The Journal of Economic Inequality (2023) 21:537–572
https://doi.org/10.1007/s10888-023-09566-w

Beyond tax-survey combination: inequality and the blurry household-firm border

Mauricio De Rosa[1,2] · **Joan Vilá**[1]

Received: 25 May 2022 / Accepted: 2 February 2023 / Published online: 24 April 2023

Abstract
Inequality evidence based on surveys, tax records, or their combination often result in divergent trends, fueling the distributional debate in Latin America. Beyond the strengths and weaknesses of these sources and their combination, tax-survey data face two shortcomings: they are unable to account for aggregate household or national income, and they are affected by firm owners' decisions about the distribution of profits, changing which incomes researchers can actually observe. We combine social security data, household surveys and matched personal and firm tax records, which allows us to accurately account for all income sources, particularly capital incomes at the firm and individual level. Based on these unique data, we assess inequality trends in Uruguay, showing that increasing profit-distribution by firms pushes tax-survey top shares upwards, but that this trend is offset when undistributed profits are accounted for. These results call for caution when using tax-survey data without considering changes in profit-distribution.

Keywords Firms · Income inequality · National accounts · Latin America · Tax records

JEL Classification D31 · D33 · E01

1 Introduction

Survey and tax data are the most extensively used sources in the study of income inequality worldwide, and they stand at the epicenter of the debate on the recent evolution of inequality in Latin America. Yet, even if we assume that survey and tax data can be effectively combined—a big *if*—are they sufficient to assess trends in inequality?

There are at least two issues that should be kept in mind. First, tax-survey inequality estimates may be detached from key variables such as growth. The data sources upon

✉ Mauricio De Rosa
mauricio.derosa@fcea.edu.uy

Joan Vilá
joan.vila@fcea.edu.uy

[1] IECON, FCEA, Universidad de la República, Montevideo, Uruguay

[2] Paris School of Economics, Paris, France

which most research is based are not consistent, since growth is measured using macroeconomic aggregates from national accounts, while inequality estimates are based on tax-survey micro-data. Income reported in household surveys is usually subject to underreporting and undercoverage (particularly at the upper tail of the distribution), while tax records only include taxable sources of income. This causes micro-macro inconsistency between national accounts and micro-data sources which not only makes it difficult to properly address the question of how economic growth is distributed among income groups, but also may lead to biased trends if the gaps between sources change over time.

Second, even if all micro-macro gaps remain unchanged, and the micro-data captures a constant share of household income, tax-survey-based personal inequality estimates depend on decisions about the allocation of income between firms and households, affecting what can actually be observed by the researcher. If firm owners decide —because of the economic cycle, tax policy changes, or another reason—to withdraw more of their incomes from the businesses they run (i.e., they increase the distribution of profits, observed capital incomes at the tax-survey level mechanically increase, pushing inequality estimates upwards.

Capital is the single most challenging income source underlying these two issues. Alvaredo et al. (2022) show a large micro-macro gap in Latin American, mostly explained by capital incomes, both at the household and national income levels. This has consequences in the measurement of inequality and its changes over time, given the potential distributive impact of capital incomes kept at the firm level (De Rosa et al. 2022). Moreover, distinguishing capital incomes from the rest is difficult even at the tax-survey level—let alone imputing unobserved ones—and it depends on a firm's legal status and its owner's decisions (see e.g. Kopczuk and Zwick 2020; Smith et al. 2019). Adequately accounting for capital incomes therefore requires detailed data on firms and owners (WIL 2021), which is very rarely available (Fairfield and Jorratt De Luis 2016; Alstadsæter et al. 2017). Thus, the micro-macro gap and the blurriness of household-firm borders both impose major challenges when drawing conclusions about levels of inequality, and more importantly, about inequality trends, from tax-survey data alone. Yet going beyond tax-survey data entails heavy assumptions unless sufficient additional information is gathered.

In this paper, we attempt to overcome these challenges based on unique data that matches records from social security, household surveys, personal income taxes, and firm taxes, combined with national accounts. These data allow us not only to provide detailed personal capital income estimates, but also to match owners' and firms' administrative data to account for the complex interplay between owners and firms. We close micro-macro gaps – particularly sensitive to undistributed profits– to provide a national income inequality series, which mechanically pushes the income concentration upwards. However, we show that as firms distribute more dividends throughout the period, tax-survey based top shares increase, and this trend is offset when (decreasing) undistributed profits –i.e. capital incomes which were not re-invested nor paid as dividends– are accounted for. Including undistributed profits, thus increases the income concentration level but tempers its trend, while at the same time enables us to jointly study inequality and national income growth.

We aim to contribute to the inequality-trends debate in Latin America, which cannot be separated from the data controversy. Household surveys and tax data are a key input for any distributional study, yet they have significant drawbacks. They do not include all income sources and, in the case of tax data, do not account for the entire income distribution. Household surveys allow for a correct estimation of the incomes of most of the population, but might be subject to underreporting and undercoverage at the top of the income distribution (Bourguignon 2015; Lustig and et al. 2019). Conversely, the increasing use of tax records

to measure income inequality has resulted in improvements in terms of coverage of top incomes (Atkinson et al. 2011), but also has important caveats. For instance, changes in the tax system may create incentives to alter reported income through income shifting or deferment, tax avoidance, or tax evasion, problems that may be particularly relevant in the short term (Burkhauser et al. 2012; Goolsbee 2000; Piketty 2003). Not surprisingly, different institutions that produce inequality estimates report heterogeneous and often divergent results. Ferreira et al. (2015) and Lustig et al. (2016) review the main international information sources that analyze the evolution of inequality[1] and conclude that results differ across databases, both in levels and in trends, even when the welfare concept and inequality measures are held constant. This divergence increases when the estimate refers to a specific country and a short time frame.

Some of the drawbacks of both household surveys and tax data can be tackled by considering the totality of national income, which does not depend on the definition of taxable income and represents a standardized income concept, precisely defined by the System of National Accounts (SNA) and internationally accepted (United Nations 2008). Yet the task of accounting for all remaining incomes not included in tax or household surveys is challenging, since the gap between micro- and macro-based income estimates is large (Deaton 2005; Alvaredo et al. 2022). Given this important micro-macro gap, the potential improvements in the distributive results obtained depend, to a large extent, on the imputation assumptions used to distribute the missing income at the household level (Zwijnenburg 2022).

Efforts to obtain income inequality estimates consistent with macroeconomic aggregates have been performed for Latin American countries in the past (Altimir 1987), showing the difficulties and pitfalls of such an exercise. More recently, following the Distributional National Accounts (DINA) methodology (WIL 2021), an increasing number of DINA-based estimates for both developed (Blanchet et al. 2019; Piketty et al. 2018; Garbinti et al. 2018) and developing countries (Piketty et al. 2017; Piketty and Chancel 2017; Novokmet et al. 2018; Morgan 2017; De Rosa et al. 2022) have emerged. We build on Burdín et al. (2022), who put together a tax-survey micro-database matching social security data (formal labor incomes and pensions), personal income tax data (detailed personal capital incomes), and firm tax data (untaxed firm income withdrawals by firm owners and incomes from pass-through corporations), accounting for over three quarters of the adult population, which is rare in Latin America. The remaining population and informal incomes were added using household survey data and a sub-sample of matched tax-survey individuals. In this paper, we supplement this tax-survey dataset with national accounts data to account for micro-macro gaps, coupled with novel firm-owner matched data to impute undistributed profits. This allows us to account not only of the incomes accounted for in the combination of administrative and survey data, but for the totality of household sector and the net national income, which prove to be critical for the trend of inequality.

The contributions of this paper are threefold. First, we document micro-macro gaps for the Uruguayan case based on recently-published national accounts data, showing decreasing gaps between tax-survey data and national accounts estimates, in contrast to what is found for most Latin American Countries (Alvaredo et al. 2022). More specifically, we show that this is the result of increased profit-distribution by firms, observed both in national accounts

[1] CEPALSTAT, Income Distribution Database (IDD), LIS, PovcalNet, Socio-Economic Database for Latin America and the Caribbean (SEDLAC), "All the Ginis" (ATG), the World Income Inequality Database (WIID), and the Standardized World Income Inequality Database (SWIID).

and at the microdata firm level. This increase on reported dividends pushes the top income shares in tax-survey data upwards, mirrored by decreasing undistributed profits, which offsets the surge in top income shares. This contributes to an understanding of the divergent trends between national income distribution and micro-data-based inequality. Our detailed account of the evolution of tax-survey income, household income, and national income distribution supports the overall conclusion that inequality in Uruguay has decreased, a conclusion further supported by national income estimates.

Second, our data allows us to account for capital incomes with significantly more precision than other studies for Latin America. By using detailed administrative microdata on most capital incomes (including in particular dividends and rents), we do not need to rely on survey adjustment methods which often produce divergent results (Blanchet et al. 2022; Jenkins 2017; Alvaredo 2011), thus resulting in more straightforward and credible estimates. Furthermore, the unusual owner-firm database we compiled for this paper allows us to impute undistributed profits in an almost surgical way. This contrasts with other studies for Latin America which rely on proxies based on corrected-survey data to impute this key mass of incomes (see e.g. De Rosa et al. 2022).

Third, we provide estimates of income distribution across the different steps, documenting that the top 1% income share is up to 15-20% higher in the national income series than what tax-survey estimates show. While the level of inequality is higher in the national income series, its trend is actually decreasing, as opposed to the increasing pattern of the tax-survey series. This is the results of imputing a decreasing share of undistributed profits –which are by definition not accounted for in the tax-survey data. Moreover, the micro-macro consistent income definitions allow us to perform two additional exercises. First, we show that income growth was lower for top incomes groups, only once the totality of national income is accounted for. Second, we compute effective tax rates, combining corporate and individual income taxes (Saez and Zucman 2020). The strong concentration of capital incomes, along with a dual income tax system, implies a loss of progressivity of direct income taxes for very high-income groups at the household income level. However, when firm owner data is used to impute corporate taxes, progressivity re-emerges at the national income level.

The paper is organized as follows. Section 2 describes recent inequality trends and data sources. In Section 3, our estimation procedure is presented, mapping and documenting data gaps across sources. Distributional results are discussed in Section 4, and Section 5 concludes.

2 Background and data sources

2.1 Recent trends

Although in the European context Uruguay might be considered a relatively high-inequality country, historically it has been among the least unequal countries in Latin America. After decades of unstable economic growth and recurrent economic crisis, it sustained an average annual growth rate of about 4.7% between 2004 and 2016. This economic growth, coupled with a series of relatively large labor market and tax and transfers system reforms implemented by a center-left coalition in office from 2005 to 2020, resulted in a significant decline in income inequality.

These reforms included a major increase in the minimum wage, the restoration of centralized collective wage bargaining, an expansion in both the coverage and the amount of

noncontributory cash transfers schemes, and the introduction of progressive labor income taxation (Amarante et al. 2014; Bucheli et al. 2013). Studies based on household surveys have consistently shown that income inequality experienced a rapid decline between 2008 and 2012, illustrated by a fall of about 7 points in the Gini index (see Fig. 9 in the appendix), followed by relative stagnation from 2013 to 2016 (Cornia 2014; Alvaredo and Gasparini 2015; Gasparini et al. 2018).

The use of tax data as an alternative database shows a decline in overall inequality measured by synthetic indexes such as the Gini or Theil, though less steeply and from a higher level than in survey data. Conversely, in tax data, top income shares show stability and a slight increase of about 15-16% between 2009 and 2016, but a drop from 11.6 to 8% in survey data (Burdín et al. 2022).

2.2 Administrative micro-data

2.2.1 Individual's tax records

The incorporation of a dual income tax in 2008 allows us to obtain detailed tax microdata records for the period 2009-2016, which are the main data source for this study. This high-quality database includes formal labor and capital incomes, as well as pensions. In the case of labor income and pensions, the information comes from matched tax-social security records, so it includes the whole universe of workers contributing to social security, independent of whether they are net taxpayers or not. Comparisons to household surveys and population projections show that income tax records account for approximately 75% of the adult population and 80% of workers. In the latter case, the discrepancy corresponds to informality (see Burdín et al. 2022 for details).

Most sources of labor incomes and pensions are taxed by a progressive scheme (*Impuesto a la Renta de las Personas Físicas*, IRPF-II and *Impuesto de Asistencia a la Seguridad Social*, IASS). Taxable sources of labor income include wages, salaries, commissions, overtime payments, vacation payments, annual leave, end of the year payments, and any other payments received from employers. Unemployment, illness and maternity subsidies, accident insurance, unemployment benefits, and child allowances are excluded from taxable income.[2]

The dual scheme of taxation also includes a flat personal capital income tax (*Impuesto a la Renta de las Personas Físicas*, IRPF-I) with different tax rates according to the taxable source (see Table 4).[3] Capital incomes are divided into rents from real estate and leases, and financial and profit rents. This second group includes all cash or in-kind rents coming from bank deposits and other financial assets, business profits and utilities distributed by those firms contributing to corporate income tax, and copyright, among others. Banks, real estate agencies, and institutions in charge of payments are set as withholding agents in most cases; if not, individuals must file a tax return. Capital gains, although available, are not included. On top of being the standard procedure in the literature (Atkinson and Piketty 2007), they also present a very erratic evolution and, more importantly, unlike remaining

[2]The tax rates on personal income (IRPF and IASS) are shown in Table 4.

[3]In the case of capital income, it is exempt from taxation for those individuals who have housing rents whose annual value is below USD 5.000 and public debt interest, gains obtained from private capitalization pension accounts, and business profits distributed by firms with total annual revenue lower than USD 500.000 (4 million indexed units).

incomes or even undistributed profits, which are flow variables, capital gains are closer to a stock variable, insofar it represents an asset valorization.

For all sources of income, most taxes are collected on an individual basis, and households are not identified.[4] For this reason, in this paper we use the individual as our unit of analysis. We believe that this definition is the most accurate description of reality that we can obtain given the data restrictions, but we should stress that it is insufficient. In particular, due to the nature of the tax records, we are not able to analyze household-level incomes and their distributional consequences. However, Burdín et al. (2022) showed that per-capita household and individual income inequality trends are very similar in the household survey (although their levels are not), and also mirror the tax data inequality pattern.

The usual caveats of this type of data, namely tax evasion and avoidance, may affect distributive results (Atkinson et al. 2011). In particular, if higher income individuals, who have access to more sophisticated ways of eluding taxation actually do so, tax-based inequality estimates may be biased downward. Torregrosa-Hetland (2020) for instance find evidence for Spain that indicates that evasion in capital incomes reaches up to 30-50%, and 20% for self-employment incomes. Taking this potential bias into account, the results should be considered a lower bound, especially regarding top income shares.[5]

2.2.2 Firm's tax records

As a second source of information from tax records, in this paper we use microdata from firms that pay corporate income tax (IRAE).[6] The data includes the amount of total profits firms report, which is equivalent to the sum of profits distributed, undistributed, and paid to the rest of the world. A single firm identifier allows us to merge the universe of firms with the micro-database of income earners, identifying from which firm each of the individuals receives salaries and dividends.

However, the main challenge is to be able to allocate the results of the firms that are not distributed as dividends, i.e. the undistributed profits. We use an ancillary social security records database which identifies individuals that report being firm owners—i.e., partners of limited companies and other firms, directors and owners of small enterprises between 2009 and 2015. We then use this register to identify firm owners in the merged firms-individuals database. This entails assuming that owners receive salaries or dividends from their firms, and that the owners of each firm are entitled to the same share of the firm's profits when more than one owner exists. The first one is a relatively safe assumption, but could potentially exclude owners who did not receive incomes from the firms they own (hence not appearing in the merged firms-individuals database). Regarding the second assumption, results are unchanged if the share is assumed to be proportional to the amount of incomes received by each individual.

We are able to identify the owners of 59-65% of firms with undistributed profits and impute these profits to them (see Table 1, panel A). For the rest of the firms that report

[4] Joint taxation of couples is allowed but rather rare, less than 2% of total formal workers in 2016.

[5] The assumption used below for scaling incomes assumes that most of the error comes from underreporting but not from individuals reporting zero incomes when they actually receive them, may imply a bias in the opposite direction. However, we consider that the effect of this bias is limited (see Section 3.2).

[6] Firms with annual revenues above approximately USD 500.000 are obliged to present annual balance sheets (around 60% of registered firms), and pay 25% of IRAE over their net operating surplus. Firms with annual revenues under USD 500.000 pay a lump fixed tax. For this subset of firms, it is not possible to recover the mass of undistributed profits, so they are not included in the national income series built from micro-data.

profits and for which we did not identify a shareholder or owner, we implement a probit model of the probability of ownership. Table 2 shows the marginal effects of this probit model by year, including socio-demographic characteristics, sources of income, and ranking in the overall income distribution. From the probability predicted by this model, we create a new owner for each firm with positive undistributed profits.[7] As robustness exercises, we implement different alternatives for the imputation of undistributed profits for this subgroup of firms. On the one hand, we impute this mass of income to the top wage earner in each firm, and in a second alternative we create new individuals in our database whose only source of income is the undistributed profits. As the estimates in the Fig. 18 in the appendix show, the results are not affected by the assumption made for these imputations.

Table 1 presents descriptive statistics of the result of our firm-owners merge and our preferred imputation procedure. The number of firms with positive results, and the number of individuals receiving undistributed profits, increases towards the end of the period. Based on our preferred alternative, only 14% of the total recipients of undistributed profits do not belong to our matched firm-owners and hence were newly created for the imputation of this income. The average income received by these individuals is significantly higher than the average of our matched owners (panel D vs. panel C). However, the average income of matched owners is affected by a low number of firms that compose a large part of the recipients, while the difference in the median and other statistics is considerably smaller.

In short, the matched firms-individuals data allows us to allocate the undistributed profits from micro data to individuals in the tax-survey database, for whom we already have all remaining formal and informal income sources. The possibility of identifying owners in matched firms-individuals data is very rare, giving us the opportunity to contrast the results obtained by this more precise approach with usual imputation methods in this literature (WIL 2021).

2.3 Household surveys

The second source of micro-data comes from household surveys (*Encuestas Continuas de Hogares*, ECH) for the entire period (2009-2016). These surveys collect information on socioeconomic variables and personal income for each member of the household. After-tax labor income includes cash and in-kind earnings for salaried workers, self-employed, and business owners. Information is separately recorded for the main occupation and additional ones. Salaried workers are also asked whether they contribute to the social security system, information which is used to identify informal earnings from this data source. Transfer income is collected for each individual, and survey questions disclose their origin (public/private, domestic/foreign) and the type of benefit: pensions (retirement and survival), contributory and noncontributory child allowances, unemployment insurance, accident compensation, or other benefits.

Except for profit withdrawal in the case of the self-employed and business owners, capital income is reported for the household as a whole, and hence, individual information cannot be recovered. In these cases, we split incomes equally among the adult members of the household to maintain our individual-based analysis. Interest, dividends, rents, benefits, and the imputed value of owner-occupied rental income are gathered in separate questions.

[7] The median owners on the merged of firms/owners database is 1, which justifies this assumption. We replicate the estimations creating a number of owners as the average number of owners in the firms (approximately 3 per firm), without relevant changes on main results.

Table 1 Summary Statistics of undistributed profits recipients: matched and imputed individuals (probit model)

	2009	2010	2011	2012	2013	2014	2015	2016
Panel A: Total recipients of undistributed profits (firms)								
Number of firms	17,043	17,869	19,399	20,724	20,994	21,916	21,687	20,834
Matched recipients	11,115	11,730	12,438	12,921	13,032	13,170	12,869	12,279
Imputed recipients*	5,928	6,139	6,961	7,803	7,962	8,746	8,818	8,555
Matched recipients (%)	65.2%	65.6%	64.1%	62.3%	62.1%	60.1%	59.3%	58.9%
Panel B: Total recipients of undistributed profits (individuals)								
Number of recipients	43,966	45,676	50,435	52,526	53,540	55,435	58,809	60,672
Mean income (USD)	101,913	144,614	136,564	150,571	174,775	184,326	163,446	120,768
p25	2,049	2,664	3,383	3,665	4,167	4,142	3,725	3,431
p50	9,494	11,921	14,804	14,645	15,868	16,196	13,851	12,752
p75	37,445	46,042	53,480	55,281	58,972	61,416	49,985	45,694
Panel C: Matched recipients of undistributed profits (individuals)								
Number of recipients	38,038	39,537	43,474	44,723	45,578	46,689	49,991	52,117
Mean income (USD)	70,684	95,996	100,914	117,344	119,182	132,343	105,530	74,069
p25	1,632	2,148	2,920	3,262	3,359	3,708	3,112	2,854
p50	8,970	11,271	13,690	13,093	14,805	14,178	11,967	10,972
p75	37,693	46,417	51,218	52,933	56,333	54,966	43,385	41,278
Panel D: Imputed recipients of undistributed profits (individuals)								
Number of recipients	5,928	6,139	6,961	7,803	7,962	8,746	8,818	8,555
Mean income (USD)	302,302	457,727	359,209	341,016	493,015	461,824	491,780	405,258
p25	4,556	5,676	7,295	7,149	8,401	8,666	8,592	8,265
p50	12,324	15,706	21,306	21,556	25,493	28,270	26,975	25,524
p75	36,729	44,215	68,302	71,385	84,014	99,845	101,360	90,808

Note. Own estimates based on firm tax data and individual tax records (DGI). The table presents the imputation method of undistributed profits based on matched owners-firms. Panel A depicts the total number of firms who report positive undistributed profits. Panel B displays individuals who receive undistributed profit in our final base. Panel C shows only the individuals for whom it was possible to match firms with individuals, while panel D includes the imputed undistributed profits from the probit model. Amounts in current dollars, at the average exchange rate of each year

Capital income sources are reported on an annual basis; only the imputed value of owner-occupied housing is gathered for the month previous to interview.

2.4 National accounts

National accounts estimates are provided by the Uruguayan Central Bank (BCU) and have very recently improved from a very low baseline. Uruguay's national accounts present estimates of gross national income based on the expenditure and production approaches, but not on the income approach, except for the newly available estimates for 2012 and 2016. Before this, the last time BCU updated the income generation account was 2005, and estimates by institutional sector have not been available since the late 1990s.

Thus, the full national accounts for these two years are the key macro-data inputs for our analysis. They present an adequate (though far from perfect) level of detail required to match and scale income concepts from tax-survey data to household sector incomes (see Section 3.2), and then on to national income. For years other than 2012 and 2016, a stable share of income components (both income sources and institutional sectors) is assumed, i.e., a simple backward interpolation is performed. Results do not change under alternative imputation procedures, given the relative stability of the estimates across both years. Incomes are presented gross of consumption of fixed capital, and therefore gross incomes were adjusted based on Mexican and Chilean data (i.e., share of Consumption of Fixed Capital, by income component and institutional sector, taken from Wid.World) to produce a net national income series.

3 Estimation steps

We estimate and compare inequality series based on (i) a combination of personal tax and survey data (*tax-survey series* hereafter), equivalent to the totality of income captured by these micro-data sources; (ii) a *household income* inequality series; and a (iii) *national income* inequality series (with a robustness check). These steps are depicted in Fig. 1.

By construction, aggregate incomes from the first step are conceptually equivalent to household sector incomes from the second step, with differences resulting from a measurement mismatch. In contrast, incomes from the third step are not supposed to be captured by tax-survey data, as they are accrued by other institutional sectors (government or corporate sector). Aggregate incomes corresponding to each series are depicted in Fig. 10 in

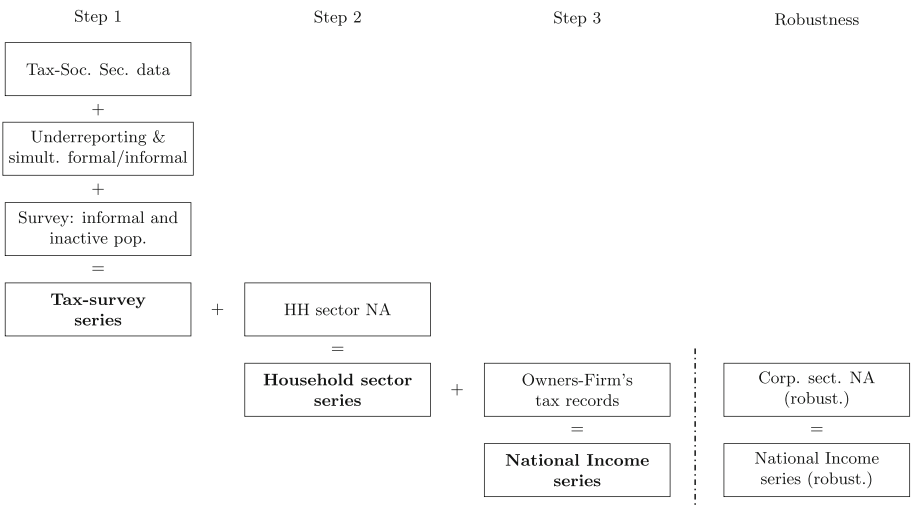


Fig. 1 Overview of the Method. ***Note.*** Own elaboration. Step 1 represents the construction of the combined tax-survey income series; Step 2 scales up to national accounts' household sector; while Step 3 uses owners-firm's administrative records to impute undistributed profits reported by firms, and scale remaining incomes proportionally to match national income. This third step is also computed –as a robustness check– based on national account's estimates of undistributed profits, imputed based on a proxy of capital incomes, i.e. the distribution of dividends plus interest from deposits

Table 2 Marginal effects of the probit model of owning a firm, by year

Marginal effect	(1) 2009	(2) 2010	(3) 2011	(4) 2012	(5) 2013	(6) 2014	(7) 2015
Male	-0.001***	− 0.000	− 0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wage earner	− 0.105***	− 0.096***	− 0.102***	− 0.094***	− 0.081***	− 0.082***	− 0.071***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Capital recipient	− 0.014***	− 0.016***	− 0.017***	− 0.021***	− 0.020***	− 0.022***	− 0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Pensioner	0.043***	0.053***	0.049***	0.068***	0.065***	0.059***	0.058***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Dividends recipient	− 0.005**	− 0.018***	− 0.023***	− 0.020***	− 0.015***	− 0.016***	− 0.014***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log of wages	0.004***	0.004***	0.004***	0.003***	0.003***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log of capital incomes	0.003***	0.004***	0.004***	0.004***	0.004***	0.005***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log of pensions	− 0.005***	− 0.006***	− 0.006***	− 0.008***	− 0.008***	− 0.007***	− 0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income percentile	− 0.000***	− 0.000***	− 0.000***	− 0.000***	− 0.000***	− 0.000***	− 0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Top 10	0.012***	0.011***	0.009***	0.007***	0.008***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Top 1	0.014***	0.014***	0.014***	0.016***	0.018***	0.018***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1,128,717	1,144,457	1,237,391	1,211,516	1,264,499	1,291,868	1,330,784

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Note. Own estimates based on firm's tax data and individual tax records (DGI). The table presents the marginal effects of a probit model of the probability of owning a firm, by year. Our ancillary database do not include information on the category of owners in 2016, so we use the marginal effects of 2015 for the imputation of that year

the appendix. The ratio of household income to net national income is relatively stable and close to 87-89%[8], which contrasts with the increasing share of the tax-survey income both in national income (almost 10 percentage points) and as a share of household income. In the following subsections, we address the estimation of each of these steps, discussing the reliability of the data and pondering alternatives.

[8]In the unadjusted national accounts, which are gross of consumption of fixed capital, the household sector represents 81 and 82% of gross national income for 2012 and 2016 respectively.

3.1 Tax-survey series

The starting point for this analysis is the tax-survey data base, which is a combination of tax, social security, and household survey data. The matched tax-social security micro-data accounts for over three quarters of the adult population, providing detailed data on total formal labor, pension, and capital incomes. Thus, on top of the typical avoidance and evasion caveats of tax data discussed in Section 2.2.1, there are three additional issues with this dataset: (i) it does not include individuals with purely informal income or with no income are all; (ii) formal incomes from the low income earners who are captured in the tax data are underreported (Flachaire et al. 2022); (iii) some formal income earners in the tax data may simultaneously earn informal incomes. To deal with them separately, we proceed in the same way as Burdín et al. (2022) and implement three adjustments to this database to build a series that is representative of the population as a whole and includes all income sources.[9]

First, individuals who lack income or who receive incomes from purely informal sources in the household survey are appended to the administrative database. The addition of this population to tax-social security database may not result in a total population that exactly matches census-based population projections. Thus, it is re-weighted assuming that individuals without earnings are correctly captured by the survey, and therefore only adjusting pure informal income earners. The re-weighting adjusts the added population about -30% on average. Second, to adjust for underreporting in the tax data, which is particularly high in the first two deciles of the income distribution –up to the median–, we use the ratios from a sub-sample of survey-tax matched households (Flachaire et al. 2022).[10] This procedure increases formal incomes of about the bottom 50% of the tax distribution. Third, corrections for simultaneous formal/informal income earners come from the household survey, using income thresholds from tax records, i.e. taking the ratio of informal-to-formal incomes in each formal income bracket in the survey, and applying those ratios to the formal incomes from the tax records.

For this article, we also add to this dataset all remaining informal and untaxed incomes that are not included in the fiscal income series but that are part of household income in the national accounts. To impute these sources of income, we use household surveys, matching both databases according to the position of individuals by income in the databases. Among the main income sources included in this step are cash transfers to households and owner-occupied rental income. Given the lower concentration of these sources with respect to the distribution of total income, the series obtained in this step show lower levels of inequality than those presented in Burdín et al. (2022).

3.2 Household income series

In order to account for all the sources considered in household sector incomes, the first step is to group tax-survey incomes in categories that match conceptually with national accounts definitions. This is done in Table 3, in which incomes are grouped in five categories: salaried work (wages), housing rent, investment income, non-salaried work (mixed), and benefits. Income components do not match exactly, especially in the cases of investment income

[9] For a full discussion of alternative methodological decisions, please see the original article Burdín et al. (2022).

[10] As mentioned above, there may also be underrporting in the tax data for higher income earners due to avoidance or evasion, even if not visible in the comparison with the survey. The implicit assumption is that tax data adequately captures higher incomes groups, but results are likely to be upwardly biased.

Table 3 Mapping households' income-concepts across data sets

	Tax-survey	Household sector national accounts	Observations
Investment income	dividends (personal tax data), interest (survey), owner withdrawals (firm tax data)	D4 = D41 + D49 (property income)	Dividends (D42) included in D49, but also rent of natural resources (D45) and investment incomes from insurance, pensions, and investment funds (D44)
Wages	Formal (tax) and informal (survey) wages	D1-D61 (compensation of employees minus net social contributions)	
Housing rent	Rent of owner occupiers (survey) + rental income (tax)	B2 (operating surplus)	Includes rental income from non-dwellings
Mixed	Self-employed income (survey + tax)	B3 (mixed income)	Does not include rental income from non-dwellings
Benefits	Pensions (tax)	D62 (social benefits)	

Note. Own elaboration based on similar table in Alvaredo et al. (2022). The first column depicts broad income concepts. The second and third describe their equivalents in the Tax-Survey data and their codes in SNA (along the SNA term for each). D4, D1-D61, B2, B3 and D62 are incomes received by household sector, named S14 in the SNA terminology. The fourth column lists incomes that do not exactly match. Sources: based on (United Nations 2009) and (OECD 2013). All incomes are gross of capital depreciation

and non-salaried work, for which mismatch is higher (for a full discussion, see (Alvaredo et al. 2022)). Nevertheless, at that level of aggregation, the correspondence is high and it is therefore possible to compare incomes from both sources.

In the case of investment incomes, household sector aggregate income is likely to include rent of natural resources and investment income from insurance, pensions, and investment funds, which do not match incomes in the tax-survey database. Tax-survey housing rent includes rental income from non-dwellings, which should be included in mixed incomes. Pensions and wages, on the other hand, can be conceptually linked without major mismatches.

Figure 2 reports the scaling factors for each type of income, i.e., the factor by which tax-survey incomes should be multiplied in order to yield SNA-household incomes. Most scaling factors are close to one, which means that tax-survey and household sector aggregates are of the same orders of magnitude. In the case of mixed incomes, the scaling factor is around 1.5 and gets close to 2 for some years, while in the case of rents, tax-survey data represents a higher value than its household income correlate. However, the scaling factor that stands out is that of investment income, which starts the period at 7-8, and slowly decreases thereafter until it stabilizes close to 3-4.

Large gaps between micro-data from administrative records or surveys and macro aggregates from national accounts are not rare in the developing world (Deaton 2005). Assuming

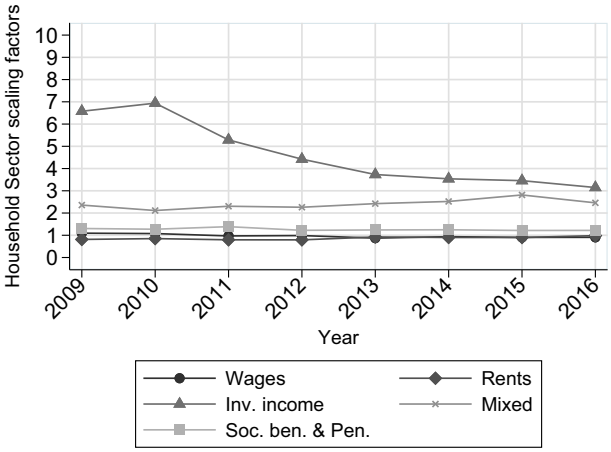


Fig. 2 Scaling factors, 2009-2016. ***Note.*** Scaling factors of tax-survey data vs household aggregates based on Table 3. Own estimates based on tax-survey data (DGI-ECH) and National Accounts 2012, 2016 (BCU). A scaling factor higher than 1 shows that the National Account's household income aggregate is larger than its counterpart in the tax-survey data. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries. See point estimates in Table 5 in the appendix

national accounts as the benchmark, such gaps could be entirely driven by underreporting in surveys and administrative records, and also by tax evasion and avoidance. However, it may also be the case that national accounts themselves are not accurately estimated, which is not easy to assess given the relative opacity of this source. Thus, we take a more agnostic stand and simply acknowledge these gaps and try to bridge them, presenting at the same time series with and without scaling.

Given the heterogeneity of scaling patterns across income sources, alternative adjustments were performed. For all but investment income, tax-survey incomes were adjusted by the corresponding scaling factor, so that aggregates are, by construction, equivalent to household sector incomes. The implicit assumption is that the gap is mainly explained by underreporting, i.e. that it does not result from individuals reporting zero incomes when they actually receive them. This assumption may overstate inequality if in some income categories the latter mechanism is at play, which was found to be true in the case of transfer programs in the United States (Meyer et al. 2015). However, given the broad income concepts from Table 3 –with the exception of investment incomes–, it is less likely that income recipients who report zero income in all of each concept sub-categories are found, thus reducing the potential bias of the assumption.

However, in the case of investment income, scaling it up would entail dramatically increasing the incomes earned by relatively few individuals, given its extreme gap. Thus, an alternative imputation procedure was implemented: the gap between tax-survey and national accounts investment income is imputed based on a proxy of capital ownership. We built this proxy from the set of capital income recipients from our tax-survey database, excluding owner occupied housing rent, but including total incomes reported in the household survey by firm-owners. This represents a conservative criterion in terms of the distributive impact of this imputation, as depicted in Fig. 11 in the appendix. The result of this procedure is to scale up the macro-aggregate by the same scaling factor, but imputing smaller incomes to a larger number of individuals, thus avoiding artificially increasing income concentration.

3.3 National income series

Of the incomes not included in household sector series, the most important one both quantitatively and for its distributional impact is undistributed profits, i.e., the net operating surplus of private financial and non-financial corporations. These incomes are one of the income sources of firm owners, who can decide to maintain them within the firm or to distribute them as dividends, due to tax incentives and other reasons. In the Uruguayan case, as dividends are taxed (see Section 2.2.1), firm owners may decide to keep part of their profits at the firm level as a form of untaxed savings.[11]

Undistributed profits are income flows in the Hicksian sense, since they can make owners wealthier (WIL 2021). Moreover, accounting for these incomes may compensate for the possible change in the series of tax-survey incomes caused by firm owners' decisions about the allocation of income, i.e., between keeping incomes at the firm level or distributing them as dividends. This is particularly relevant in the Uruguayan case, where only a small number of firms distribute dividends (De Rosa et al. 2018), resulting in a level of undistributed profits that is in the upper bound of available Latin American estimates (De Rosa et al. 2022).[12] In the remainder of this section, we discuss two alternative procedures to estimate the quantity of undistributed profits and, more importantly, to impute these profits to individuals.

Undistributed profits can be directly calculated based on firm tax records, which are equivalent to their aggregate accounting surplus (i.e., before any tax-related adjustments), net of distributed profits and capital incomes paid to the rest of the world. Aggregate distributed profits are calculated based on individual tax records, while capital income to the rest of the world is computed based on the balance of payments (see Fig. 17 in the appendix, more on this below). Figure 3 compares both alternative undistributed profit aggregates in terms of national income, showing that the tax record-based aggregate is 1-3 percentage points higher. It is worth noting, however, that in years with observed national accounts estimates (2012 and 2016), the results are very similar.

In this alternative, we distribute this mass of undistributed dividends from the identification of firm owners from the social security microdata, as described in Section 2.2.2. The possibility of matching owners with firms allows us to build a national income series based fundamentally on micro-data, which is quite uncommon even for developed countries. Some precedents, which achieved estimates of top incomes by incorporating retained profits by the firms, highlight the importance of this source in determining the levels, and in many cases the evolution, of inequality based on these indicators (Fairfield and Jorratt De Luis 2016; Alstadsæter et al. 2017; Kopczuk and Zwick 2020; Wolfson et al. 2016).

Most of previous works imputed the mass of undistributed profits reported on the National Accounts from capital income observed in the tax-survey base. We implement this procedure constructing an alternative National Income series derived directly from the SNA, to test the impact on the levels and evolution of inequality of the usual assumptions

[11] Moreover, owners sometimes use bank accounts shared by owners and firms, out of which owners can withdraw money. This procedure, which is registered as an asset for the firm and a liability for the owner, is a tax avoidance mechanism used by firm owners. See details in Burdín et al. (2022).

[12] In most countries, the share of undistributed profits is between 4-10% WIL (2021), and there is evidence that it is growing Flores (2018). The reasons for this difference are beyond the scope of this study, which has the more modest aim of analyzing its distributional impact. Explanations may include dividends being taxed in Uruguay (which is not the necessarily case in all remaining Latin American countries) and there were no real penalties for not distributing profits up to 2016, since the personal income tax was relatively new in the period under analysis.

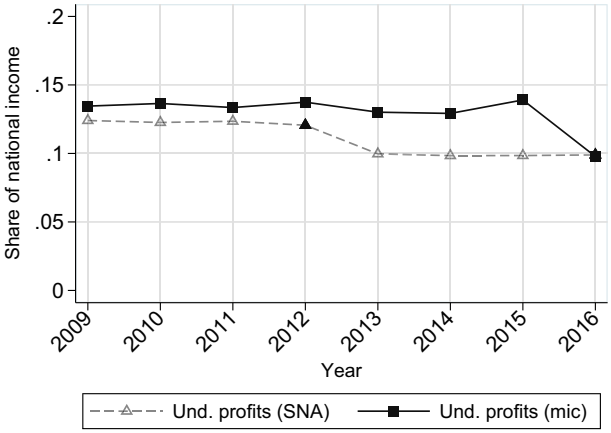


Fig. 3 Undistributed profits imputation: alternatives. ***Note.*** Own estimates based on firm tax data (DGI), National Accounts 2012, 2016 (BCU), and Balance of Payments (BCU). Undistributed profits calculated based on national accounts are equivalent to B5n-S11/12 (net undistributed profits of the corporate sector). We use information from BCU for 2012 and 2016, and we extrapolate the rest of the series from these two points (alternative ways of extrapolating do not affect the results). Undistributed profits computed based on firms' tax files are computed directly based on the micro-data provided by DGI, after subtracting rents paid to the rest of the world by the private sector (from Balance of Payments). All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries (undistributed profits from firms' tax files are already net of depreciation)

based on taxable capital (WIL 2021). On this alternative series, these undistributed profits are imputed following the same criterion used to scale up investment income in Section 3.2, i.e., using a proxy of capital ownership based on tax and survey data (see Fig. 11 in the appendix). In our case, as in the rest of Latin American countries (Alvaredo et al. 2022), given the very low share of dividends and remaining investment income in tax-survey data, the choice of the imputation method is crucial in explaining the results obtained.

The share of SNA's undistributed profits and the remaining residual incomes are depicted in Fig. 14 in the appendix. The bulk of the incomes to be imputed are from undistributed profits, while the gap to reach net national income is only 1-2%. This residual income is imputed proportionally to individuals, so by construction, it has no distributional impact. Undistributed profits, on the other hand, represent 10-12% of national income and one fourth of total capital income, which amounts to 38-39% of national income, as shown in Fig. 13 in the appendix.[13] The figure also includes the amount of investment income captured in the tax-survey micro-database as a reference. The first thing to note is that the shares of both investment income and undistributed profits decrease throughout the period, which is partially offset by an increase in the operating surplus of households (i.e., owner-occupied rental income). It is important to note that the share of investment income in the tax-survey database is increasing throughout the period, but still represents less than a third of the total investment income of national accounts at the end of the period.

[13]The overall functional distribution of income is presented in Fig. 12 in the appendix. It depicts household incomes from Table 3, as well as private undistributed profits and other incomes, particularly public undistributed profits. The figure shows the labor-capital split based on a simple 70-30% mixed-incomes distribution rule, which allocates income to labor and capital (WIL 2021). The labor share represents 61-62% of national income, of which 54-55% represents the wages component. It is worth pointing out that this is the share of net national income, including taxes net of subsidies.

4 Results

4.1 The evolution of income distribution

The evolution of pre-tax income shares in the three imputation steps is depicted in Fig. 4, i.e., the tax-survey, household sector, and national income series detailed in Sections 3.1 to 3.3. The national income series is based on our preferred method of imputation for the mass of undistributed profits using the individual/firm matched micro-database for owner identification. In turn, in all cases we show as a reference the results based on national accounts data for private net undistributed profits.

The first thing to note is that at each imputation step, period-average inequality increases, as both scaling up to household sector income and allocating undistributed profits increase the relative importance of capital income, regardless of the way it is imputed. Recalling the scaling factors from Fig. 2, capital income is scaled up in greater proportion than other incomes and is imputed based on the distribution of dividends and interest (Fig. 11 in the

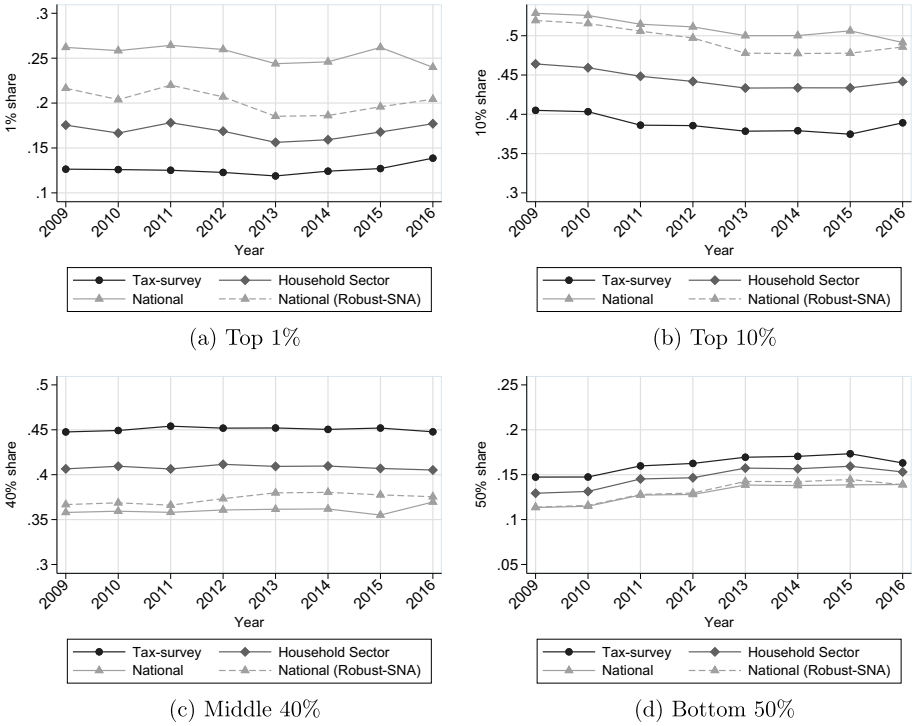


Fig. 4 Pre-tax income shares by imputation step, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts (see point estimates in Table 6 in the appendix). irst step estimates (Tax-survey series) are the result of the combination of tax data and household surveys. Second step estimates (Household sector series) include imputed undistributed profits and taxes, and in third step estimates (National series) incomes are scaled up to National Income aggregates by income source. National series uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). We also depicts the series based on SNA as robustness. All estimates refer to pre-tax personal income distribution. Top 1, 10, middle 40 (p51-90) and bottom 50%'s shares depicted in panels a, b, c and d respectively

appendix), which allocates it to top 10 and especially the top 1%. Our preferred national income series which allocates undistributed profits to individuals who report firm ownership (or to individuals created for firms with no matched owners), results in higher concentration levels than the SNA series. This difference is not ot only as a result of the imputation rule, but also because the quantity of firm-based net undistributed profits is 1-2 points higher on average (Fig. 3). Moreover, it is interesting to note that the effect of these alternatives is only visible when considering the top 1%'s share, where the top share is around 5 percentage points higher in the owner-firm matched series, but less so in the remaining ones, and virtually undetectable when considering the overall Gini index (see Fig. 15 in the appendix).

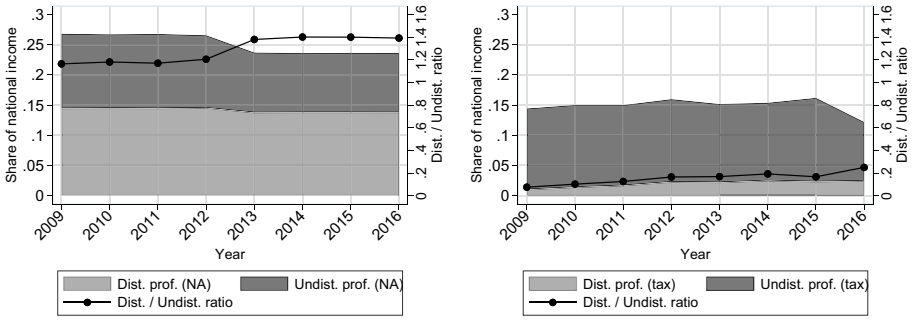
Aside from the importance of the alternative imputation procedure used for undistributed profits, another dimension of the series deserves to be highlighted. Although it is true that inequality trends appear to be rather similar across all imputation steps, while tax-survey and household series stay remarkably close, national income series present a slightly different trend. In fact, as depicted in Table 6 in the appendix, while the top 1%'s share increases for the tax-survey series between 2009 and 2016 (from 12.6% to 13.9%), it remains relatively stable in the household income series, and it decreases in the national income series (from 26.2 to 24.0%). The origin of this changing trend is discussed in Section 4.2.

4.2 The effect of (un)distributed profits on inequality

The increasing trend of the top 1%'s share in tax-survey data, unaffected by undistributed profits, is consistent with similar estimates from Burdín et al. (2022), which were based on the same data and imputation procedures. However, the increase in the top 1%'s share by the end of the period in the tax-survey series is somewhat neutralized by the imputation of undistributed profits in the third step. The explanation lies in the changing size of the undistributed profits vis á vis the quantity of distributed profits. As dividends are taxed, they appear in an individual's tax records, pushing top incomes' shares upwards; however, this increase is mirrored by a decrease in undistributed profits. Therefore, when undistributed profits are imputed, the top 1%'s increasing share is offset and even slightly reversed. This finding highlights the importance of considering both distributed and undistributed profits in inequality analysis, since what may appear to be a surge in inequality may only reflect a change in the decisions of firm managers to either distribute dividends or keep them at the firm level.

To dig into this increase in the share of capital income captured in the tax-survey data, we present pre-tax profits produced at the firm level and their distribution into distributed profits (the bulk of investment income), undistributed profits, and profits distributed abroad. This last component is taken from the Balance of Payments and is depicted in Fig. 17 in the appendix. Although it is not, by definition, a component of net national income, it is informative for how firm profits are split between the country and the rest of the world. Profits sent abroad represent close to 10% of net national income, while distributed profits represents less than half of total profits. In Fig. 5, distributed and undistributed profits are portrayed, as well as the ratio between the two, using SNA data and tax data, i.e., undistributed profits from firm tax records and dividends from individual tax records. Despite different levels, which result from the previously discussed large gap between dividends observed in individual tax data and investment income from national accounts, both data sources indicate that throughout the period, firms have increased their distributional share.

Thus, Figs. 5 and 16 in the appendix indicate that two effects are at play: (i) firms increased their share of distributed profits; and (ii) a higher share of dividends is captured in the tax-survey data. These two combined effects result in the increase in tax-survey top



(a) SNA (b) Firm's and individual tax data

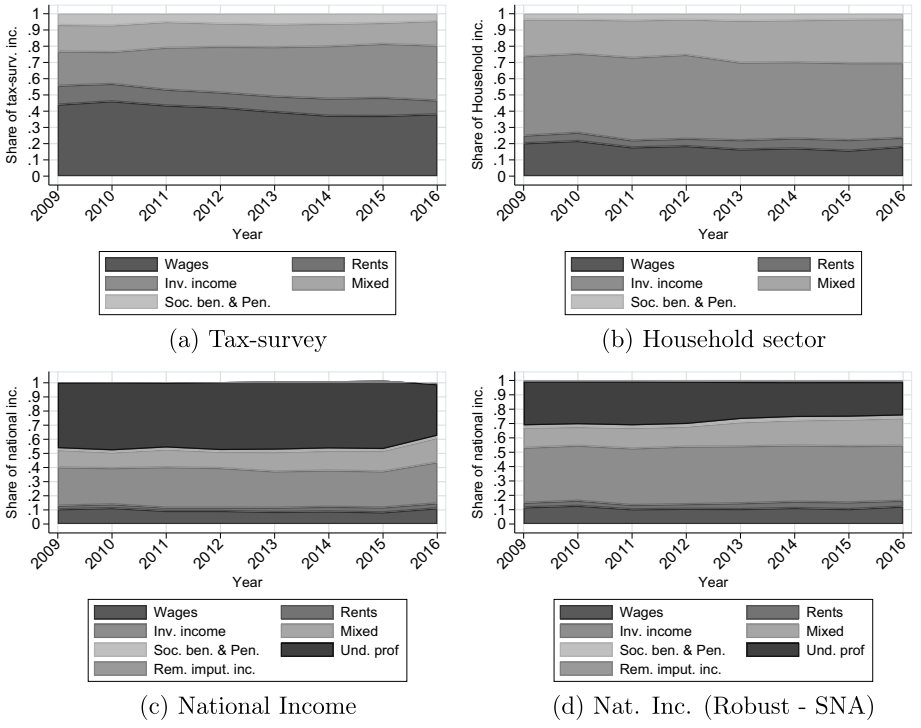
Fig. 5 Distributed and undistributed profits by source, 2009-2016. ***Note.*** Own estimates based on firm tax data (DGI), tax-survey data (ECH-DGI), National Accounts 2012, 2016 (BCU), and Balance of Payments (BCU). Both panels depict distributed and undistributed profits, as well as their ratios. In panel a, undistributed profits are equivalent to B5n-S11/12 (net undistributed profits of the corporate sector), while in panel b they come from balance sheets net of private capital incomes paid to the rest of the world (based on Balance of Payments). Distributed profits from panel a come from investment incomes excluding interest received by households (D41-S14 in SNA, see Table 3), while in panel b they represent aggregate dividends from individual tax records. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries (undistributed profits from panel b are already net of depreciation)

income shares shown in Figure 4 and documented by Burdín et al. (2022). The increase in the distributional share of the firms also lowers the undistributed profits to be allocated in the national income series, decreasing the gap between the different series towards the end of the period.

The incorporation of undistributed profits into this last step also has implications for the composition of income in the upper tail of the distribution. Figure 6 shows the income composition of the top 1% in the three estimation steps, while the composition for the other income groups is included in Figs. 19, 20 and 21 of the appendix. Between the first two estimation steps, the top 1% experienced significant growth in its share of investment income, explained by the large percentage of this income not observed in the tax-survey database. A similar increase is observed in the top 10% of the income distribution. On the other hand, the strong concentration of undistributed profits implies a clear change in the income composition of the top 1% in the national income series. Depending on the imputation method, this source of income represents between 25 and 40% of the total income of the top 1%. At this step, capital income (investment income + rents + undistributed profits) represents at least two thirds of total income for this group. Finally, the downturn in the quantity of undistributed profits in the micro-database towards the end of the period is also evident in its decline as a share of the total income of the top 1%.

4.3 The distribution of growth

One of the most important advantages of this exercise is that in the last estimation step, the national income series provide full micro-macro consistency. This is relevant, in particular, for the analysis of growth and its distribution, since growth is typically measured in macroeconomic terms while inequality is analyzed from a microeconomic perspective. Thus, our national income inequality series allow us to analyze growth and inequality consistently.



(a) Tax-survey (b) Household sector

(c) National Income (d) Nat. Inc. (Robust - SNA)

Fig. 6 Top 1% income composition, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates (panel a) are the result of the combination of tax data and household surveys. Second step estimates (panel b) include imputed undistributed profits and taxes, and in third step estimates (panels c and d), incomes are scaled up to National Income aggregates by income source. Panel c uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). Panel d shows the series based on SNA as robustness

Figure 7 depicts the growth incidence curves, i.e., the growth rate by percentile over the 2009-2016 period, for the three imputation steps (panels a, b, c) and the robustness national income series (panel d). Broadly speaking, the slopes of the curves are negative, meaning that income grew faster for the bottom 50% and the lower half of the middle 40% (51st to 90th percentile) than it did for top earners, hence fueling the decrease in inequality. This negative slope is less pronounced in the tax-survey-based series (panel a) compared to the series from the other two steps. Up to the sixth decile, real income growth is above 40% in real terms, which is consistent with the fact that both economic growth and the wage policy resulted in job creation and rapid labor income growth at the bottom of the distribution. Income growth falls thereafter, with the exception of the top 10%, which shows heterogeneous trends.

On the tax-survey income series the spike in growth for the top 1% is noticeable, which is consistent with the increase in the income share of this group towards the end of the period. In the rest of the series (panels b and c), this increase is less pronounced, but it is also observed in other percentiles of the distribution within the top 10%. The fall in the trend in the national income series is mostly due to the reduction in the quantity of undistributed profits towards the end of the period. Figure 22 in the appendix shows the same growth incidence curves for the national income series but for the period 2009-2015. In this case,

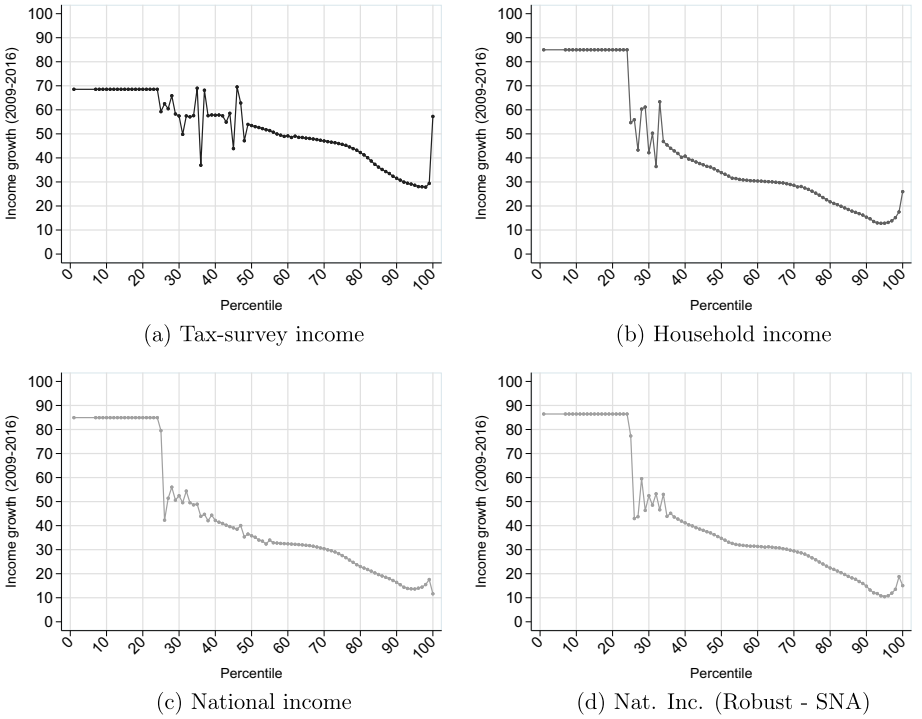


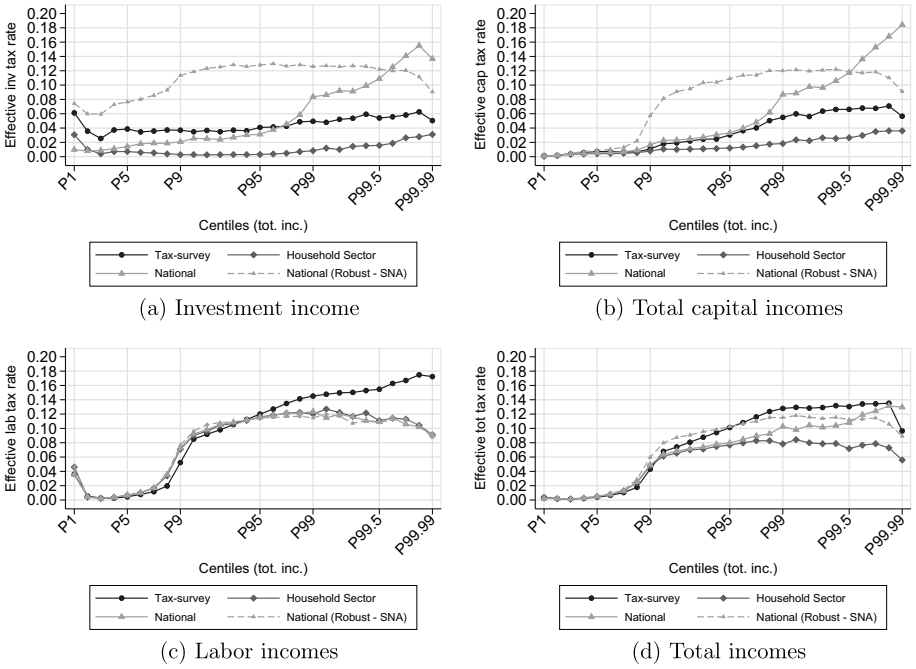
Fig. 7 Growth Incidence Curves (GIC) by imputation step, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates (panel a) are the result of the combination of tax data and household surveys. Second step estimates (panel b) include imputed undistributed profits and taxes, and in third step estimates (panels c and d), incomes are scaled up to National Income aggregates by income source. Panel c uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). Panel d shows the series based on SNA as robustness

the trend reverses, with the top 1% having the largest growth within the highest percentiles. Therefore, changes in the aggregate of undistributed profits can generate significant annual variations in the right tail of the distribution, resulting in noisy estimates.

4.4 Effective direct tax rates

The blurry line dividing firms and their owners has consequences for income, but also for taxes paid as observed in the tax records, and therefore also for the effective tax rates estimated using these sources of information. Thus, our three-step estimation procedure allows us to calculate effective tax rates while accounting for differences that may emerge from these imputation decisions. Corporate taxes were imputed following the same criteria as undistributed profits. In this way, the different income taxes on individual incomes (taxes on both labor and capital) are combined with the corporate tax (see Saez and Zucman 2020 for similar procedures).

Figure 8 shows the effective tax rates paid by income fractile for the three steps and the two alternatives corresponding to the national income series for 2016. Given the concentration of capital income and undistributed profits, we provide greater detail for the top 10 and 1%. The progressiveness of income taxes implies an effective rate close to zero up to



(a) Investment income (b) Total capital incomes

(c) Labor incomes (d) Total incomes

Fig. 8 Effective tax rates, 2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates are the result of the combination of tax data and household surveys. Second step estimates include imputed undistributed profits and taxes, and in third step estimates, incomes are scaled up to National Income aggregates by income source. National series uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). We also depicts the series based on SNA as robustness. All estimates refer to pre-tax personal income distribution. Investment income (panel a) is included in total capital incomes (panel b). Panel d (total incomes) is the sum of panels b and c, plus all remaining incomes

the median income (panel d of Fig. 8), with an increasing incidence of taxes throughout the distribution at least up to the top 1% in all estimates.

Series comparisons indicate that the scaled-up household income series, which scales incomes but not taxes since they are reported in tax records and assumed to be an accurate depiction of total revenue, results in a reduction in the average effective rate from 13 to 8% for the top 1%. The inclusion of corporate income tax (CIT) entails an increase in effective rates to levels similar to those corresponding to the tax-survey series. This last step implies the incorporation of highly concentrated income, which is in turn taxed at a flat rate of 25%. The effect of the introduction of taxes on the corporate sector is more evident in the series for capital income (panel a of Fig. 8), and in particular in our preferred national income series, which translates into a growing effective rate even in the highest income fractiles.

However, this result should be considered an upper bound of progressivity, insofar the implicit incidence assumption of this exercise is that all CIT is payed by firm owners. Evidence in turn suggests that a significant fraction may actually be paid by workers. Causal estimates show that workers bear half of the tax burden in Germany (Dwenger et al. 2019) as well as other European countries (Arulampalam et al. 2012) and between 35% (Suárez Serrato and Zidar 2016) and 60% (Liu and Altshuler 2013) in the United States.

Finally, in all the series, a reduction in effective rates is observed in the right tail of the distribution. The combination of a dual income tax system that taxes capital at lower average rates than labor along with the concentration of capital income in the top 1% results in a reduction in average taxes for the top income groups. The drop is evident in the top 0.1%, particularly for the tax-survey series. The regressiveness of the set of taxes at the very top of the distribution is similar to that found by Saez and Zucman (2020) for 2008 in the United States, explained by the ability of high-income individuals to avoid personal income taxes and obtain their income from direct participation in their firms.

5 Concluding remarks

In this paper, we highlight the difficulty of assessing inequality trends, not only as a result of the challenges inherent in combining different data sources to close measurement gaps, but also stemming from what can actually be observed and how economic decisions affect it. We tackle these challenges using a rare combination of survey, social security, personal income tax, and corporate tax micro-data, combined with national accounts. We presented distributive estimates for the Uruguayan case based on this unique data in three different steps: tax-survey series, household income series, and national income series in order to document their differences.

Thus, this article points out the need to consider different income aggregates, and to track changes in inequality based on both what we can see in our tax records and surveys, and what remains hidden within firms and, more generally, within national income as a whole. We have shown that the imputation of these incomes does not have a mechanical effect on inequality trends, and may change our understanding of their evolution. However, imputing undistributed profits has massive effects on the level of income inequality, which implies that income concentration could be considerably underestimated, hence calling for more ambitious redistributive policies.

Appendix

Table 4 Income categories and tax rates of IASS and IRPF (cat. I and II)

Panel a) IRPF: Labor income

2009-2011		2012-2016	
Annual income in BPC	Tax rate	Annual income in BPC	Tax rate
0 - 84	0%	0-84	0%
84 - 120	10%	84 - 120	10%
120 - 180	15%	120 - 180	15%
180 - 600	20%	180 - 600	20%
600 - 1200	22%	600 - 900	22%
1200 or more	25%	900-1380	25%
–	–	1380 or more	30%

Panel b) IASS: Pensions

Annual income in BPC	Tax rate
0 - 96	0%

Table 4 (continued)

Panel a) IRPF: Labor income

2009-2011	2012-2016
96 - 180	10%
180-600	20%
600 or more	25%

Panel c) IRPF: Capital income

Capital income category	Tax rate
Interest on bank deposits in Uruguayan currency or UI (one year length or less)	3%
Interest on bank deposits in Uruguayan currency or UI (one year length or less)	3%
Interest, obligations and other securities (3 years or more length)	5%
Copyrights	7%
Profits, dividends and benefits	7%
Sports rights	12%
Participation certificates (issued by financial trusts)	7%
Remaining financial and mobiliary capital	12%
Real-estate capital	12%
Capital gains	12%
Dividends or benefits from IRAE contributors	7%
Imputed rents by non-resident entities	12%

Note. Own elaboration based on DGI

Table 5 Scaling factors, 2009-2016

Year	Wages	Rents	Inv. Income	Mixed	Soc. ben & Pen.
2009	1,09	0,81	6,58	2,36	1,31
2010	1,08	0,85	6,94	2,11	1,27
2011	0,97	0,80	5,29	2,30	1,39
2012	0,99	0,79	4,42	2,26	1,22
2013	0,87	0,92	3,73	2,42	1,24
2014	0,94	0,90	3,54	2,52	1,25
2015	0,90	0,90	3,46	2,81	1,22
2016	0,91	0,98	3,15	2,46	1,22

Note. Own elaboration based on tax records, household surveys, and national accounts. See Note of Fig. 2

Table 6 Income shares, 2009-2016

	Tax-survey	Hous. Sector	National Inc.	Nat. Inc. (SNA)
Panel A: Top 1%				
2009	12.6%	17.5%	26.2%	21.7%
2010	12.6%	16.7%	25.8%	20.4%
2011	12.5%	17.8%	26.4%	22.0%
2012	12.3%	16.9%	26.0%	20.7%
2013	11.9%	15.6%	24.4%	18.5%

Table 6 (continued)

	Tax-survey	Hous. Sector	National Inc.	Nat. Inc. (SNA)
2014	12.4%	15.9%	24.6%	18.6%
2015	12.7%	16.8%	26.2%	19.6%
2016	13.9%	17.7%	24.0%	20.4%
Panel B: Top 10% (p91-100)				
2009	40.5%	46.4%	52.9%	51.9%
2010	40.3%	45.9%	52.6%	51.6%
2011	38.6%	44.8%	51.5%	50.6%
2012	38.6%	44.2%	51.1%	49.7%
2013	37.9%	43.3%	50.0%	47.8%
2014	37.9%	43.4%	50.0%	47.7%
2015	37.5%	43.4%	50.6%	47.8%
2016	38.9%	44.2%	49.1%	48.6%
Panel C: Middle 40% (p51-90)				
2009	44.8%	40.6%	35.8%	36.7%
2010	44.9%	40.9%	35.9%	36.9%
2011	45.4%	40.6%	35.8%	36.6%
2012	45.2%	41.2%	36.1%	37.3%
2013	45.2%	40.9%	36.1%	38.0%
2014	45.0%	41.0%	36.2%	38.0%
2015	45.2%	40.7%	35.5%	37.8%
2016	44.8%	40.5%	37.0%	37.5%
Panel D: Bottom 50% (p1-50)				
2009	14.7%	12.9%	11.3%	11.4%
2010	14.7%	13.1%	11.5%	11.6%
2011	16.0%	14.5%	12.7%	12.8%
2012	16.3%	14.7%	12.8%	13.0%
2013	16.9%	15.7%	13.8%	14.2%
2014	17.0%	15.7%	13.8%	14.2%
2015	17.3%	15.9%	13.9%	14.5%
2016	16.3%	15.3%	13.9%	13.9%

Note. Own elaboration based on tax records, household surveys, and national accounts. First step estimates are the result of the combination of tax data and household surveys (tax-survey series). Second step estimates include imputed undistributed profits and taxes (Household sector series), and in third step estimates, incomes are scaled up to National Income aggregates by income source. National series uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). We also depicts the series based on SNA as robustness. All estimates refer to pre-tax personal income distribution

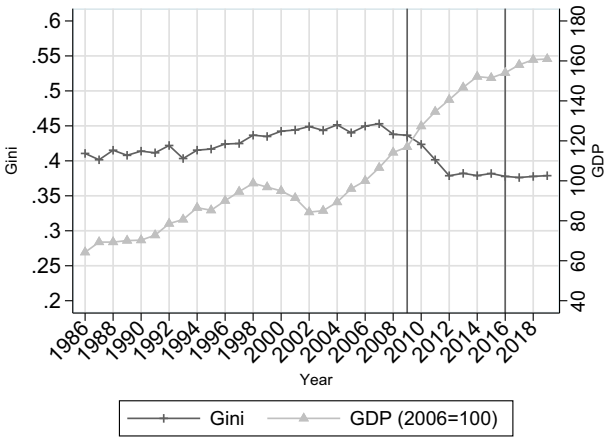


Fig. 9 GDP and income inequality 1986-2019. ***Note.*** In the primary axis GDP is presented with GDP 2005=100, whilst percapita household income gini index (estimated based on the household survey) is depicted on the secondary axis. During the period 2009-2016 (between vertical lines, period with tax data available), gini index dropped by about 7 points, and National Income grew at a 5.5% rate

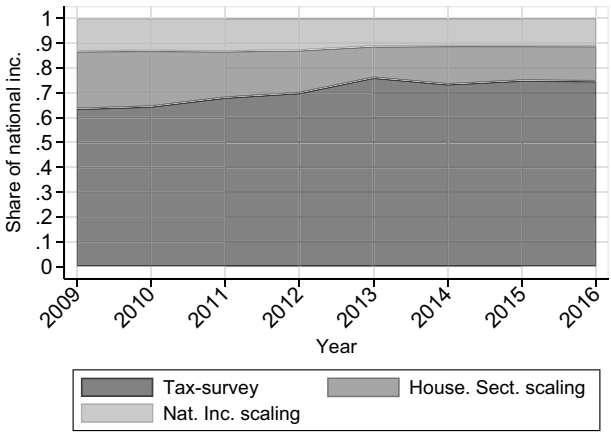


Fig. 10 Income shares by estimation step, 2009-2016. ***Note.*** Own estimates based on tax-survey data (DGI-ECH) and National Accounts 2012, 2016 (BCU). The figure depicts aggregate income by estimation step: the dark-green area is the sum of tax-survey incomes, the orange area are incomes added during scaling to household sector based on scaling factors depicted in Fig. 2, while the blue area represents reaming imputed incomes of Fig. 14. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries

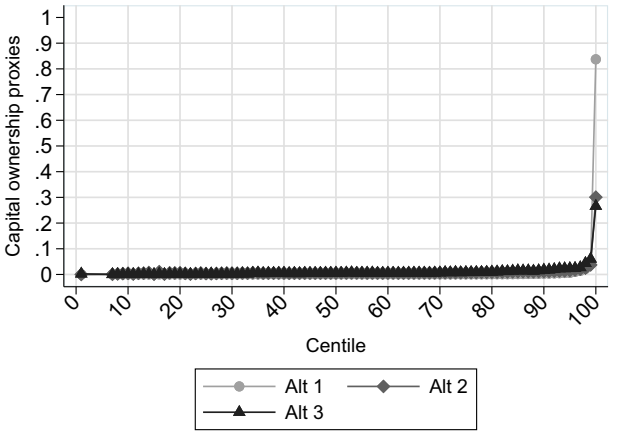


Fig. 11 Proxies of firm ownership, 2016. ***Note.*** Own estimates based on tax-survey data (ECH-DGI). *Alt. 1* refers to the distribution of taxable capital incomes from DGI. *Alt. 2* refers in turn to the sum of all taxable and non-taxable capital incomes, including rents and owner occupied housing rents. The preferred alternative (*Alt. 3*), is equivalent to the second one, but excludes owner occupied housing rent and includes total incomes reported in the household survey by firm-owners

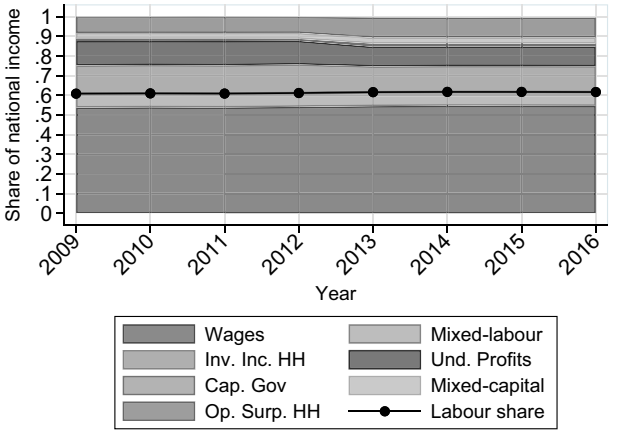


Fig. 12 Functional income distribution, 2009-2016. ***Note.*** Own estimates based on tax-survey data (DGI-ECH) and National Accounts 2012, 2016 (BCU). The figure presents the distribution of net national incomes in capital and labor shares and their components. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries

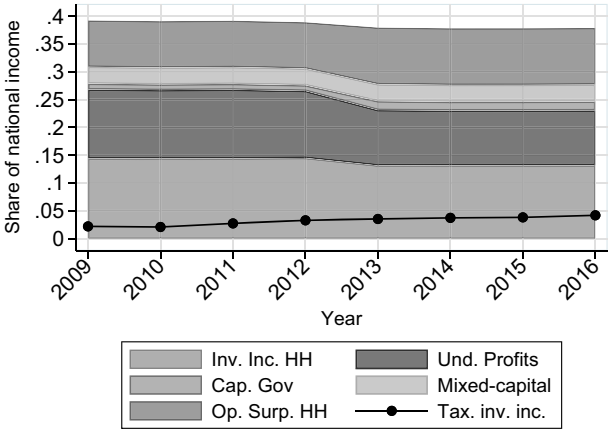


Fig. 13 Capital incomes composition. ***Note.*** Own estimates based on firm tax data (DGI), tax-survey data (ECH-DGI) and National Accounts 2012, 2016 (BCU). Solid filled areas represent national account's aggregates, while doted line depicts aggregate investment incomes (dividends, interest, etc.) from tax-survey data. This line is conceptually consistent with national account's investment income received by households (light blue area), D4-S14. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries

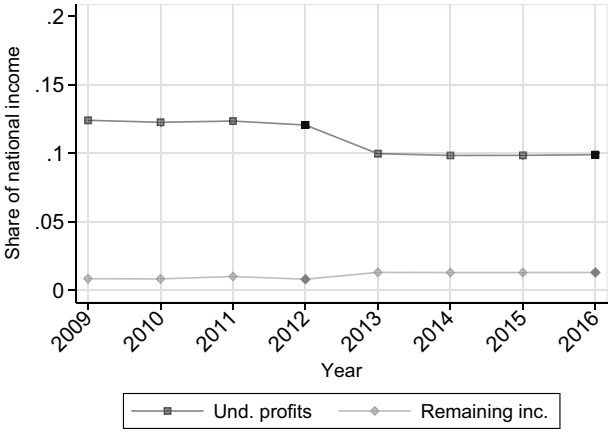


Fig. 14 Income aggregates of non-household sector. ***Note.*** Own estimates based on National Accounts 2012, 2016 (BCU). Dots in dark represent actually observed data points in national accounts. Undistributed profits are allocated based on the capital ownership proxy, while remaining components of national income are distributed proportionally to total incomes from tax-survey data. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries

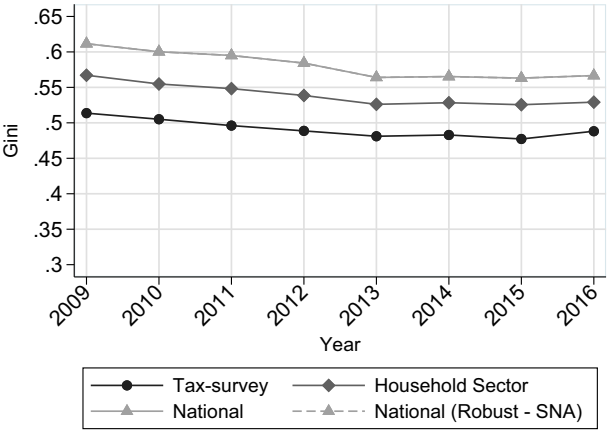
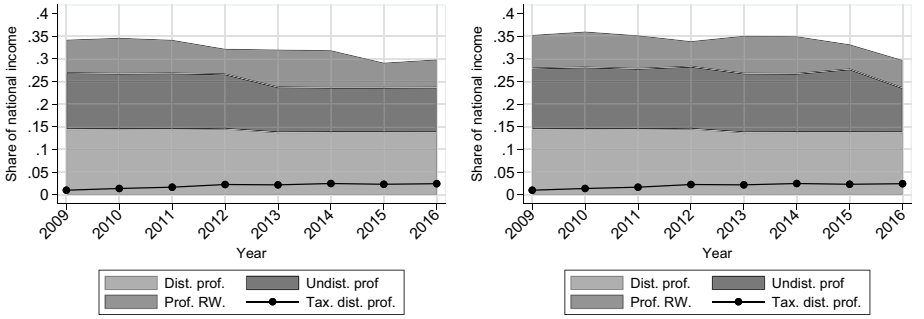


Fig. 15 Pre-tax Gini index by source and imputation step, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates (panel a) are the result of the combination of tax data and household surveys. Second step estimates (panel b) include imputed undistributed profits and taxes, and in third step estimates (panels c and d), incomes are scaled up to National Income aggregates by income source. National series uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). We also depicts the series based on SNA as robustness. All estimates refer to pre-tax personal income distribution



(a) SNA (b) Firm's micro data

Fig. 16 Firm's profits by alternative, 2009-2016. ***Note.*** Own estimates based on firm tax data (DGI), tax-survey data (ECH-DGI), National Accounts 2012, 2016 (BCU), and Balance of Payments (BCU). Both panels depict observed dividends observed in tax-survey data, investment incomes of households excluding interest, undistributed profits and capital incomes sent abroad (computed based on Balance of Payments). All but undistributed profits are equivalent in both panels. In Panel a, undistributed profits are calculated based on national accounts, while Panel b presents undistributed profits computed based on firms' tax files. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries (undistributed profits from panel b are already net of depreciation)

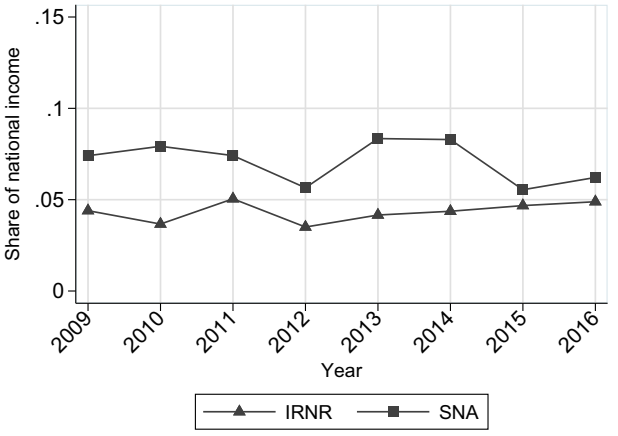
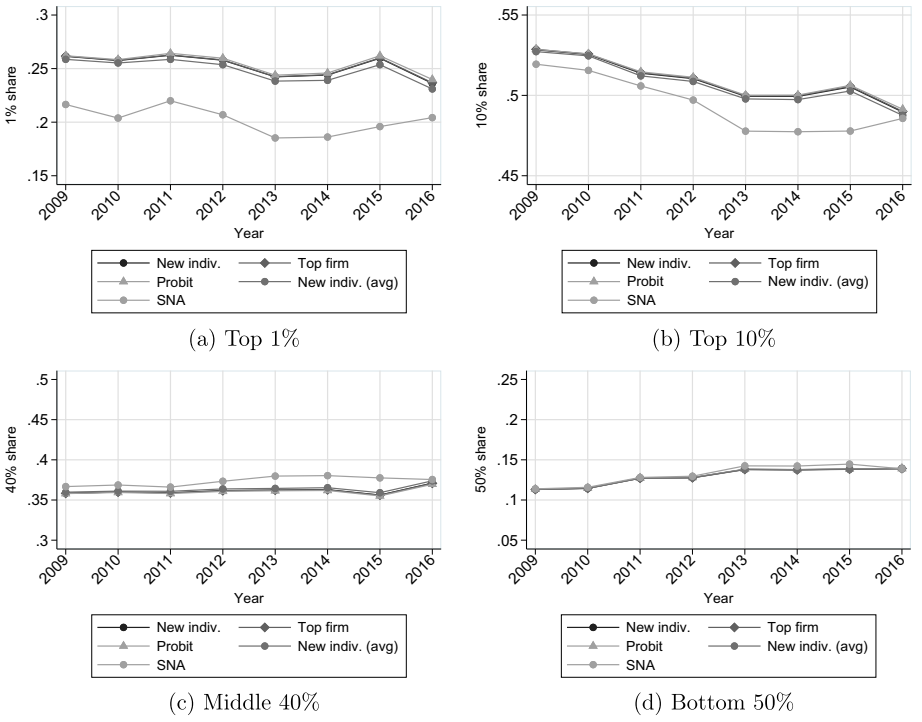


Fig. 17 Private capital incomes paid to the rest of the world. ***Note.*** Own estimates based on Balance of Payments (BCU) and *Impuesto a la Renta de los No Residentes* (IRNR) series (DGI). Balance of payments series is constructed based on Central Bank data for two periods: 2009-2012 and 2013-2016. The latter series has an updated methodology but has not been matched with the previous one, resulting in higher private primary income (1.B-credit), i.e., capital incomes paid to the rest of the world by the private sector. The 2009-2012 series was thus adjusted by the ratio of the two period averages. IRNR series is constructed by dividing IRNR aggregate taxes collected by its main flat rate (7%)



(a) Top 1% (b) Top 10%

(c) Middle 40% (d) Bottom 50%

Fig. 18 Pre-tax income shares of National Income by imputation method, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. All estimates refer to pre-tax personal income distribution. Top 1, 10, middle 40 (p51-90) and bottom 50%'s shares depicted in panels a, b, c and d respectively.The 5 series show alternatives for the allocation of undistributed profits. Our preferred series uses the matched base of individuals/firms to identify owners, and for the firms for which this identification is not posible, imputes through a probit model (probit series). The new individuals and new individuals (avg) series creates new perceivers for the unmatched firms (1 individual per firm or the average number of individuals per firm of the matched base). The top firm series allocates the non-distributed dividends to the recipient of the highest income of the firm, while the last alternative uses the SNA for the imputation. For more details see Section 2.2.2

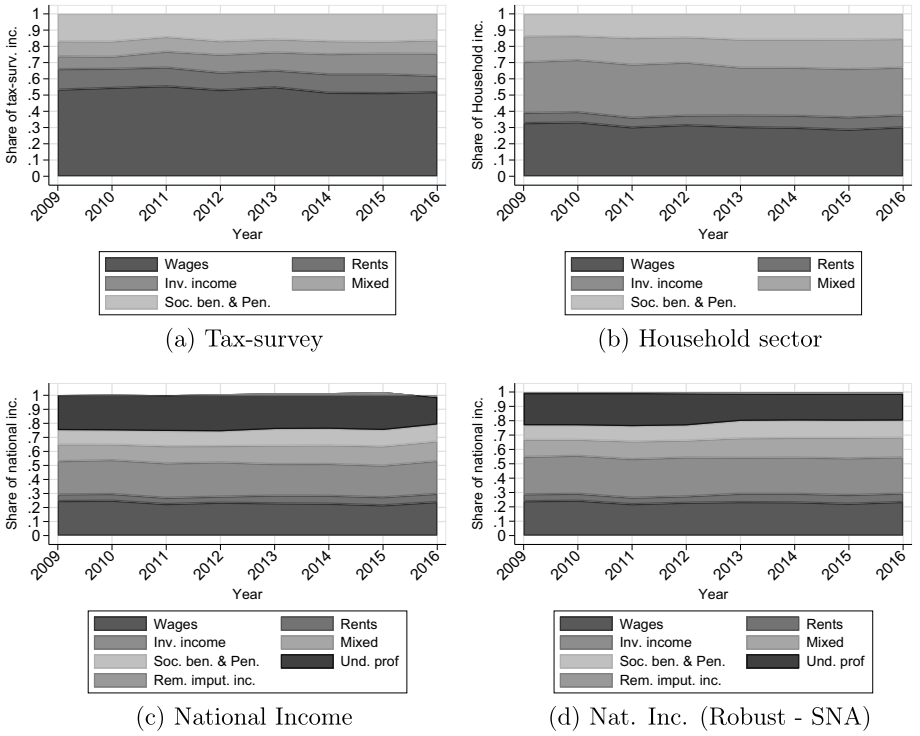


Fig. 19 Top 10% income composition, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates (panel a) are the result of the combination of tax data and household surveys. Second step estimates (panel b) include imputed undistributed profits and taxes, and in third step estimates (panels c and d), incomes are scaled up to National Income aggregates by income source. Panel c uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). Panel d shows the series based on SNA as robustness

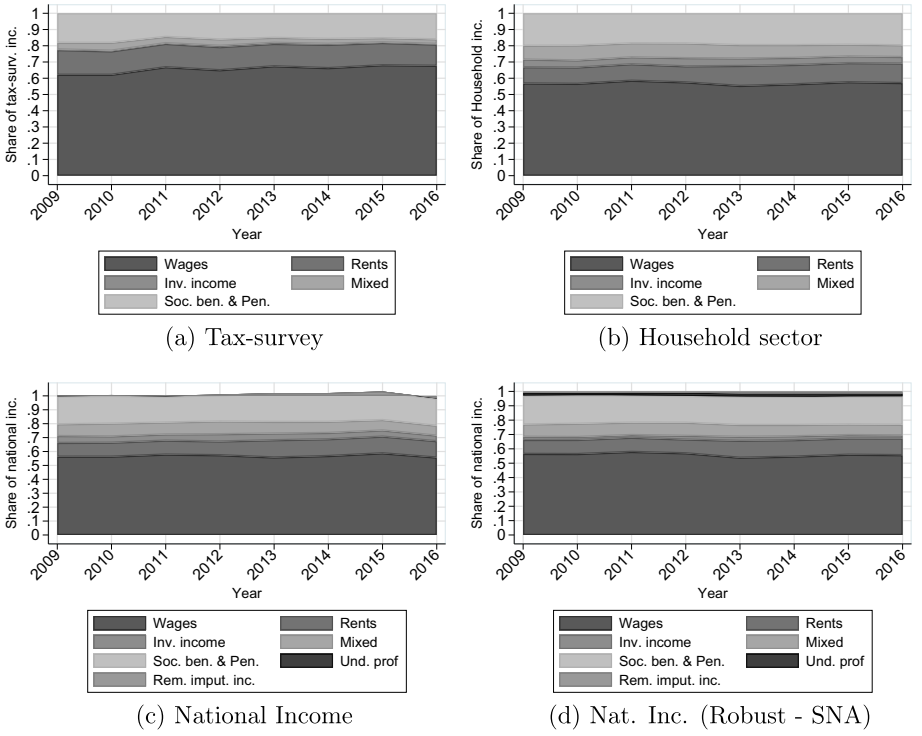


Fig. 20 Middle 40% income composition, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates (panel a) are the result of the combination of tax data and household surveys. Second step estimates (panel b) include imputed undistributed profits and taxes, and in third step estimates (panels c and d), incomes are scaled up to National Income aggregates by income source. Panel c uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). Panel d shows the series based on SNA as robustness

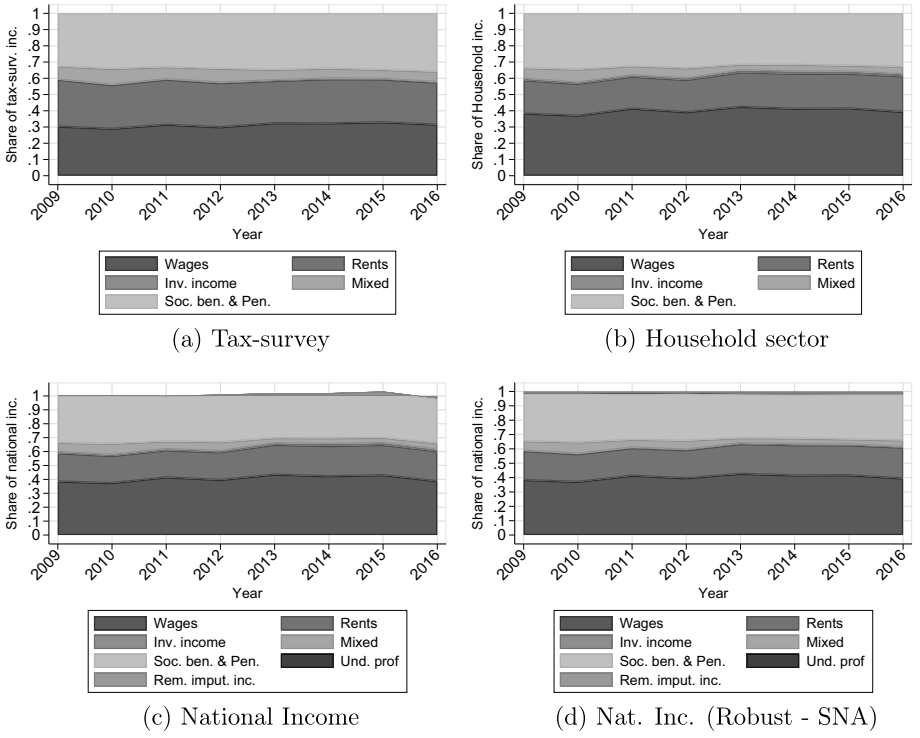


Fig. 21 Bottom 50% income composition, 2009-2016. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. First step estimates (panel a) are the result of the combination of tax data and household surveys. Second step estimates (panel b) include imputed undistributed profits and taxes, and in third step estimates (panels c and d), incomes are scaled up to National Income aggregates by income source. Panel c uses the micro database of firm owners, and our preferred imputation method (based on a probit model, see Section 2.2.2). Panel d shows the series based on SNA as robustness

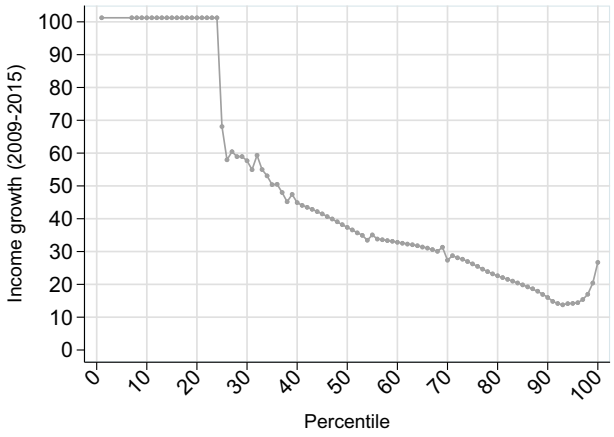


Fig. 22 Growth Incidence Curves (GIC), 2009-2015. ***Note.*** Own elaboration based on tax records, household surveys, and national accounts. Preferred imputation method for undistributed profit based on micro database of firm owners and on a probit model (see Section 2.2.2). All estimates refer to pre-tax personal income distribution

Acknowledgements We are grateful to the inequality research group *Instituto de Economía-UdelaR* for their inputs and contributions. We would like to especially thank Andrea Vigorito for her insights and suggestions throughout the whole process of writing this article; we also extend our thanks to Martín Leites and Luciana Méndez, coordinators of the group, and to the participants of the seminar in which we presented this article. We are also very grateful to the World Inequality Lab's team for their continuous help and invaluable comments at different stages of this process: in particular, we would like to thank Facundo Alvaredo, Ignacio Flores, and Marc Morgan. We are also grateful to the participants in the session on "DINA in developing countries" at the First WIL Conference in December 2017, and the participants in the session "Income Distribution" at the IX Conference of the Network of Inequality and Poverty (Uruguayan session, November 2017), who commented on early versions of this paper. We thank Nora Lustig, with whom we discussed in depth many of the key aspects of this effort. Finally, we thank Gustavo González, former coordinator of *Asesoría Económica* at *Dirección General Impositiva*, and to his staff, Fernando Peláez and Sol Mascarenhas, for the invaluable help and comments received during this research. Any errors remain our own.

Data Availability Social Security and Tax micro-data (both for individuals and firms) are not available since they were provided to *Instituto de Economía-Universidad de la República* by the uruguayan tax authority *Dirección General Impositiva* under a non-sharing with third parties agreement. These data may be requested directly to *Dirección General Impositiva*.

Survey and National Accounts data is publicly available at *Instituto Nacional de Estadísticas* and *Banco Central del Uruguay* web-pages, and may be downloaded directly without any specific authorization.
Survey data: https://www.ine.gub.uy/encuesta-continua-de-hogares1.
National Accounts data: https://www.bcu.gub.uy/Estadisticas-e-Indicadores/Paginas/Cuentas-Nacionales-e-Internacionales.aspx.

Declarations

Conflict of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Alstadsæter, A., Jacob, M., Kopczuk, W., Telle, K.: Accounting for business income in measuring top income shares. In: Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association, volume 110, pp. 1–39. JSTOR (2017)

Altimir, O.: Income distribution statistics in latin america and their reliability. Rev. Income Wealth **33**(2), 111–155 (1987)
Alvaredo, F.: A note on the relationship between top income shares and the gini coefficient. Econ. Lett. **110**(3), 274–277 (2011)
Alvaredo, F., De Rosa, M., Flores Beale, I., Morgan, M., et al: The inequality (or the growth) we measure: data gaps and the distribution of incomes. CEPR discussion papers (2022)
Alvaredo, F., Gasparini, L.: Recent trends in inequality and poverty in developing countries. In: Atkinson, A., Bourguignon, F. (eds.) Handbook of Income Distribution, vol. 2, pp. 697–805. Elsevier (2015)
Amarante, V., Colafranceschi, M., Vigorito, A.: Uruguay's Income Inequality and Political Regimes over the Period 1981–2010. In: Cornia, A. (ed.) Falling Inequality in Latin America. Policy Changes and Lessons, WIDER Studies in Development Economics. Oxford University Press (2014)
Arulampalam, W., Devereux, M.P., Maffini, G.: The direct incidence of corporate income tax on wages. Eur. Econ. Rev. **56**(6), 1038–1054 (2012)
Atkinson, A.B., Piketty, T.: Top incomes over the twentieth century: a contrast between continental european and english-speaking countries. oup Oxford (2007)
Atkinson, A.B., Piketty, T., Saez, E.: Top incomes in the long run of history. J. Econ. Lit. **49**(1), 3–71 (2011)
Blanchet, T., Chancel, L., Gethin, A.: How unequal is europe? evidence from distributional national accounts, 1980-2017. WID. world Working Paper, vol. 6 (2019)
Blanchet, T., Flores, I., Morgan, M.: The weight of the rich: improving surveys using tax data. J. Econ. Inequal. **20**(1), 119–150 (2022)
Bourguignon, F.: Appraising income inequality databases in latin america. J. Econ. Inequal. **13**(4), 557–578 (2015)
Bucheli, M., Lustig, N., Rossi, M., Amábile, F.: Social spending, taxes and income redistribution in Uruguay. The World Bank (2013)
Burdín, G., De Rosa, M., Vigorito, A., Vilá, J.: Falling inequality and the growing capital income share: reconciling divergent trends in survey and tax data. World Dev. **152**, 105783 (2022)
Burkhauser, R.V., Feng, S., Jenkins, S.P., Larrimore, J.: Recent trends in top income shares in the united states: reconciling estimates from march cps and irs tax return data. Rev. Econ. Stat. **94**(2), 371–388 (2012)
Cornia, G.A.: Falling inequality in Latin America: Policy changes and lessons. OUP Oxford (2014)
De Rosa, M., Flores, I., Morgan, M.: More unequal or not as rich? On the missing half of latin american income. Technical report (2022)
De Rosa, M., Siniscalchi, S., Vilá, J., Vigorito, A., Willebald, H.: La evolución de las remuneraciones laborales y la distribución del ingreso en Uruguay; futuro en foco Cuadernos Sobre Desarrollo Humano: Montevideo, Uruguay (2018)
Deaton, A.: Measuring poverty in a growing world (or measuring growth in a poor world). Rev. Econ. Stat. **87**(1), 1–19 (2005)
Dwenger, N., Steiner, V., Rattenhuber, P.: Sharing the burden? empirical evidence on corporate tax incidence. German Econ. Rev. **20**(4), e107–e140 (2019)
Fairfield, T., Jorratt De Luis, M.: Top income shares, business profits, and effective tax rates in contemporary c hile. Rev. Income Wealth **62**, S120–S144 (2016)
Ferreira, F.H., Lustig, N., Teles, D.: Appraising cross-national income inequality databases: an introduction. J. Econ. Inequality **13**(4), 497–526 (2015)
Flachaire, E., Lustig, N., Vigorito, A.: Underreporting of top incomes and inequality: a comparison of correction methods using simulations and linked survey and tax data. Rev. Income Wealth (2022)
Flores, I.: Income under the carpet: what gets lost between the measure of capital shares and inequality. http://precog.iiitd.edu.in/people/anupama (2018)
Garbinti, B., Goupille-Lebret, J., Piketty, T.: Income inequality in france, 1900–2014: evidence from distributional national accounts (dina). J. Public Econ. **162**, 63–77 (2018)
Gasparini, L., Bracco, J., Galeano, L., Pistorio, M.: Desigualdad en países en desarrollo:¿ ajustando las expectativas? Documentos de Trabajo del CEDLAS (2018)
Goolsbee, A.: What happens when you tax the rich? evidence from executive compensation. J. Polit. Econ. **108**(2), 352–378 (2000)
Jenkins, S.P.: Pareto models, top incomes and recent trends in uk income inequality. Economica **84**(334), 261–289 (2017)
Kopczuk, W., Zwick, E.: Business incomes at the top. J. Econ. Perspectives **34**(4), 27–51 (2020)
Liu, L., Altshuler, R.: Measuring the burden of the corporate income tax under imperfect competition. Natl. Tax J. **66**(1), 215–237 (2013)
Lustig, N., et al.: The missing rich in household surveys: causes and correction approaches technical report, Tulane University, Department of Economics (2019)

Lustig, N., Teles, D., et al: Inequality convergence: how sensitive are results to the choice of data? Technical report (2016)
Meyer, B.D., Mok, W.K., Sullivan, J.X.: The under-reporting of transfers in household surveys: its nature and consequences (2015)
Morgan, M.: Extreme and persistent inequality: new evidence for Brazil combining national accounts, surveys and fiscal data, 2001-2015. World Inequality Database (WID. org) Working Paper Series **12**, 1–50 (2017)
Novokmet, F., Piketty, T., Zucman, G.: From soviets to oligarchs: inequality and property in Russia 1905-2016. J. Econ. Inequal. **16**(2), 189–223 (2018)
OECD: OECD framework for statistics on the distribution of household income, consumption and wealth. OECD Publishing (2013)
Piketty, T.: Income inequality in france, 1901–1998. J. Political Economy **111**(5), 1004–1042 (2003)
Piketty, T., Alvaredo, F., Assouad, L.: Measuring inequality in the middle east 1990-2016: the world's most unequal region? (2017)
Piketty, T., Chancel, L.: Indian income inequality, 1922-2014: from british raj to billionaire raj? (2017)
Piketty, T., Saez, E., Zucman, G.: Distributional national accounts: methods and estimates for the united states. Quarterly J. Econ. **133**(2), 553–609 (2018)
Saez, E., Zucman, G.: The rise of income and wealth inequality in America: evidence from distributional macroeconomic accounts. J. Econ. Perspectives **34**(4), 3–26 (2020)
Smith, M., Yagan, D., Zidar, O., Zwick, E.: Capitalists in the twenty-first century. Quarter. J. Econ. **134**(4), 1675–1745 (2019)
Suárez Serrato, J.C., Zidar, O.: Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. Amer. Econ. Rev. **106**(9), 2582–2624 (2016)
Torregrosa-Hetland, S.: Inequality in tax evasion: the case of the spanish income tax. Appl. Econ. Anal. **28**(83), 89–109 (2020)
United Nations: System of national accounts. https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (2008)
United Nations: System of National Accounts 2008 (2009)
WIL: Distributional national accounts guidelines: methods and concepts used in the world inequality database (2021)
Wolfson, M.C., Veall, M.R., Brooks, W.N., Murphy, B.B.: Piercing the veil: private corporations and the income of the affluent (2016)
Zwijnenburg, J.: The use of distributional national accounts in better capturing the top tail of the distribution. J. Econ. Inequal, pp. 1–10 (2022)