.7	7.4	
	2000s	60.4	43.4	58.7	53.5	55.3	8.5	1.5
	2010s	57.0	44.5	56.0	49.6	52.2	8.4	1.6
Costa Rica	1980s	47.5	34.4	47.5	44.0	45.7		
	1990s	47.7	41.9	47.7	44.0	45.7	3.0	1.0
	2000s	51.8	43.9	51.8	46.7	49.2	4.3	0.9
	2010s	53.3	45.9	50.8	47.9	48.9	5.8	1.5
Dominican Republic	1980s	50.5	43.0	50.5	47.8	49.1		
	1990s	51.4	45.6	51.4	47.2	48.7	1.7	0.3
	2000s	59.8	40.9	54.0	48.1	50.4	12.2	1.2
	2010s	56.5	37.1	48.4	41.9	45.3	10.4	0.9
Ecuador	1980s	50.5	43.9					
	1990s	58.6	46.1	58.6	53.4	56.9	8.0	
	2000s	60.4	42.6	60.1	48.4	52.6	7.3	1.2
	2010s	49.0	41.4	49.0	44.4	45.9	3.9	0.2
El Salvador	1980s							
	1990s	55.9	42.3	54.5	49.9	52.3	6.4	
	2000s	53.2	39.1	51.9	45.2	48.7	6.6	1.0
	2010s	45.4	35.7	45.4	38.0	41.4	5.0	1.7
Guatemala	1980s							
	1990s							
	2000s	63.6	39.1	63.6	50.4	55.0	8.6	6.2
	2010s	56.6	33.0	55.4	45.0	50.4	14.2	6.0
Honduras	1980s							
	1990s	59.5	47.4	57.4	49.5	53.4	6.5	2.3
	2000s	60.1	44.9	59.5	50.2	55.9	7.1	0.9
	2010s	56.3	40.9	55.2	46.9	50.4	5.8	1.5
Mexico	1980s	55.0	38.3	51.6	47.3	49.7	13.1	1.2
	1990s	58.0	42.1	54.4	51.7	53.0	12.5	1.5
	2000s	55.3	39.8	53.4	48.9	50.7	8.0	1.0
	2010s	51.1	37.9	51.0	45.4	48.4	10.0	2.3
Nicaragua	1980s							
	1990s	57.4	45.0	57.4	54.2	55.6	8.9	0.6
	2000s	56.8	40.5	56.8	43.9	49.7	10.1	3.4
	2010s	49.5	44.2	49.5	46.2	47.8	5.2	3.4
Panama	1980s							
	1990s	58.2	48.5	58.2	54.4	56.5	6.8	2.4
	2000s	57.5	46.6	57.5	51.8	54.6	6.7	1.3
	2010s	55.4	44.1	52.8	49.2	50.7	6.1	0.9
Paraguay	1980s	45.1	45.1					
	1990s	59.4	38.9	58.2	40.8	54.3	5.9	0.8
	2000s	65.5	43.4	60.5	49.1	53.7	8.8	2.3
	2010s	60.0	42.6	54.5	45.6	49.6	8.2	2.5
Peru	1980s	57.0	31.0				14.6	
	1990s	55.1	39.9	55.1	53.3	54.4	8.7	0.1
	2000s	57.1	37.1	56.1	47.0	51.0	11.1	2.3
	2010s	52.5	33.6	49.6	41.5	44.4	8.2	2.1
Uruguay	1980s	44.6	37.8				1.6	
	1990s	46.8	38.4	42.5	40.7	41.4	5.0	0.1
	2000s	49.2	38.2	47.6	42.9	45.7	5.2	1.3
	2010s	44.6	31.3	44.6	39.0	40.4	7.3	1.1
Venezuela	1980s	55.6	37.3				7.1	
	1990s	65.8	37.5	48.1	41.3	46.1	11.7	0.9
	2000s	49.5	37.9	49.5	37.9	43.8	4.8	6.2
	2010s							

Figure 5. Inequality Band widths using HHS estimates
Average 1980-2020



Across the countries with the longest continuous inequality series shown in Figures 4 and 5 (Argentina, Brazil, Chile, Colombia, Mexico, and Uruguay), the dynamics of inequality can be characterized by three 'episodes', the centerpiece of which is an inverted U. This is easily seen from the bands in Figure 4, but Table 2 below also lists the maxima and minima (or peaks and troughs) for each of the four harmonized series contained within Band 2. The first episode begins in the 1970s or early 1980s and is marked by a notable increase in income inequality which peaked in the late 1990s or early 2000s, depending on the country. This rise in inequality was sustained, despite short-run fluctuations. This episode includes the effects of the debt crisis and the subsequent "lost decade" in the 1980s, the high-inflation and hyperinflation years in Argentina, Brazil, Peru, and Uruguay over 1989-1991, as well as the period characterized by sweeping market-oriented reforms (i.e.: trade liberalization, deregulation, and privatization) in the 1990s.

Table 2. Gini coefficient across comparable series. Maxima and minima, peaks and troughs 1980-2021

		Peaks / Max				Troughs / Min				
		ECLAC	LIS	POVCALNET	SEDLAC	ECLAC	LIS	POVCALNET	SEDLAC	Notes
Argentina	Gini coefficient	50.5		53.8	53.3	38.8		41.1	40.7	(a)
	year	*2001*		*2002*	*2002*	*2017*		*2017*	*2015*	
Bolivia	Gini coefficient	63.5		61.6	61.5	41.8		41.7	40.9	(b)
	year	*2000*		*2000*	*2000*	*2021*		*2019*	*2021*	
Brazil	Gini coefficient			63.3	62.6	51.1	49.0	51.9	51.2	(c)
	year			*1989*	*1989*	*2015*	*2013*	*2015*	*2015*	
Chile	Gini coefficient			57.3	57.1	45.3	47.6	44.4	44.3	
	year			*1990*	*1990*	*2015*	*2015*	*2015*	*2015*	
Colombia	Gini coefficient	57.2	57.9	58.7	56.2	51.1	49.7	49.7	49.6	(d)
	year	*2008*	*2007*	*1999*	*2001*	*2017*	*2017*	*2017*	*2017*	
Costa Rica	Gini coefficient	51.2		51.8	51.7			45.3	44.0	
	year	*2009*		*2002*	*2002*			*1990*	*1990*	
Dominican Republic	Gini coefficient	52.5	54.0	52.1	52.0	39.5		39.6	38.5	
	year	*2006*	*2007*	*2003*	*2003*	*2021*		*2020*	*2021*	
Ecuador	Gini coefficient	60.1		58.6	58.5	44.4		44.7	44.6	
	year	*2003*		*1999*	*1999*	*2017*		*2017*	*2017*	
El Salvador	Gini coefficient	51.4		54.5	51.7	39.9		38.0	38.0	
	year	*2001*		*1998*	*2002*	*2017*		*2017*	*2017*	
Guatemala	Gini coefficient				56.3					
	year				*2002*					
Honduras	Gini coefficient	56.8		59.5	59.5	46.9		48.2	48.1	
	year	*2005*		*2005*	*2005*	*2015*		*2019*	*2019*	
Mexico	Gini coefficient		54.4	53.6	53.4	45.2		45.4	45.4	
	year		*1994*	*1996*	*1996*	*2020*		*2020*	*2020*	
Panama	Gini coefficient			58.9	56.5	49.5	49.5	49.2	49.2	
	year			*1989*	*1997*	*2016*	*2016*	*2018*	*2018*	
Paraguay	Gini coefficient	58.4	60.5	58.2	57.9	44.7			52.9	(e)
	year	*2002*	*2002*	*1995*	*1995*	*2021*			*2021*	
Peru	Gini coefficient			55.1	55.1	42.6		41.6	40.2	(f)
	year			*1998*	*1998*	*2021*		*2019*	*2021*	
Uruguay	Gini coefficient	47.6	47.2	46.4	46.4	39.2	39.9	39.5	39.5	
	year	*2004*	*2004*	*2007*	*2007*	*2012 & 2017*	*2012*	*2017*	*2017*	
Venezuela	Gini coefficient	43.3		49.5	47.9	36.4		42.1	41.3	
	year	*2005*		*2005*	*1997*	*2010*		*1992*	*1992*	

(a) There is a local maximum in 1991 according to POVCALNET (46.8) and SEDLAC (46.5).
The troughs identified for 2015-2017 are above the first Gini (34.4) of the SEDLAC series for 1974.
(b) 2021 is the last observation of the series, so the minimum Gini is not necessarily a trough.
(c) There is a local maximum in 1997 according to POVCALNET (59.8) and SEDLAC (58.9). There is a local trough in 2020 according to POVCALNET (48.9) and SEDLAC (48.8).
(d) There is a local maximum in 2020 according to ECLAC (55.2), POVCALNET (54.2), LIS (54.1) and SEDLAC (53.3).
(e) ECLAC and LIS series start later. All four sources point to a local maximum in 2002.
(f) 2021 is the last observation of the ECLAC and SEDLAC series, so the minimum Gini is not necessarily a trough.

The second episode took place from roughly the turn of the 21st century until the second half of the decade of 2010 in Argentina, Brazil, Chile, Colombia, Ecuador and Mexico, and until 2021 (the end of our sample) in Bolivia, Dominican Republic, Paraguay, and Peru. It was marked by a significant and widespread decline in inequality: over a shorter or longer period, every country in the region – with the notable exception of Costa Rica – experienced a reduction, even if of varying

magnitude, duration and speed.[24] The remarkable performance during the second episode is further illustrated in Figure 6 for fourteen countries, as the availability of information increased markedly for this episode. The Gini coefficient dropped 20 percentage points (pp) in Bolivia, and 15 in Peru and Paraguay; 12 pp in (urban) Argentina; 10pp in Chile; 8 pp in Brazil and Mexico; and 5pp in Uruguay (urban).

Finally, a third episode begins in the years between 2015 and 2019, in which the picture is mixed. Inequality has been increasing in some countries – relatively mildly in Argentina, Brazil, and Chile, but more markedly in Colombia since 2017. In others, inequality has remained roughly constant (Costa Rica and Uruguay) and, in some, as mentioned, inequality continued to decline but at a slower rate than previously (Bolivia, Paraguay, and Peru). The last three countries have their series minima fall in 2021 – the last year in the available series – according to one or two of the data sources that constitute Band 2 (see Table 2).

The inequality estimates for 2020 already include the first effects of the COVID pandemic, as well as of the measures taken to offset its economic impact. Inequality rose in 2020 in many countries (Bolivia, Chile, Colombia, Ecuador, Peru) but not in all, a notable case being Brazil, which saw a decline in the Gini coefficient of 4.5 points, owing largely to the strength of the mitigating social protection response.[25] Other cases of more moderate decline or stability are Argentina, Chile, Dominican Republic, Mexico, Paraguay, Uruguay.[26] In 2021 Brazil completely reversed the 2020 drop, while Bolivia reversed the 2020 increase.[27]

Naturally, despite the remarkable commonality of the broad dynamic patterns, at least during the first two episodes, there are clear differences in timing across countries.[28] For the first episode of increasing inequality, the Gini coefficient peaked earlier in Brazil (in 1989) and Chile (1990), and later in Argentina and Uruguay (2002, 2007, at least according to SEDLAC and PovcalNet); see Table 2.[29] Moreover, if the end of the second episode, that of declining inequality, is to be marked by a trough in the series and by the reversal of the previous trend (excluding the policy responses

[24] Many authors have tried to understand the second episode of inequality decline. See, for instance, Cornia (2014), Gasparini and Cruces (2022), Gasparini, Cruces and Tornarolli (2011), López-Calva and Lustig (2010), Rodriguez et al. (2022), and Székely and Mendoza (2016). Ferreira et al. (2008) provide an early description of the inverted U for the case of Brazil, and Alvaredo, Cruces and Gasparini (2018) for Argentina.
[25] For a discussion of what happened to inequality during the pandemic see, for example, Acevedo et al. (2024).
[26] See Lustig et al. (2023).
[27] While we believe there is some signal in the statistics for 2020, one cannot but fear that the difficulties with fielding the surveys during the pandemic – including, in many cases, moving some of the data collection to phone interviews – must have introduced additional noise for that year.
[28] It is possible that the difficulty in characterizing the third episode may reflect the fact that it is just too early to identify clear patterns. Alternatively, there may be a real divergence in trajectory among different countries in the region. It is simply too early to tell.
[29] The 1989 inequality peak in Brazil is somewhat suspect since it is likely to reflect mechanical effects of that year's hyperinflation on the measurement of (rapidly changing) nominal incomes; see Ferreira, Leite and Litchfield (2008). This is why Brazil's second period in Figure 6 is taken from the local maximum in 1997.

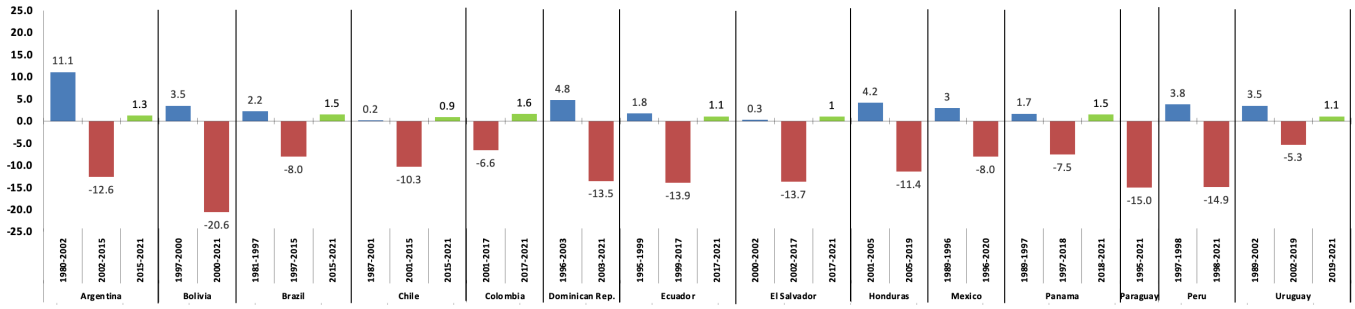
to COVID), this happened in 2015 in Brazil and Chile, and in 2017 in Colombia and Argentina. At the same time, even if the Gini coefficient has not moved much in Uruguay since 2012, this has occurred at the lowest levels of the region.

Overall, there are three main take-aways from our examination of existing household survey-based statistics summarizing inequality within Latin American countries since the late 1970s. First, given the many varieties of sources of non-comparability, there is an unsurprisingly wide range of estimates, even for the same country in the same year. Second, however, this is mostly due to the use of different welfare indicators and units of analysis, and different treatments of the data. Once one restricts attention to the four main 'harmonized' series drawing on primary household survey data, the inequality band is reassuringly narrow, typically less than two Gini points on average. If one were content to make cross-country comparisons on the basis of these data alone, one might reasonably consider three groups of Latin American countries: at one end, Brazil, Colombia, Guatemala, Honduras, and Panama are extremely high-inequality nations, with average Gini coefficients in the 2010s above 0.50. At the other end, urban Argentina, El Salvador, Uruguay and Venezuela have Gini coefficients slightly above 0.40 – still very high inequality in global terms, but less extreme. The other countries in Table 1 lie somewhere in between.[30]

Finally, regardless of which of the two 'bands' one chooses to focus on, there is quite a lot of consistency in inequality trends in the region, marked by three detectable episodes, the first two of which describe inverted U curves, albeit of different magnitudes and with peaks and troughs on different years. In the next section we ask whether and how introducing additional data sources – particularly tax and national accounts data – changes these conclusions.

[30] Two of those four harmonized data compilations (LIS and PovcalNet-PIP) have data for many countries in most other regions of the world. Once adjusting for the differences between income and consumption inequality discussed above, international comparisons using these data suggest that Latin America and sub-Saharan Africa are the world's two most unequal regions (but it is difficult to rank one against the other). In any case, such a comparison relies only on household survey data, and it is possible, of course, that the adjustments discussed in Section 4 below would affect other regions quite differently. Given the uncertainties we have been discussing, reaching reliable conclusions about inequality levels across global regions is difficult at present and might require, as a first step, that the exercise we undertake here be replicated elsewhere.

Figure 6. The rise and fall of income inequality in fourteen Latin American countries
Changes in Gini coefficient (percentage points)



Source: Authors' calculations based on SEDLAC.
Notes: This figure relies on the series published by SEDLAC. The time coverage pre-1990s and pre-2000s varies across countries; consequently, the starting point is different in each case, making the increase in inequality during the first episode quantitatively not comparable (as an extreme case, there is no information for Paraguay before 1995).

4. Introducing information from administrative data and national accounts aggregates

There is reason to believe that household surveys do not capture the rich well. That household surveys in LAC may suffer from mismeasurement of incomes (and consumption) has been recognized long ago by, for example, Altimir (1987). While measurement errors may affect the entire distribution, the fact that household surveys may measure the relevance of top incomes imperfectly is of particular importance for documenting inequality levels and trends. Moreover, the problem that household surveys do not capture incomes at the top well is not just a Latin American phenomenon (e.g. Atkinson and Piketty, 2008, 2010). How do we know that very high incomes are not captured in household surveys? By inspection, one can observe that survey top incomes are at most close to the earnings of a well-paid manager; additionally, capital incomes as measured by surveys are a fraction of what is reported in tax registries and National Accounts.

Given that top incomes (especially the very rich) might be mismeasured or altogether missing from household surveys (even in the harmonized series), it is pertinent to ask whether the levels and dynamics of inequality based on household surveys, which we have discussed so far, are robust to adjustments made with administrative data (tax or social security records). Indeed, question

marks over the reliability of household surveys have grown, as evidence on top incomes from tax records in country studies accumulates.[31]

However, partially replacing, adjusting, or complementing survey data with administrative information is a statistically complex matter, and not at all trivial. First, there are issues regarding the data themselves. In no country have the tax and other administrative data been primarily designed to provide comprehensive statistical information on income or wages (which surveys are).[32] They were designed as a means to inform the collection of fiscal revenue. As a result, inequality series based exclusively on tax data, or series based on income-tax adjustments to survey-based data, are always affected by the definitions of the tax code and the behavioral responses to taxation. For example, interest income is usually withheld at the source (when taxable), and dividends are in many cases taxed through the corporate tax and not reported in individual income tax files.[33] But these rules are not universal.

Evasion and elusion may render some high incomes and capital incomes not observable through tax data either. It is thus not uncommon in Latin American countries to see that the income tax information includes very little income from capital sources. In addition, the income definitions recorded for tax purposes display significant variations across countries, highlighting once again, in this new context, that details are important. Also, income tax data almost always excludes informal sector workers, who are a very sizable proportion of most labour forces in the region.

It must also be remembered that, although surveys have been increasingly affected by non-response and under-reporting in many countries around the world (as described, for instance, by Meyer, Mok and Sullivan (2015) for the US and Campos-Vazquez and Lustig (2020) for Mexico), the intuition that the absence of the rich from the surveys must necessarily – mathematically -

[31] With the increasing availability of income tax and social security records providing distribution information in LAC over the last two decades, many studies have analyzed the distribution of top incomes and compared the results with household surveys. See for instance, for Argentina, Alvaredo (2010); for Colombia, Alvaredo and Londoño-Vélez (2013, 2014) and Díaz Bazán (2015); for Dominican Republic, Alvaredo et al. (2022); for Uruguay, Burdín et al. (2022) and Flachaire, Lustig and Vigorito (2022); for Mexico, Bustos and Leyva (2017); for Ecuador, Cano (2015, 2017) and Rossignolo, Oliva and Villacreces (2016); for Brazil, Ferreira de Souza and Medeiros (2017), Morgan (2018) and Morgan and Souza (2019); for Chile, Fairfield and Jorrat (2016), López, Figueroa and Gutiérrez (2013), and Flores et al. (2020); for Costa Rica, Zuñiga-Cordero (2018).

[32] And in almost no country (an exception being the Nordic countries) do administrative registries provide today a reasonably complete coverage of the population targeted in inequality studies. For this reason, the 'top incomes' literature has been restricted to estimating shares accruing to high income groups only.

[33] Saez and Zucman (2016) estimate that capital incomes reported to the income tax in the US represent only 1/3 of capital incomes measured by National Accounts, the gap being a combination of imputed interest, imputed rents, exemption and deductions in the income tax, earnings retained in trusts, accounting differences, income paid to pension funds and life insurance funds, and tax evasion.

reduce measured inequality is false. Whether it does so or not is an empirical matter and depends on the underlying distributions.[34] As we will see, in most if not all cases, it does have that effect.

Those data caveats notwithstanding, most Latin American countries have once again fully participated in top incomes research and, over the last fifteen years, an increasing number of fiscal authorities have been providing income tax and social security distributional information in a systematic way, adding to those countries which had already been doing so for a long time.

Second, there are important methodological issues: various different methods have been suggested to adjust household surveys using tax and other administrative data but, unfortunately, these methods tend to produce very different results and none should be applied mechanically.[35] One influential method, proposed by Blanchet, Flores and Morgan (2022, henceforth BFM), is based on the estimation of an income threshold – or 'merging point' – above which the information from tax data is incorporated into the survey.[36] This is followed by a rescaling of the survey weights below the merging point using the ratio of the survey to tax data densities at the merging point, while keeping the representativeness of the survey in terms of socio-demographic variables (e.g., age, gender, among others). Finally, survey observations above the merging point are also reweighted, and the survey incomes replaced with observations that reproduce the distribution observed in the tax data.

Although this approach is conceptually appealing – in that it tries to combine the advantages of tax data for high incomes with the strengths of surveys elsewhere along the distribution – the results it generates may not be robust to the choice of the merging point. Inequality estimates can be very sensitive to the selected threshold. Flachaire, Lustig and Vigorito (2022) show, in the context of simulations of synthetic data, that "if one chooses a threshold that is not close to the true one, corrected inequality measures may be significantly biased" (p.1).

Nor is it clear which other applied method – if any – does a much better job. De Rosa, Lustig and Martinez Pabon (forthcoming) find that there does not appear to be a robust ranking across methods: that is, a particular method may result in the highest (corrected) inequality with one data

[34] See for instance Deaton (2005).
[35] See Lustig (2019) for a survey of adjustment approaches applied in the literature.
[36] The merging point is selected as follows. The first step is to assess the 'trust region' in administrative tax records: that is, above which the tax data seems reliable. This step is recommended especially in developing countries given the extent of informality that prevails. The latter creates a lot of noise at the bottom end of tax data, where the starting point of the "bottom" depends on the country. The tax data below the trusted threshold is removed. The next step starts by comparing the cumulative distribution functions and the density functions from the survey and tax records to identify the threshold or merging point. The merging point is the maximum point at which the ratios of the cumulative density function in survey and tax records and the ratio of the density function in survey and tax records are equal. Selecting the threshold as this merging point helps ensure continuity between the two functions. The method chooses the maximum point at which these ratios coincide (there can be other points) to preserve the survey below the top incomes as much as possible.

set but not with another. These authors combine household survey and tax data for seven Latin American countries using various different approaches, and find that the ranges of estimates are wide, and that rankings vary. For example, for Argentina (2017), the unadjusted (i.e. survey-based) Gini is 47.7; the highest adjusted Gini is 52.9 (BFM method), and the lowest is 44.3 (replacing survey's incomes with tax-based incomes above an ad-hoc threshold (percentile 90), as in Jenkins, 2015) – lower than the unadjusted one. For Chile (2013), on the other hand, the unadjusted Gini is 52.4; the highest adjusted Gini is 63.2 (replacing incomes with threshold at percentile 90), and the lowest is 58.4 (replacing incomes above percentile 99). For Uruguay (2009), the unadjusted Gini is 50.6; the highest adjusted Gini is 68.4 (replacing incomes above percentile 90), and the lowest is 56.7 (BFM).

The fact that inequality estimates obtained from combining household survey and tax data appear to be so sensitive to the specific methodological approach that is used – and that there is no clear conceptual reason so far to prefer one to another – poses a serious challenge. Particularly since it now seems clear that household surveys alone do a poor job of capturing the top of the distribution, while tax data (in countries with large informal sectors) are not good at capturing the bottom – or even the middle – of the distribution. Combining information from the two sources would seem to offer a very promising path – if only an accurate, reliable and robust approach based on general statistical principles could be found. Unfortunately, it seems that we are not there yet.

Naturally, these data and methodological limitations have not stopped people from trying – quite rightly, as this is how progress is made – and, in the process, they have generated additional variation in inequality measures for many Latin American countries. To shed some light on that additional variation, in what follows we use the series from De Rosa, Flores, and Morgan (2022), which applies a modified version of the BFM method described above. This modified version skips the final step, where survey observations above the merging point are replaced with observations that reproduce the distribution in the tax data. Keeping the incomes reported to the survey is equivalent to assuming (among other things) that there is no underreporting in the upper tail, and that the survey's weights are incorrect.

We use the estimates from De Rosa, Flores, and Morgan (2022) for three main reasons. First, they include ten of the eighteen countries considered in Figures 3 and 4, harnessing most of the income tax information available for the region between 2000 and 2020.[37] Second, they allow for a cross-country comparison of the effects of the income tax-based adjustments in a context where exactly the same methods and assumptions are applied in each country. This is particularly important given the aforementioned lack of robustness across methods (including this one). Third, besides assessing the effects of incorporating distributional information from registries to surveys, these

[37] By necessity, the period covered can only start in 2000, when administrative data are more widely available.

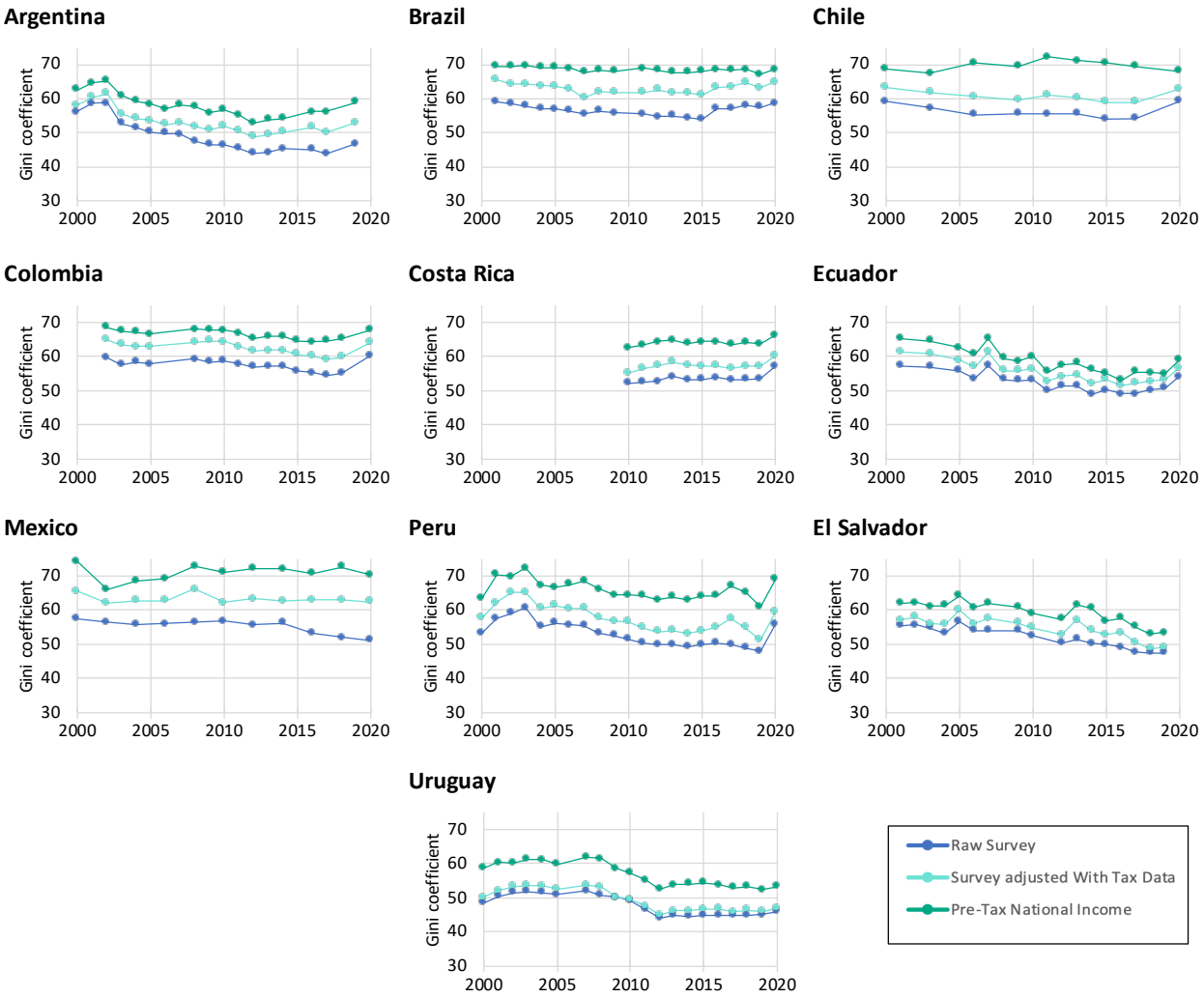
authors go a step further and estimate inequality after imputing the gap to the National Accounts totals, also following comparable methods. That said, in light of the previously discussed limitations, one would have preferred to present the range of inequality indicators generated by several correction methods over the same time span, and show bands rather than a single figure. That exercise must be left for future research.

Figures 7 to 11 show the Gini coefficients and the income shares accruing to the top 1% (P99-100), the top 10% (P10-100), the "middle" 40% (P50-90), and the bottom 50% (P0-50) of the population (in that order) from De Rosa, Flores and Morgan (2022), between 2000 and 2020. In these figures, the series at the bottom of each plot are always based on household surveys only and, in this sense, they are comparable to those discussed in section 3 and shown in Figure 4: they refer to household per capita income and are built on the surveys homogenised by ECLAC. The second series in each plot is based on the tax-adjusted correction by De Rosa, Flores, and Morgan (2022).

Inspection of the figures gives rise to two main general observations. First, the measured *level* of inequality increases in all cases, but the magnitude of the change varies across countries. It is the smallest in Uruguay (circa 2 Gini points on average across the period), and the largest in Brazil and Mexico (c. 8 pp. on average but more than 10 pp in some years). Note that the magnitude of the gap before and after the adjustment in each case should not be taken as indication of the quality of the surveys, at least not as the only cause. The reason may well also pertain to the quality and coverage of the tax information, which, as mentioned above, is not homogenous. For instance, personal tax data in Brazil include dividend incomes to a much larger extent than in any other LAC country, partially explaining the size of the adjustment. In any case, the adjustment produces substantial changes, where the Gini increases following the gains in the top 10% (and the top 1% in particular), while the shares of the middle 40% and the bottom 50% decline.[38]

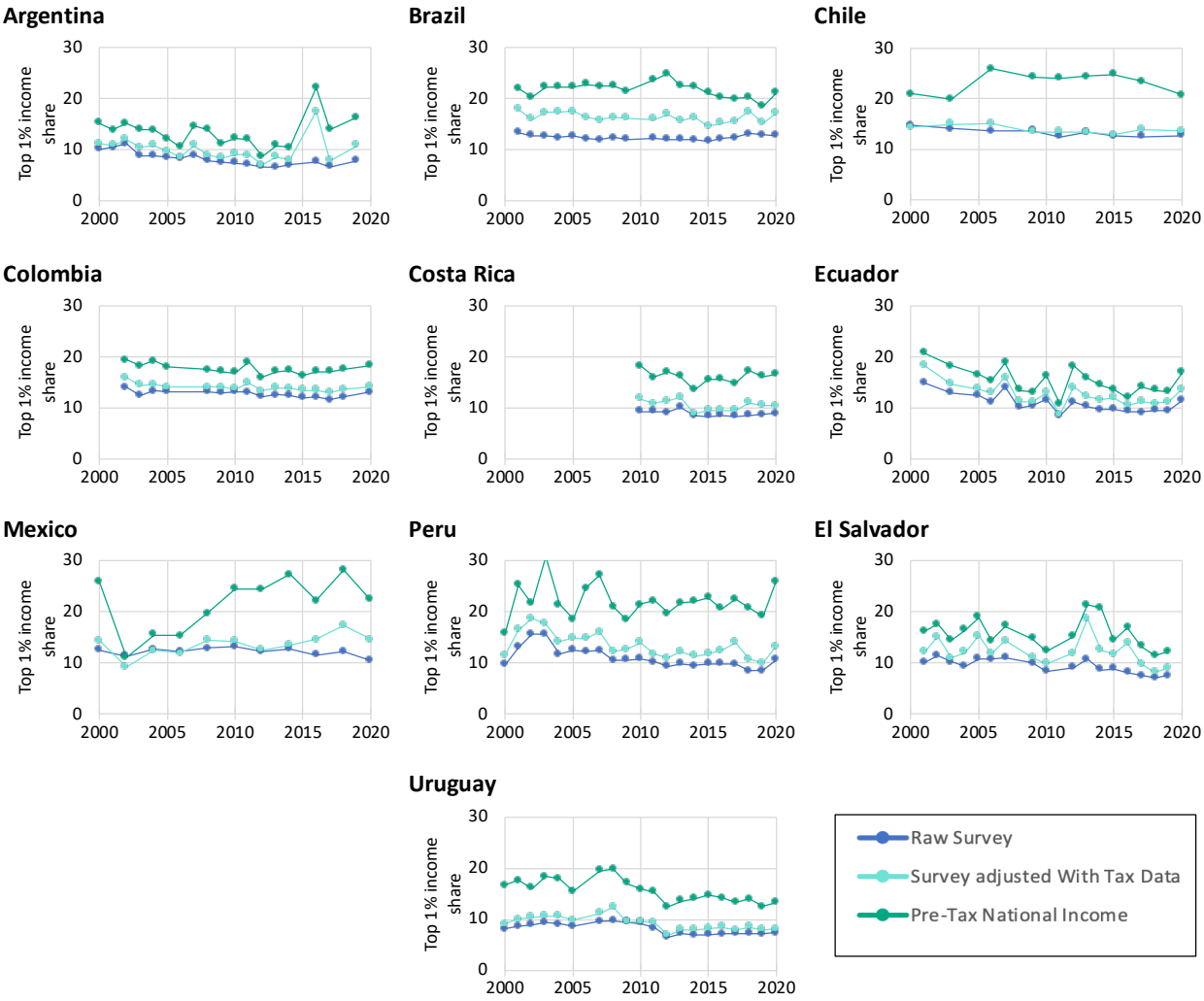
[38] It should be noted that the applied literature focusing on this type of adjustments has so far abstracted from the effects on poverty rates.

Figure 7. Gini coefficients in Latin America 2000-2020. Household surveys and the effects of adjustments with administrative data and national accounts



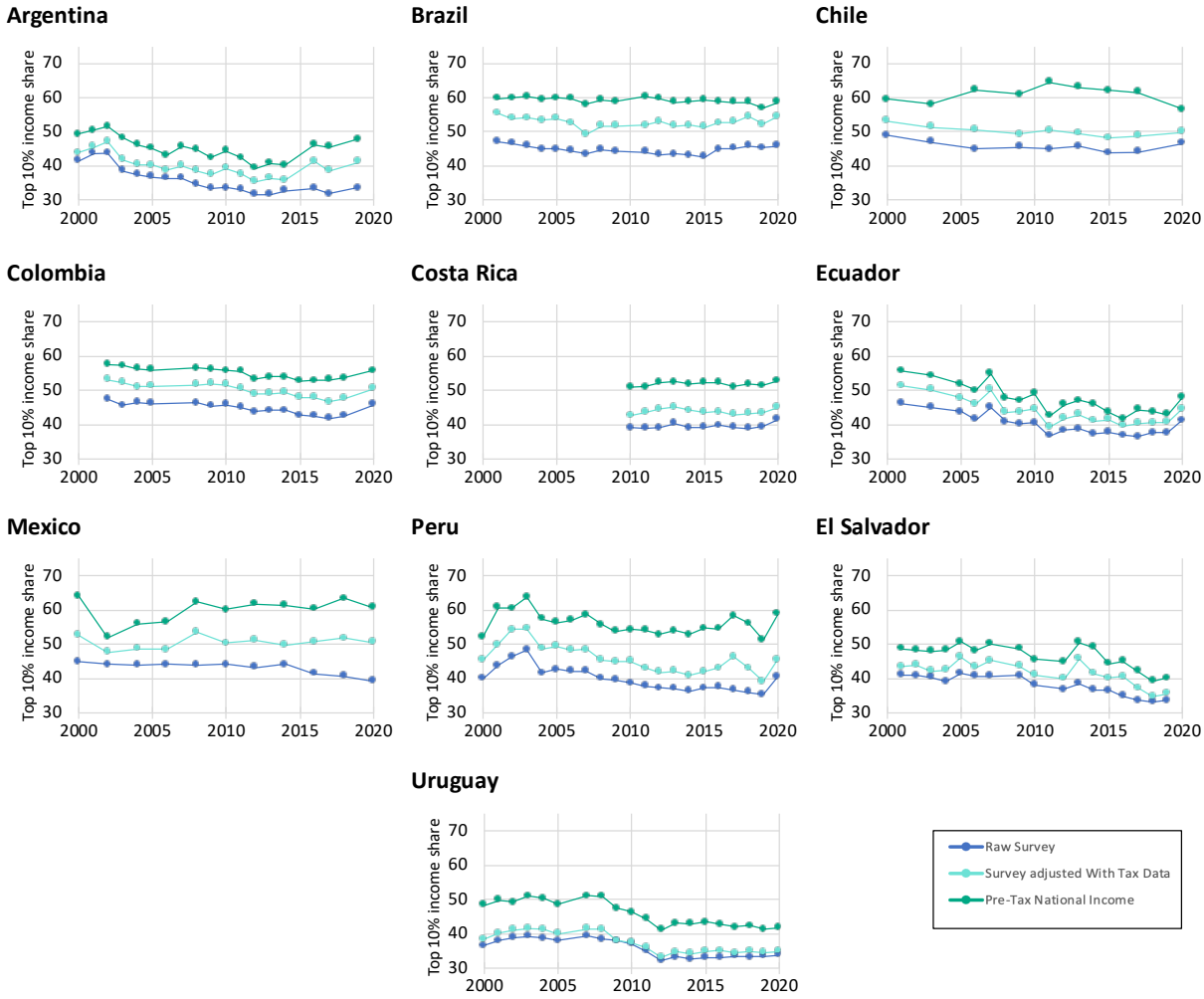
Source: Series from De Rosa, Flores, and Morgan (2022)

Figure 8. Top 1% income share 2000-2020: household surveys and the effects of adjustments with administrative data and national accounts



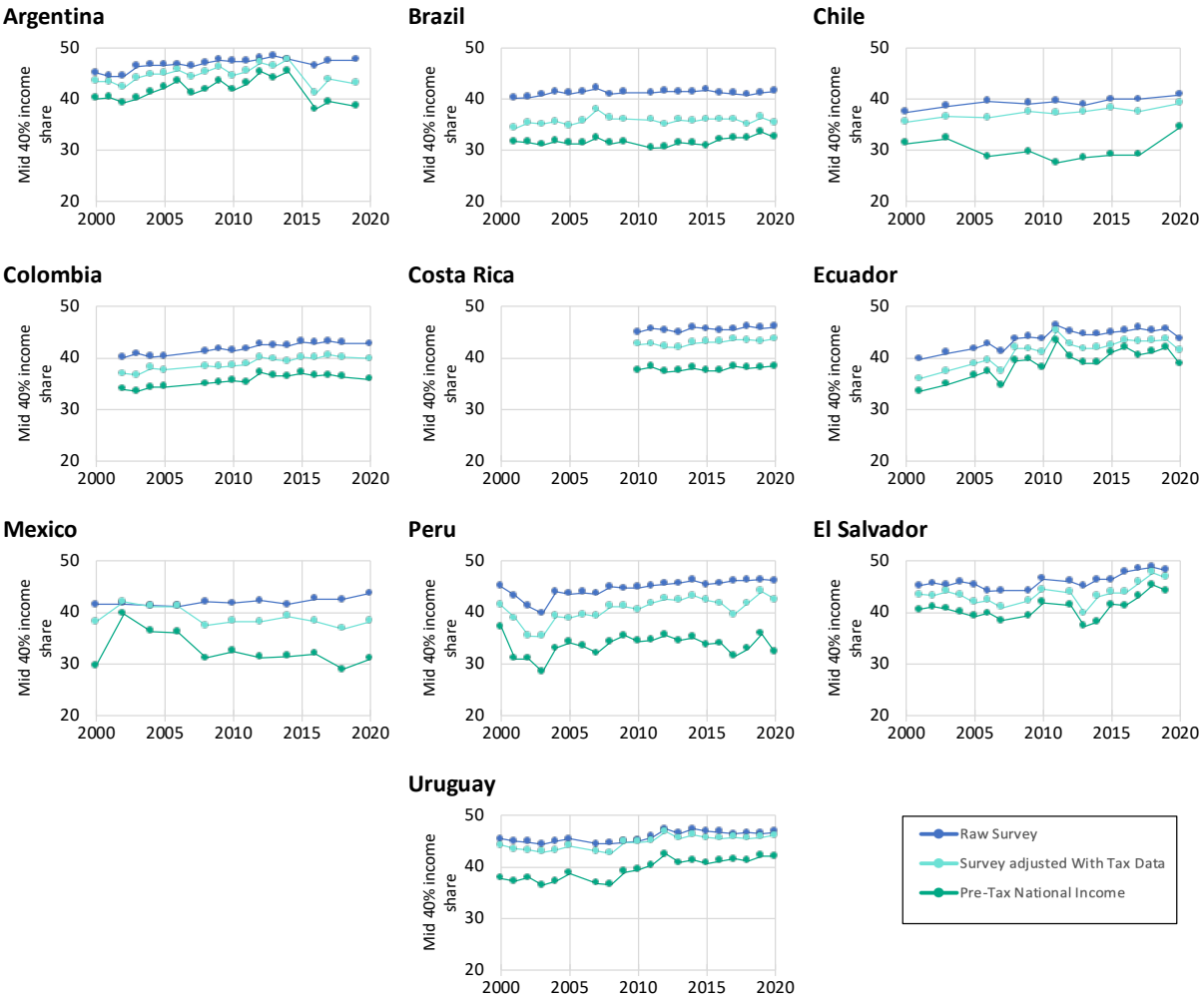
Source: Series from De Rosa, Flores, and Morgan (2022)

Figure 9. Top 10% income share 2000-2020: household surveys and the effects of adjustments with administrative data and national accounts



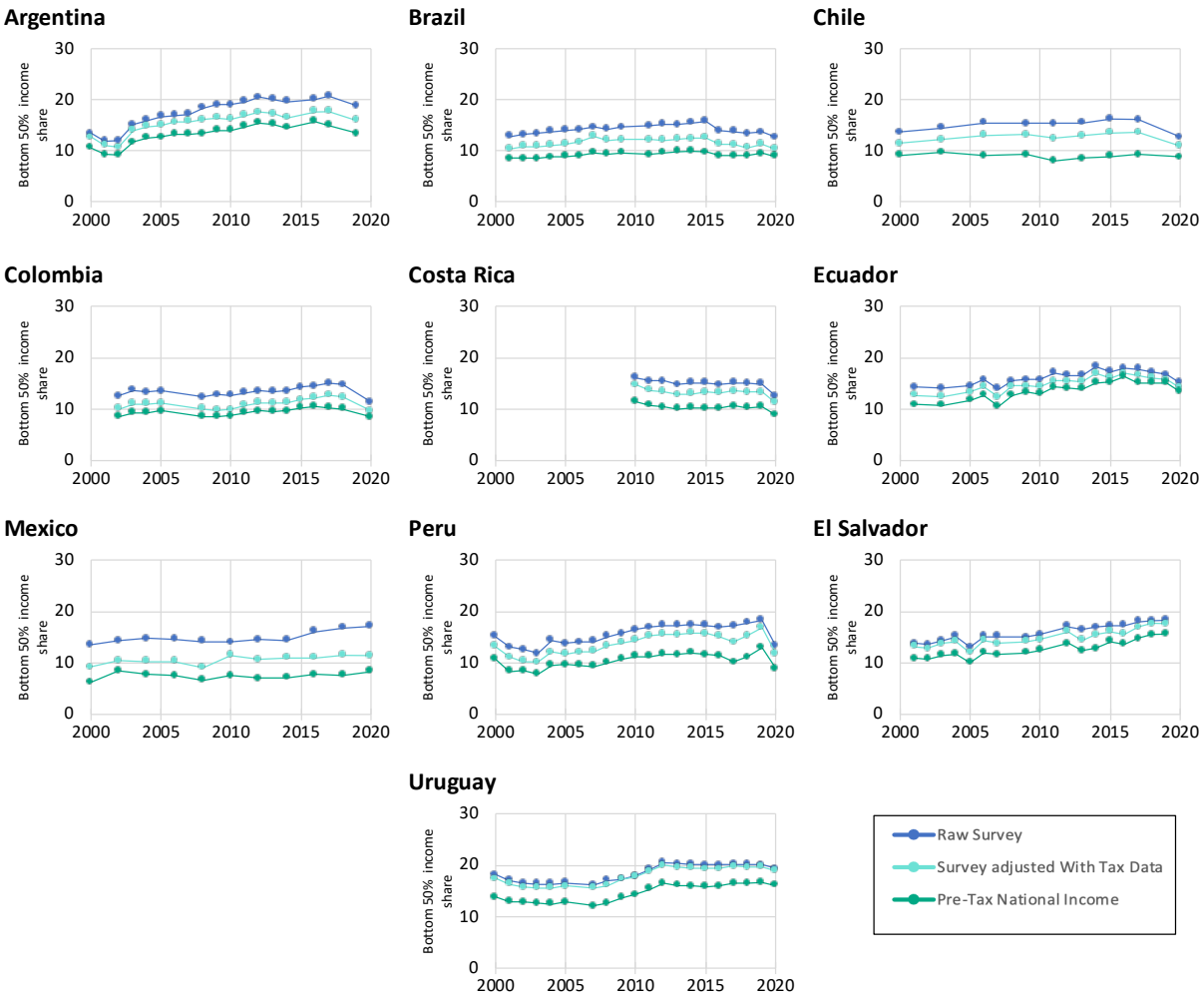
Source: Series from De Rosa, Flores, and Morgan (2022)

Figure 10. Middle 40% income share 2000-2020: household surveys and the effects of adjustments with administrative data and national accounts



Source: Series from De Rosa, Flores, and Morgan (2022)

Figure 11. Bottom 50% income share 2000-2020: household surveys and the effects of adjustments with administrative data and national accounts



Source: Series from De Rosa, Flores, and Morgan (2022)

Second, and most importantly, with the exception of Mexico in the second half of the 2010s, the adjustment does not alter the dynamics given by surveys when attention is focused on the top 10%, middle 40% and bottom 50% percent shares. Indeed, and perhaps surprisingly, the introduction of tax information does not seem to significantly alter the story in the case of the top 1% share either, even if, this being a small group (for which survey estimates are always particularly affected by the sample size), the adjustments present a higher degree of variability, as in Argentina, El Salvador, and Peru. Mexico may represent an exception after 2014, as the adjusted top one percent share increases, while the survey-based share decline.

The third (and uppermost) series in each plot in Figures 7-11 incorporates the imputation of the gap with National Income from the SNA, including an estimate of undistributed profits, which are retained by firms. In Figure 7, the scaling up to NA produces a substantial upward shift in the series. Adding these "missing" incomes to households in the surveys has first and foremost a level effect: they add between 5 and 10 percentage points to the Gini coefficient. But again, the dynamics of inequality as depicted by the survey do not change substantially, with three main exceptions. In Mexico, the decline in inequality given by surveys during the second half of the 2010s disappears, with inequality staying more or less constant. In Brazil, something similar happens. In Chile, perhaps surprisingly, the adjustment turns the survey and tax-based decline in inequality over 2000-2010 into an increase, and the increase over 2010-2020 into a decline.

Why is this final adjustment made? Even taking into consideration the adjustments of surveys with tax information, there is still a large – and sometimes increasing – gap between aggregates from inequality studies based on microeconomic data and the income totals from the SNA. The discrepancies can be seen in the levels of income, as well as in their growth rates (see, for example Ravallion (2003); Deaton (2005); Bourguignon (2015); Nolan et al. (2019)) and can attain particularly high levels in developing countries. While it may not be surprising that national income is larger than the income concepts traditionally used to study inequality, it has sometimes been growing faster too. It has been argued that these discrepancies make it hard to assess how macroeconomic growth is distributed across income groups, and to what extent existing distributional statistics are a proper representation of the income flows in an economy.

Implicitly rooted in the tradition of inequality research in the 1940s, 1950s and 1960s (mentioned in section 2), as well as in the methods developed by ECLAC and Altimir (1986, 1987), recent work has embarked on a process of scaling up the various available data sources (surveys, administrative records, rich lists) – through further imputations – to SNA totals. These include, among others, the World Inequality Lab (Alvaredo et al., 2020), Fixler et al. (2017) and the DNA-Distributional National Accounts project coordinated by the OECD (Zwijnenburg, 2019). While the existing gaps have sometimes fed feelings of uncertainty about inequality measurement, these

new approaches have taken for granted the numbers provided by the national accounts, a practice that does not always contribute to diminish those feelings, at least in the case of developing countries.[39]

The situation in Latin America concerning the comparison of aggregates across different sources is discussed in Alvaredo et al. (2022) and in De Rosa, Flores, and Morgan (2022). They show that the gap between the total income reported in household surveys and the national income estimate from the NA is very large, with total survey income accounting for between 40% and 60% of national income. Mexico appears as an extreme case, with total survey income representing only 30% of national income. According to Alvaredo et al. (2022), these gaps arise from conceptual issues as much as from measurement issues. Following Deaton (2005), and Deaton and Kozel (2005) in the case of India, it is important to reckon that if the discrepancy between household consumption expenditures or disposable income in surveys and in NA is partly due to some conceptual or measurement flaws in the latter, then other components of NA in developing countries are also affected, including GDP. This adds still more uncertainty in the correction of survey-based inequality estimates by NA. Be it as it may, the correction leads to a dismal message: scaling up incomes from the surveys, even after adjusting with administrative tax data, to close the gap with the SNA aggregates requires allocating to households approximately half of total national income (as measured in the SNA), with this half mostly composed of items about whose distribution we know the least: capital incomes and retained earnings. This is not a minor detail. Imputations rely on fragmentary information as much as on researcher judgement.[40]

The gaps between inequality estimates from tax and other administrative data and household surveys – particularly when the former is scaled up to match National Accounts aggregates – dwarf the inequality bands among estimates based exclusively on household surveys. Table 3 below presents the width of three inequality bands, but now only for the ten countries shown in Figures 7-11. First, we re-state the inequality Bands 1 and 2 from Section 2, and then we report additional bands between the lower bound of Band 2 and the two estimates from Figure 7: the survey adjusted with tax data/BFM estimate (Band 3), and the estimate scaled up to match the SNA total (Band 4).

[39] While the definition of national income certainly includes an estimate of undistributed profits (which could be considered as a proxy for shareholders' capital gains), it excludes other items which may be very relevant for the groups in the middle of the distribution, such as capital gains from real estate. This points to the fact that distributing the national income does not close the conceptual and practical questions on the most suitable definition of income for inequality studies.

[40] Although we refer only to the situation in Latin America, scope for disagreement on these imputations also exists in countries with richer data constellations. See, for example, the debate about the United States between Piketty, Saez and Zucman (2018) and Auten and Splinter (forthcoming).

Table 3. Four inequality bands for LAC, 2000-2020

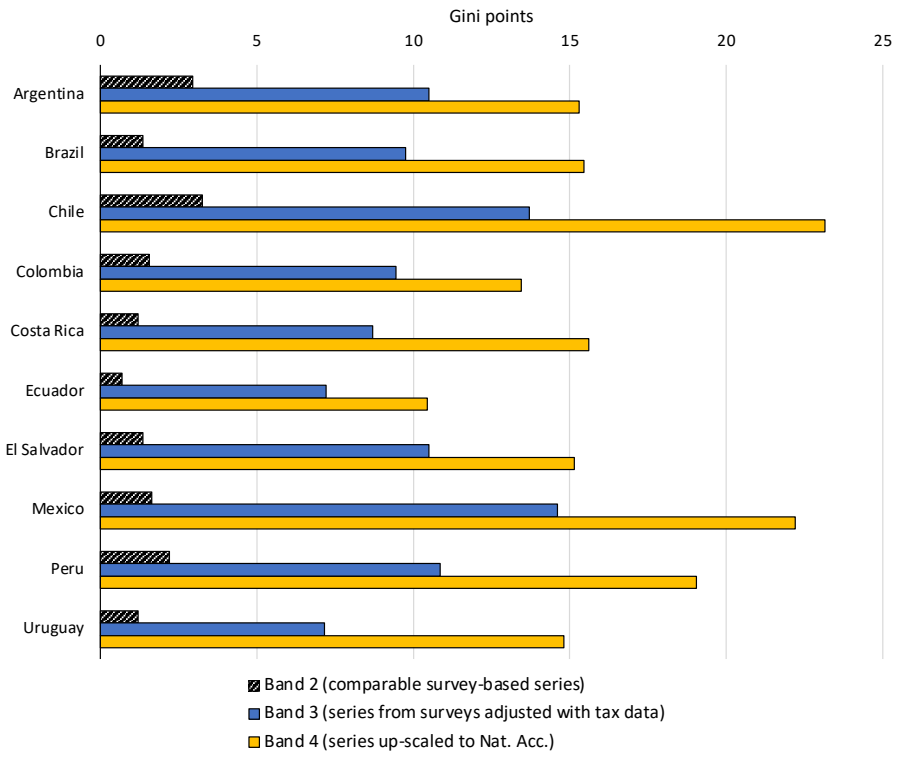
		Band 1	Band 2	Band 3	Band 4
		Annual range of estimates across all survey-based series (idem Table 1)	Annual range of estimates across SEDLAC / LIS / ECLAC new / PovcalNet-PIP (idem Table 1)	Annual range of estimates between the lower bound of Band 2 and Gini coefficient including BFM/tax adjustment	Annual range of estimates between the lower bound of Band 2 and Gini coefficient up scaled to National Accounts
		decadal average	*decadal average*	*decadal average*	*decadal average*
	Years	*Gini points*	*Gini points*	*Gini points*	*Gini points*
Argentina	2000s	4.2	3.5	9.9	14.7
	2010s	2.9	2.4	11.1	15.9
Brazil	2000s	9.1	1.3	8.4	14.2
	2010s	6.9	1.4	11.1	16.7
Chile	2000s	9.6	3.3	12.7	20.4
	2010s	9.3	3.2	14.7	25.9
Colombia	2000s	8.5	1.5	9.3	13.1
	2010s	8.4	1.6	9.6	13.8
Costa Rica	2000s	4.3	0.9		
	2010s	5.8	1.5	8.7	15.6
Ecuador	2000s	7.3	1.2	6.9	10.6
	2010s	3.9	0.2	7.5	10.3
El Salvador	2000s	6.6	1.0	9.1	13.8
	2010s	5.0	1.7	11.9	16.5
Mexico	2000s	8.0	1.0	13.6	19.9
	2010s	10.0	2.3	15.6	24.5
Peru	2000s	11.1	2.3	10.8	17.7
	2010s	8.2	2.1	10.9	20.4
Uruguay	2000s	5.2	1.3	7.5	15.4
	2010s	7.3	1.1	6.8	14.2

Band 1, in the first column of Table 3, was already discussed in Section 3. It includes estimates from concepts other than household per capita incomes, so the additional variation with respect to Band 2 is accounted for by differences in the concepts of welfare indicator and units of analysis (Variety 2). We now focus on a comparison across the other three inequality bands, which start from the lower bound of Band 2 – that is, the lowest inequality estimate (for household per capita income from the household surveys) for a country/year. With that lower bound, Band 3 incorporates the tax adjustments reported by De Rosa, Flores, and Morgan (2022), and Band 4 also accounts for the scaling up to SNA carried out by the same authors. The average widths (over time) of Bands 2, 3 and 4 for each country are shown in Figure 12.

Across the whole period and all ten countries, the average width for Band 2 is 1.7 pp, while it is 10.4 pp for Band 3 and 16.9 pp for Band 4. [41] In other words, incorporating adjustments from the tax data in LAC adds an average of 8.7 percentage point to the variation in inequality measures, and attempting to scale up to the SNA adds another 6.5 pp.

[41] The width of Band 3 is based entirely on the De Rosa, Flores and Morgan (2022) tax-cum-survey estimates. Since there is considerable uncertainty about those numbers and non-negligible variation across estimates from other methods (see above), this width is likely an underestimate.

Figure 12. Average width of inequality bands over 2000-2020: among HHS-based measures; including tax adjustments; and including scaling up to SNA aggregates.



Given the strength of the assumptions that must be made to make these adjustments, and the variation in outcomes across methods, which we have summarized above, we find it difficult to treat these numbers as many appear to have done: as unambiguous indicators of how much higher inequality really is in the region. We believe that these results do confirm that household survey-based measures underestimate inequality in the region. But they do not yet represent credible, robust point estimates of the true, higher level. Given the available information, we believe looking at the three bands described above gives us a sense of the range where true income inequality in household per capita income is likely to lie.

This range is typically quite wide – around 10 pp for the tax adjustment only, and closer to 17 pp when scaling up to SNA – but the levels are very high across the entire plausible bands. In other words, even under the almost certainly underestimated household surveys, LAC has Gini coefficients that are almost always greater than 0.40. Depending on the adjustments, they can sometimes exceed 0.70. This is very high inequality indeed, although precisely how high remains, for the moment, difficult to say.

Fortunately, as documented above, the dynamics and, in particular, the rough inverted-U pattern between episodes 1 and 2, are much more robust. They appear to hold – albeit with different timing and in different magnitudes – for all but one or two countries, across all three bands. In the next section, we briefly consider some of the economic – and political – factors likely to have played a role in underpinning these dynamics. Given all the uncertainty, the discussion should be read as largely speculative.

5. Interpreting the inequality dynamics in LAC: a bird's eye view

Despite the difficulties associated with measuring income inequality and the various – not necessarily consistent – data sources used to do so, inequality in LAC countries seems to have followed a reasonably clear pattern over the last few decades. First, based on inequality estimates grounded on household surveys and judging from the countries for which such estimates have been available now over 40 years or more, inequality was higher in the mid- or late-1990s than in the preceding decades over which data are available. The actual path bears a degree of uncertainty. The rise was continuous in Argentina (urban) and Chile – except for the Allende years where inequality declined for a few years – whereas there are not enough observations to say whether this has also been the case in Brazil or Mexico. Second, for most countries in the region, inequality has steadily declined over the 2000-2015 period, and the cumulative decline has been substantial, reaching in some instances more than 10 Gini points. Third, this common trend seems to have broken down in most countries in the mid-2010s, with inequality moving up again in some countries (Brazil, before Covid), flattening (Argentina, Chile, Colombia, Panama, Uruguay) or still declining, possibly at a slower rate (Bolivia, Peru and Paraguay)

At the risk of over-simplifying the picture for the pre-1990 decades (especially for those countries with few pre-2000 observations), a schematic representation of the evolution of LAC inequality is that of an inverted-U, with a peak in the late 1990s or early 2000s, and a flattening of the descending arm around 2015. While such a description is precise for countries like Argentina and Colombia, it is more approximative for Brazil and Mexico. For countries such as Guatemala and Nicaragua where there are no inequality estimates before 2000, the inverted-U should be treated as a hypothesis, where only the descending part of the inverted U is observed.

Based on that simple representation of the evolution of inequality, the question we ask in this section is how to interpret such a long and pronounced inequality cycle. When faced with an inverted-U inequality pattern over time, it is almost a reflex for economists to think of the Kuznets curve. Yet, the mechanisms emphasized by Kuznets (1955) for the historical patterns he discerned in his data do not seem to fit the phenomena we observe in Latin America over this more recent period: the last fifty years were not, in the main, a period of rapid industrialization and structural change marked by a constant flow of workers from a low-wage to a higher-wage sector, with

between-sector inequality first rising and then falling, largely as a result of changing population shares. On the contrary, the 2000s were, if anything, marked by de-industrialization and a reversal of the previous decline in primary commodity exports in most of the Latin American countries where inequality fell.

With a traditional Kuznets process seemingly not a suitable explanation for this particular inverted-U trajectory, we turn instead to two alternative – or perhaps ultimately complementary – types of interpretation, which depend on the nature of the dynamic process behind the inverted-U shape and the recent flattening of the trajectory.

In the first scenario, the evolution of inequality is a kind of random process with a sequence of temporary shocks, with some similarity across countries, and some of them with a long memory. Then the ascending part would simply correspond to a succession of shocks with a net positive impact on inequality, the most recent ones tending to reinforce the effects of the previous ones. Then the process reverses in the early 2000s when some opposite shock, or a sequence of shocks, occurs, which progressively erase the effects of the preceding sequence and triggers a virtuous evolution.

The bad sequence of common shocks that hit Latin America from the late 1970s to the early 1990s seems rather obvious: the oil price rises in the 1970s; the subsequent increased indebtedness; growing inflationary pressures and slowdown of economic activity; the debt crisis triggered by the rise of interest rates in 1982; the severe recession that followed; and the IMF-led macroeconomic adjustments. Most of these shocks affected the bottom of the income scale more severely than the top as, for instance, with the reduction in real labor earnings arising from a devaluation responding to an adverse external position. It is true that capital was also severely hurt, which could have moderated the impact on inequality. However, this cannot be checked due to estimates in that period relying mostly on surveys and without a simple solution to generate alternative estimates that would include top incomes and their possible losses.

The symmetric trend to the preceding sequence of bad shocks was the substantial improvement in terms of trade after 2000, largely driven by rising commodity prices and, to some extent, by a gradual shift in the origin composition of manufactured imports into Latin America from OECD sources to China. The progressive decline in interest rates, global growth at unusually high rates, and a substantial drop in inflation rates all around the world also contributed. It is also the case that left-leaning governments came to power in several countries – something that may not be totally exogenous with respect to mounting inequality in the preceding decades and suggests that, overall, rising inequality may generate its own counteracting forces through social and political

channels.[42] These governments typically adopted policies more favorable to the poorest segments of the population, such as cash transfers, minimum wage increases, or increasing coverage of pension schemes.[43,44] Interestingly, this strong equalizing trend stops, or becomes more heterogeneous across countries, in the mid 2010s, precisely at a time where the price of the commodities exported by LAC countries peaks and starts to decline, after a close-to-15-year rally. Chinese growth slows down, with strong spillover effects on the rest of the world, and the US Federal Reserve starts increasing interest rates (temporarily, as it turned out), signaling the coming end of an era of particularly abundant capital inflows into LAC financial markets.

According to this first interpretation of the evolution of income inequality in LAC, inequality changed because of shocks to the economy, with no permanent impact even though shocks may affect the economy during several years after they hit. It would thus be because there has been a succession of inegalitarian shocks in the 1980s and the 1990s that the ascending part of the inverted U was observed. The same applies on the way down but, in this view, there is some stationarity of inequality; after a shock, effects will progressively weaken and the economy as well as inequality will return to their 'normal' level. This may partly correspond to the downward part of the inverted U, a process reinforced by the positive shocks mentioned earlier. In short, inequality would be following a 'stationary random process' that would return to a long-run equilibrium value if the economy were not constantly hit by shocks. The correlation across countries arises essentially from global shocks affecting them in a similar way. But of course, there may also be purely domestic shocks governing the random process. For instance, Colombia was affected less strongly by the debt crisis than other countries, and inequality was declining there in the early 1980s, while it was rising elsewhere.

The second interpretation sees inequality changes instead as arising from permanent shocks, resulting from structural changes taking place in the economy, themselves the result of deep factors like education, demographics, technology, preferences, or policies. According to this view, inequality was generally high in the late 1990s not so much because of the succession of bad temporary shocks, but because certain structural changes had taken place in LAC economies or new policies were at work. And inequality started to decline in the 2000s not only because of higher commodity prices, faster growth, or improved employment, but under the pressure of

[42] A point stressed by Cornia (2010) and Lopez-Calva and Lustig (2010). Feierherd et al. (2023) give evidence that, even though inequality declined in practically all countries, the decline was more pronounced under left-leaning governments in the region. This has (tentatively) been linked to policies such as increasing minimum wages and introducing (or expanding) noncontributory pensions.

[43] See López-Calva and Lustig (2010), Cecchini and Madariaga (21011), and Azevedo et al. (2013).

[44] On the equalizing effect of an increase of the minimum wage, see Maurizio (2014), Maurizio and Vázquez (2016), Borraz and Pampillon (2017), Engbom and Mozer (2022), and Ferreira, Firpo and Messina (2022).

another type of structural change, sometimes in addition to policy changes implemented by inequality-averse governments.

An obvious structural change that took place during the 1980s in most LAC countries was the structural adjustment promoted by the Bretton Woods institutions at the time of the debt crisis. The adjustment was concerned not only with macroeconomic policy – e.g., fiscal and monetary policies, exchange rate management, and so on – but also with the whole structure of these economies, especially their trade orientation; the private/public nature of monopolies in sectors like energy, transport, or banking; and the whole tax/subsidy system. This market-oriented set of reforms meant a big change for the economies in LAC. It was not a change that would operate in a few years, but an adjustment to a new economic regime that would take a long transition period before accelerating income growth. It was found that this profound redefinition of the economic system produced more and more inequality, in the absence of countervailing policies – which, as a matter of fact, would have gone against the market orientation of the reforms.[45] On the other hand, the favorable terms of trade shock of the 2000s may have given governments the political space and the means to adapt the new economic regime and reduce its inequality bias through various types of social policies, possibly engineering another structural change.

Another example of structural changes that may have helped the descending arm of the inverted U are to be found in the dynamics of education and its returns in the labor market. Here again, several authors have emphasized this point.[46] The entry into the labor force in the 2000s of young cohorts of workers with a higher level of education than their predecessors caused significant changes in the distribution of labor incomes. There are two channels for such a change. On the one hand, more educated workers have an ambiguous effect on inequality, uneducated workers become poorer with respect to the mean labor income, whereas the contrary occurs for educated workers who become less rich in comparison to the mean worker. On the other hand, an equalizing effect may arise from a drop in the wages of more educated workers relative to those of less educated. This requires that demand for the former not to increase faster than the supply and these dynamics appear to have been at play in the region during the first two decades of this century.

Other factors have the capacity to generate structural changes in the distribution of income. Changes in labor force participation, especially of women, fertility and more generally household composition are example of such factors. They have been present in LAC countries over the last

[45] Structural adjustment policies have long been criticized for the heavier burden they imposed on the poorest part of the population. The impact of structural adjustment programs on rising inequality is tested in Forster et al. (2019) for developing countries, and in de Janvry and Sadoulet (2005) for LAC countries. More generally, see Morley (1995) and Lustig (1995), especially the chapters on Argentina, Brazil, Chile, Mexico, Peru and Venezuela.
[46] See, in particular, the analysis in Lopez-Calva and Lustig (2010).

decades, even though their distributional impact has possibly been hidden by other long-run trends and a succession of stronger shocks.

Coming back to our initial question, what explains the inverted U shape of the evolution of inequality and then the more recent disruptions of the declining trend in LAC countries? The multiplicity of causes of change mentioned in the preceding paragraphs, and the partial evidence collected on practically all of them by researchers and analysts for specific countries and periods, illustrate the difficulty of the question. We have suggested that a number of factors have cumulatively contributed to expanding inequalities in the last two decades of the 20th century: trade liberalization, privatization and other market liberalization policies, the macroeconomic adjustment caused by the debt and adverse terms of trade. Together, they largely overcame equalizing forces like changes in labor force participation or in fertility. Then, the trend reversal in the last two decades may be explained by the weakening of the previous inegalitarian forces, changed external conditions, policy initiatives and, in some countries, a clear equalizing impact of the increasing educational level of the labor force, itself the result of past educational efforts. Again, there is a summation of multiple factors moving in the same direction.

Much harder would be to ascertain the *magnitude* of the contribution of each of these components to the overall change in inequality, some of these effects being transitory and others permanent. This would be a formidable challenge, likely beyond present analytical capacity, not least because the data necessary to answer the question are simply not available. They obviously go much beyond income data, as they must allow for an analysis of the micro and macroeconomic mechanisms that ultimately determine household incomes. It seems to us that, at the present stage, we must be satisfied with identifying some of the main factors behind the inverted U curve, with having partial evidence on some of them, but without being able to quantitatively disentangle their exact contributions to the overall change.

6. Conclusions

Drawing on more over 64,000 estimates of inequality and quantile shares for Latin America and the Caribbean, between 1948 and 2021, this paper has sought to summarize what can be said with some confidence about income inequality in the region, while acknowledging the unavoidable uncertainty that arises from the use of multiple data sources and from our present methodological inability to combine them precisely and robustly. We have reached three main conclusions.

First, when one restricts attention to household surveys only and, furthermore, to the four main harmonized series (ECLAC, LIS, PovcalNet/PIP, and SEDLAC) of measures based on household

per capita income, there is relatively little variation. The width of that narrowest inequality band (Band 2) is typically less than two Gini points, and averages 1.7pp for the ten countries in Table 3. Nevertheless, it is widely recognized that these surveys are likely to underestimate income inequality, for various reasons, including under-coverage of top incomes and underreporting of capital incomes.

Second, recent attempts to better capture those capital incomes and richest households lead to substantially higher inequality estimates. Gini coefficients are ten percentage points higher on average (from the lower bound of Band 2) when information from tax data in Latin America is combined with household surveys, and more than fifteen percentage points higher when attempts are made to scale up to National Accounts estimates of national income.[47] Unfortunately, the methods themselves follow different approaches that even with the same data may yield strikingly different corrected inequality estimates. Moreover, these methods involve a variety of assumptions and decisions which must be made on the basis of limited information, and to which results are quite sensitive. This gives rise to a situation of considerable uncertainty about the exact levels of income inequality in the region: household surveys are almost certain to underestimate inequality, but corrections to them are not robust and may, in some cases, may overestimate it. So far as levels are concerned, we live in a world of some uncertainty, which we represent by showing means of a set of inequality bands.

Third, in terms of dynamics, it turns out that these inequality bands do paint a generally consistent picture. The information available for the first 25 years in our data span – from the late 1940s to the early 1970s – is too scant and difficult to compare to generate a reliable dynamic pattern. From the mid- 1970s onwards, those countries in the region with sufficient data generally experienced an inverted-U curve of rising inequality until the mid- to late 1990s and declining inequality thereafter, up until the mid-2010s, with divergent paths in the last few years. For some countries, such as Guatemala and Nicaragua, where household surveys started later and other sources were not available, we only observe measures since the 2000s. In those cases, the observed trend is consistent with the declining part of the inverted U. It is important to note that this is a broad pattern only. Different countries experienced peaks and troughs in different years and the magnitude of the rises and declines were different.

It is difficult to tell whether this broad inverted-U pattern reflects a stochastic, mean-reverting process, where different short-lived shocks drive changes or, instead, more permanent changes that reflect deeper structural transformations. Or, indeed, a combination of both. As discussed in

[47] The magnitude of these gaps is of comparable size in other developed and developing countries, as revealed by the comparison of survey-based estimates, income tax-based estimates, and the figures arising from the distribution of National Income by the OECD-DNA or the World Inequality Database.

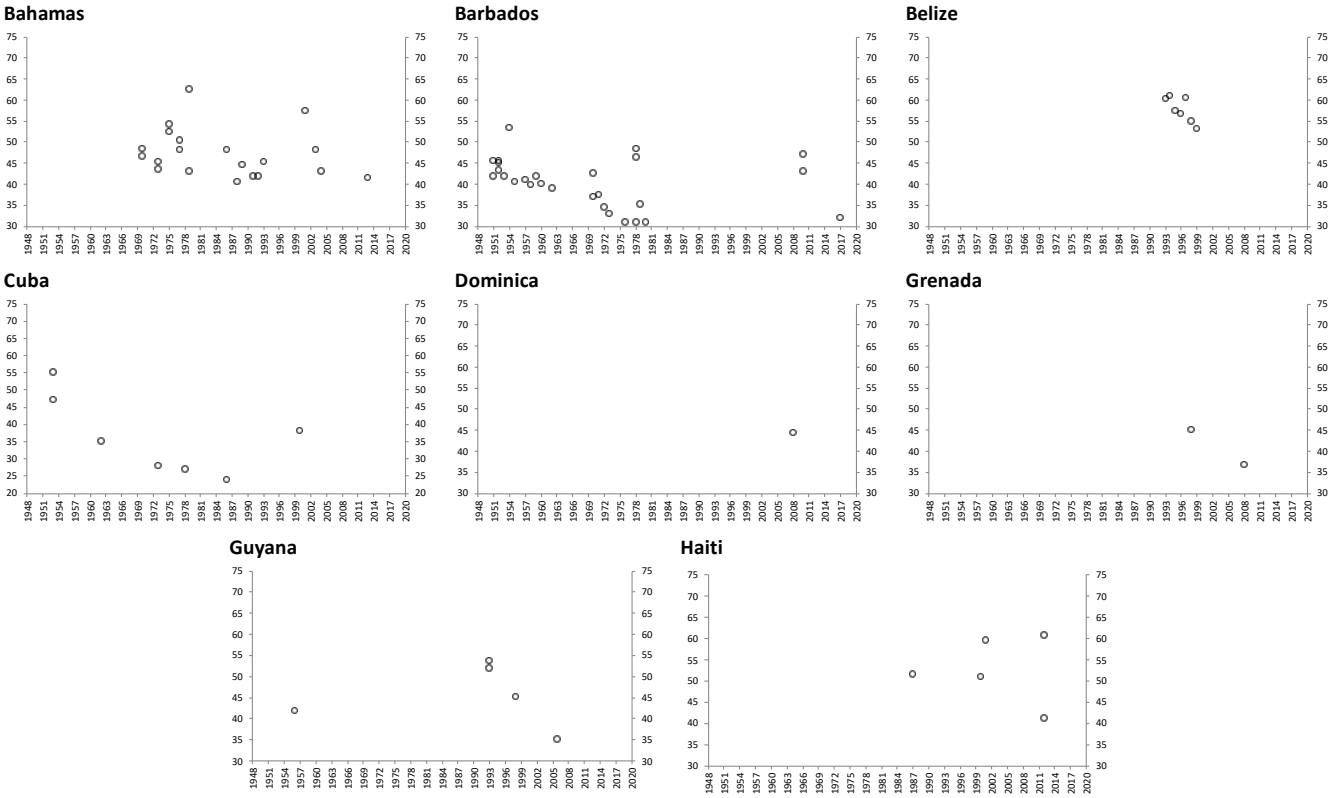
Section 5, there are plausible candidates for both interpretations. Oil price shocks, rising interest rates, the debt crisis, and sweeping market-oriented reforms are plausible suspects during the period of rising inequality, while rising commodity prices and the surge of left-leaning governments may have contributed on the descending side of the inverted U. These global phenomena will appear as correlate shocks for Latin American countries, while country specificities will account for differences in the exact timing and magnitudes. There are also plausible structural factors at work, such as declining returns to schooling (and in some cases, experience) during the 2000s, as educational expansions outpaced the demand for skills. While it is difficult to causally decompose the overall inequality changes into the effects of each of these forces, it is likely that all of them played some role.

Enormous progress has been made in the region over the years in data collection to measure inequality. Some of this progress, particularly as relates to capital incomes and the incomes of high-earning households, remains incomplete. While this generates a humbling sense of uncertainty about exact levels of inequality, it is also a promise of further progress ahead, as data collection improves across all types of instruments – surveys, administrative records, and national accounts estimates. In the meantime, the existence of a relatively robust broad dynamic pattern – with country specific peculiarities – provides researchers with plenty of challenging questions about drivers and mechanisms.

Some final thoughts are in order. This paper has underscored the uncertainty surrounding existing statistics on economic inequality. There are a number of actions that could be undertaken to reduce the uncertainty. Here we mention a couple of salient ones. First, to reduce the uncertainty in measuring inequality requires improving the quality of surveys and their harmonization. In particular, questionnaires should be designed and implemented in such a way to remove the ambiguity of the gross/net status of incomes reported by surveyed people. Otherwise, the uncertainty surrounding something as basic as the welfare indicator will subsist. Second, the underrepresentation of top incomes in household surveys could be addressed by oversampling high-income individuals. But this would solve the issue only in part since, even if included, rich individuals may underreport their income. The latter could be addressed if governments made the information from (anonymised) tax records available and allow for the linking through personal identification numbers between surveys and registries. More broadly, governments, international organisations and the scholarly community need to be committed to transparency and to make information publicly available in ways that facilitate the measurement and analysis of economic inequality.

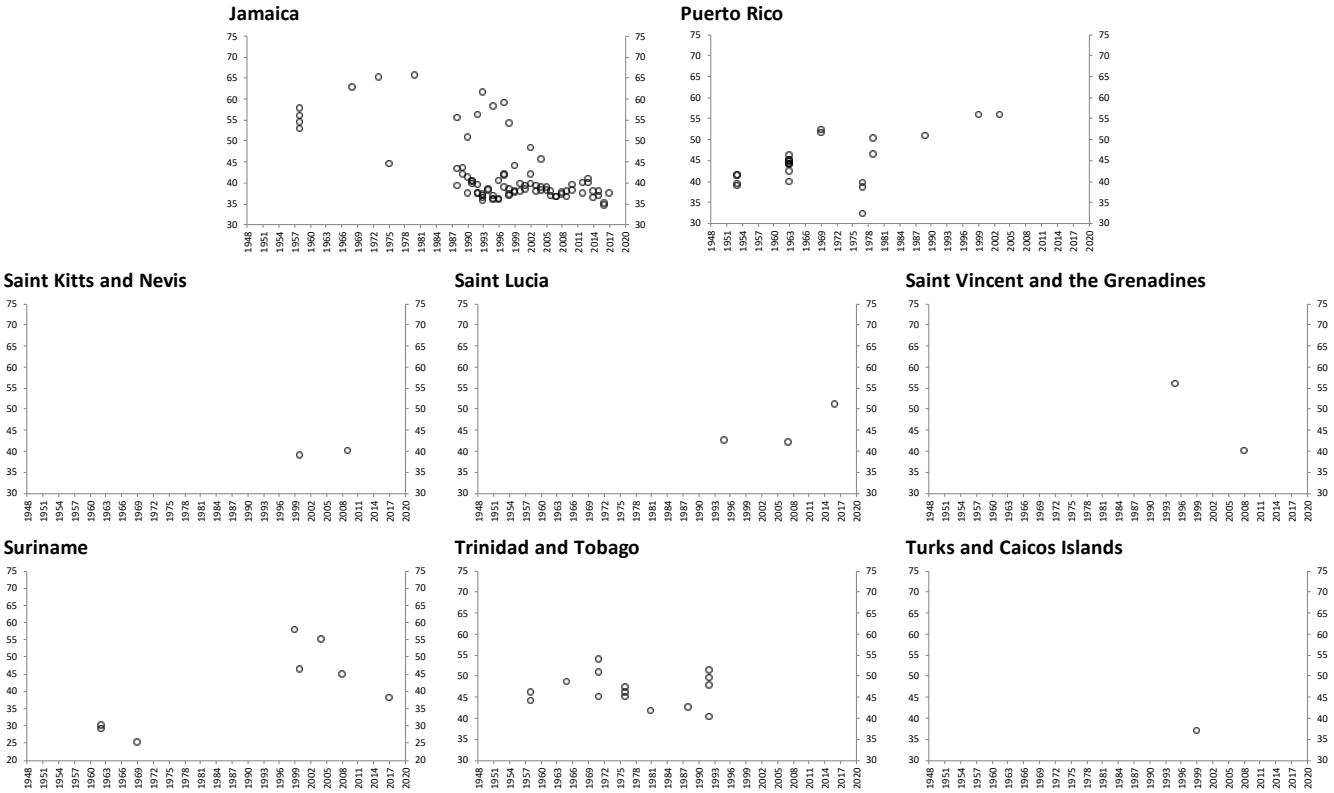
Appendix

Figure A1. Gini coefficients for sixteen smaller countries in Latin America and the Caribbean (complementing Figure 3)



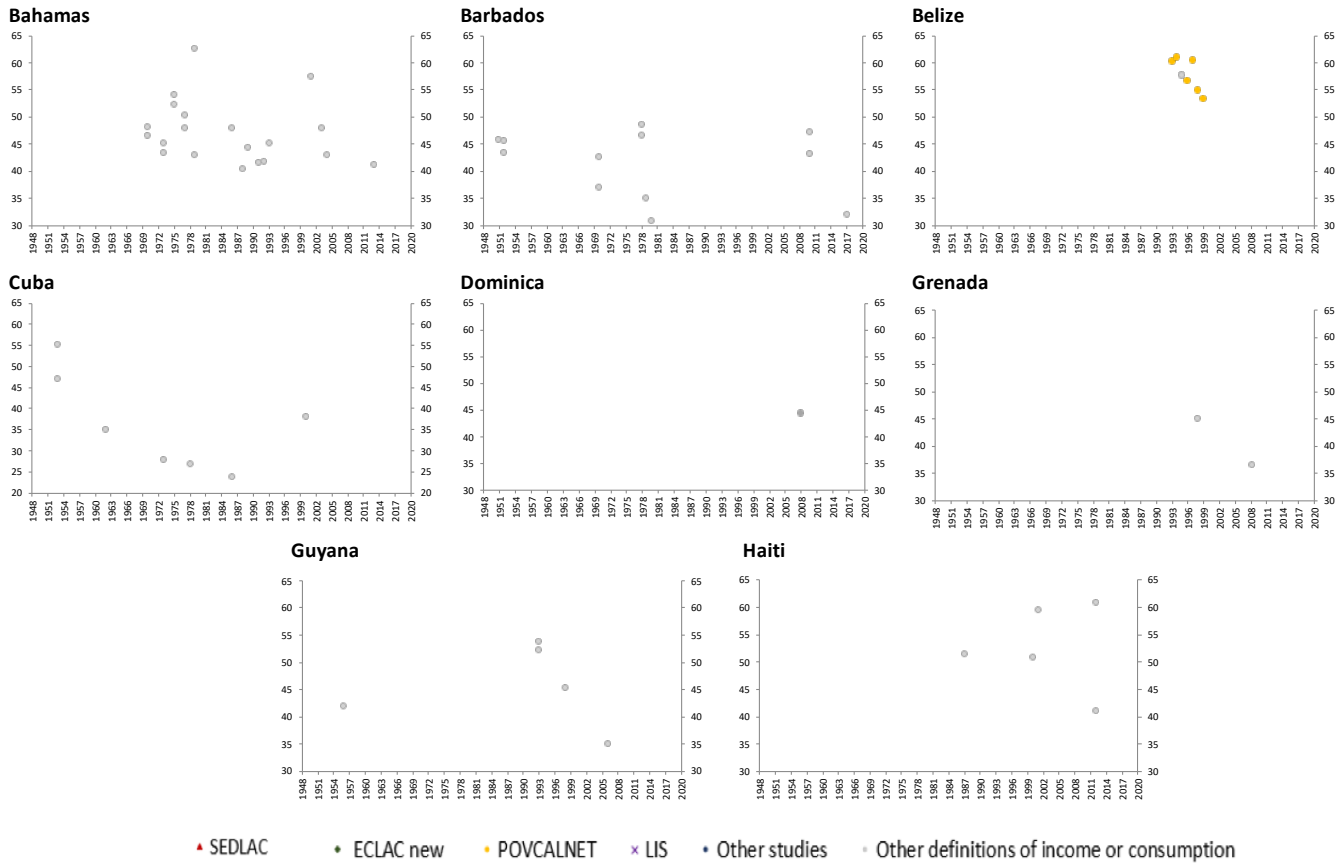
Note: The figure shows all the Gini coefficients from WIID, from De Rosa, Flores, and Morgan (2022), and from additional historical series and studies collected by the authors.

Figure A1. Gini coefficients for sixteen smaller countries in Latin America and the Caribbean (complementing Figure 3) (continued)



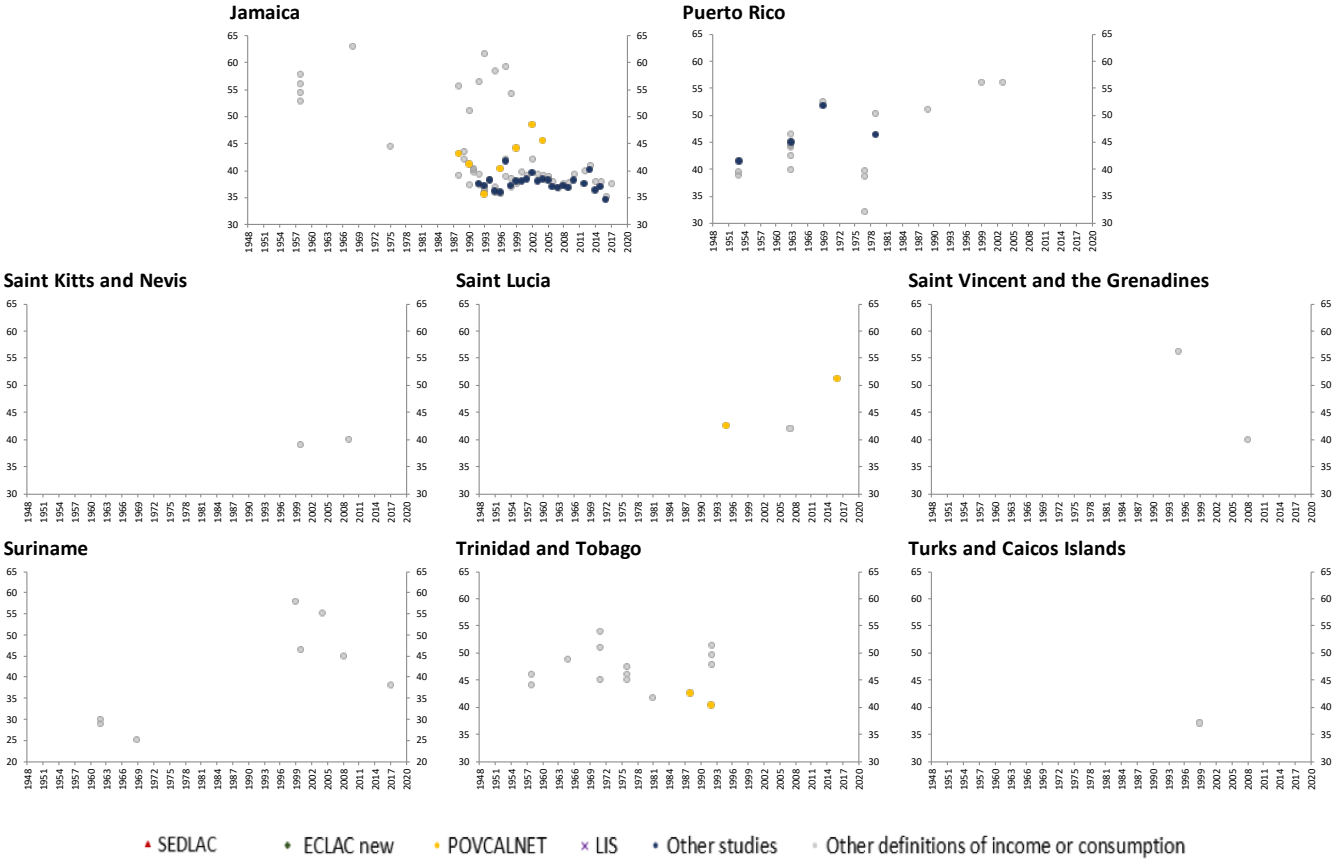
Note: The figure shows all the Gini coefficients from WIID, from De Rosa, Flores, and Morgan (2022), and from additional historical series and studies collected by the authors.

Figure A2. Predominantly HHS-based Gini coefficients for sixteen smaller countries in Latin America and the Caribbean (complementing Figure 4)



Note: The figure shows all the Gini coefficients from WIID, and from additional historical series and studies collected by the authors.

Figure A2. Predominantly HHS-based Gini coefficients for sixteen smaller countries in Latin America and the Caribbean (complementing Figure 4) (continued)



Note: The figure shows all the Gini coefficients from WIID, and from additional historical series and studies collected by the authors.

REFERENCES

Acevedo, I., Castellani, F., Cota, M.J., Lotti, G., and Székely, M. (2024). Higher Inequality in Latin America: A Collateral Effect of the Pandemic. *International Review of Applied Economics*. 38(3): 280-304.

Altimir, O. (1982). La distribución del ingreso en México, 1950-1977. In *Distribución del ingreso en México: Ensayos. Cuaderno 2, Tomo 1*. México, DF: Banco de México.

Altimir, O. (1986) Estimaciones de la distribución del ingreso en la Argentina, 1953-1980. *Desarrollo Económico*. 25(100): 521–566.

Altimir, O. (1987). Income distribution statistics in Latin America and their reliability. *Review of Income and Wealth*. 33(2): 111–155.

Alvaredo, F. (2010) The rich in Argentina over the twentieth century, 1932-2004. In A.B. Atkinson and T. Piketty (eds.), *Top Incomes. A Global Perspective*. Oxford: Oxford University Press. Ch. 6, pp. 253–298.

Alvaredo, F. and Gasparini, L. (2015). Recent trends in inequality and poverty in developing countries. In A.B. Atkinson and F. Bourguignon (eds.) *Handbook of Income Distribution*. North Holland. Ch. 9, pp. 697–805.

Alvaredo, F. and Londoño Vélez, J. (2013). High Incomes and Personal Taxation in a Developing Economy: Colombia 1993-2010. *CEQ Working Paper* 12.

Alvaredo, F. and Londoño Vélez, J. (2014). Altos ingresos e impuesto de renta en Colombia 1993-2010. *Revista de Economía Institucional*. 16(31): 157–194.

Alvaredo, F., Atkinson, A. B., Blanchet, T., Chancel, L., Estevez Bauluz, L., Flores, I., Fisher-Post, M., Garbinti, B., Goupille-Lebret, J., Martínez-Toledano, C., Morgan, M., Neef, T., Piketty, T., Robilliard, A., Saez, E., Yang, L., and Zucman, G. (2020). Distributional National Accounts Guidelines: Methods and Concepts Used in the World Inequality Database. World Inequality Lab.

Alvaredo, F., Cruces, G. and Gasparini, L. (2018). A Short Episodic History of Income Distribution in Argentina. *Latin American Economic Review*. 27(7)

Alvaredo, F., De Rosa, M., Flores Beale, I. and Morgan, M. (2022). Desigualdad del ingreso en la República Dominicana 2012-2020. Una revisión a partir de la combinación de fuentes de información. Publicación de las Naciones Unidas/CEPAL.

Alvaredo, F., De Rosa, M., Flores, I., and Morgan, M. (2022). The inequality (or the growth) we measure. Data gaps and the distribution of incomes. CEPR DP 17135.

Atkinson, A. B. (2019). Measuring Poverty Around the World. Princeton University Press: Princeton, NJ.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*. 2(3): 244-263.

Atkinson, A. B. and Brandolini, A. (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): 771–799.

Atkinson, A. B. and Piketty, T. (2007) *Top Incomes over the Twentieth Century. A Contrast between Continental European and English-Speaking Countries.* Oxford: Oxford University Press.

Atkinson, A. B. and Piketty, T. (2010) *Top Incomes. A Global Perspective*. Oxford: Oxford University Press.

Auten, G. and Splinter D. (forthcoming). Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends. *Journal of Political Economy*.

Azevedo, J. P., Inchauste, G., Sanfelice, V. (2013). Decomposing the recent inequality decline in Latin America. World Bank Policy Research Working Paper 6715.

Blanchet, T., Flores, I. and Morgan, M. (2022) The weight of the rich: improving surveys using tax data. *Journal of Economic Inequality*. 20: 119-150.

Borraz, F. and Gonzalez-Pampillon, N. (2017). Assessing the distributive effects of minimum wage. *Review of Development Economics*. 21(4): 1081-1112.

Bourguignon, F. (2015). Appraising income inequality databases in Latin America, *Journal of Economic Inequality*. 13(4): 557–578.

Bracco, J., Gasparini, L., Galeano, L. and Pistorio, M. (2021). Slowdown of monetary inequality reduction in developing countries. *Journal of Income Distribution*. 29(1–2).

Burdín, G., De Rosa, M., Vigorito, A. and Vilá, J. (2022). Falling inequality and the growing capital income share: Reconciling divergent trends in survey and tax data. *World Development*. 152.

Bustos, A. and Leyva, G. (2017). Towards a more realistic estimate of the income distribution in Mexico. In *United Nations Economic Commission for Europe, Conference of European Statisticians, Expert meeting on measuring poverty and equality, 26-27 September 2017*. Budva, Montenegro.

Campos Vazquez, R. and N. Lustig. (2020). Labour income inequality in Mexico: Puzzles solved and unsolved. *Journal of Economic and Social Measurement*, 44(4): 203-219. May.

Cano, L. (2015). Income mobility in Ecuador: new evidence from individual income tax returns. UNU-WIDER Working Paper 2015/040.

Cano, L. (2017). Personal income tax and income inequality in Ecuador between 2007 and 2011. *Cepal Review*. 123: 56-75.

Cechini, S. and Madariaga, A. (2011). Conditional Cash Transfer Programs: the recent experience in Latin America and the Caribbean, ECLAC, Santiago, Chile

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1986a). Antecedentes estadísticos de la distribución del ingreso. Brasil 1960-1983. Serie Distribución del Ingreso 2.

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1986b). Antecedentes estadísticos de la distribución del ingreso. Colombia 1960-1983. Serie Distribución del Ingreso 1.

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1987a). Antecedentes estadísticos de la distribución del ingreso. Chile 1940-1982. Serie Distribución del Ingreso 3.

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1987b). Antecedentes estadísticos de la distribución del ingreso. Argentina 1953-1982. Serie Distribución del Ingreso 5.

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1988a). Antecedentes estadísticos de la distribución del ingreso. Mexico 1950-1977. Serie Distribución del Ingreso 7.

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1988b). Antecedentes estadísticos de la distribución del ingreso. Venezuela 1957-1985. Serie Distribución del Ingreso 6.

Comisión Económica para América Latina y el Caribe CEPAL/ECLAC (1989). Antecedentes estadísticos de la distribución del ingreso. Peru 1961-1982. Serie Distribución del Ingreso 8

CONADE-CEPAL (1965). Distribución del ingreso y cuentas nacionales en la Argentina. Tomos I-IV. Buenos Aires.

Cornia, G. A. (2010). Income distribution under Latin America's New Left regime. *Journal of Human Development and Capabilities*. 11(1): 85-114.

Cornia, G. A. (editor.) (2014). *Falling Inequality in Latin America: Policy Changes and Lessons*. Oxford University Press.

Coulter, F. A. E., Cowell, F. A. and Jenkins, S. P. (1992). Equivalence scale relativities and the extent of inequality and poverty. *Economic Journal*. 102, 1067-1082.

Cowell, F. (2000). Measurement of inequality. In Atkinson and Bourguignon (eds.), Handbook of Income Distribution, Elsevier, chapter 2, pp. 87-166.

Cowell, F. (1980). On the structure of additive inequality measures. *Review of Economic Studies*. 47(3): 521-531.

De Janvry, A. and E. Sadoulet (2005). Growth, poverty and inequality in Latin America: A causal analysis, 1970–94. *Review of Income and Wealth*. 46(3): 267-287.

De Rosa, M., Flores Beale, I. and Morgan, M. (2022). More unequal or not as rich? Revisiting the Latin American exception. Stone Center Working Paper 53.

De Rosa, M., Lustig, N., and Martinez Pabonn, V. (forthcoming). Fiscal redistribution when accounting for the rich in Latin America. CEQ Working Paper Series.

Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in a poor world). *Review of Economics and Statistics*. 81(1): 1–19.

Deaton, A. and V. Kozel, V. (2005), Data and dogma: the great Indian poverty debate, *The World Bank Research Observer*, 20(2):, 177-198

Deininger, K. and Squire, L. (1996). A new data set measuring income inequality. *World Bank Economic Review*. 10(3): 565–591.

Díaz Bazán, T. (2016). Measuring inequality from top to bottom. *World Bank Policy Research* Working Paper 7237.

Duclos, J-Y. and Araar, A. (2006). Poverty and Equity. New York: Springer.

Engbom, N. and Moser, C. (2022). Earnings inequality and the minimum wage. Evidence from Brazil. *American Economic Review*. 112(12):3803-47.

Fairfield, T. and Jorrat, M. (2016). Top income shares, business profits, and effective tax rates in contemporary Chile. *Review of Income and Wealth*., 62(S1): 120–144.

Feierherd, G., Larroulet, P., Long, W. and Lustig. N. (2023). The pink tide and income inequality in Latin America. *Latin American Politics and Society*. 65(2): 110-144.

Ferreira de Souza, P. and Medeiros, M. (2017). The concentration of income at the top in Brazil 2006-2014. IPEA-Institute for Applied Economic Research Working Paper 163.

Ferreira, F. Lakner, C. and Silwal, A. (unpublished) Is Income Inequality Rising Everywhere? Unpublished manuscript, The World Bank.

Ferreira, F., Firpo, S., and Messina, J. (2022). Labor market experience and falling earnings inequality in Brazil: 1995-2012. *World Bank Economic Review*. 36(1): 37-67.

Ferreira, F., Leite, P and Litchfield (2008), The Rise and Fall of Brazilian Inequality: 1981-2004. *Macroeconomic Dynamics* 12 (S2): 199-230.

Ferreira, F., Lustig, N. and Teles, D. (2015) Appraising cross-national income inequality databases: An introduction. *Journal of Economic Inequality*. 13: 497–526.

Fishlow, A. (1972). Brazilian Size Distribution of Income. *American Economic Review*, 62(1/2): 391–402.

Fixler, D., Johnson, D., Craig, A., and Furlong, K. (2017). A consistent data series to evaluate growth and inequality in the national accounts. *Review of Income and Wealth*. 63: S437–S459.

Flachaire, E., Lustig, N. and Vigorito, A. (2022) 'Underreporting of top incomes and inequality: A comparison of correction methods using simulations and linked survey tax data.', *Review of Income and Wealth* [Preprint].

Flores, I., Sanhueza, C., Atria, J. and Mayer, R. (2020) Top incomes in Chile: A historical perspective on income inequality, 1964–2017. *Review of Income and Wealth*. 66(4): 850–874.

Forster, T., Kentikelenis, A., Reinsberg, B., Stubbs, T., King.L. (2019). How structural adjustment programs affect inequality: A disaggregated analysis of IMF conditionality, 1980–2014. *Social Science Research*. 80: 83-113.

Gasparini, L. and Cruces, G. (2022). The changing picture of inequality in Latin America. Evidence for three decades. UNDP LAC Working Paper 1.

Gasparini, L., Bracco, J., Galeano. L. and Pistorio, M.. (2018). Desigualdad en países en desarrollo: ¿Ajustando las expectativas? CEDLAS Documento de Trabajo 224.

Gasparini, L., Cruces, G. and Tornarolli, L. (2011). Recent trends in income inequality in Latin America. *Economía LACEA Journal*. 11(2): 147–201.

Gradín, C. (2021). WIID companion (May 2021): integrated and standardized series. UN WIDER.

Griliches, Z. (1986). Economic data issues, in Zvi Griliches and Michael D. Intriligator, eds., Handbook of Econometrics. Amsterdam: Elsevier, vol. 3.

Heskia, I. (1980). Distribución del ingreso en el Gran Santiago 1957-1979. Documento Serie de Investigación 53, Universidad de Chile, Departamento de Economía.

Hoffman, R. (1971). Contribuiçao a analisis de distribuição de renda e de posse de terra no Brasil. Tese de livre docência a Escola de Agricultura da Universidade de S. Paulo. Piracicaba.

Jenkins, S. (2015). World income inequality databases: an assessment of WIID and SWIID. *Journal of Economic Inequality*. 13(4): 629–671.

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

Langoni, C. (1973) Distribuição da Renda e Desenvolvimento Econômico do Brasil. Rio de Janeiro: Expressão e Cultura.

Lopez-Calva, L. F. and Lustig, N. (2010). Declining Inequality in Latin America: A Decade of Progress? Brookings Institution Press.

López, R., Figueroa, E. and Gutiérrez, P. (2013). La parte del león: Nuevas estimaciones de la participación de los súper ricos en el ingreso de Chile. Serie Documentos de Trabajo de la Universidad de Chile. SDT379.

Lustig, N. (2019). The "missing rich" in household surveys: causes and correction approaches. Commitment to Equity (CEQ) Working Paper Series 75.

Lustig, N. (ed.) (1995). Coping with austerity: poverty and inequality in Latin America. The Brookings Institution, Washington DC.

Lustig, N., Martínez Pabón, V., Neidhöfer, G. and Tommasi, M. (2023). Short and Long-Run Distributional Impacts of COVID-19 in Latin America. *Economía LACEA Journal*. 22(1): 96–116.

Mahler, D., N. Yonzan, and C. Lakner (2022). The impact of COVID-19 on Global Inequality and Poverty. World Bank Policy Research Working Paper 10198.

Maurizio, R. (2014). El impacto distributivo del salario mínimo en la Argentina, el Brasil, Chile y el Uruguay. *Políticas Sociales* 194, CEPAL

Maurizio, R. and Vázquez, G. (2016). Distributive effects of minimum wages in Latin America, *International Labour Review*. 155(1): 97-131.

Meyer, B., Mok, W. and Sullivan, J. (2015). Household surveys in crisis. *Journal of Economic Perspectives*, 29(4): 199–226.

Morgan, M. (2018). Falling inequality vs persistent concentration: reconciling recent evidence for Brazil from surveys, administrative data and national accounts. PhD Thesis. Paris School of Economics/École des Hautes Études en Sciences Sociales.

Morgan, M. and Souza, P. (2019). Distributing growth during late-development: New evidence on long-run inequality in Brazil. In *CLADHE IV*. Universidad de Santiago, Chile.

Navarrete, I. (1960). La distribución del ingreso y el desarrollo económico de México. Instituto de Investigaciones Económicas, Escuela Nacional de Economía, Universidad Autónoma de México. México DF.

Nolan, B., Roser, M. and Thewissen, S. (2019) GDP per capita versus median household income: What gives rise to divergence over time? *Review of Income and Wealth* 65(3): 465–494.

Piketty, T., Saez, E. and Zucman, G. (2018). Distributional National Accounts: Methods and Estimates for the United States. *Quarterly Journal of Economics*, 133 (2): 553-609.

Ravallion, M. (2003). Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree? *Review of Economics and Statistics*. 85(3): 645–652.

Rodríguez-Castelán, C., L. F. López-Calva, N. Lustig, and D. Valderrama. (2022). Wage inequality in the developing world: Evidence from Latin America. *Review of Development Economics*. 26(4): 1944-1970.

Rossignolo, D., Oliva, N. and Villacreses, N. (2016). Cálculo de la concentración de los altos ingresos sobre la base de los datos impositivos: un análisis para el Ecuador. Serie Macroeconomía del Desarrollo, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).

Saez, E. and Zucman, G. (2016). Wealth inequality in the United States since 1913. Evidence from capitalized income tax data. *Quarterly Journal of Economics*. 131(2): 519-578.

Székely, M. and Hilgert, M. (2007). What's Behind the Inequality We Measure? An Investigation Using Latin American Data. *Oxford Development Studies*. 35(2): 197-217.

Székely, M. and P. Mendoza. (2016). Declining Inequality in Latin America: Structural Shift or Temporary Phenomenon? *Oxford Development Studies*. 45(2): 204-221.

World Bank (2017). Monitoring Global Poverty. Report of the Commission on Global Poverty. World Bank Group: Washington DC.

Zuñiga-Cordero, A. (2018). Income inequality in Costa Rica: An inquiry towards distributional national accounts. Master's dissertation, Paris School of Economics/École des Hautes Études en Sciences Sociales.

Zwijnenburg, J. (2019). Unequal distributions: EG DNA versus DINA Approach. *American Economic Review Papers and Proceedings*. 109: 296–301.