IARIW 2024

IARIW 2024

Thursday, August 22 – Friday, August 30

Extreme Income Inequality in Ecuador 2007 to 2021 – Dollarization, Commodity Price Boom, and Citizen Revolution

Markus Karl Nabernegg

(University of Siegen)

Paper prepared for the 38th IARIW General Conference
August 26-30, 2024

Plenary Session 5-2: Further Improving Household Distributional Results with Particular Reference to Adjusting the Top of the Distribution

Time: Thursday, August 29, 2024 [11:00-12:30 GMT]

Extreme income inequality in Ecuador - Dollarization, Commodity Price Boom and Citizen Revolution in Ecuador

Markus Nabernegg*[1]

[1] *School of Economic Disciplines, University of Siegen, Germany*

July, 2024

Abstract: This study introduces the first comprehensive Distributional National Account (DINA) series for Ecuador, covering the period from 1990 to 2022, and provides detailed pretax and posttax income indicators. Utilizing an innovative integration method that merges extensive household and tax microdata, this research offers a nuanced analysis of income inequality across three decades characterized by economic booms and crises. The results indicate that Ecuador's income distribution was notably unequal in 1990, with extreme concentration at the top persisting into the present. In 2022, the top 10% of earners captured 58% of pretax income, and Ecuador stands out among Latin American countries for having the highest income share held by the top 0.1%. Contrary to previous studies suggesting increasing inequality during the banking crisis and subsequent dollarization in 2000, this study finds that both the bottom 50% and the middle 40% of the income distribution experienced gains at the expense of the top 10%. After dollarization and during the commodity price boom (2004-2014), income growth was markedly favorable for lower income groups, while the income of the middle 40% stagnated. Recent economic downturns, including the political crisis and the COVID-19 pandemic, have uniformly impacted all income groups without significantly altering the overall distribution of income. Although not designed for causal inference, this study offers significant insights into the dynamics of income inequality during various policy regimes. It demonstrates that income concentration at the top remained stable despite economic shocks and reveals that the banking crisis did not adversely affect the income of the bottom 50%. Furthermore, it highlights the effectiveness of progressive policies implemented during the "Citizen Revolution" government in reducing income inequalities in the short term.

*markus.nabernegg@uni-siegen.de, Kohlbettstraße 17, 57072 Siegen, +49 271 74012899.

Keywords: Distributional National Accounts, middle class, Ecuador, dollarization, commodity price boom, redistribution

JEL Keywords: D3 Distribution, O4 Economic Growth and Aggregate Productivity

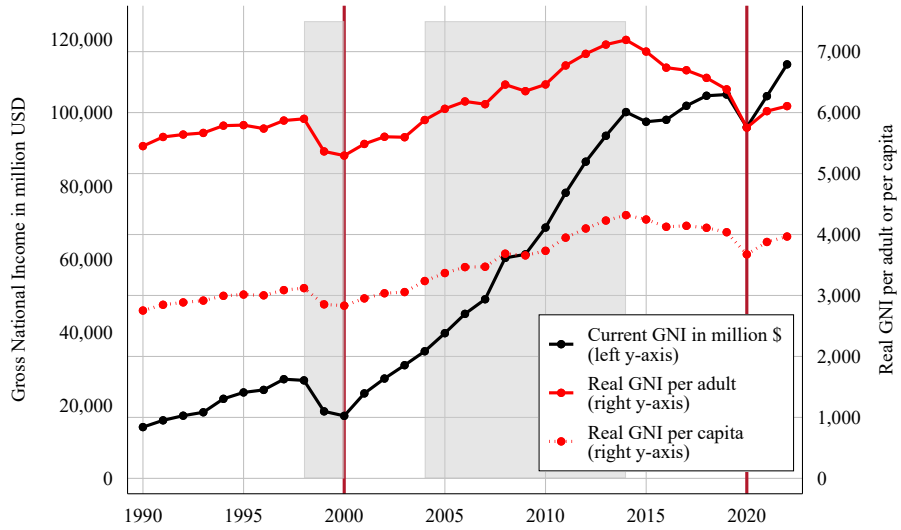
1. Introduction

The past three decades in Ecuador have been characterized by substantial economic volatility, marked by cycles of booms and busts. The weak growth experienced in the 1990s was abruptly halted by a severe banking crisis, which led to the abolition of the national currency. In 2000, the US dollar was adopted as the exclusive legal tender - an unprepared measure introduced within a few weeks, but with profound implications for fiscal and trade policies. This transition coincided with a global commodity price boom in the early 2000s. During the fifteen years between 2000 and the end of the commodity price boom in 2015, the Ecuadorian economy expanded five-fold in nominal terms. This period saw improvements in all social indicators, including public spending on education and health, the development of productive infrastructure, and poverty reduction. The end of high commodity prices marked the onset of economic and political instability, culminating in a recession that was exacerbated by the COVID-19 pandemic in 2020. Figure 1 illustrates the evolution of national income in nominal and real terms over the period between 1990 and 2022. Despite relatively low inflation, average yearly real national income growth *per capita* between 1990 and 2022 was only 1.4%, and average growth *per adult* - an indicator that distributes income to the population earning national income - did not exceed 0.4%.

However, these very low real growth rates over three decades do not necessarily imply that all income groups had stagnant income levels. High political instability up to 2007, marked by two coup d'états and nine presidents between 1990 and 2006 (including a three-day presidency in 1997), was followed in 2007 by significant policy interventions during the "Citizen Revolution" government - such as progressive tax reforms, minimum wage increases, investment in infrastructure, among others. These reforms were accompanied with a considerable increase in resources stemming from primary product exports during the commodity price boom (2004-2014). However, some of these reforms were gradually reversed after the end of the Citizen Revolution government in 2016. Political and economic instability, in combination with the COVID pandemic and an increasing penetration of narco-terrorist groups in the public sector and in the society in general, have maintained the Ecuadorian economy in recession since then.

Despite ongoing statistical data collection since the late 1980s, Ecuador lacks systematic information on long-term income distribution trends. Several limitations hinder a comprehensive understanding of the country's inequality evolution: the absence of detailed national accounts before 2007, changes in household survey methodologies, and difficulties in capturing income data from wealthy individuals (the so-called "missing rich", Lustig, 2020) in household surveys. Scholars have partially tackled these problems for the Ecuadorian distribution estimates, but either by only using tax register for a short periods (Cano, 2015; Rossignolo et al., 2016), or by including Ecuador in a regional study for Latin America, without the use of detailed tax

Figure 1: Evolution of National Income 1990-2022



Red vertical lines: 2000 Dollarization, 2020 COVID pandemic, Shaded areas: 1998-2000 Bank crisis, 2004-2014 Commodity price boom. The share of adult population (20 years and older) increased from 50% in 1990 to 64% in 2022, reducing the gap between GNI per adult and GNI per capita over time.
Source: Ecuadorian Central Bank.

data(De Rosa et al., 2022).

The study at hand aims to provide a comprehensive series of the distribution of national income over the period 1990-2022, by combining a wide range of survey and tax microdata. I have for the first time exclusive access to historical individual tax declarations from the Internal Revenue Service in Ecuador and integrate this information to the national employment survey by expanding the existent integration method proposed by the Distributional National Accounts methodology (DINA, Blanchet et al., 2020). The granular dataset is linked to national accounts to cover 100% of macroeconomic indicators and allows to analyze the evolution of income for different groups.

Using the official inequality estimates published by the INEC and studies based on household surveys (Gachet et al., 2019; Ponce and Vos, 2014), synthetic inequality measures such as the Gini index have shown a steady decline from 2000 to 2015, followed by a slight increase — a trend consistent with the broader Latin American region during and after the commodity price boom (Flechtner and Middelanis, 2024). However, my findings indicate high and persistent income inequalities, with significant concentration at the top of the income distribution. In 2022, the bottom 50% of the population captured only 14% of posttax national income, while the top 1% captured 20%. Notably, Ecuador ranks first in Latin America in terms of income concentration among the top 0.1%. The analysis further reveals that, although the period following dollarization and during the commodity boom was characterized by pro-poor growth, the middle-income group did not significantly benefit from the economic expansion.

This paper makes two main contributions: First, it produces for the first time a distributional

series covering over three decades and improves the inequality estimates by applying Distributional National Accounts and by introducing an extension to the integration method for survey and tax microdata. Second, While not explicitly designed to draw causal inferences, the paper contributes to the discourse on the political economy of currency regimes, the reliance on primary product exports, and the distributional impacts of state interventions in developing countries.

The remainder of this paper is organized as follows: section 2 reviews the literature on Distributional National Accounts and on the political economy associated with dollarization, commodity boom and progressive state intervention. In section 3, I describe the data and the steps applied to produce the inequality estimates for Ecuador, which are presented in section 4. Section 5 concludes.

2. Literature Review

This paper contributes to two strands of literature: it introduces a new series of income inequality measures for Ecuador, based on the Distributional National Accounts framework, and it examines the evolution of national income distribution in a developing country dependent on primary goods exports, particularly during periods of external price shocks and changes in currency regimes.

2.1. Distributional National Accounts Literature

Over the last decade, scholars have made significant strides in estimating income distribution in ways that align with macroeconomic aggregates and are comparable over time and across countries. The World Inequality Lab (WID) has facilitated many country studies that offer valuable insights into the distribution, composition, and growth of income through the application of Distributional National Accounts (DINA). This study follows the general DINA guidelines (Blanchet et al., 2020) and extends the integration method to incorporate both survey and tax microdata.

One of the contributions of DINA is the inclusion of the top income groups into inequality estimates to correct the underestimation of traditional inequality measures stemming from household surveys. Existing studies for developing countries often face difficulties to access tax data that would allow to account for the top incomes, and analyze relatively short periods.[1] One solution to this lack of information and access is to apply "simplified" DINA (Piketty et al., 2019) based on survey or tax tabulations and regional or global averages (De Rosa et al., 2022; Assouad et al., 2018; Novokmet et al., 2018). Such studies can shed light on distributional questions even in the absence of high quality data. In addition, standardized programs to convert tax tabulations into distributions (Blanchet et al., 2022c) or to reweight

[1]Exceptions to the short period restriction are studies for emerging or peripheral European countries including Flores et al. (2020), Koutentakis and Chrissis (2020), Novokmet et al. (2018), and Chancel and Piketty (2019), but they also face limited access to microdata.

household survey observations to resemble incomes at the top (Blanchet et al., 2022b) have contributed to more country studies with limited tax data (Chatterjee et al., 2021; De Rosa et al., 2022).

Tax data from Ecuador has been used before to produce top income shares: Cano (2015) for 2008-2011 and Rossignolo et al. (2016) for 2012-2014 estimated top income shares exclusively relying on tax data for short periods. De Rosa et al. (2022) use the results form the former studies as source to correct household surveys by reweighting the importance of the richest individuals in the survey, and extrapolating to the years before 2008 and after 2014. The study from De Rosa et al. (2022) focuses on comparability between different Latin American countries and therefore mobilizes the harmonized household survey database from Source the Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank) for a great number of Latin American countries.

This study is one of the rare cases, where a researcher has access to confidential tax microdata in a developing country, spanning over two decades.[2]. In section 3, I provide a detailed methodology to integrate survey and tax observations through an extension to the existing Stata command `bfmcorr` (Blanchet et al., 2022b), and produce an income distribution series for Ecuador for three decades.

2.2. Political Economy Literature

This paper also contributes to the political economy literature by examining the evolution of income distribution in various economic and political contexts within a primary goods exporting country. Over the past three decades, Ecuador has experienced different fiscal and monetary institutional arrangements, significantly impacting its economy. The banking crisis at the turn of the century, which led to the dollarization of the economy, has been extensively discussed in the early 2000s (Alesina and Barro, 2001; Beckerman and Solimano, 2002; Jameson, 2003). From a theoretical point of view, this harsh measure dollarization can be viewed as the most effective "commitment device" to stable monetary policies (Alesina and Barro, 2001, p. 382), providing a "good housekeeping seal of approval" for international money inflow into the country (Jameson, 2003, p. 653). Despite several assessment studies on dollarization in Ecuador (Soto, 2009; Özyurt and Cueva, 2020), there remains a gap in the literature concerning how income distribution evolved with the adoption of the new currency.

A few years after the crisis, the global economic environment shifted favorably for primary commodity-exporting countries. The commodity price boom from the early 2000s until 2014 expanded fiscal policy space for many Latin American nations. These additional resources, depending on the institutional arrangements, could be allocated in various ways. In Ecuador, Flechtner and Middelanis (2024) found a sustained increase in public social spending linked to the commodity boom, which potentially altered the income distribution significantly. This period coincided with a shift from a chaotic era of short-term presidencies to a relatively stable decade under the progressive "Citizen Revolution" government. During this time of economic

[2]For Latin America, I am only aware of Burdin et al. (2022) which studies Uruguayan income inequality with tax microdata between 2009 and 2016.

growth, public investment and transfers increased alongside public debt, resulting in a more equitable income distribution in the short term, though with uncertain long-term effects. Conversely, Cachanosky et al. (2022, p.432) argue that dollarization served as a constrained optimal choice within Ecuador's institutional framework, providing credibility by limiting the state's ability to implement populist policies..

The end of the commodity price boom in 2014 is also associated with a downturn of the Ecuadorian economy, given its strong dependence on primary product exports. García-Albán et al. (2021) have recently shown that oil revenue shocks drove output above or below trend during the years 2004-2019. Finally, the economic recession was associated to a political crisis and the COVID pandemic in the last years of the study period, which also requires an assessment of the impact in the national income distribution.

The study is not designed to answer the question of causality between different policy measures and income inequality, but the results provide insights on whether certain policy episodes redced or increased income inequality and enhanced growth for different income groups.

3. Data and Methodology

This study follows the methodology of "Distributional National Accounts" (DINA), a method to estimate the distribution of income consistent with macroeconomic accounts and comparable across time and countries. Developed by the World Inequality Lab (WIL), the concepts and methods to produce the World Inequality Database (WID), were formalized in the DINA guidelines (Blanchet et al., 2020). The focus of the DINA method lies on the inclusion of incomes at the top of the distribution, which are often not captured by household surveys due to a systematic undercoverage and underreporting of income from very wealthy individuals. This problem of the "missing rich" Lustig (2020, p. 2) causes official inequality estimates based on household surveys to omit the concentration at the top and therefore to potentially underestimate income inequalities. The solution to this lack of information proposed by the DINA method is the inclusion of data sources that better capture high incomes, such as tax records or rich lists. A second feature of the DINA method is the distribution of the entire national income, including the income generated by the corporate and government sector. This inclusion accounts for income sources traditionally overlooked by inequality studies (like the undistributed profits of corporations or indirect taxes) and allows to construct distributional series that account for 100% of macroeconomic growth coming from national accounts.[3]

For Ecuador, I use three main data sources: national accounts, household surveys, and tax records. The steps to produce the DINA series with these data sources for Ecuador are (a) constructing a national accounts series, (b) harmonization of income concepts, (c) integration of tax and survey data on the micro level, and (d) imputation of national income concepts. In what follows I outline the characteristics of the data sources and the principal features of the

[3]For a comparison with the EG-DNA method to inequality estimation in combination with national account total see Zwijnenburg et al. (2021). The EG-DNA method is focused on OECD countries and their harmonized household surveys.

DINA steps. Appendix A contains detailed information and discussions for each issue.

3.1. Data sources

In this study I have privileged access to data sources in Ecuador to establish a distributional series over a relatively long period for a country in the Global South - 33 years between 1990 and 2022. The main sources comprise national accounts from the Ecuadorian Central Bank (*Banco Central del Ecuador*, BCE), employment survey microdata from the National Statistical Office (*Instituto National de Estadistica y Censos*, INEC), and tax record microdata from the Internal Revenue Service (*Servicio de Rentas Internas*, SRI). National income totals from national accounts are available for all years, but detailed income item and institutional sector disaggregations only exist between 2007 and 2022. The national accounts series therefore requires assumptions for the years before 2007 and adjustments due to methodological ruptures for the last years, described in subsection 3.2.

The employment survey (*Encuesta Nacional de Empleo, Desempleo y Subempleo*, ENEMDU) has been made available by the INEC since 1989 andconducted annually thereafter. Given the change in currency and methodological improvements over such a long period, several harmonization steps were necessary, as detailed in Appendix **??**These steps included harmonizing income concepts[4] and converting income from Sucres to US dollars for the years before dollarization in 2000. Additionally, the survey was adjusted for the lack of rural data in the years 1990-1999 and 2002. For this correction I mobilized the data from a second national survey that covered urban and rural areas, the *Encuesta de Condiciones de Vida* (ECV), conducted by the INEC in 1995, 1998, 1999, 2006, and 2014. The population weights of each income group[5] in the ENEMDU survey were multiplied with the share of rural population in the ECV survey of each comparable group. Finally, the same ECV survey was used to account for imputed rents, which is a notional income source, specifically important for lower income groups.

Tax records are available from the SRI since 2004. Individuals declare their income with one of two tax forms (exclusive income from employment - F107, and comprehensive income declaration F102), and the information is consolidated by the SRI, whereby all income items are comparable throughout the entire series based on a conversion matrix accounting for all changes in tax forms. However, data quality is poor for the first years Cano (2015) and De Rosa et al. (2022), which can be observed by an increase in tax declarations between 2004 and 2007 that cannot be explained by any economic phenomenon. I therefore use the dataset starting in 2008 where the number of declarations stabilized. For the years before 2004, unfortunately no tax information is available, which requires to make assumptions on the distribution of tax income for this period. I use the distribution of income from the tax records in 2008 and extrapolate this distribution to earlier years, anchoring absolute income of each register to the evolution of national income totals. Consequently, the results of the integration of the survey and tax register (see subsection 3.4) for the years 1990-2007 will primarily reflect differences in

[4]The questionnaire in the 1990s was less extensive than those from the 2000s onward. However, the coverage of total household income compared to national accounts did not change drastically.

[5]I use percentiles and more granular groups at the top, which result in 127 groups, called "fractiles"

distribution driven by household surveys and changes in the income of the household sector in national accounts.

3.2. Construction of national accounts series

The Ecuadorian Central Bank publishes to series of national accounts containing dissagregations for institutional sectors (households, government, corporations, rest of the world) and income items (e.g. wage income, capital income, transfers, taxes, etc.): The first series extends to the period 2007-2020, with 2007 as a fixed base year, while the second period include the years 2018-2022, with mobile base years. The pivotal period is the longer 2007-2020 period, and I use the 2007-based GNI (provided by the World Development Indicators based on the data from the BCE), and adjust the dissagregations from the years 2021 and 2022 to this GNI total. Some of the changes in the structure of income due to the methodological change are presented in appendix A.2. To extend the series to years before 2007, I first take the relative size of each sector and income item from 2007 and extrapolate to the years prior to 2007. In a second step, I adjust the income items of direct and indirect taxes, as well as social contributions for the years 1990-2006 with information from the OECD tax data[6]. Third, given the importance of remittances after the banking crisis and dollarization in 2000, I correct the series with data from the remittances statistics from the BCE. With these adjustments, I count with a series of macroeconomic totals by sector and income item between 1990 and 2022.

3.3. Unit of observation and harmonization of income concepts

I use a variation of the individualistic adult as the unit of observation. Adults are defined as individuals aged 20 years and older, in line with most studies in the DINA context, which assume that nearly all income is earned by adults. Individualistic means that there is no distribution of income among adults in the household - in contrast to equal-split methods where income is shared between couples or groups of adults within a household. Given that tax declarations in Ecuador are filed individually (no tax exemptions for married couples, etc.), and that income in surveys is available for households and individual members, the individualistic adult provides a logical starting point. However, some income sources (especially important for the population at the bottom of the distribution), cannot be fully attributed to individuals, such as targeted transfers for children (which are paid to parents), remittances, or imputed rents. I therefore make use of the improvements in the integration method (see section 3.4) and distribute non-separable household incomes equally to all adult members within the household.

The income concepts applied in this study follow again the DINA Guidelines (Blanchet et al., 2020). For the integration of survey and tax data on the micro level I produce the *pretax income of the household sector*, or simply pretax household income (PRTHHI), which contains only income sources that appear in both micro sources. Scaling and imputations for

[6] https://data-explorer.oecd.org/vis?tenant=archive&df[ds]=DisseminateArchiveDMZ&df[id]=DF_REVECU&df[ag]=OECD

income in the household sector, as well as the government and corporations sector applied to this integrated micro dataset then allow to produce the national income before and after taxes and transfers, which are named *pretax national income* (PRTNI) and *posttax national income* (POTNI). Distinguishing between both income concepts allows to assess the impact of redistribution policy from public intervention.

3.4. Integration of tax and survey data

The central methodological contribution of this paper is the integration of survey and tax microdata at the micro level. While previous studies have combined tax and survey information (e.g., Bach et al., 2020; Bartels and Metzing, 2019; Blanchet et al., 2022a; De Rosa et al., 2020), this study uniquely applies this integration at the individual level, addressing a gap in the literature. The methodological limitations in previous research are often due to differing objectives and data availability. Studies focusing on income or wealth concentrations beyond household survey results typically use registers from countries with long-standing tax records (Piketty and Saez, 2003; Garbinti et al., 2018). In these cases, survey data is mainly used to correct for demographic inaccuracies and undercoverage, without needing detailed income information from survey microdata.[7] However, household surveys have only been widely available in recent decades, limiting long-term analyses to countries with comprehensive historical tax records.

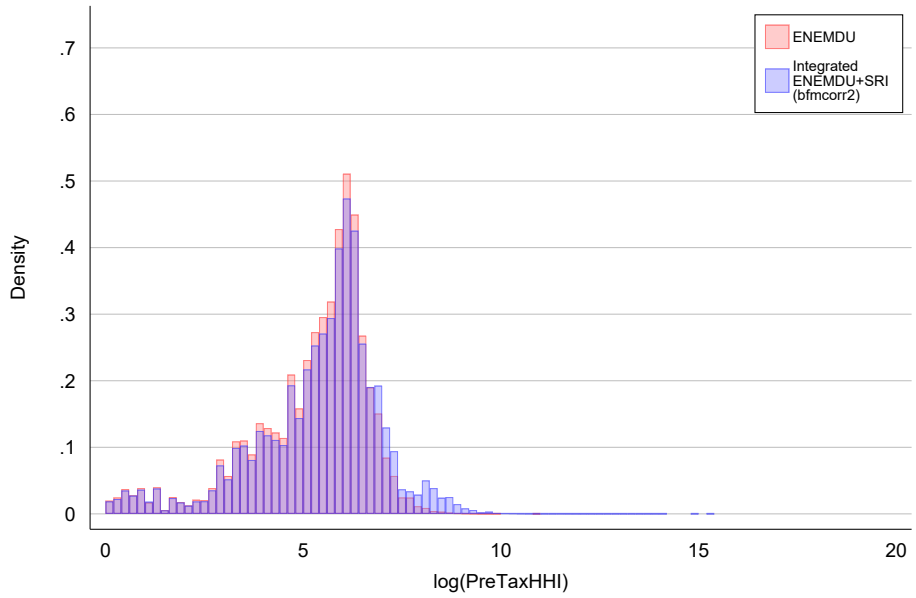
In developing countries, where comprehensive historical tax records may be lacking or restricted, studies often rely on household surveys to address top-income data limitations (De Rosa et al., 2022; Assouad et al., 2018; Alvaredo et al., 2019). Research by Bach et al. (2020), Bartels and Metzing (2019), and Blanchet et al. (2022c) shows that simple tax tabulates can approximate income distributions derived from detailed tax registers when interpolation methods are used. This approach is supported by evidence that top income and wealth distributions often follow a Pareto distribution or power law (Atkinson, 2007; Clementi and Gallegati, 2005; Piketty and Saez, 2003). Recent contributions have standardized correction methods for household surveys with simple tax tabulates. For example, the R package `gpinter` (Blanchet et al., 2022c) and the Stata command `bfmcorr` (Blanchet et al., 2022b) reweight survey microdata to align with tax information, facilitating the production of distributional national accounts and enabling regional comparisons of inequality and concentration levels.[8]

However, interpolation methods have limitations. Reweighting survey data based on interpolated tax data preserves household income composition but fails to represent the income characteristics of the wealthiest individuals, often found in tax data. Specifically, higher property income or undistributed rents, significant among the wealthiest but underrepresented in

[7] For example, estimating the income concentration of the top 1% can be achieved by comparing exclusively register data to national accounts totals.

[8] Examples include studies on the most unequal region of the world (Alvaredo et al., 2019), income concentrations in Africa (a continent with limited distributional information Chancel et al., 2019), and inequality comparisons in Latin America (De Rosa et al., 2022).

Figure 2: Integration of ENEMDU and SRI on the microlevel (2022)



Sources: ENEMDU, SRI, BCE.

surveys, are not accurately captured in the reweighted dataset.

The study leverages privileged access to tax and survey microdata for Ecuador to extend the `bfmcorr` command. This enhancement overcomes the limitations of traditional reweighting approaches, facilitating a more comprehensive analysis of income dynamics, composition, and redistribution.

The process of integrating real tax and survey microdata involves reweighting survey observations with `bfmcorr`, followed by expanding the weighted observations from rich households. For Ecuador, this means one observation from the highest income decile in the survey represents about 200 individuals. These individuals are then replaced by actual observations from the tax register. The integration is done at the individual level. As illustrated in figure 2, the resulting integrated dataset has lower densities for income groups at the bottom and and higher densities for groups at the top, now including individuals with incomes beyond the survey's range.[9]

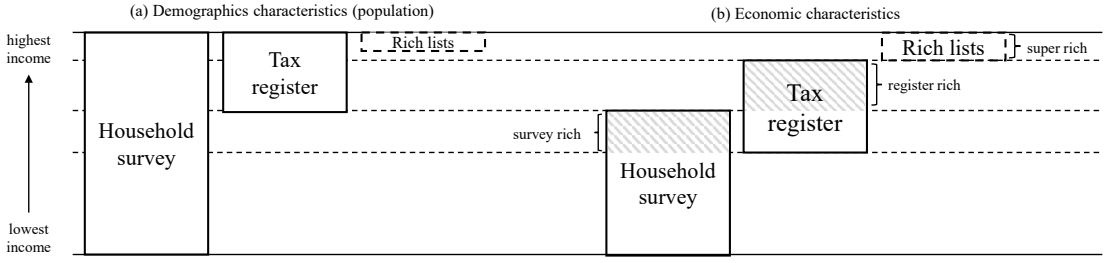
One critical assumption for integrating tax and survey data is maintaining the ranking order between individuals from both sources. This means that the "poorest" individual from the tax register replaces the "poorest" individual from the survey, and so on, up to the "richest". This rank stability is crucial when integrating individuals' income characteristics into the household structure from the survey, as it assumes that demographic properties (e.g., household size, age, and sex structure) are preserved and that the household composition of the wealthiest individuals in the survey is similar to that of those in the tax register. The credibility of this assumption depends on whether the the household survey is be representative for *demographic* characteristics of the entire population, including the "missing rich", and whether the richest

[9]Figure 2 includes extremely high incomes, exemplified by log values exceeding 15 (over 3.3 million dollars).

individuals from the tax register have similar demographic characteristics as the rich individuals from the survey. Figure 3 illustrates the representativeness of demographic characteristics in panel (a), and of economic characteristics in panel (b). We have argued that a considerable part of the upper *income* distribution is not represented in the survey, indicated by the modest height of the household survey box in panel (b). Next to the household survey, the tax register box indicates an overlapping part between both sources for the survey rich, and a part which is only covered by the tax register (register rich).

However, in terms of demographic characteristics, the box in panel (a) is represents higher representativeness for income strata that are not captured by the survey. The tax register box therefore overlaps for these characteristics with the survey. The assumptions here can be sustained by the fact that demographic characteristics have relatively narrow boundaries. While average household size decreases from more than 5 members in decile one to slightly above 3 members in decile seven for the year 2022, this number fluctuates between 2.5 and 3 members for the highest two deciles, given the strict lower bound of one member per household. The share of households with female household heads also slightly increases with income from 22% to 25%, but remains at this level for the highest deciles. Finally, the survey covers the entire geographic area of Ecuador, and I assume that the tax rich live in areas where the survey rich were enumerated (mainly the major cities).

Figure 3: Demographic and economic coverage of data sources



Note: Panel (a) in this figure illustrates the coverage of demographic characteristics of the entire population in different data sources. The household survey is designed to capture these characteristics to a great extent. In panel (b) the figure indicates that the economic characteristics of the entire population is partially covered in the survey, given limitations of undercoverage and underreporting. The inclusion of the tax register increases this coverage. In both cases a lack of information remains for the super-rich.

Another consideration is determining the point at which replacement of survey data with tax data should begin. The *merging point*—where the density of the tax data distribution surpasses that of the survey data—is identified through a data-driven approach in `bfmcorr`. However, the starting point for replacement must be manually defined. The extension of this paper for `bfmcorr` searches for an optimal replacement point by minimizing the difference between the total value from national accounts and the integrated method's result.

The procedure of this study additionally enhances the coverage and reliability of income distribution data after the micro integration by incorporating information from the survey into the highest income strata, particularly for income items that were dropped due to the replacement of individuals in the survey by tax individuals, but are better captured in the

survey, like imputed rents or non-pension social benefits. Additionally, informal income, often only recorded in surveys, may be overlooked, particularly income from own account production, which is part of mixed income (B3R). Therefore, the method imputes information from operating surplus and mixed income (B2R and B3R), property income (D4R), social benefits (D62R), and other current transfers (D7R) when the value from the survey exceed those observed in individual tax declarations (for details see appendix A.4).

Finally, all income items are scaled up to 100% of the corresponding income totals of the household sector in national income. Assumptions for the imputation of social contributions in the household sector and negative income items are explained in appendix A.4.

3.5. Imputation of national income components

The integration of sources so far produces pretax income that covers 100% of the household sector in national accounts. To distribute incomes from the government and the corporate sector, this study imputes income components from other sectors to arrive at national income.

First, *pretax factor income* (PRTFI) adds undistributed corporate income an income from the government to pretax household income (PRTHHI). To produce *pretax national income* (PRTNI), social benefits and contributions are added for each individual. In addition, I include remittances (other transfers in national accounts) to PRTNI, considering the importance of this item in the Ecuadorian economy. Finally, *posttax national income* (POTNI) deducts corporate taxes from individual income (assuming that capital owner pay this tax), and indirect taxes (mainly value added tax and importation tax, paid by individuals following their pretax household income). In a second step, government spending is distributed to individual income, whereas health is distributed via lump sum, and other individual and collective consumption items are distributed like pretax household income. For details on the different assumptions see appendix A.5.

4. Results

This section provides the first continuous historical series of the Ecuadorian income distribution for the 33 years between 1990 and 2022. While the study is not designed to provide causal inference for changes in the distribution, the development of income shares of different groups is strongly associated with historical events, elaborated in subsection 4.1. The evolution of real income is described in subsection 4.2, and provides insights on which groups benefited and lost during the three decades. Finally, subsection 4.3 puts the results from Ecuador in context with other countries in the region.

4.1. The distribution of income 1990-2022

The first insights about Ecuador's national income distribution are illustrated in tables 1 (for 1990) and 2 (for 2022). Both tables indicate the the distribution of pretax and posttax income of the adult population. In 1990, average national income for the slightly more than 5 million adults in Ecuador in 1990 was about US$ 2,800. The bottom 50% earned only US$ 147 on

Table 1: Distribution of Income in Ecuador 1990

Income group	Number of adults	Pretax national income			Posttax national income		
		Income threshold	Average income	Income share	Income threshold	Average income	Income share
Full adult population	5,220,013		2,807	100.0%		2,807	100.0%
Bottom 50%	2,609,487		147	2.6%		424	7.6%
Middle 40%	2,088,519	1,037	2,488	35.5%	1,180	2,514	36.0%
Top 10%	522,007	5,167	17,380	61.9%	4,907	15,782	56.5%
incl. Top 1%	52,303	31,807	86,134	30.7%	28,831	76,886	27.6%
incl. Top 0.1%	5,225	137,901	295,501	10.5%	120,056	265,103	9.5%
incl. Top 0.01%	521	375,919	858,496	3.1%	340,247	761,534	2.7%
incl. Top 0.001%	52	914,958	4,568,392	1.6%	807,367	4,039,990	1.4%

Own elaboration based on ENEMDU, SRI, BCE.

average before taxes and transfers, which is barely 5% of average income and only 2.6% of total pretax income. The next group is the middle 40%, who earned about US$2,500 or 35.5% of total pretax income. This "middle class" of the population therefore captured a share of income that is close to their share within population (40%). Moving to the richest income groups, the top 10% earn US$17,400, which is 5 times the average income, and this group accounted for almost 62% of total national income. The concentration becomes even more remarkable when the top 10% is split up in smaller groups: Only the top 1% of the population (52,000 adults) had an income 30 times the average income of the country and held 31% of pretax income. The top 0.1% (5,000 individuals) still earned more than 10% of pretax income; the top 0.01% (500 adults) captured an average income of US$858,000 and a share that was higher than the one from the entire bottom 50%; and the richest 52 adults made more than US$4,000,000 yearly and captured 1,6% of pretax income. Their income was therefore 1,600 times the average income and 31,000 times the income of the bottom 50%. When moving to posttax national income, one can observe that state intervention via taxes and transfers produces a slightly more equal distribution: The posttax income of the bottom 50% almost triplicated to US$424 or 7.6% of total national income. However, this is still only 15% of average income. The change for the middle 40% is negligible: Posttax income remains at about US$2,500, which indicates that taxes and transfers cancel themselves out. The top 10% only marginally lose income after state intervention (about 10% of their pretax income) and still account for more than 56% of total posttax income.

The picture changes only moderately for the most recent year: In 2022, average income for the now 11,5 million adult Ecuadorians has increased to more than US$10,000 in nominal terms. The pretax income for the bottom 50% is almost US$1,800 or 20% of average income, which represents a share of 8,4% of total national income. The middle 40% now captures a third of total pretax income, which is a smaller share than in 1990, and the top 10% remains at a very

Table 2: Distribution of Income in Ecuador 2022

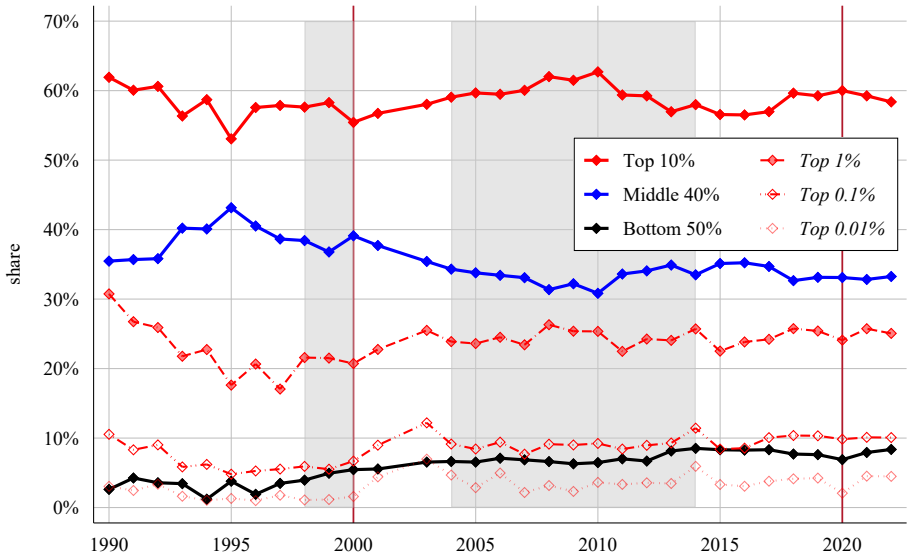
Income group	Number of adults	Pretax national income			Posttax national income		
		Income threshold	Average income	Income share	Income threshold	Average income	Income share
Full adult population	11,452,587		10,230	100.0%		10,230	100.0%
Bottom 50%	5,723,470		1,712	8.4%		2,924	14.3%
Middle 40%	4,583,476	4,887	8,500	33.3%	5,834	9,058	35.4%
Top 10%	1,145,641	18,290	59,701	58.4%	17,892	51,448	50.3%
incl. Top 1%	115,379	104,986	254,488	25.1%	88,629	206,951	20.2%
incl. Top 0.1%	11,457	412,617	1,030,029	10.1%	332,819	821,336	8.0%
incl. Top 0.01%	1,145	1,491,958	4,586,576	4.5%	1,153,252	3,655,543	3.6%
incl. Top 0.001%	115	6,821,985	24,011,959	2.4%	5,456,637	19,273,517	1.9%

Own elaboration based on ENEMDU, SRI, BCE.

high participation with 58.4%. Table 2 also demonstrates, that the concentration at the top increased in these 33 years, where the top 0.01% (1,100 adults) now account for 4.5% and the top 0.001% (115 adults) for 2.4% of pretax income. The progressive tax and transfer system increases the income of the bottom 50% to US$2,900 - a value that is in *nominal terms* now close to the average of the year 1990. The middle income group slightly benefits from taxes and transfers overall by increasing its share in income to 35.4%, and the top 10% reduces its income share from 58% to 50% after redistribution policy. This reduction for all top income groups is more progressive than in 1990, but the posttax shares are still higher than in 1990, where the richest 0.001% of the population captured 1.4% and now captures 1.9%. In section 4.3, I will show that this concentration of income is not only much higher than in official sources, but also the highest in the Latin American region.

I now turn to analyze the distribution of income over the entire period between 1990 and 2022. Figure 4 shows the share of the pretax national income that goes to the different groups of the population. Starting at the bottom of the population, the share of the bottom 50% (black line) was below 5% in the 1990s and increased slowly after dollarization and during the commodity price boom to about 9% in 2014, where this level is constant since then. The middle 40% (blue line) increased its share in the middle of the 1990s from 36% to 42% (which means that its income share was higher than its population share), but steadily lost participation in pretax income in subsequent years, plunging to 30% in 2010. Since then, this group has slightly recovered participation to levels above 33%. The top 10% (red solid line), on the other hand captured more than 60% of the total pretax income in 1990. Despite a reduction in their share 58% in 2000, this group could increase the participation in the distribution of pretax income to 63% again in 2010. After a drop to 58% again in 2015 its participation accounts now for 59% of total pretax income. Figure 4 also indicates the concentration for smaller groups at the top (dashed red lines): The share of the top 1% decreased from 30% to 20% in the 1990s,

Figure 4: Pretax income shares in Ecuador (1990-2022)



Red vertical lines: 2000 Dollarization, 2020 COVID pandemic, Shaded areas: 1998-2000 Bank crisis, 2004-2014 Commodity price boom. The year 2002 was excluded due to inconsistent data.

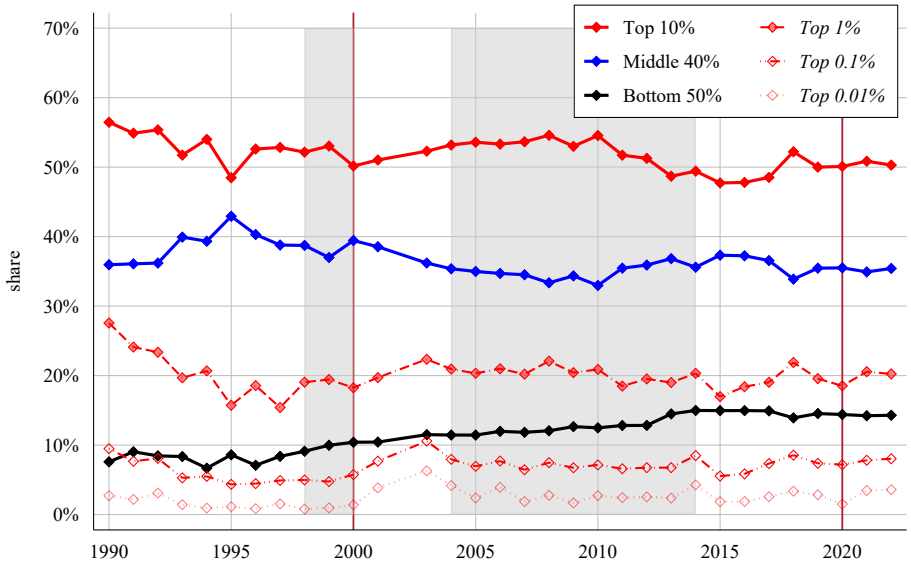
but recovered to 25% only three years later, being constant around this number since then. A parallel movement can be observed for the top 0.1% (top 0.01%), who's share fell from 10% (2%) to 7% (1%) before 2000, but maintains a level of around 10% (5%) again since 2003. Interestingly, the COVID pandemics, which represented a similar shock to the economy as the 2000 crisis, did not affect the distribution of income significantly.

Comparing the series for posttax national income in figure 5 to the pretax series, one observes the expected differences in the shares caused by redistribution policies. The bottom 50% captured already about 10% of posttax income in the 1990s and increased their share continuously to 15%. On the other hand, while the share of the top 10% fluctuated around 60% in the pretax series, its posttax income share is around 50%. Also for higher income groups, their share decreases from pretax to posttax income. The series for the middle 40% confirms what was already observed in tables 1 and 2: taxes and transfers have no net effect on income for this group. A general observation is that both pretax and posttax income series develop similar over time, although at different levels. Another insight form the figures is that the evolution of the richest decile's share is almost the price mirror of what happens to the middle 40%, indicating that income shifts primarily between these groups.

The combination of figures 4 and 5 allow us to analyze (a) how income groups change their participation in national income over time, and (b) how redistribution policies influence the distribution of income in different moments. The bottom 50% of the population increased its pretax share continuously over the entire period (+6 percentage points), and was also the group that increasingly benefited most by taxes and transfers (starting at +5 percentage points and increasing up to +7points difference between pretax and posttax income in the last years).[10] The

[10]See figure B.1.

Figure 5: Posttax income shares in Ecuador (1990-2022)



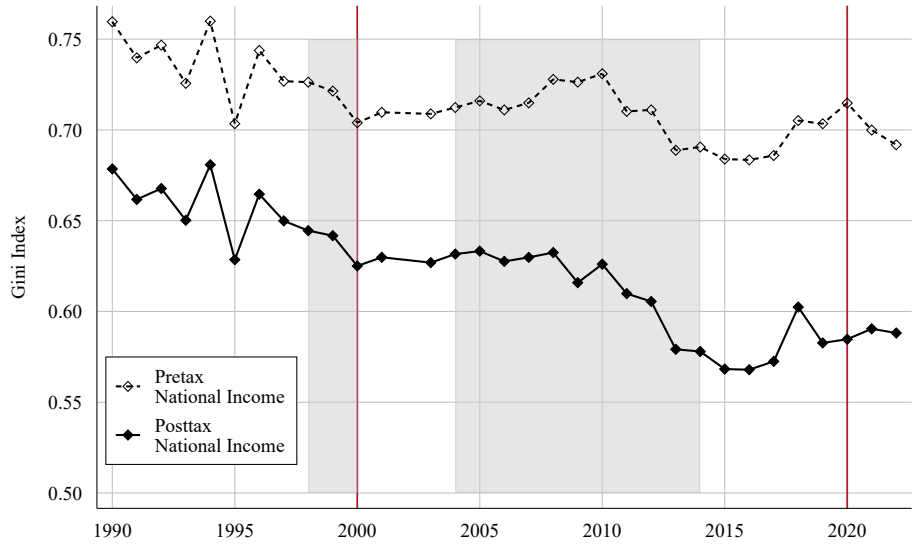
Red vertical lines: 2000 Dollarization, 2020 COVID pandemic, Shaded areas: 1998-2000 Bank crisis, 2004-2014 Commodity price boom. The year 2002 was excluded due to inconsistent data.

middle 40% strongly lost participation in pretax and posttax income until 2010 and recovered only marginally afterwards. In addition, the net effect of taxes and transfers is negligible for this income group over the whole period, only accounting for +2 percentage points in the most progressive years. On the other hand, the share of the top income groups decreased before 2000, and increased later. In terms of redistribution polices, one can observe an increasing burden of the top 10%: starting with a reduction of 5 percentage points in the income share in the 1990s, this reduction was between 8 and 10 points since 2010. For the top 1% and richer groups, the difference between pretax and posttax also increases over time, although at a more moderate pace.

One way to summarize the findings so far is to look at the overall development of the Gini index for pretax and posttax national income in figure 6. This synthetic inequality indicator - ranging form 0 for total equality to 1 for total inequality - decreases during the 1990s and between 2010 and 2015 for both pretax and posttax income.[11] In addition, the gap between pretax and posttax income Gini - representing the strength of redistributional policy - is 8 points until the mid 2000s, and widens 12 points in the COVID pandemics, before contracting to 10 points in 2022. Despite being informative for an overall assessment of the inequality evolution, figure 6 also calls for caution when relying exclusively on a synthetic inequality indicator, such as the Gini index. The index cannot inform about where in the distribution income growth or redistribution happens.

[11] For a comparison to the official Gini index stemming from the ENEMDU, see figure B.2. The level and evolution of inequality is different in this figure, as the official method only uses household sector income and produces *per capita* income estimates.

Figure 6: Pretax and posttax income Gini index



Red vertical lines: 2000 Dollarization, 2020 COVID pandemic, Shaded areas: 1998-2000 Bank crisis, 2004-2014 Commodity price boom. The year 2002 was excluded due to inconsistent data.

4.2. Growth rates by income group

Figure 1 in the introduction illustrated the volatile evolution of real income per adult, which finally results in a poor overall average real growth rate of 0.4% per year. On the other hand, the pretax and posttax series of subsection 4.1 suggest increasing shares in national income for the bottom 50% and relatively constant shares for the middle 40% and the top 10% between the first and the last years. Given that the series produced in this study follows the Distributional National Account methodology, I can distribute national income growth from national accounts to different population groups. In this way, I can explain which groups benefited most from growth in these 33 years.

To analyze how income changed over time in real terms, I set the base value of real income for each group to 1 in 1990. This normalization to an index allows for a clear comparison of income changes over time, with values above 1 indicating real income growth and values below 1 indicating a decline relative to the base year. Figure 7 draws the index of pretax income for each group and year in panel (a), while panel (b) shows the index for posttax income. For pretax national income, the bottom 50% experimented high volatility in the 1990s but finally earned twice the income in 2000 compared to 1990. The middle 40% earned 20% more in 1995 than in 1990, but lost this gain in the banking crisis 1998-2000. Also the richest groups lost income in real terms. After dollarization, and during the major period of the commodity price boom, all income groups experimented income growth, especially the bottom 50% which doubled again its real pretax income between 2000 and 2014. The index for the top 10% increased from 0.8 to 1.2 in 2014, and for the top 1% from 0.6 to 1. The index of the middle 40% only recovered marginally from 1 to 1.2. After the commodity price boom all income groups lost income in

real terms until the year 2020, with a major shock from the COVID pandemic. The only group recovering its growth after 2020 was the bottom 50% which closes with an index of 3.5, while all other groups close with an index around 1.

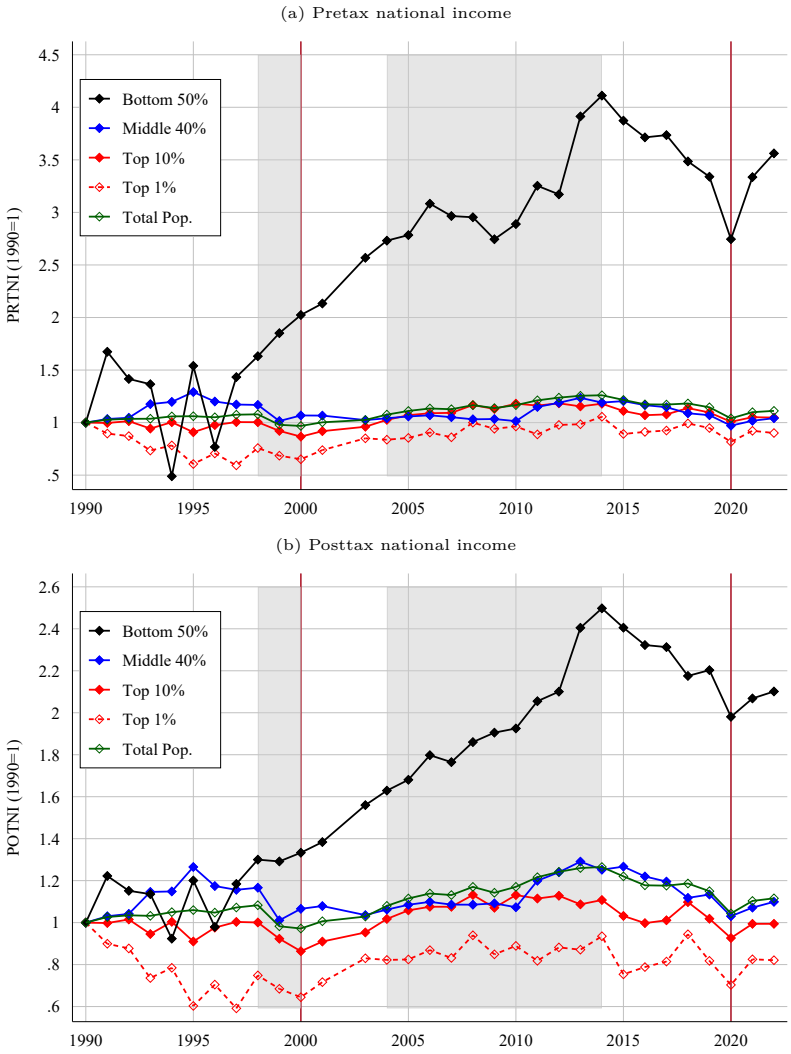
For posttax national income, the evolution is similar, but smoother. The bottom 50% had less sharp movements in ups and downs before 2000, but also reached 1.3 times the income of 1990. The middle class could only maintain its income in real terms, the rich lost as sharp as in pretax terms. Growth for the bottom was then steady (without drops in 2006 and 2011 as in pretax) but also more moderate, reaching an index of 2.5 times the income of 1990, and fell back to an index of 2 in the year 2020. The middle class did not evolve until 2010 in real terms, increased its income to 1.3 times in 2013 and fell again to an index of 1 in the year 2020. The rich increased income after dollarization and culminated with almost 1.2 in 2010, since then around the same level as in 1990. Top 1 percent could never recover the 1990 income in real terms.

4.3. International comparison

The DINA method produces internationally comparable inequality estimates for Ecuador, and I can put the result in context with countries of the Latin American region that have data from the WID. I first observe the Gini index in figure 6 for a group of Latin American countries: For pretax income (panel (a)), Ecuador (red) performs similar to its neighbors Colombia and Peru (black), in both levels (around 0.7) and changes (significant reductions between 2008 and 2015, followed by an increase in the early 2020s). Chile, Argentina and El Salvador also experimented an overall decrease in the pretax Gini index, Brazil stagnated, and Mexico even faced increasing inequality over the last 20 years. Turning to posttax income (panel (b)), Gini estimates place Ecuador between the three most equal countries, after Argentina and El Salvador. The tendency is still close to the neighbors Colombia and Peru, but the Gini level lies about 0.05 points below.

In terms of the Gini index, Ecuador could be considered as an average scholar in terms of inequality, outperforming its neighbors especially when taxes and transfers are included in the estimate. However, as I highlighted in section 4.1, only looking at a synthetic indicator can obscure developments for more detailed income groups. I therefore compare the shares of the different income groups of the same countries in figure 9. In panel (a) and (b), the share of pretax and posttax income for the bottom 50% increases in the Latin American region, and Ecuador has the second highest share (14%) after Argentina (16%). The middle 40% income share (panels (c) and (d)) are more similar to the other countries in the Andean region at around 35% for posttax income, while in Argentina and El Salvador this group captures more then 40% and 45% percent of national income. For the top 10% (panels (e) and (f)), the pretax income share is again similar to Colombia and Peru, and this group has a lower share than its neighbors in posttax income (around 50%), undercut only by Argentina and El Salvador (40%). Up to this point, the analysis of shares resembles the results from the Gini index at more detail. Turning to smaller and richer groups, the picture changes: The pretax income share of the top 1% (panel (g)) is already among the three highest with 25%, but state intervention moves Ecuador again towards the regional average for the posttax share (panel (h)). Finally, the share

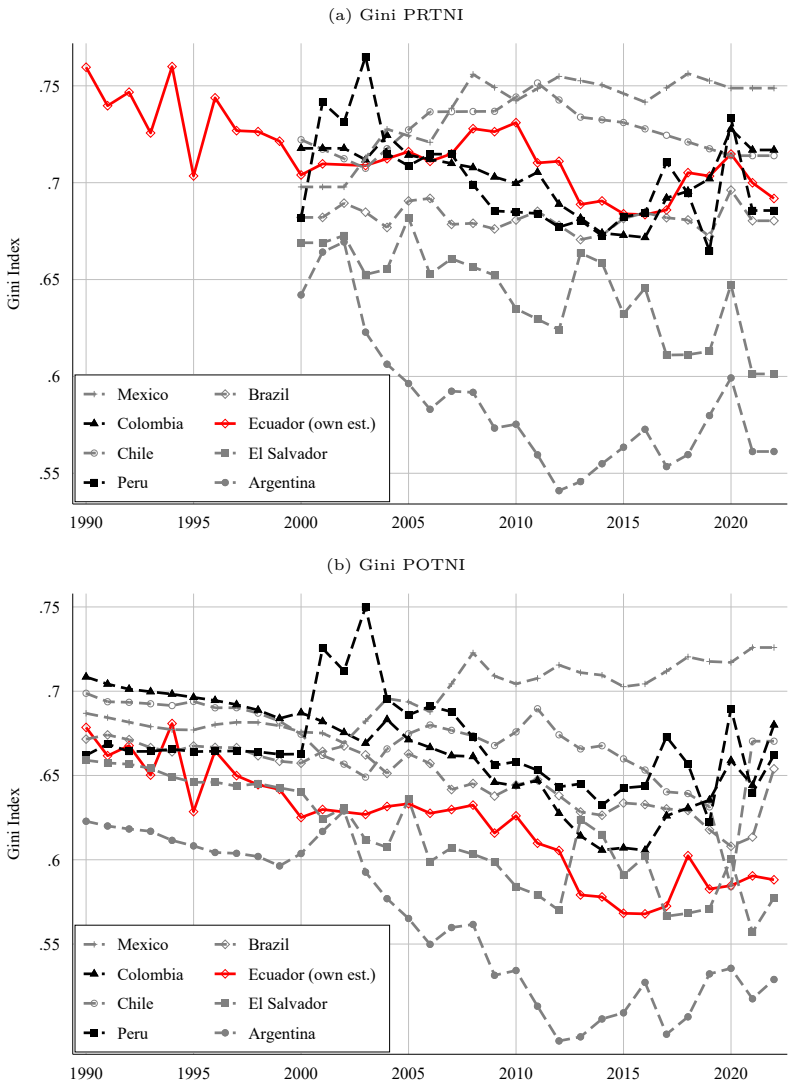
Figure 7: Real growth of different income groups



Red vertical lines: 2000 Dollarization, 2020 COVID pandemic, Shaded areas: 1998-2000 Bank crisis, 2004-2014 Commodity price boom.

The index represents the real pretax and posttax income levels per adult in comparison to the base year of 1990. For instance, the index for the bottom 50% of the population in the year 2000 is about 2, which indicates that their real income in 2000 was twice as high as their income in 1990.

Figure 8: Gini index for countries in the region



Ecuador based on own estimates, all other from WID.

of the top 0.1% (panel (i)) has fluctuated at around 10%, which is at the top of all countries of the region, and despite reductions after taxes and transfers, Ecuador is still the country with the highest share of income going to the richest 10% of the richest percentile (panel (j)).

5. Conclusion

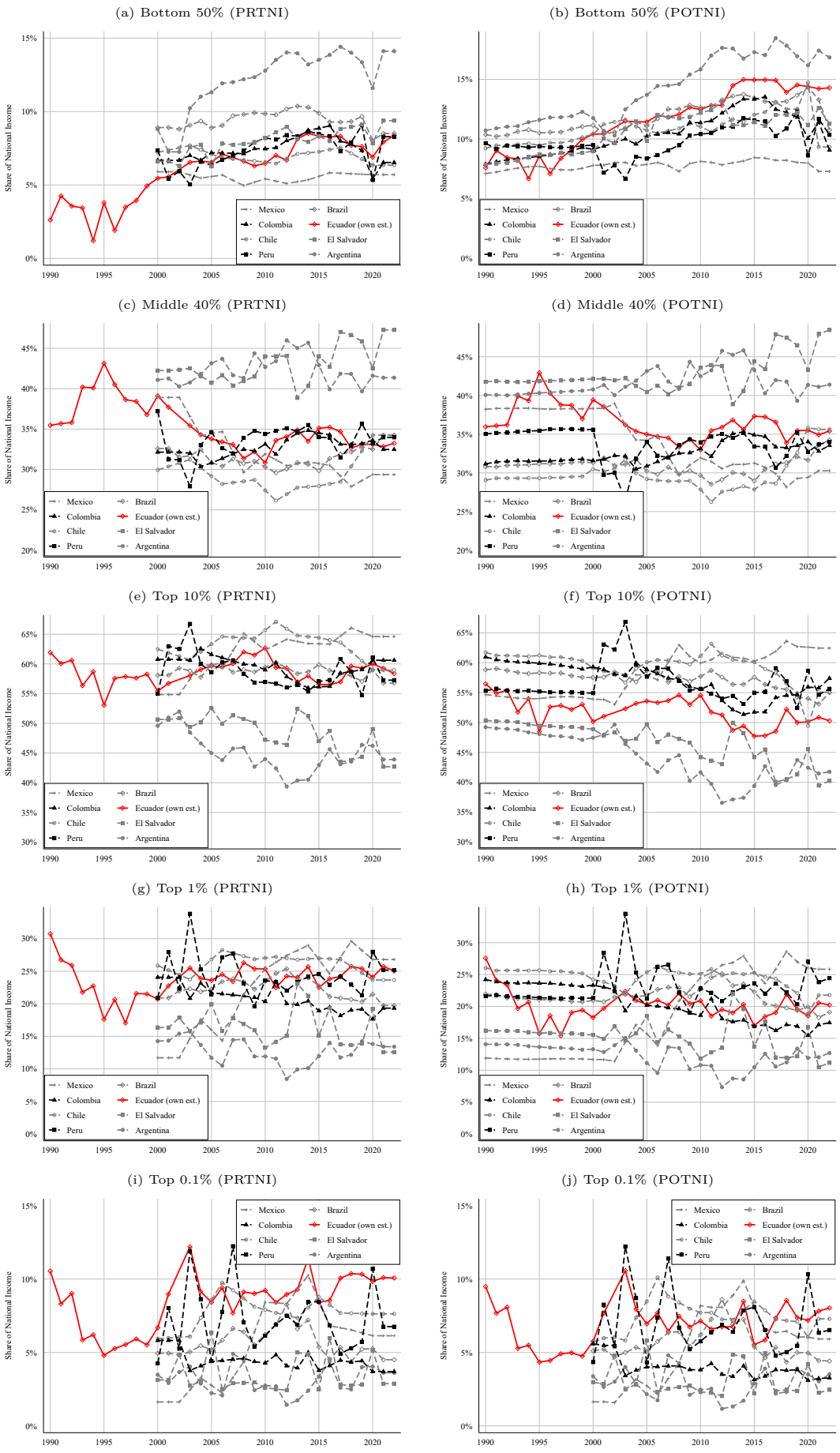
This study mobilizes for the first time survey and tax microdata, national accounts and additional data sources to produce distributional national accounts for Ecuador covering the period since 1990. This new series addresses the underrepresentation of wealthy households in surveys and accounts for 100% of income from national accounts. Furthermore, the study introduces a novel integration method for merging survey and tax microdata, contributing to the literature on income distribution measurement.

I find that the distribution of income in Ecuador was highly unequal in 1990, with extreme concentrations at the top persisting today, despite some improvements for lower income groups. The share of income earned by the bottom 50% of the population increased from 2.6% in 1990 to 8.4% in 2022 for pretax income, and from 7.6% to 14.3% for posttax income. The middle 40% of the income distribution maintained a relatively stable share of around 35% in both years, with state interventions via taxes and transfers having a negligible net effect on this group. In contrast, the top 10% captured 62% of pretax income in 1990 and 58% in 2022. The top 1% accumulated 25% of income in 2022, while the top 0.001% (115 individuals) earned 2.4% of national pretax income. The average income for this top 0.001% group is 2,300 times the average income of the total adult population, and even after state interventions, their posttax income is 1,900 times the average. Compared to other Latin American countries using DINA data, Ecuador has the highest share of income held by the top 0.1%.

Contrary to earlier studies, I do not observe a rise in inequality during the 1990s and the banking crisis of 1998-2000. Instead, the participation of both the bottom 50% and the middle 40% increased at the expense of the top 10%. Following dollarization until 2010, the bottom 50% gained higher shares of income, while the middle 40% experienced a significant decrease, and the top 10% only slight increased their share. These shifts resulted in a relatively stable Gini index. From 2010 to 2015, during the final years of the commodity price boom and the Citizen Revolution government, both the bottom and middle income groups increased their share further. However, estimates for recent years indicate a resurgence of increasing inequality since then.

Given that the annual average real growth rate per adult over the entire period was just 0.4%, I disaggregated this growth across different income groups. The findings reveal a clear pro-poor growth pattern: the bottom 50% saw their real income increase 2.5 times from 1990 to 2015. The middle 40% experienced a 1.2-fold increase in real income during the same period, while the top 10% only grew by a factor of 1.1. Since 2015, all income groups have faced real income declines. As of now, the bottom 50% earns twice the post-tax income compared to 1990, the middle income group earns about 1.1 times the 1990 level, but the top income group remains at the same real income level as in 1990.

Figure 9: International comparison of income shares by group



Ecuador based on own estimates, all other from WID.

Although this study was not designed for causal inference, it provides important insights into the relationship between policies and income inequality. First, income concentration at the top remained notably stable throughout the studied period, showing little fluctuation in response to monetary and fiscal shocks. Second, The banking crisis, which affected top incomes and somewhat impacted the middle-income group, did not reduce the income of the bottom 50%. This is consistent with the observation that low-income households typically do not hold financial assets exposed to volatile exchange and interest rates. Third, in the period following dollarization and during the commodity price boom, the economy experienced pro-poor growth. During this time, the richest groups recovered both real income and income shares, while the middle class did not benefit significantly from the economic expansion. Fourth, the recent political crisis and the COVID pandemics had a uniform impact across all income groups, resulting in no significant changes to the overall distribution of income.

Bibliography

Alesina, Alberto and Robert J Barro (2001). "Dollarization". *American Economic Review* 91.2, pp. 381–385.

Alvaredo, Facundo, Lydia Assouad, and Thomas Piketty (2019). "Measuring lnequality in the Middle East 1990–2016: The World's Most Unequal Region?" *Review of Income and Wealth* 65.4, pp. 685–711.

Assouad, Lydia, Lucas Chancel, and Marc Morgan (2018). "Extreme Inequality: Evidence from Brazil, India, the Middle East, and South Africa". *AEA Papers and Proceedings*. Vol. 108, pp. 119–23.

Atkinson, Anthony B (2007). "Measuring top incomes: methodological issues". *Top incomes over the twentieth century: A contrast between continental European and English-speaking countries*, pp. 18–42.

Bach, Stefan, Charlotte Bartels, and Theresa Neef (2020). "Distributional National Accounts: A Macro-Micro Approach to Inequality in Germany". *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*. University of Chicago Press.

Bartels, Charlotte and Maria Metzing (2019). "An integrated approach for a top-corrected income distribution". *The Journal of Economic Inequality* 17.2, pp. 125–143.

Beckerman, Paul and Andrés Solimano (2002). *Crisis and Dollarization in Ecuador"¯: Stability, Growth, and Social Equity*. World Bank/World Bank Group.

Blanchet, Thomas, Lucas Chancel, Ignacio Flores, and Marc Morgan (2020). "Distributional National Accounts (DINA) guidelines: Concepts and methods used in the World Inequality Database". *updated and extended version of the DINA Guidelines first published as WID Working Paper 2016/2*.

Blanchet, Thomas, Lucas Chancel, and Amory Gethin (2022a). "Why is Europe more equal than the United States?" *American Economic Journal: Applied Economics* 14.4, pp. 480–518.

Blanchet, Thomas, Ignacio Flores, and Marc Morgan (2022b). "The weight of the rich: improving surveys using tax data". *The Journal of Economic Inequality* 20.1, pp. 119–150.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2022c). "Generalized pareto curves: Theory and applications". *Review of Income and Wealth* 68.1, pp. 263–288.

Burdin, Gabriel, Mauricio De Rosa, Andrea Vigorito, and Joan Vila (2022). "Falling inequality and the growing capital income share: Reconciling divergent trends in survey and tax data". *World Development* 152, p. 105783.

Cachanosky, Nicolás, Alexander W Salter, and Ignacio Savanti (2022). "Can dollarization constrain a populist leader? The case of Rafael Correa in Ecuador". *Journal of Economic Behavior & Organization* 200, pp. 430–442.

Cano, Liliana (2015). *Income mobility in Ecuador: New evidence from individual income tax returns*. Vol. 2015. UNU-WIDER. isbn: 978-92-9230-925-1.

Chancel, Lucas, Denis Cogneau, Amory Gethin, Alix Myczkowski, et al. (2019). "How large are African inequalities? Towards distributional national accounts in Africa, 1990-2017". *World Inequality Database (WID) Working Papers*, pp. 1–71.

Chancel, Lucas and Thomas Piketty (2019). "Indian Income Inequality, 1922-2015: From British Raj to Billionaire Raj?" *Review of Income and Wealth* 65, S33–S62.

Chatterjee, Aroop, Léo Czajka, and Amory Gethin (2021). "Can Redistribution Keep Up with Inequality? Evidence from South Africa, 1993-2019". *World Inequality Lab – Working Paper N° 2021/20*.

Clementi, Fabio and Mauro Gallegati (2005). "Pareto's law of income distribution: Evidence for Germany, the United Kingdom, and the United States". *Econophysics of wealth distributions*. Springer, pp. 3–14.

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2020). "Inequality in Latin America revisited: insights from distributional national accounts". *World Inequality Lab – Technical Note* (02).

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2022). "More Unequal or Not as Rich? Revisiting the Latin American Exception". *World Inequality Lab – Working Paper N 2022/13*.

Flechtner, Svenja and Martin Middelanis (2024). "The role of the commodity price boom in shaping public social spending: Evidence from Latin America". *World Development* 182, p. 106717. issn: 0305-750X.

Flores, Ignacio, Claudia Sanhueza, Jorge Atria, and Ricardo Mayer (2020). "Top Incomes in Chile: A Historical Perspective on Income Inequality, 1964–2017". *Review of Income and Wealth* 66.4, pp. 850–874.

Gachet, Ivan, Diego F Grijalva, Paúl A Ponce, and Damián Rodríguez (2019). "Vertical and horizontal inequality in Ecuador: The lack of sustainability". *Social Indicators Research* 145, pp. 861–900.

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2018). "Income inequality in France, 1900–2014: evidence from distributional national accounts (DINA)". *Journal of Public Economics* 162, pp. 63–77.

García-Albán, Freddy, Manuel González-Astudillo, and Cristhian Vera-Avellán (2021). "Good policy or good luck? Analyzing the effects of fiscal policy and oil revenue shocks in Ecuador". *Energy Economics* 100, p. 105321. issn: 0140-9883.

Jameson, Kenneth P (2003). "Dollarization in Latin America: Wave of the future or flight to the past?" *Journal of economic issues* 37.3, pp. 643–663.

Koutentakis, Franciscos and Kostas Chrissis (2020). "Top Income Shares in Greece From Dictatorship to Crisis: 1967–2017". *Review of Income and Wealth*.

Lustig, Nora (2020). "The "Missing Rich" in Household Surveys: Causes and Correction Approaches". *ECINEQ Working Paper Series 2020 520* 520, pp. 1–37.

Novokmet, Filip, Thomas Piketty, and Gabriel Zucman (2018). "From Soviets to oligarchs: inequality and property in Russia 1905-2016". *The Journal of Economic Inequality* 16.2, pp. 189–223.

Özyurt, S and S Cueva (2020). "Twenty years of official dollarization in Ecuador: a blessing or a curse?" *Agence Française de Développement–Macroéconomie et développement, Working paper No. 31*.

Piketty, Thomas and Emmanuel Saez (2003). "Income inequality in the United States, 1913–1998". *The Quarterly journal of economics* 118.1, pp. 1–41.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2019). "Simplified Distributional National Accounts". *AEA Papers and Proceedings* 109. issn: 2574-0768.

Ponce, Juan and Rob Vos (2014). "Redistribution without structural change in Ecuador: rising and falling income inequality in the 1990s and 2000". *Falling Inequality in Latin America. Policy Changes and Lesson, Oxford, United Kingdom*.

Rossignolo, Dario, Nicolas Oliva, and Nestor Villacreses (2016). "Cálculo de la concentración de los altos ingresos sobre la base de los datos impositivos". *CEPAL: Serie Macroeconomia del Desarrollo* (184).

Soto, Raimundo (2009). "Dollarization, economic growth, and employment". *Economics letters* 105.1, pp. 42–45.

Zwijnenburg, Jorrit, Sophie Bournot, David Grahn, and Emmanuelle Guidetti (2021). "Distribution of household income, consumption and saving in line with national accounts: Methodology and results from the 2020 collection round". *OECD Statistics Working Papers 2021/01 No. 108*.

APPENDIX

Appendix A Supplementary information on data and methodology

A.1 Characteristics of data sources

This study combines three types of data: national accounts, survey microdata, and tax microdata. Each data source offers unique advantages and disadvantages, as summarized in table A.1. First, national accounts provide a benchmark for the total income of each income component, institutional sector, and year. They are considered to comprehensively represent the income of the entire economy, offering a broad overview. However, their level of disaggregation is limited, which can hinder detailed analyses. Second, microdata from nationally representative household surveys contain information on household and individual incomes from major income sources, covering most of the Ecuadorian total income up to a certain point in the distribution (e.g., the lower 90%, or group A in the table), but lack information for the richest households (group B in the table). Household surveys usually include tens of thousands of observations featuring a high level of disaggregation and additional contextual variables beyond income. Finally, administrative records, such as tax registers, complement survey data by providing detailed information on higher-income individuals (group B), who are often underrepresented in household surveys. Tax microdata, while not covering the lower deciles of the income distribution (group A), offer detailed data on individual incomes, particularly for formal employees, business owners, and those declaring property income. Tax records are valuable for constructing income profiles of the wealthiest individuals in the country[12], offering a high level of disaggregation at the individual level and sometimes including contextual variables from tax declarations or integrated registers.

However, these three data sources also have limitations, such as changes in methodologies and missing information, especially when studied over an extended period. The following subsections describe the preparation of these sources to make them compatible with the Distributional National Accounts (DINA) methodology, ensuring a comprehensive and accurate analysis of income distribution in Ecuador:

National accounts: The Ecuadorian Central Bank (*Banco Central del Ecuador, BCE*) publishes the "Cuadros Economicos Integrales", which contain national accounts following the United Nations' System of National Accounts (United Nations, 2009) for the years 2007-2022[13]. These tables contain the internationally standardized income components for all institutional sectors. For the years before 2007, the BCE only publishes macro-aggregates (GDP,

[12]Some studies include information from rich lists to try to cover a very reduced number of individuals at the top, but this information is primarily used to correct wealth inequality, and scarcely for income inequality estimates (Alvaredo et al., 2019).

[13]Retrieved from `https://contenido.bce.fin.ec/documentos/PublicacionesNotas/Catalogo/CuentasNacionales/Anuales/Dolares/indicecn1.htm` (years 2007-2020) and `https://contenido.bce.fin.ec/documentos/informacioneconomica/cuentasnacionales/ix_cuentasnacionalesanuales.html` (years 2018-2022).

Table A.1: Strengths and limitations of data types

Data type	Example	Completeness/ representativeness	Level of disaggregation	Context information
Macro aggregate	National Accounts, Input-Output tables	high	low	low
Household survey	Labor market surveys, Expenditure surveys	middle (high for group A, low for group B)	high	high
Administrative record	tax record, social security records, social assistance records	middle (high for group B, low for group A)	high	middle

The table presents strengths and limitations of different data sources. Groups in column *Completeness/representativeness* can be interpreted as follows: Group A covers low and middle income observations, while group B covers high income observations.

GNI) without detailed information on institutional sectors and income components (Ecuador, 2012)[14]. Additionally, during the process of updating the national accounts, the BCE changed the fixed reference year from 2007 to a mobile reference year starting in 2018. For the years 2018-2020, both series (the fixed base year 2007 and the mobile reference year) are available. To disaggregate total GNI for the years before 2007, two additional sources are utilized: the remittances statistics published by the BCE, which include yearly data since 1989 (Ecuador, 2001 for years before 2000 and online for current years[15] ; and the "Details of Tax Revenue" dataset from the OECD[16]. The harmonization of these different sources to produce a series for detailed national account items is elaborated in section A.2.

Household surveys: The National Statistical Office of Ecuador (*Instituto Nacional de Estadistica y Censos*, INEC) has conducted the National Employment Survey (*Encuesta Nacional de Empleo, Desempleo y Subempleo*, ENEMDU) under different names since 1989 Ponce and Vos (2014). This study primarily relies on this long-term survey as a key source of household income data. However, it is important to note that the survey's coverage and methodologies have evolved over this long period. Initially, the definitions and disaggregations related to income concepts were less clear, and the questionnaires were relatively brief in the 1990s. Since 2003, the survey has employed more detailed questionnaires, which have enhanced the clarity and depth of the data collected. Despite these changes, all significant income items have been consistently included in the survey from its inception. The coverage of total household income in the national accounts was approximately 40% in the early years, increasing to around 50% in the last decade. This level of coverage is considered reasonable for the purposes of this study.

[14]Retrieved from https://contenido.bce.fin.ec/documentos/PublicacionesNotas/Catalogo/CuentasNacionales/Anuales/Dolares/indicecn1.htm

[15]Retrieved from https://contenido.bce.fin.ec/documentos/Estadisticas/SectorExterno/BalanzaPagos/Remesas/indice.htm.

[16]Retrieved from https://data-explorer.oecd.org/vis?tenant=archive&df[ds]=DisseminateArchiveDMZ&df[id]=DF_REV&df[ag]=OECD&dq=...&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false&vw=tb.

Second, the abolition of the Sucre as Ecuador's currency presents a challenge on the comparability of *absolute* income of households for the years around 2000. The period was marked by galloping inflation and significant discrepancies between Ecuador's official and unofficial exchange rates, particularly from the late 1990s. These factors complicate comparisons of household income, especially between 1999 and 2000.[17] While these economic conditions introduce volatility in absolute income comparisons for the years surrounding 2000, they do not necessarily affect the comparison of income *distribution*. To ensure general comparability across the entire series, I use the official national accounts series in US dollar and the official exchange rates to standardize the data before and after the currency change.

Third, the ENEMDU survey is nationally representative for most years, but from 1990 to 1999 and again in 2002, it only includes data from urban areas. This lack of rural data poses challenges for comparing the series before and after 2000, despite most income being generated in urban areas, particularly at the top income levels. To address this issue, I use data from of a second survey in Ecuador, the Living Condition Survey (*Encuesta de Condiciones de Vida*, ECV, INEC, 2015). The survey was conducted in the years 1995, 1998, 1999, 2006 and 2014 and contains information on both urban and rural households. The ECV contains similar income-related questions to those in the primary ENEMDU survey, allowing to construct pretax household income for all income fractiles for the years 1995, 1998, 1999, and 2006, and therefore to derive the share of rural individuals within each fractile.[18] To estimate rural population shares for the years 1990-2002, I use the 1995 shares for the years 1990-1995 and interpolate the years between 1995 and 2002 using data from 1995, 1998, and 2006. For 2002, the rural shares are derived from ENEMDU data from 2001 and 2003, as both these years already include rural households. Figure A.1 illustrates that the participation of rural population is generally lower for higher income fractiles in both surveys.[19] By multiplying the weight of each observation in the ENEMDU survey with the share derived from the ECV, the ENEMDU survey is corrected for the missing rural population. For example, if the rural share in fractile 50 in the ECV is 30%, I adjust the weights of observations in fractile 50 in the ENEMDU by a factor of 1.3. Figure A.2 illustrates the series of share of rural population.

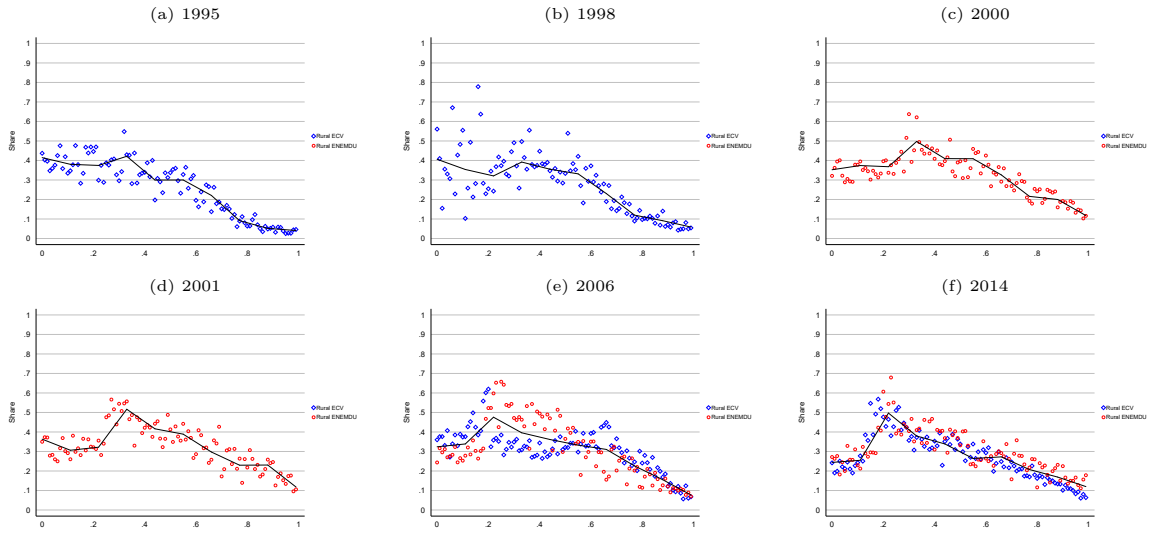
Fourth, the ENEMDU survey does not account for imputed owner-occupied rents (income concept B2R1 in national accounts), a notional income source that is crucial for making national accounts comparable across countries. For many Ecuadorian households, particularly in rural areas, imputed rents represent a significant portion of their total income, even though

[17] The exchange rate Sucre-Dollar increased from 7.000 in January to almost 18.000 in December of 1999, and was finally frozen at 25.000 in January 2021 (`https://www.bce.fin.ec/cotizaciones`). Accumulated inflation in 1999 was 60.7% (`https://www.ecuadorencifras.gob.ec/historicos-ipc/`.

[18] Results for income from 1999 deviate sharply from results in 1998 and 2006, likely caused by difficulties in the recollection during the crisis and currency change. I therefore exclude the 1999 ECV survey from the series.

[19] Note that in later years the rural population share is highest around the percentile 20. This shift indicates that - at least looking at income from household surveys - over time more urban households belong to the poorest income groups.

Figure A.1: Rural population shares 1995-2014 (ENEMDU and ECV)



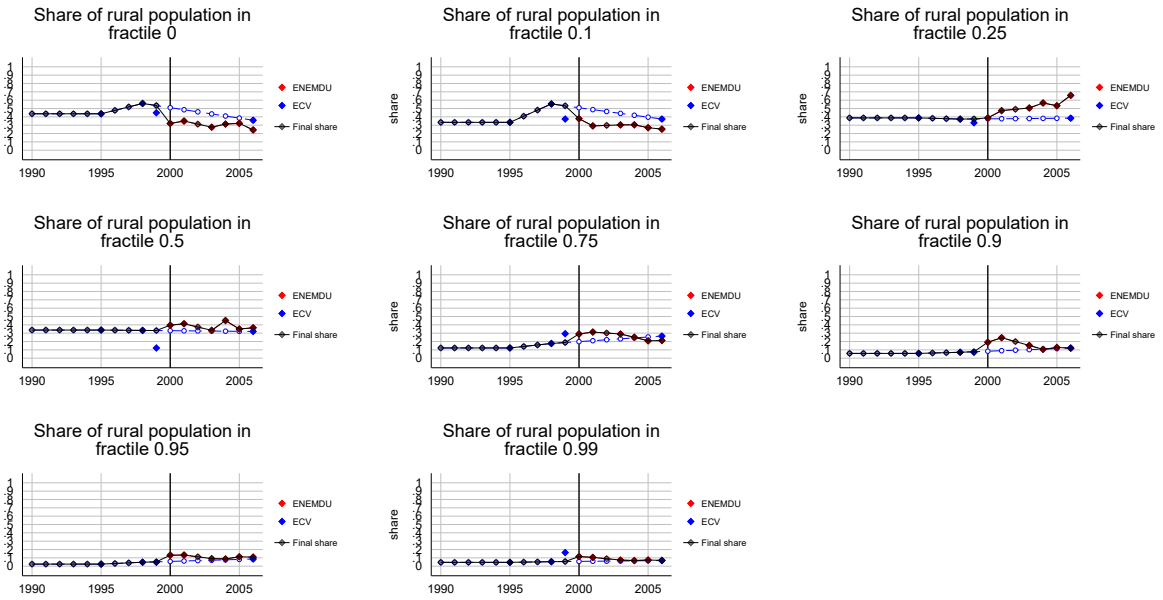
Dots indicate population share from rural area in each fractile. Sources: ENEMDU, ECV.

they are notional. Therefore, including this income source is essential for accurate national income estimates (see figure A.3. To address this, I utilize data from the ECV survey, which includes information on imputed rents for the years 1995, 1998, 1999, 2006, and 2014. I first derive absolute values of imputed rents by fractile[20] for these years. Then, I interpolate and extrapolate these values for all years between 1990 and 2022. For the years before 1995 and after 2014, I maintain the share of imputed rents in total pretax household income and adjust the absolute values according to the growth rates of income. This approach was chosen over a pure extrapolation method, which appeared problematic, particularly for the period before 1995 (see figure A.4).

Tax records: Historically, tax collection in Ecuador was handled by the Ministry of Finance, which had limited administrative capacity and was highly susceptible to corruption. Significant improvements in tax administration only began in the early 1990s, following a series of legal reforms (Arias et al., 2008). Greater autonomy granted in 1994 and the establishment of the *Servicio de Rentas Internas* (SRI) in 1997 were pivotal in enhancing tax collection efficiency, leading to a more than doubling of revenues from less than US$600 million to almostUS$1.5 billion between 1990 and 1998.Throughout the study period, Ecuador implemented several tax reforms in response to changing economic conditions. For instance, the deteriorating macroeconomic environment in the late 1990s prompted short-term measures, such as the introduction of a tax on the circulation of capital (*Impuesto a la Circulacion de Capitales*, ICC) in December 1998 replacing the income tax. However, this measure was discontinued after just 12 months. Despite representing a significant shift in the tax structure, the fiscal impact was moderate:

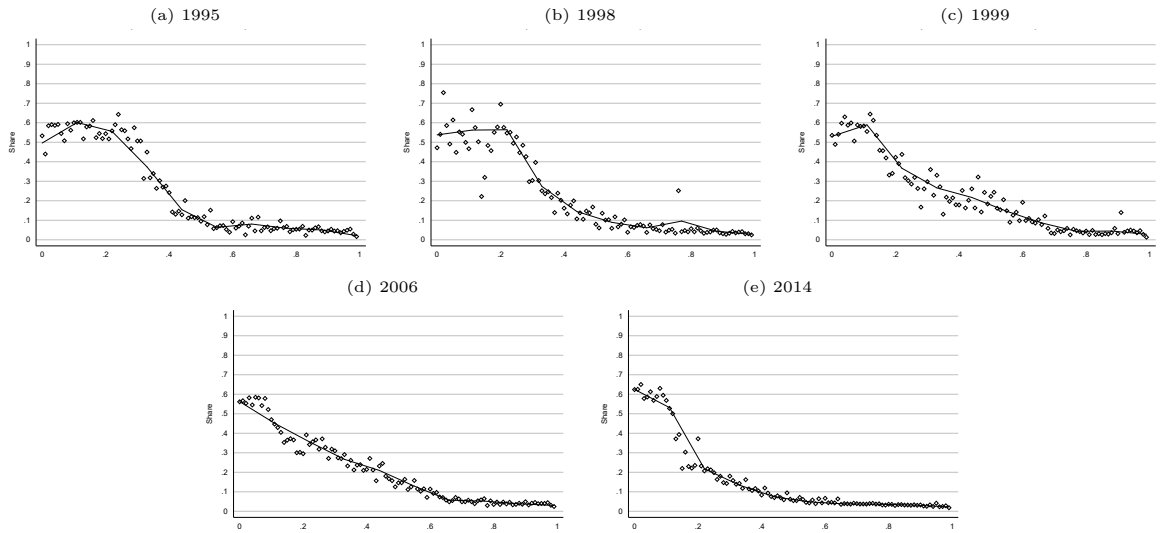
[20]Fractiles in the ECV survey are calculated excluding imputed rents, to ensure comparability with PreTaxHHI fractiles from ENEMDU.

Figure A.2: Comparison: Share of rural population series for fractiles



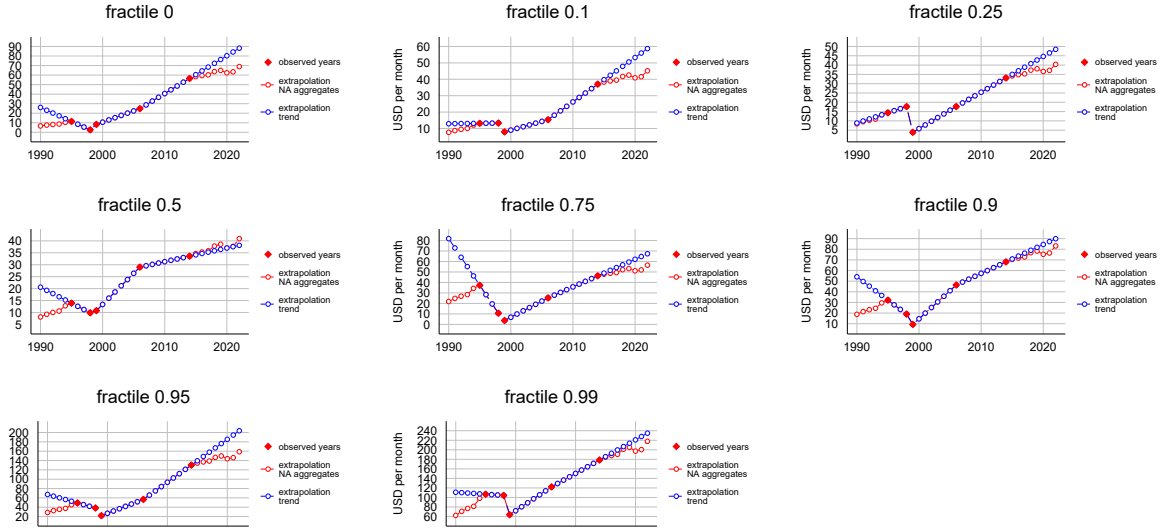
Sources: ECV, ENEMDU, BCE.

Figure A.3: Share of imputed rents per fractile



Source: ECV.

Figure A.4: Imputed rents in USD (monthly) for fractiles



Sources: ECV, BCE

income and wealth taxes accounted for 30% of overall tax revenues on average in the years 1995-1998, while the ICC contributed 35% in 1999 (Ramírez and Carrillo Maldonado, 2020). Following dollarization, the share of income tax in total tax revenue recovered gradually from 25% to 32% in 2007. With the Constitution of 2008 entering into force, tax policies were focused on a more progressive income tax collection (progressive income tax scale, targeted incentives and benefits, stricter tax controls, etc.). These measures increased the contribution of income taxes to over 35% of total tax revenues after the Global Financial Crisis (Ramírez and Carrillo Maldonado, 2020). [21].

Given these ruptures in tax policies and the capacity of tax collection, it is crucial to assess the content and quality of statistical information derived from individual tax records before incorporating them into this study. The first electronic and digitized income tax declarations date back to 2004. Since then, individuals have either declared exclusive labor income using tax form 107 or their comprehensive income—including labor, capital, mixed income, inheritance, etc.—using tax form 102. The SRI consolidates these forms using the individual identifying tax number (*Registro Unico de Contribuyente*, RUC), and maintains a rigorous conversion matrix for all subsequent years to account for changes in tax forms. This approach allows for the consistent construction of income items for the entire period from 2004 to 2022. However, the quality of the digitized information in the initial years following the introduction of digital tax declarations likely suffers from significant issues. The sharp increase in the number of declarations between 2004 and 2008 suggests that the observed data may be more reflective of improvements in tax administration rather than actual economic developments. Cano (2015,

[21] The share is higher if the tax on foreign outward foreign-exchange transfers (*Impuesto a la Salida de Divisas*, ISD), introduced in 2009 a is excluded. This tax accounted for up to 10% of total tax revenue in 2012, but was gradually reduced since 2019.

p. 11) notes that changes in income inequality patterns before 2007 were mainly to "a reinforcement of tax collection and [to] an expansion of the fiscal data, rather than by an increase in income inequality". I therefore use tax microdata for the years 2008-2022, and assume that the income distribution of tax declaration remained constant from 1990 to 2007, using the structure observed in 2008 as the baseline.

This assumption implies that the share of income covered by the tax dataset in 2008 (41.2% of pretax household sector income) and the composition of wage, property and other incomes were consistent across the years before 2008. As a result, the integration of survey and tax data (see section A.4) will primarily reflect differences in distribution driven by household surveys and changes in national accounts totals. To ensure robustness, I also consider tax declarations from 2004 to 2007, using the 2004 data as a starting point for extrapolation back to 1990-2003 (see section **??**). This robustness check addresses the possibility that the increase in total tax revenue during these early years was not only caused by data quality increases but by a real economic phenomenon.

A.2 National accounts series

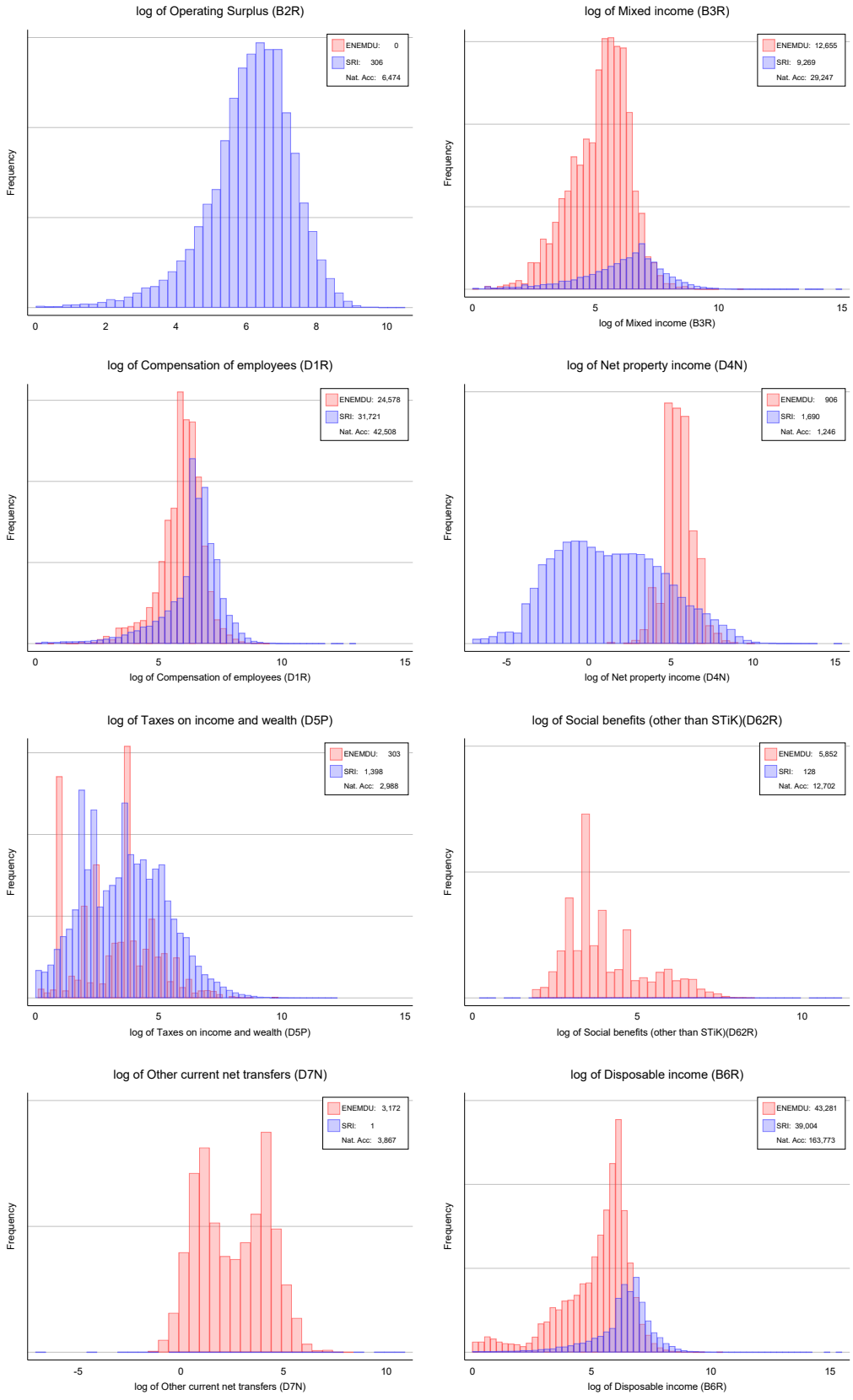
For the purposes of this study, the information from the Ecuadorian Central Bank (BCE) regarding national accounts can be categorized into three distinct periods, though not in chronological order:

2007-2020: The period from 2007 to 2020 represents the longest span with a consistent methodological base in the national accounts, using 2007 as the fixed base year. This "fixed series" provides detailed information for the institutional sectors (households, government, financial/non-financial corporations, rest of the world), and dissagregations of the different income items. However, the BCE limited the publication of disaggregations for the years 2017 to 2020, prior to the change of the base year.

2018-2022: In 2021, the BCE introduced a significant update to the national accounts, implementing a new base year methodology that transitioned from a fixed base year approach to a "mobile series" that creates a new base for each year (BCE, 2023). For consistency with the longer 2007-2020 series, I use the traditional GDP and GNI totals that were produced simultaneously for the years 2018-2020 and published by the BCE, and the totals from the World Development Indicators (WDI) for 2021-2022, when the BCE ceased publishing comparable numbers.[22] This approach ensures that the structure of institutional sectors and income items from the "fixed series"is applied to the total national income, allowing for deflation and comparison to the 2007 base year. One advantage of the new series is the increased disaggregation for certain items, such as operating surplus and mixed income (see figure A.7 (a)). Notable changes include an increase in employers' social contributions as a share of total compensation

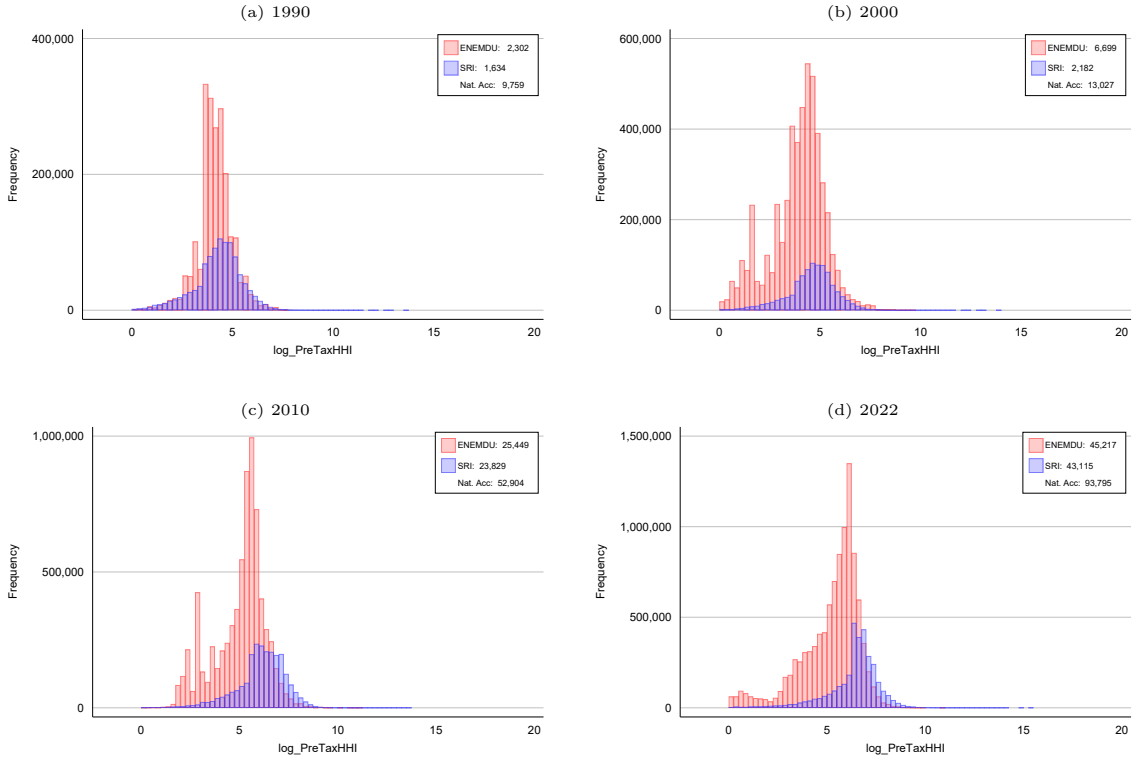
[22]The numbers from the WDI allign with BCE figures for the series 1990-2020. It is therefor assumed that the WDI figures for 2021 and 2022 were estimated using meethod from the "fixed series".

Figure A.5: Histograms for ENEMDU and SRI for different income components (2022)



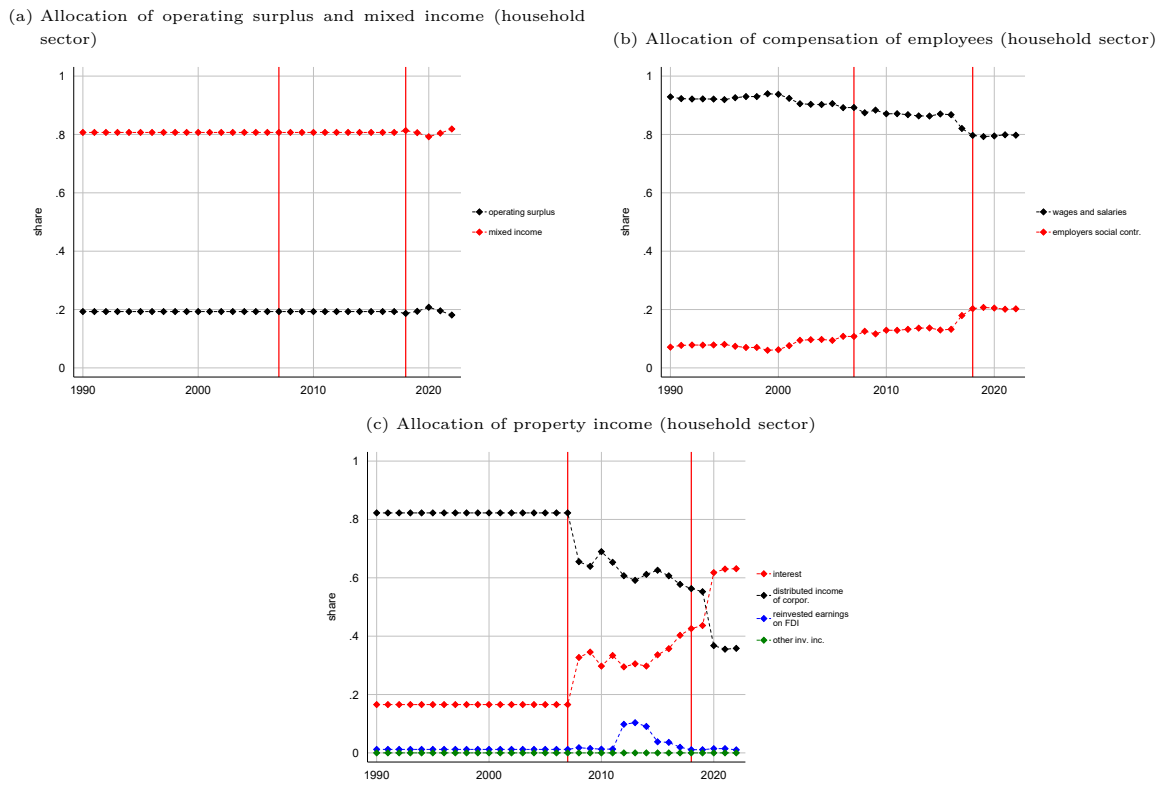
Sources: Own elaboration based on ENEMDU and SRI.

Figure A.6: Histograms for ENEMDU and SRI for pretax household income - different years



Sources: Own elaboration based on ENEMDU and SRI.

Figure A.7: Comparison of income components in three periods



Before 2007: Extrapolation of share of components. 2018: Change of base year.

of employees (see figure A.7 (b)), and a shift in the importance of interest income relative to distributed income from corporations within the property income component, a trend that began in 2015, as illustrated in figure A.7 (c)).

1990-2007: For the years before 2007, the BCE published only aggregate figures for GDP and GNI. To develop disaggregated indicators for these years, three main steps were undertaken:

- **Reproducing Component Ratios**: The relative sizes of each component and sector from the year 2007 were applied to the years prior to 2007. This involved using the sectoral and income component breakdowns from 2007 as a reference to estimate the distribution of income components for earlier years.

- **Adjusting with OECD Data**: OECD tax data, available since the 1980s, were used to adjust the income items for direct and indirect taxes, as well as social contributions. The four panels in figure A.8 illustrate that for the years after 2007, national accounts and OECD datasets report similar levels of taxes and social contributions. This consistency supports the use of OECD data for the years before 2007. The importance of incorporating OECD data becomes evident in the years before 2000, where a simple extrapolation of 2007 tax and contribution shares would overestimate state intervention in the economy. Specifically, taxes on products and imports (panel (a)), taxes on income and wealth (panel

(b)), institutional sectors (panel (c)), and social contributions (panel (d)) were often only half as significant before 2000 compared to 2007.

- **Incorporating Remittances**: An additional adjustment was made to address the economic impact of remittances, especially significant after the 2000 crisis. Figure A.9 compares the income concept "other current transfers" (D7) and the foreign remittances statistics from the BCE. Following the 2000 crisis, remittances became a major income source for many Ecuadorian families.[23] The similarity between "other transfers" and remittances series, particularly around the pivotal year 2007, suggests that a significant portion of "other transfers" are remittances. Since the two sources report the same amount for the year 2000, directly extrapolating the importance of "other transfers" from 2007 would overestimate remittances' role in the Ecuadorian economy. I therefore use the remittances series to extrapolate "other transfers" for years before 2007, which reduces the amount of this income source. The resulting discrepancy in income distribution, where less income is assigned to the household sector, is adjusted by reallocating it to the corporate sector, while the government sector figures have already been adjusted using OECD tax statistics.

A.3 DINA concepts

A.3.1 Unit of observation

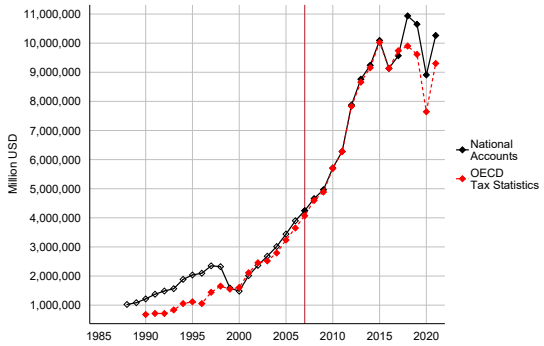
In the benchmark scenario, the unit of observation is the individualistic adult, defined as individuals aged 20 years and older, consistent with the World Inequality Lab (WIL) definition (Blanchet et al., 2020). This approach aligns with the observation that nearly all income is earned by adults. Within households, there are two primary methods for distributing income: individualistic adult method, where each adult is considered separately, and income is attributed to individuals based on their own earnings; and equal-split adult method, where income is divided equally among all adult household members, regardless of their individual earnings. Ecuadorian tax declarations are filed individually, which does not facilitate the representation of household income structures directly. Conversely, household surveys provide detailed individual income data for each member, therefore allowing for individual, narrow equal-split (spouses), broad equal-split (all adult members of the household), and per capita (all members, including children) income distribution.

Given the individualist tax register declarations, the individualistic adult is the starting point for the unit of observation. To account for certain income sources that are shared between members of the household I make use of the methodological improvement of this paper: the integration of the richest individuals from the tax register in the household structure of the survey (see assumptions for the integration of sources in section A.4). This method enhances the individualistic adult scenario and allows for the calculation of indicators for the equal-split adult scenario. However, some income sources primarily affecting the lower segments of

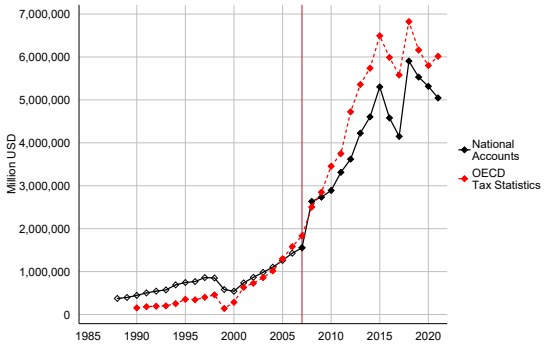
[23] Remittances arrived mainly from Spain, Italy, and the United States, and contributed up to almost 5% of national income after 2000.

Figure A.8: Comparison retropolation and OECD tax database

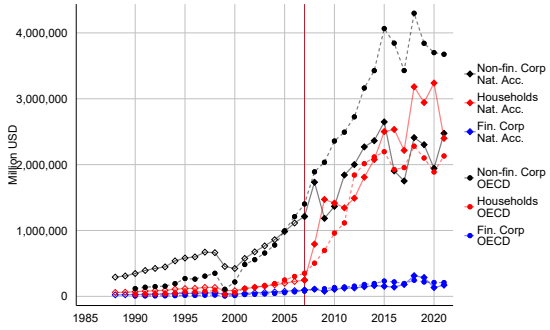
(a) Taxes on products and imports (D2).



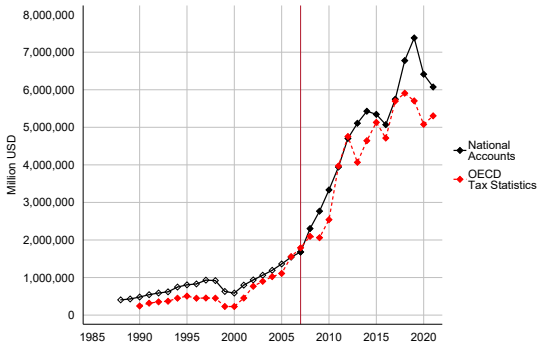
(b) Current taxes on income and wealth (million USD) (D5).



(c) Current taxes on income and wealth - different sectors (D5).

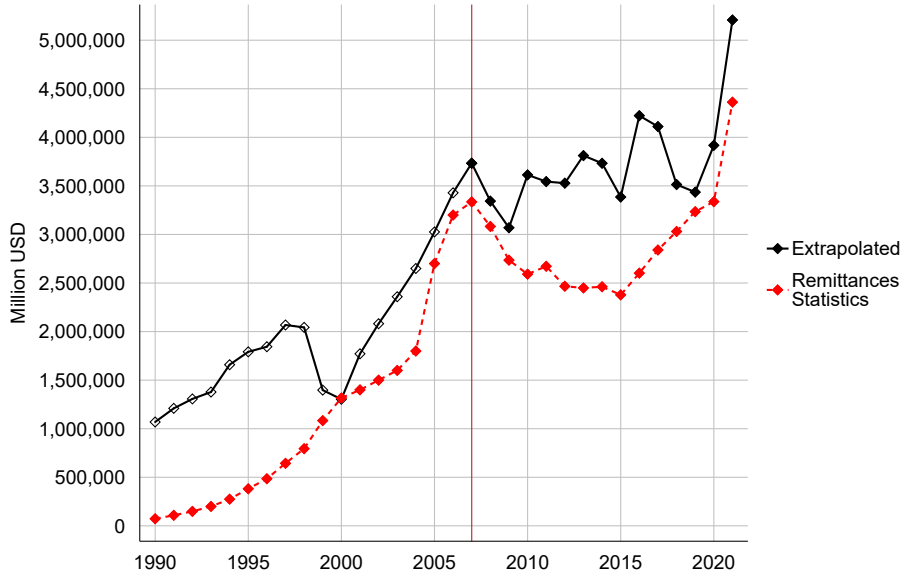


(d) Social contributions (to Gov. sector) (D61).



For National Accounts the values before 2007 are extrapolated from the distribution of GDP of 2007. Sources: BCE, OECD

Figure A.9: Comparison: Remittances and other transfers received (household sector)



For National Accounts the values before 2007 are extrapolated from the distribution of GDP of 2007.

Table A.2: Income concepts from National Accounts

Income component		N.A. code	sector
operating surplus	+	B2R	HH (S14)
mixed income	+	B3R	HH (S14)
wages and salaries	+	D11R	HH (S14)
property income received	+	D4R	HH (S14)
social benefits others than social transf.in kind	+	D62R	HH (S14)
other current transfers received	+	D7R	HH (S14)
pretax income of households	=	**PRTHHI**	**HH (S14)**

Income component		N.A. code	sector
operating surplus	+	B2R	HH (S14)
mixed income	+	B3R	HH (S14)
wages and salaries	+	D11R	HH (S14)
property income net	+	D4N	HH (S14)
primary income of households	=	**B5**	**HH (S14)**
other current transfers	+	D7N	HH (S14)
social contributions	-	D61P	HH (S14)
social benefits others than social transf.in kind	+	D62R	HH (S14)
primary income of corporations	+	B5	Corp. (S11+S12)
primary income of government	+	B5	Gov. (S.13)
pretax national income	=	**PRTNI**	
net taxes on production (taxes minus subsidies)	-	D2-D3	Gov. (S.13)
taxes on income and wealth	-	D5	HH (S14)
taxes on income and wealth	-	D5	Corp. (S11+S12)
collective consumption expenditure of the government	+	P32P	Gov. (S.13)
primary government surplus	+	D2-D3+D5-D63P-P32P	Gov. (S.13)
social transfers in kind	+	D63R	HH (S14)
posttax national income	=	**POTNI**	

Source: United Nations (2009)

the distribution cannot be fully attributed to individuals. For example, targeted transfers for low-income households or those with members with disabilities are often household-based and may not be directly received by the intended beneficiaries. Similarly, remittances reported at the household level in surveys might be distributed across the entire family or exclusively to children, depending on whether parents work abroad. In the benchmark scenario, the income is calculated for individualistic adults but social transfers and remittances are distributed equally among all adults within the household.

It is important to note that inequality indicators derived from the individualistic approach will, by design, show higher levels of inequality compared to official indicators. Official measures of inequality are typically based on *per capita* calculations, where household income is distributed equally among all household members, regardless of their age.

A.3.2 Income concepts

Table A.2 indicates the construction of the income variables used in this study. The integration process is applied exclusively to income from the institutional sector of the household (S14). To integrate the most comparable sources of income that are available in the survey and the tax register, the *pretax income of households* (PRTHHI) only aggregates positive income items from wage, labor, social benefits and transfers (including remittances). National income indicators extend beyond household income to include earnings from other sectors. *Pretax national income* (PRTNI) encompasses primary income of households, net remittances, the net of social benefits/contributions, and the primary income of the corporate sector (S11+S12) and the government (S13). *Posttax national income* (POTNI) further adjusts for taxes levied on households and the corporate sector, government expenditure, and social transfers in kind.

A.4 Integration of tax and survey data

The central methodological contribution of this paper is the integration of survey and tax microdata at the micro level. While the present study is not the first to combine tax and survey information (e.g. Bach et al., 2020; Bartels and Metzing, 2019; Blanchet et al., 2022a; De Rosa et al., 2020), this study uniquely applies this integration to the individual level, a gap that has not been addressed in the literature so far. The methodological shortcoming in literature can be explained by the different objectives and by the lack of data availability: On the one hand, studies that historically focus on the concentrations of income or wealth beyond official household survey results derive their results primarily from registers for countries with a long history of official tax records (Piketty and Saez, 2003; Garbinti et al., 2018). In these studies, survey information is only used to correct for demographic fallacies and undercoverage of the tax register, without the need to derive detailed income information from survey microdata. For example, research aiming to estimate the income concentration of the top 1% of adult population can use the register information to compared to the totals of national accounts. the availability of household surveys has been limited to recent decades (including the industrialized countries), which constrains long-term analyses to countries with comprehensive historical tax records.

Studies focusing on developing countries, on the other hand, have used different strategies to include the income of the richest individuals in the analysis of inequalities. Many of these countries either lack comprehensive historical tax records or have recently established them, often with restricted access for researchers. Consequently, studies in these contexts frequently rely on household surveys and focus on correcting for income data at the top of the distribution (De Rosa et al., 2022; Assouad et al., 2018; Alvaredo et al., 2019).Research by Bach et al. (2020), Bartels and Metzing (2019), and Blanchet et al. (2022c) suggests that simple tax tabulates - which are easier to obtain than micro data - can effectively approximate the income distribution derived from detailed tax registers when interpolation methods are applied. This approach is supported by empirical evidence showing that top income and wealth distributions often follow a Pareto distribution or power law (Atkinson, 2007; Clementi and Gallegati, 2005; Piketty and Saez, 2003). Recent contributions have proposed standardized and automated correction programs for household surveys with simple tax tabulates. the R package `gpinter` (Blanchet et al., 2022c) generates a microdataset with a continuous income distribution based on tabulated tax data, while the Stata command `bfmcorr` (Blanchet et al., 2022b) reweights survey microdata to align with tax information. This standardized approach not only facilitate the production of distributional national accounts across many countries with limited data but also prove effective for comparing inequality and concentration levels regionally.[24]

However, interpolation methods face a notable limitation: Using the information from the

[24]Examples are studies on the most unequal region of the world (Alvaredo et al., 2019), the income concentrations for Africa (a continent where there is very few distributional information Chancel et al., 2019), or the comparison of inequality in Latin America (De Rosa et al., 2022).

interpolated tax data to reweight the observations in the survey maintains the composition of income in each household and only creates more of the *survey rich* to compensate for the total income not covered by the survey and found in the tax tables. This means, that income characteristics of the *register rich* are not represented in the reweighted dataset. Specifically, higher property income or undistributed rents, which are often significant among the wealthiest individuals recorded in tax data but underrepresented in survey data, are not accurately reflected in the reweighted dataset.

In this study, I leverage privileged access to both tax and survey microdata for Ecuador to extend the `bfmcorr` command. This enhancement addresses the limitations inherent in traditional reweighting approaches, enabling a more comprehensive analysis of income dynamics, composition, and redistribution.

A.4.1 The extended `bfmcorr` command

The aim of the proposed integration method is to reduce the disadvantages of existing methods and to exploit as much information as possible from the available sources. Instead of using interpolated synthetic microdata, I use the real tax microdata to replace the upper tail of the distribution. This replacing procedure is achieved by first reweighting the observations from the survey with `bfmcorr`, and then by expanding the weighted observations from the rich households[25], so that the *individuals* living in these households can be replaced by the individual observations from the tax register. Up to this step the integration is exclusively undertaken on the individual level, meaning that the units of observation from the survey are individuals above the age of 20 years, and also tax declarations stem from individuals. The result of the integration procedure is observable in figure A.10, which exemplifies the procedure for the year 2022: The red histogram constitutes the log-distribution of the pretax household income from the survey, while the blue histograms reproduces the distribution from the integrated dataset, including the individuals from the tax register. By reducing the weight for the observations in the lower part of the distribution, the density for these income groups falls, while the introduced individuals from the tax register increase the density in the upper part of the distribution. Most importantly, the integrated distribution now includes observations with incomes that are out of the range in the survey.[26]

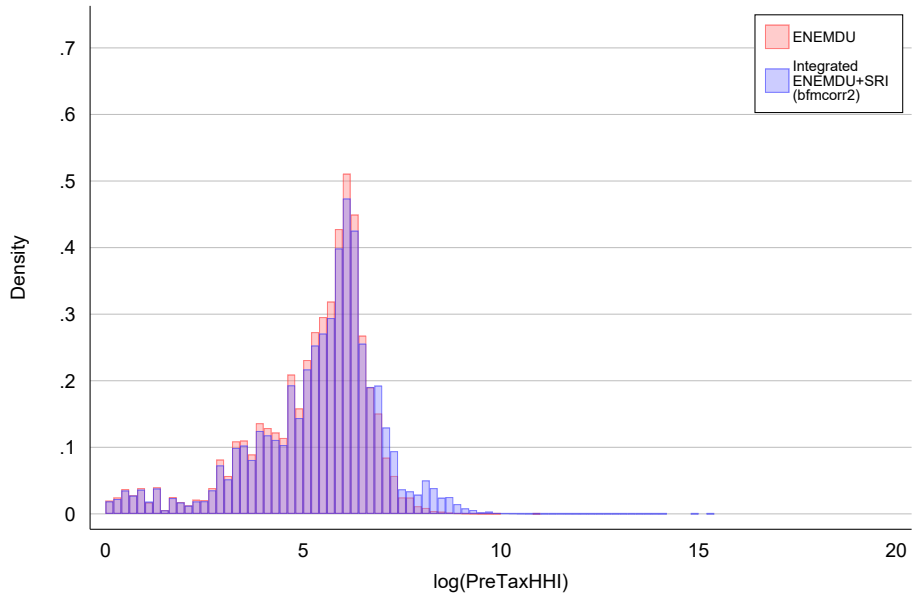
One important assumption for this integration process is that the ranking between individuals from survey and individuals from register has to be maintained[27]. This rank stability is not so much an important assumption on the individual level, but it needs to be made transparent that it is important at the moment of integrating individuals (with their corresponding

[25] For the case of Ecuador, one observation from the highest household income decile in the survey represents about 200 individuals.

[26] Note that the highest incomes have log values greater than 15, which is more than 3.3 million dollars in absolute terms and little increases in this histogram mean huge increases in absolute terms (the log of 17 is almost 25 million dollars).

[27] This means that the "poorest" individual taken from the tax register will replace the "poorest individual" from the survey that has to be replaced, and this is also true for the individuals at any position, up to the "richest" individual from both sources.

Figure A.10: Integration of ENEMDU and SRI on the microlevel



Sources: ENEMDU, SRI, BCE.

income characteristics) into the household structure from the survey. What is supposed here is that certain demographic properties of the individuals from the survey are "inherited" to the replacing individuals from the tax register. To put it in an example, the top 1% of the *register rich* will live now in the households of the top 1% *survey rich*, replace them as wives, husbands, cousins and children. Although this sounds like a strong assumption, it simply says that the household structure (number of people living within one household, age and sex structure) of the *register rich* is assumed to be the same as in the *survey rich* (see figure **??**. As the household surveys of the INEC are based on master samples derived from population census, demographic characteristics of the upper tail should suffer far less from underrepresentation than income characteristics.[28]

A second question when replacing observations at the top, is the point at which this replacing should start. Whereas the *merging point* (Blanchet et al., 2022b), the point at which the density of the distribution of tax register is higher than the density of the survey data, has been programmed in `bfmcorr` to be found by a data driven approach, the point at which replacing starts has to be defined manually. One extension implemented in `bfmcorr2` is the search for an optimal replacing point. The algorithm searching for the replacing point minimizes the difference between the total value from national accounts and the result of the integration method. The brackets for the potential points is given by the interval from the *merging point* up to at least 3% of the population replaced (percentile 0.97). The algorithm chooses the replacing point that results in minimizing the sum of differences for the strictly positive values of operating surplus and mixed income (B2R and B3R), wages and salaries (D11R), and property income

[28]Compare figure **??** in the main manuscript.

Figure A.11: Optimization results for selected years

Figure shows the results of equation 1 for replacing observations in the survey by tax register observations in the percentile interval [0.93,0.96]. The gap between national accounts and the integrated dataset is minimized at percentile 0.93.

received (D4R), as illustrated in equation 1. The formula does not include components that are added after primary income, as their quality in both sources decreases and imputations become more important.

$$\text{replacing point} = min\Bigg(\sum_{\text{merging point} \leq i \leq 0.97} \Bigg(\left| \frac{(B2+B3)_{BCE} - (B2+B3)_{bfmcorr2}}{(B2+B3)_{BCE}} \right| + \left| \frac{D11_{BCE} - D11_{bfmcorr2}}{D11_{BCE}} \right| + \left| \frac{D4R_{BCE} - D4R_{bfmcorr2}}{D4R_{BCE}} \right| \Bigg) \Bigg) \tag{1}$$

A.4.2 Use of survey information to improve coverage of integrated dataset

The integration procedure improves the coverage and reliability of the distribution of income. By integrating, nevertheless, some information from the survey is replaced by the tax register in the highest income strata, which is designed for income items better captured by the register (e.g., property income). Nevertheless, income items are better captured in the survey can get lost by the integration (e.g., imputed rents or social benefits others than pension payments). In addition, any income that is at least partially informal will only be captured in the survey. Within the SNA, this refers mainly to income from own account production, which is part of mixed income (B3R). Unfortunately, when applying the reweighting method, well captured information from the register does not find its way into the integrated dataset. On the other hand, our integration method `bfmcorr2` proposed so far replaces the information from the survey with tax registers, therefore also cutting out valuable information from well captured income components of the *survey rich*.

Second, recalling the representativeness of household surveys in demographic characteristics (see figure **??**), I assume that these characteristics of rich households are relatively well captured in the survey. On the other hand, the whole exercise of correcting distributional data is based on the assumption that underreporting and undercoverage of economic characteristics increases with higher deciles. It is therefore realistic that inheriting some economic components (the ones better captured in the survey than in the register) from the *survey rich* to the *register rich* will not create any upward bias of income in these higher income groups.

The method therefore imputes information from operating surplus and mixed income (B2R and B3R), property income (D4R), social benefits (D62R) and other current transfers (D7R) when their values are greater than the observed in the individual tax declaration.

A.4.3 Scaling to 100% of national accounts' household sector income

Some further assumptions are made to scale the income from the integrated data to the totals of the household institutional sector of national accounts. In a first step, general imputations based on legal information and accounting identities (such as D611=D121) are imputed. This affects only labor income, where the social contributions are determined by regulations (e.g. in 2022, employers pay 11.45% social contributions). Finally, imputed rents from the ECV survey are assigend to each fractile (see section A.1), and negative property income (D4P) and negative current transfers are distributed like pretaxhhi, given the lack of information about interests paid in surveys or tax data. All other concepts in HH income (S14) are directly scaled to 100% of the corresponding national account totals.

A.5 Imputation of National Income concepts

A.5.1 From pretax household income to pretax factor income

The integration of sources so far produces pretax income that covers 100% of the household sector in national accounts. To distribute incomes from the government and the corporate sector, this study imputes income components based on different assumptions. The first national income concept is pretax factor income (PRTFI), which includes undistributed corporate incomes (B5 from S11S12), income from government (see detail), and incomes from NPISH (very small).

- For the gov, primary income consists of B2 operating surplus (by definition close to zero, but as we have gross B2 and not net B2 - excluding depreciation - this value is always positive and small), D4 (property rent), and indirect taxes (D2R $\rightarrow$ value added taxes mainly, based on consumption) minus subsidies (D3R).

- Primary income from corporations B5 (which are not distributed) has to be distinguished between corporations owned by private and public. We use the share of D4N (property income) that goes to gov and to corp as a proxy for how much is private and public. I will ask the SRI for more detailed information on this distribution. The private share of B5 is distributed like capital income from households, as we view capital income as a proxy for stock ownership. For the public share we use the PreTaxHHI as proxy for the distribution.

- Primary income from the government is distributed differently: For operating surplus and for property income we distribute it like PreTaxHHI EXCLUDING IMPUTED RENTS (simplest solution and only level shifter, such as in the public share of B5 in S11S12). Also D2D3 (indirect taxes) on the Pretax level (for Posttax see below) is distributed following PreTaxHHI (uniformly) to reflect the purchasing power of pretax income. To sum up,

gov primary income is distributed like PreTaxHHI in the PRTNI level, to not disturb too much.

- Income from nonprofit institutions serving households (NPISH) is distributed via lump-sum.

A.5.2 From pretax factor income to pretax national income

Pretax national income (PRTNI) adds social benefits and subtracts social contributions.[29] Considering the importance in years after economic crises, it includes other transfers - which mainly contain remittances. Therefore, the sum of PRTNI is not exactly the same as PRTFI.

A.5.3 From pretax national income to posttax disposable income

Arriving at Posttax disposable income (POTDI) requires assumptions about how public policy affects the income distribution: First, all direct taxes (D5) and indirect taxes (D2) from PRTNI are deducted. For the distribution of corporate taxes (direct taxes of institutional sector S11S12) I assume that the owners pay the tax and use capital income, as suggested in Blanchet et al. (2020). For indirect taxes (which is mainly value added tax and importation tax), I use pretax household income as a proxy for how consumption is distributed.

A.5.4 From pretax national income to posttax national income

Posttax national income adds government activity to POTDI, by making explicit assumptions about government spending: health expenditure is distributed via lump sum, all other spending (no matter if individual or collective) are distributed following the pretax income distribution. In addition, government surplus is assumed to follow the distribution of property income.

Appendix B Supplementary graphs

References

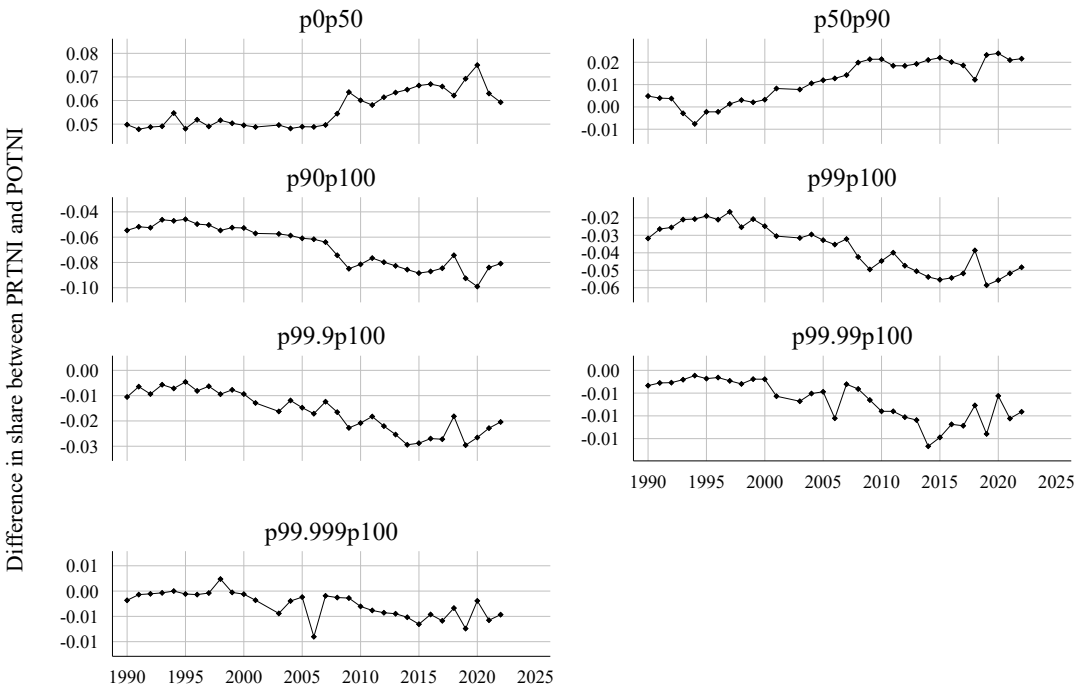
Alvaredo, Facundo, Lydia Assouad, and Thomas Piketty (2019). "Measuring Inequality in the Middle East 1990–2016: The World's Most Unequal Region?" *Review of Income and Wealth* 65.4, pp. 685–711.

Arias, Diana, Edwin Buenaño, Nicolás Oliva, and José Ramírez (2008). "Historia del Sistema Tributario Ecuatoriano 1950-1999". *SRI: Fiscalidad, Edición 2*.

Assouad, Lydia, Lucas Chancel, and Marc Morgan (2018). "Extreme Inequality: Evidence from Brazil, India, the Middle East, and South Africa". *AEA Papers and Proceedings*. Vol. 108, pp. 119–23.

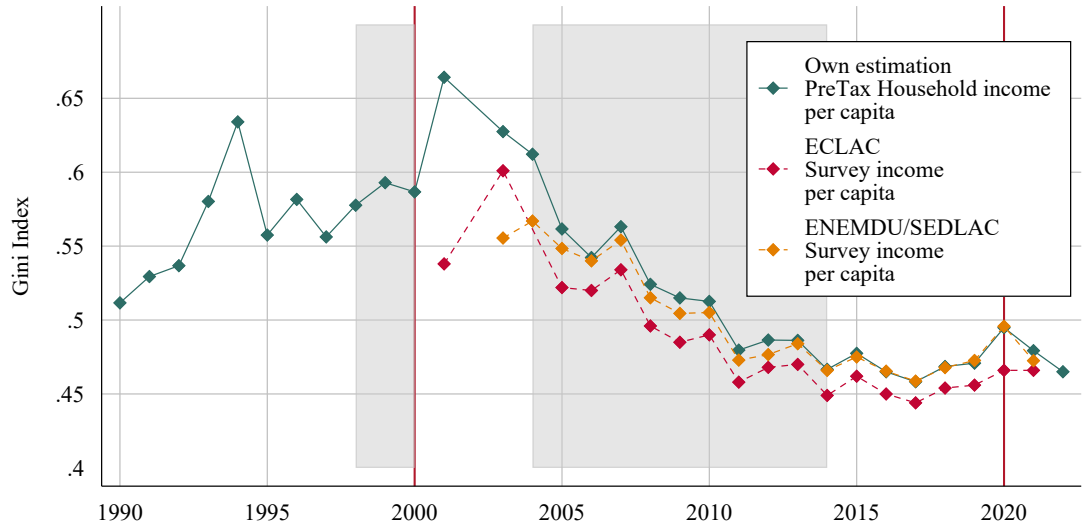
[29] The aggregate surplus of social contributions less social benefits is distributed neutrally (similar to government property income).

Figure B.1: Incidence of taxes and transfers



Source: Own elaboration based on ENEMDU, SRI, BCE.

Figure B.2: Comparison different sources based ENEMDU



ENEMDU: Labour force survey from the Ecuadorian NSO,
SEDLAC: Socio-Economic Database For Latin America and the Caribbean,
ECLAC: Economic Commission for Latin America and the Caribbean.
Pretax Household Income includes: Operating surplus and mixed income, compensation of employees, net property income, net social security benefits, and other current transfers.
Survey income for SEDLAC/ENEMDU and ECLAC slightly vary from these national accounts concepts.
Red vertical lines: 2000 Dollarization, 2020 COVID pandemic,
Shaded areas: 1998-200 Bank crisis, 2004-2014 Commodity price boom.

Note: Gini Index of income per capita (household sector).

Atkinson, Anthony B (2007). "Measuring top incomes: methodological issues". *Top incomes over the twentieth century: A contrast between continental European and English-speaking countries*, pp. 18–42.

Bach, Stefan, Charlotte Bartels, and Theresa Neef (2020). "Distributional National Accounts: A Macro-Micro Approach to Inequality in Germany". *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*. University of Chicago Press.

Bartels, Charlotte and Maria Metzing (2019). "An integrated approach for a top-corrected income distribution". *The Journal of Economic Inequality* 17.2, pp. 125–143.

BCE (2023). "Metodología de cálculo de los índices de volumen encadenados de las cuentas nacionales de Ecuador". *Banco Central del Ecuador:Cuaderno de Trabajo No. 138*.

Blanchet, Thomas, Lucas Chancel, Ignacio Flores, and Marc Morgan (2020). "Distributional National Accounts (DINA) guidelines: Concepts and methods used in the World Inequality Database". *updated and extended version of the DINA Guidelines first published as WID Working Paper 2016/2*.

Blanchet, Thomas, Lucas Chancel, and Amory Gethin (2022a). "Why is Europe more equal than the United States?" *American Economic Journal: Applied Economics* 14.4, pp. 480–518.

Blanchet, Thomas, Ignacio Flores, and Marc Morgan (2022b). "The weight of the rich: improving surveys using tax data". *The Journal of Economic Inequality* 20.1, pp. 119–150.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2022c). "Generalized pareto curves: Theory and applications". *Review of Income and Wealth* 68.1, pp. 263–288.

Cano, Liliana (2015). *Income mobility in Ecuador: New evidence from individual income tax returns*. Vol. 2015. UNU-WIDER. isbn: 978-92-9230-925-1.

Chancel, Lucas, Denis Cogneau, Amory Gethin, Alix Myczkowski, et al. (2019). "How large are African inequalities? Towards distributional national accounts in Africa, 1990-2017". *World Inequality Database (WID) Working Papers*, pp. 1–71.

Clementi, Fabio and Mauro Gallegati (2005). "Pareto's law of income distribution: Evidence for Germany, the United Kingdom, and the United States". *Econophysics of wealth distributions*. Springer, pp. 3–14.

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2020). "Inequality in Latin America revisited: insights from distributional national accounts". *World Inequality Lab – Technical Note* (02).

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2022). "More Unequal or Not as Rich? Revisiting the Latin American Exception". *World Inequality Lab – Working Paper N 2022/13*.

Ecuador, Banco Central del (2001). “"Las remesas de ecuatorianos en el exterior"”. *Cuaderno de Trabajo No. 130*.

– (2012). “"Retropolación de las Cuentas Nacionales 1965-2007"”. *Presentacion*.

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2018). “Income inequality in France, 1900–2014: evidence from distributional national accounts (DINA)”. *Journal of Public Economics* 162, pp. 63–77.

INEC (2015). “Metodología de la Encuesta de Condiciones de Vida ECV”. *Documento Metodologico*.

Piketty, Thomas and Emmanuel Saez (2003). “Income inequality in the United States, 1913–1998”. *The Quarterly journal of economics* 118.1, pp. 1–41.

Ponce, Juan and Rob Vos (2014). “Redistribution without structural change in Ecuador: rising and falling income inequality in the 1990s and 2000”. *Falling Inequality in Latin America. Policy Changes and Lesson, Oxford, United Kingdom*.

Ramírez, José and Paul Carrillo Maldonado (2020). “Indicator of the efficiency of value added tax and income tax collection in Ecuador”. *Revista CEPAL* 131.

United Nations (2009). *System of National Accounts 2008*. United Nations, European Communities, International Monetary Fund, Organisation for Economic Co-operation, Development, and World Bank: New York.