École des Hautes Études en Sciences Sociales

PARIS SCHOOL OF ECONOMICS
(Doctoral School 465 - Économie Panthéon Sorbonne)

Ph.D. Thesis Submitted for the Degree of Doctor of Philosophy in Economics

Economic Sciences

MAURICIO DE ROSA

Essays on Capital and Inequality in Latin America

Supervised by: Facundo Alvaredo

Defense date: May 22, 2023

Examiner: 1 Prof. Francisco Ferreira, London School of Economics

2 Prof. Salvatore Morelli, Università Degli Studi Roma Tre

Jury
1. Prof. Facundo Alvaredo (supervisor), Paris School of Economics/École des Hautes Études en Sciences Sociales
2. Prof. Francisco Ferreira (examiner), London School of Economics
3. Prof. Nora Lustig, Tulane University
4. Prof. Branko Milanovic, City University of New York
5. Prof. Salvatore Morelli (examiner), Università Degli Studi Roma Tre
6. Prof. Thomas Piketty, Paris School of Economics/École des Hautes Études en Sciences Sociales

École des Hautes Études en Sciences Sociales

École d'Économie de Paris
(École Doctorale 465- Économie Panthéon Sorbonne)

Thèse pour l'obtention du titre de Docteur en Sciences Économiques

Discipline : Sciences Économiques

MAURICIO DE ROSA

Essais sur le capital et les inégalités en Amérique Latine

Thèse dirigée par: Facundo Alvaredo

Date de soutenance : le 22 mai 2023

Rapporteur: 1 Prof. Francisco Ferreira, London School of Economics

2 Prof. Salvatore Morelli, Università Degli Studi Roma Tre

Jury
1. Prof. Facundo Alvaredo (directeur de thèse), Paris School of Economics/École des Hautes Études en Sciences Sociales
2. Prof. Francisco Ferreira (rapporteur), London School of Economics
3. Prof. Nora Lustig, Tulane University
4. Prof. Branko Milanovic, City University of New York
5. Prof. Salvatore Morelli (rapporteur), Università Degli Studi Roma Tre
6. Prof. Thomas Piketty, Paris School of Economics/École des Hautes Études en Sciences Sociales

Acknowledgements

Writing this section is at the same time the easiest and the hardest task of the whole thesis. On the one hand, the list of people to whom I am deeply grateful comes to mind immediately, and so do the feelings of gratitude for all the little and huge things they did, which made this possible. Still in a long standing project such as a PhD dissertation, it quickly becomes obvious just to what extent everything we do is a collective effort, even the things that at first glance appear to be the result of individual drive or creativity. Thus the list grows exponentially, the task becomes increasingly difficult and the words harder to find. Faced with this challenge, anything I could possibly write will come up as insufficient, sloppy or corny, yet here I go.

The first people I want to thank are Facundo Alvaredo and Thomas Piketty. They took turns to supervise this dissertation, the first half by Thomas and the second by Facundo, so I will proceed in that order. Thomas and his *Capital in the twenty first century* were the reason I fell in love with economics again. I had grown increasingly estrange from economics for a few years, under the feeling that there was very little there that resonated with the things I was worried about. But the book changed that and I decided to go full in, with dramatic consequences. I never told him this (it felt somewhat inappropriate) but when I decided to leave Uruguay to study abroad, I only applied to Paris School of Economics, because my only goal was to study with him. It was not an easy process but things turned out alright, and I got the chance to enjoy not only his classes, but also his advice, insights and hospitality (even welcoming us to his home multiple times –thank you Julia for that too!). He is for me, above all, the proof that we can do high quality research and at the same time be engaged in the struggles of our time, which are huge and with fearsome foes. Good research can indeed be a powerful weapon, and they know it. Thomas stood at the front-line, and he will always have my admiration and respect for that.

Facundo is *the* reason why all of this turned out just fine. I can't even count the times when he was absolutely decisive, since the very beginning. Facundo accepted to be on the jury of my Master thesis in Uruguay back in 2016, without even knowing me. He then invited me to PSE in 2017 for three months, which was without doubts a turning point for me, after which everything changed. He was also decisive in me being accepted to the Masters and the PhD. He was my Master thesis supervisor (this time at PSE) and also my supervisor in this last half of the PhD. Besides these formal positions, he was always a constant and uncompromising support in all possible ways. His generosity knew no boundaries, and I will be forever grateful. On top of this, working with him and getting his feedback were key in my training as a researcher. In the last few years, the political economy project discussion group (and this goes also to Pilar and Ingrid!) was of the utmost importance for me. It allowed me to channel my increasing concerns with current economic doctrines and open a gate to rediscover the discussions with which I "grew up" as an economist. This will have, I am now sure, immense consequences. I'm sure he'll hate reading this, but I am one (of many) Alvaredistas.

I'd also like to thank the rest of the Jury and Thesis Committee. I really feel my dissertation was read by researchers who I admire and been following for quite a while now, so I couldn't ask for more. With Nora we've been in touch for many years now, even before she accepted to be a part of the Thesis Committee. She was always incredibly generous and kind with me, and her vast knowledge was indeed a great contribution during the whole process of writing this thesis. The members of the Jury, Branko, Salvatore and Chico, took the time to go over a very long dissertation, and still provided extremely valuable, detailed and insightful comments that really helped me improve the thesis. Salvatore, who before reading the chapters had already given me plenty of feedback in congresses and one-to-one conversations on my wealth research; Chico, undoubtedly one of the top experts of inequality in Latin America with whom I share the worry of being overconfident of the inequality estimates we produce, was the best possible discussant of the main income inequality chapters; and Branko, who above his expertise in inequality also shares the worry of the theoretical foundations of our research, not only gave me some amazing feedback on the thesis' *epilogue*, but was also generous enough to include that discussion in his famous blog.

Next I would like to thank all my coauthors. This thesis has six chapters, out of which four are joint work with Facundo, Andrea Vigorito, Gabriel Burdín, Joan Vilá, Marc Morgan and Ignacio Flores. The thesis, by definition, would have been impossible without them. I'm sure I learned more from them than I could bring to the table. This is true but also slightly generic, so let me go in more detail, and I will start with Joan, Marc and Nacho. With these three I have spent more hours than I can recall working in really hard and demanding projects, but in a way, it never quite felt like work. Turns out these guys are my friends, and really good ones, so it's been an incredibly fun ride. I feel extremely lucky to have been able to spend as much time as I did with them. They are some of the best people I know, as simple as that. And on top of that they are (and I don't feel I'm exaggerating) the finest researchers you can find. They are just incredibly intelligent and hard working. I thought I worked hard until I saw these people in action, it's just ridiculous... The good thing about these three is that I am sure they will be around for a long time. As I write these words I imagine their faces reading them, it is hilarious how differently they would react. I love you guys.

Then comes Andrea. This is not the first time I write acknowledgements about her, and I'm sure I will fall short once more, no matter how much I write. Andrea has been by far the most important person in my training as a researcher. The funny thing is that I'm positive I'm not the only one who can say this, she has played this role for a crowd of people. You realise that someone has made an indelible stamp in you when you ponder about what would they think of what your are writing, even if they are not working on that particular project. She wouldn't like me to write this, but I am sure she is the best economist of this side of the world. Her knowledge about basically everything is limitless. If this feels like an overstatement, I dare anyone to talk to her. And please don't let her modesty fool you, she will outsmart you any day. As our friend Pablo Messina usually says, "she has read all the books". I know that's not possible (I really do), yet somehow I believe she actually has. She is just an outstanding person, one of a kind.

Time for the institutional acknowledgements. They may seem one of those thing that you kind of *have to* do, but you really don't mean to. Yet I actually do. I come from the far end of the world, where there are very few who can get even near a PhD, and in my case, as in many others, it is all because of Uruguay's public and tuition-free education system. I did primary and secondary school, as well as University (including the Masters) for free. I would probably not been able to do that if I were born anywhere else. I got, on top of that, a grant from *Agencia Nacional de Investigación e Innovación* to do my PhD, and kept my position as research assistant at the *Instituto de Economía*. Without these two key funding sources, this PhD probably wouldn't have been possible. I'm still working here, at *Universidad de la República*, my only true home, and I will be forever thankful for all which she has given me, which I'm sure I'll never be able to repay.

I'd like to thank all my friends from the World Inequality Lab and surrounding areas. They were some of the most committed and loud researchers I ever met. These people made my time spend in PSE far better and more interesting that I thought it could be. I'm thinking of Li (+ Lisa!), Richard, Lydia, the two Toms, Álvaro, Morten, Santi, Flor, Olivia, Nitin, Mark, Amory, Ander, Theresa, Yaz, Paolo and so many more. Moreover, to my non-economists friends, that beautiful gang that became my second parisian family: Mike and Eva (and the glorious Random Resistance Group), Ernesto, Sabrina, Paolo, Marcela and Gray. To you all, my deepest gratitude and love, nothing would've been the same without you.

I thank my family, the whole extended one! Their unconditional support was more important than they think. Everything was so dizzying that I guess I could never fully thank them. I remember before leaving for France in 2018 that I was deeply moved by just how many people were actively helping me in the most varied ways. The feeling stuck with me and every-time I looked at other international students I couldn't help but wonder just how many people were behind all of them, all the sacrifice and effort, invisible for the rest. I think of Andrea and Bernardo who above many other things, took care of my dog Vasili, Paio who went to visit and with whom we had the most amazing trip, my mom and sister, who went to spend the coldest new year of their lives with us in our little shoe-box apartment (¡qué cogote!), the whole family plotting to throw me a surprise welcome-home party, and the list goes on and on. My dad didn't get the chance to be a part of this, but he was around somehow (I did a final chapter on Marx after all, right?).

Anyone who knows me even a little bit will not be surprised by this last paragraph. It is obviously for Magu. Just to what extent the fact this dissertation even exists is the direct result of we being together, I don't think she'll ever know. She changed my life, and this is only one of the many things she did for me. She never stopped encouraging me to get as far as I could possibly go, not once. Even when absolutely nothing was clear and we both knew things could get really complicated really fast, she held the line. She left her life on pause to join me, and kept me going in the long periods we had to be apart. Now they look distant, like everything happened in a single second, but I've never missed anyone like I missed her in those months. Still life goes on, even when you are doing a PhD, and so did we. Then came Valentín, and all of the sudden the PhD, and all the rest of the world's affairs for that matter, stopped feeling as important as they used to. I wrote the final chapters of this thesis with Valen sleeping on me. He was my companion, and has been for the last year and a half (so long already!). This thesis is for you, mamarracho.

To Valentín.

So one day may
"cada niño un poco, todos tomarán,
de la misma leche y del mismo pan,
de la misma leche y del mismo pan"
Gurisito, 1969

Contents

List of Figures

List of Tables

Summary

This is a thesis about capital. It is intended to follow a logical sequence to which capital is the structuring concept. This came to be in part by design, but also because empirical research on inequality in Latin America brought us once and again to this dimension, as a gravity center that pulled us in. The thesis is made of three parts, with a total of five essays plus an epilogue.

The first part is dedicated to income inequality in Latin America. It is an effort to understand what actually happened with income inequality in the region in the first two decades of this century, which was a period in which most research points at a decline. In Chapter 1, written with Facundo Alvaredo, Marc Morgan and Ignacio Flores, we systematically compare income aggregates from household surveys, tax and social security records as well as national accounts for ten Latin American countries, accounting for over 80% of the region's population. We map micro-macro data discrepancies, which are very heterogeneous across countries and result in micro-data sources only accounting for about half of national income on average. By distinguishing between measurement and conceptual gaps, we document that half of the overall miss-match is the result of the former, and that capital incomes are the main source of discrepancy. Moreover, we find that the gap between survey and tax's top tails is large and increasing, especially when capital incomes are considered. The article is the first to carefully map and measure the data discrepancies and discuss its likely implications for inequality analysis in the region.

Chapter 2, written with Ignacio Flores and Marc Morgan, builds on the first and moves one step forward to actually combine the different data sources.[1] We correct household surveys with tax and social security data and then scale up to household income and national income. We document inequality trends for 2000-2019 in the region, discussing the differences between income series. We find that each adjustment stage moves inequality level upwards, but the

[1]In this thesis, cross-Chapter references are made not to other chapters within the document, but to the homonyms articles and working papers available online.

effect on trends is much more heterogeneous. The current narrative of decreasing income inequality in the region does hold for the post-tax bottom 99% and earnings inequality, but is sometimes neutralized or even reversed once the top 1% and capital incomes –which is to say the same thing– are accounted for. Capital incomes at the top thus are the key variable to explain changes both in levels and trends. While we do not intend to re-write the recent history of inequality in the region –and actually provide evidence that confirms it– we do claim that there is more to it and highlight the limits of Latin America's redistributive process.

The second part zeros-in the Uruguayan case, which may be interesting insofar it is, by all accounts, the least unequal country in the region. It is also made of two essays. Chapter 3 is written with Gabriel Burdín, Andrea Vigorito and Joan Vilá. It is an essay that provides a detailed combination of tax, social security and survey data to account for personal income inequality. Unlike Chapter 2, in which surveys are corrected to better account for top incomes, the departure point is the tax-social security micro-data, which accounts for three-quarters of the adult population. It is carefully supplemented with household survey to account for informal and untaxed personal incomes based on a sub-sample of matched tax-survey data. We find that although overall inequality did fall, top income shares remained stable and even slightly increased by the end of the period under analysis. We reconcile these trends across sources, documenting that the bulk of the divergence is located at the very top and is explained by capital incomes.

Chapter 4 is written with Joan Vilá and builds on the previous one. We use recently available national accounts series to scale up to national income. What distinguishes the essay is the use of novel owners-firms micro matched data to impute undistributed profits, which is always the trickiest part and usually relies on heavy assumptions for lack of better data. This allows us to account for the distributed/undistributed profits dynamics, which increases during the period, both from a national accounts and at firm's micro level. We show that, somewhat surprisingly, scaling up to national accounts has the effect of neutralizing and even reversing the upwards trend of top incomes shares found in the tax-survey data. Moreover, we are able to discuss effective taxation at the tax-survey, household and national income levels. Again, decisions by firm-owners of capital incomes' management at the firm level heavily drives not only the level of inequality, but also what the researches can actually observe, even if all micro-macro gaps remain constant.

Finally, the third part is made of a single essay, which is the result of a long-standing project which intended to dig into the source of these problematic capital incomes, i.e. wealth distribution. Thus, Chapter 5 is a slightly longer than the rest, since it intends to provide a complete depiction of wealth level, distribution and inheritance in Uruguay. It was a long process given the absence of adequate data, which had to be gathered and harmonized. The country's balance sheet and the resulting wealth to income ratio were estimated from scratch. Based on it and the capital incomes microdata of part two, I estimated wealth distribution based on the capitalization method, which was the work-horse methodology I chose for the main estimates. I was able to characterize top wealth holders and document a high level of wealth concentration at the top. As these results heavily depend on key assumptions such as the homogeneous rates of return and, more importantly, on the estimated capital incomes distribution itself, I systematically compared the main estimates with orthogonal data sources and methodologies. I used virtually *all* of the available strategies, i.e. a wealth household survey, real estate wealth tax micro-data, *Forbes* billionaires list and the estate multiplier method. These results, although still imperfect, do represent a unique effort to provide wealth estimates consistent both at the micro-macro levels, as well as with incomes distribution for a Latin American country.

After that fifth chapter, the thesis ends. However, I decided to include an Epilogue in 6, which ponders on the theoretical foundations of the empirical work of the previous five essays. These afterthoughts are the result of my increasing discomfort with the neoclassical growth theory, which I find extremely useful to support an evidence-based historical narrative in some cases, but entailing crippling limitations nonetheless. I discuss the concept of capital and its theory in light of past controversies and try to build a bridge with classical political economy concepts, in an attempt to provide different lenses through which to look at our current estimates. In the best case scenario, it is a contribution to better understand the key concept of capital; in the worst, it represents nothing more than my notes on the readings of the last few years. Either way, including this epilogue is what feels more honest with the process I went through during my PhD.

Résumé

Celle-ci est une thèse sur le capital. Elle entend suivre une logique dont le capital est le concept structurant. Cela s'est fait en partie par construction, mais aussi parce que la recherche empirique sur les inégalités en Amérique latine nous oblige à graviter vers cette dimension. La thèse est composée de trois parties, avec un total de cinq essais plus un épilogue.

La première partie est consacrée aux inégalités de revenus en Amérique latine. Il s'agit d'un effort pour comprendre ce qui s'est réellement passé avec l'inégalité des revenus dans la région au cours des deux premières décennies du siècle courant, période au cours de laquelle la plupart des recherches indiquent un déclin. Dans le chapitre 1, écrit avec Facundo Alvaredo, Marc Morgan et Ignacio Flores, nous comparons systématiquement les agrégats de revenus des enquêtes auprès des ménages, des registres fiscaux et de sécurité sociale ainsi que des comptes nationaux pour dix pays d'Amérique latine, représentant plus de 80% de la population de la région. Nous comparons les écarts de données micro-macro, très hétérogènes d'un pays à l'autre, qui se traduisent par des micro-données qui ne représentent qu'environ la moitié du revenu national en moyenne. En distinguant les écarts de mesure et les écarts conceptuels, nous montrons que la moitié de la difference globale est expliquée par des problèmes de mesure et que les revenus du capital sont la principale cause. De plus, nous constatons que l'écart entre les queues supérieures de la distribution entre l'enquête et les sources administratives est important et croissant, en particulier lorsque les revenus du capital sont pris en compte. Cet article est le premier à cartographier et à mesurer soigneusement les écarts de données et à discuter de ses implications probables pour l'analyse des inégalités dans la région.

Le chapitre 2, écrit avec Ignacio Flores et Marc Morgan, s'appuie sur le premier, faisaint un pas vers la combinaison des différentes sources de données pour corriger des biais de mesure.[1] Nous corrigeons les enquêtes auprès des ménages avec des données fiscales et de

[1] Dans cette thèse, les références interchapitres sont faites et aux articles homonymes et documents de travail disponibles en ligne.

sécurité sociale, puis nous repondérons les valeurs pour qu'elles soient consistantes avec le revenu macroéconomique des ménages ; nous imputons aussi d'autres revenus pour étudier la distribution de la totalité du revenu national. Nous documentons les tendances des inégalités pour 2000-2019 dans la région, en discutant des différences entre les séries de revenus. Nous constatons que chaque étape d'ajustement fait monter le niveau d'inégalité, mais l'effet sur les tendances est hétérogène. Le récit actuel de la diminution de l'inégalité des revenus s'applique au 99% inférieur de la population si on parle de revenus après impôts, mais la baisse des inégalités est parfois neutralisée ou même inversée une fois que les 1% supérieurs et les revenus du capital sont pris en compte. Les revenus du capital au sommet sont donc la clé pour expliquer à la fois les changement de niveaux et de tendances. Nous n'avons pas l'intention de réécrire l'histoire récente des inégalités dans la région, maisn nous affirmons qu'il y a des nuances au processus et nous en soulignons les limites.

La deuxième partie se focalise dans le cas de l'Uruguay, ce qui peut être intéressant dans la mesure où il s'agit sans doute du pays le moins inégalitaire de la région. Elle est également composée de deux essais. Le chapitre 3 est co-écrit avec Gabriel Burdín, Andrea Vigorito et Joan Vilá. Il s'agit d'un essai qui fournit une combinaison détaillée de données fiscales, de sécurité sociale et d'enquêtes pour tenir compte de l'inégalité des revenus personnels. Contrairement au chapitre 2, dans lequel les enquêtes sont corrigées pour mieux tenir compte des hauts revenus, le point de départ est la micro-donnée administrative qui représente les trois quarts de la population adulte. Elle est soigneusement complétée par une enquête auprès des ménages pour tenir compte des revenus personnels informels et non imposés sur la base d'un sous-échantillon de données d'enquêtes fiscales appariées. Nous constatons que bien que l'inégalité globale ait diminué, les parts des revenus les plus élevés sont restées stables et ont même légèrement augmenté à la fin de la période analysée. Nous réconcilions ces tendances entre les sources, documentant que l'essentiel de la divergence se situe tout en haut et s'explique par les revenus du capital.

Le chapitre 4 est co-écrit avec Joan Vilá et s'appuie aussi sur le précédent. Nous utilisons des séries de comptes nationaux récemment disponibles pour passer au revenu national. Ce qui distingue l'essai, c'est l'utilisation de nouvelles micro-données appariées propriétaires-entreprises pour imputer les bénéfices non distribués, ce qui est toujours la partie la plus délicate et repose généralement sur des hypothèses lourdes faute de meilleures données. Cela nous permet de rendre compte de la dynamique des bénéfices distribués/non distribués, qui augmente au cours de la période, tant à partir des comptes nationaux qu'au niveau micro de l'entreprise. Nous montrons que, de manière quelque peu surprenante, le passage à l'échelle des comptes nationaux a pour effet de neutraliser et même d'inverser la tendance à la hausse des parts des revenus les plus élevés constatée dans les données des enquêtes fiscales. De plus, nous sommes en mesure de discuter de la fiscalité effective au niveau de l'enquête fiscale, des ménages et du revenu national. Encore une fois, les décisions des propriétaires d'entreprise sur la gestion des revenus du capital au niveau de l'entreprise influencent fortement non seulement le niveau d'inégalité, mais aussi ce que les recherches peuvent réellement observer, même si tous les écarts micro-macro restent constants.

Enfin, la troisième partie est constituée d'un seul essai visant à creuser sur l'origine des revenus du capital qui sont à la source de toutes ces problématiques. Il s'agit d'une étude sur la répartition des richesses. Le chapitre 5 est légèrement plus long que le reste, car il fournit une description complète du niveau de richesse, de sa répartition et de la distribution des héritages en Uruguay. Ce fut un long processus étant donnée l'absence de données adéquates, qui ont dû être rassemblées, harmonisées et parfois estimées pour la première fois. Sur la base de ces données et des micro données sur les revenus du capital de la deuxième partie, j'ai estimé la répartition de la richesse selon la méthode de capitalisation, que je considère comme méthode de réference. J'ai pu caractériser les principaux détenteurs de la richesse ainsi que documenter un niveau élevé de concentration de la richesse au sommet. Comme ces résultats dépendent fortement d'hypothèses clés, telles que les taux de rendement homogènes et, plus important encore, de la distribution estimée des revenus du capital elle-même, j'ai systématiquement comparé les principales estimations avec des sources de données et des méthodologies orthogonales. J'ai utilisé pratiquement *toutes* les stratégies disponibles, c'est-à-dire une enquête auprès des ménages sur le patrimoine, des micro-données sur l'impôt sur la fortune immobilière, la liste des milliardaires *Forbes* et la méthode du multiplicateur basée sur des données d'héritage. Ces résultats, bien qu'imparfaits, représentent un effort unique pour fournir des estimations de richesse cohérentes à la fois aux niveaux micro-macro, ainsi qu'avec la distribution des revenus pour un pays d'Amérique latine.

Après ce cinquième chapitre, la thèse se termine. Cependant, j'ai décidé d'inclure un épilogue dans 6, qui s'interroge sur les fondements théoriques du travail empirique des cinq essais précédents. Ces réflexions après coup sont le résultat de mon incommodité croissante avec la théorie néoclassique de la croissance, que je trouve extrêmement utile pour soutenir un récit historique empiriquement fondé dans certains cas, mais qui comporte néanmoins des limitations considerables. Je discute du concept de capital et de sa théorie à la lumière des controverses passées et j'essaie de construire un pont avec les concepts classiques d'économie politique, dans une tentative de fournir différentes perspectives à travers lesquelles examiner nos estimations actuelles. Dans le meilleur des cas, ce chapitre est une contribution à la compréhension conceptuelle du capital; au pire, il ne s'agit pas plus que de mes notes sur les lectures des dernières années. Quoi qu'il en soit, l'inclusion de cet épilogue me semble cohérente avec le processus que j'ai traversé pendant mon doctorat.

General Introduction

Although inequality is one of the oldest concerns in economics, during the twentieth century its study became increasingly marginalised (Atkinson, 2015). In recent years, however, these issues have returned to the spotlight alongside the increasing income inequality in developed countries (Piketty, 2014), the significant increase in the wealth to income ratio (Piketty and Zucman, 2014) and, in some cases, also of wealth inequality (Saez and Zucman, 2016; Kopczuk and Saez, 2004; Alvaredo, Atkinson, and Morelli, 2018). For the rest of the world, although research is making fast progress, inequality trends are still far from clear.

Within the highly heterogeneous part of the globe usually called *developing world*, Latin America represents a significant proportion with over 600 million inhabitants, and has been the object of an intense discussion in light of the significant social, political and economical changes that the region has witnessed in the last few decades. In particular, income inequality seems to have experienced a downturn since the early 2000s, in the context of vigorous economic growth and redistributive public policies (Cornia, 2014b). Yet, it is still very high compared with European countries, and the decreasing trend has stopped or even reversed in recent years (Gasparini, Bracco, Galeano, and Pistorio, 2018). The question of whether these trends are the result of poor performance of household surveys is still ongoing, with mixed results depending on the country, as the evidence based on income tax records accumulates (Alvaredo and Londoño Velez, 2014; Alvaredo, 2010; Morgan, 2017b; Flores, 2021).

Although relatively recent tax-based estimates allow researchers to look at income inequality and especially top income groups under new light, these estimates still face many challenges. First, the difficulty to adequately combining them with survey data to account for low income groups and, especially the informal sector. Second, there are a number of income concepts that may still not be captured by tax-survey data (WIL, 2020). Indeed, income inequality trends need to be studied in a framework that allows to analyse it simultaneously and consistently with economic growth. These variables do not necessarily refer to the same income definitions, nor they coincide in aggregate terms, making growth and inequality studies inherently inconsistent. Moreover, in order to understand trends in income inequality, significant progress still needs to be made to understand wealth distribution in the region, for which estimates are extremely scarce (see e.g. Torche and Spilerman (2006); Gandelman, Lluberas, et al. (2022)) and based on very limited data.

The aim of this dissertation is to engage in the conversation of inequality trends in the region, based on the careful gathering and combination of surveys, social security and tax administrative data and national accounts for the whole region. The specific question that is addressed is to what extent inequality actually fell in the region, and what this might tell us about the current survey-based narrative. To answer it, the first two chapters will study the region as a whole, while chapters three to five will address the case of Uruguay in depth, both regarding income and wealth distribution. One of the main findings is that capital incomes and wealth distribution are the key to understand the level and trend in income inequality, hence the final chapter discusses the concept of capital from a theoretical viewpoint.

Part I. The distribution of income in Latin America

Chapter 1

The first part is dedicated to income inequality in Latin America. It is an effort to understand what actually happened with income inequality in the region in the first two decades of this century, which was a period in which most research points at a decline. In Chapter 1, written with Facundo Alvaredo, Marc Morgan and Ignacio Flores, we systematically compare income aggregates from household surveys, tax and social security records as well as national accounts for ten Latin American countries, accounting for over 80% of the region's population.

The macroeconomic aggregates from the System of National Accounts (SNA) are the most widely used measures of economic activity and are considered as benchmark numbers. At the same time, the applied analysis of the distribution of income has mostly relied on household surveys and administrative data, and this has usually been approached quite independently from the SNA. One of the observed results of such a disconnect has been the development of a large and sometimes increasing gap between aggregates from inequality studies based on microeconomic data, i.e. surveys and administrative records, and the SNA. The discrepancies can be seen in the levels of income, as well as in their growth rates (see, for example Ravallion (2003); Deaton (2005); Bourguignon (2015); Nolan, Roser, and Thewissen (2019)), and can attain particularly high levels in developing countries. While it may not be surprising that national income is larger than the income concepts traditionally used to study inequality, it has also been growing faster. It has been argued that these discrepancies make it hard to assess how macroeconomic growth is distributed across income groups, and to what extent existing distributional statistics are a proper representation of the income flows in an economy. Recent work has embarked on a process of combining the various available data sources (surveys, administrative records, rich lists) upscaled -with further imputations- to SNA totals, aimed to produce comparable distributional results. These include, among others, WIL (2020), Fixler, Johnson, Craig, and Furlong (2017) and a project coordinated by the OECD (Zwijnenburg, 2019). While the existing gaps have sometimes strengthened the feelings of uncertainty about inequality measurement, these new approaches have taken for granted the numbers provided by the national accounts, a practice that does not always contribute to diminish those feelings, at least in the case of developing countries.

Data availability is arguably one of the main restrictions to properly study the distributional aggregates that feed the research on income distribution. In Latin America, most of this research has used survey data to analyze the evolution of inequality. The notable finding of this research is that the region experienced a historic decline in income inequality since the twenty-first century, attributed to a mix of vigorous economic growth and redistributive public policies (López-Calva and Lustig, 2010; Cornia, 2014b; Rodríguez-Castelán, López-Calva, Lustig, and Valderrama, 2016; Messina and Silva, 2017; Bértola and Williamson, 2017; Gasparini et al., 2018). However, question marks over the reliability of household surveys persist, as evidence on top incomes from tax records accumulates (Alvaredo, 2010; Alvaredo and Londoño Velez, 2014; Alvaredo, Garriga, and Pinto, 2017; Burdín, De Rosa, Vigorito, and Vilá, 2022; Cano, 2015; Rossignolo, Oliva, and Villacreses, 2016; Morgan and Souza, 2019; Flores, Sanhueza, Atria, and Mayer, 2020; Zuniga-Cordero, 2018). Recent research has also found that capital incomes appear to be remarkably less covered than labor incomes when survey aggregates are compared to SNA aggregates (Törmälehto, 2011; Bourguignon, 2015; Flores, 2021).

We document that official inequality estimates coming from household surveys only account for around half of national income and 60% of household sector's income. Important differences in this ratio exist between some countries, ranging from 50-60% in Brazil, to 25-30% in Mexico, which result in 65-70% and 35-40% of household income coverage respectively. Of particular significance is the fact that in most countries, the total survey income is a declining share of national income over the course of the last two decades. We decompose the gap into two quantifiable components, which we call the "measurement gap" – the gap between measurable household income in surveys and equivalent household income in the SNA – and the "conceptual gap" – the portion of national income that is not directly received by households or measurable in survey questionnaires. We find that the measurement gap accounts for roughly 52% of the survey-SNA gap on average, with substantial variation between countries, and a general increasing tendency in recent years. We estimate that overwhelming majority of this gap is due to missing capital income received by households, consistent with other recent literature cited above. Moreover, assuming that administrative data better account for incomes in the right tail of the distribution, we find an increasing undercoverage of top incomes in surveys, especially when non-wage incomes are considered. Additionally, the top tail of the tax data distribution not only depicts higher concentration levels than surveys, but also a higher degree of top income inequality, meaning that the income of individuals at the top in surveys are progressively less covered as one moves up the distribution. These findings have notable implications for analyses of inequality levels and trends within and between countries, which should be given greater attention by the literature.

Chapter 2

Chapter 2, written with Ignacio Flores and Marc Morgan, builds on the first and moves one step forward to actually combine the different data sources. Numerous studies have documented and explained the apparent decline of income inequality taking place throughout Latin American countries during the first decade and a half of the twenty-first century (López-Calva and Lustig, 2010; Lustig, López-Calva, and Ortiz-Juarez, 2011; Cornia, 2014b; Alvaredo and Gasparini, 2015; Gasparini et al., 2018). This trend has even been viewed as historically unprecedented in a region characterised by extreme inequality legacies (Bértola and Williamson, 2017). However, its narrative, built on the use of publicly available household survey data, has come to be questioned by the increasing use of administrative data on upper incomes in the region (Alvaredo, 2010; Alvaredo and Londoño Velez, 2014; Morgan, 2018; Souza, 2018; Flores et al., 2020; Burdín et al., 2022), which have shown either milder reductions in top income concentration or more stable, if not increasing, trends in some countries. These doubts are compounded by the large discrepancies between incomes in micro data sources (surveys, tax data) and the national accounts.

Such gaps present us with a distributional conundrum: if they are widest in capital incomes, as has been found historically and more recently, then this would entail significant repercussions for existing inequality indicators (Chapter 1). Moreover, if these gaps are subject to changes over time, as also appears to be the case, our assessment of inequality trends would be severely compromised. With all this cumulative information at our disposal, how confident can we be in thinking that Latin America was an exceptional outlier in the global income inequality narrative? The stakes of this question are high, since accepting the survey-based narrative outright, in the context of large and growing micro-macro data gaps, could mean putting into serious question the official macroeconomic growth statistics of countries in the region. Seeing whether the conventional narrative on Latin American income inequality is robust to the reconciling of micro-macro data gaps is the focus of the chapter.

In order to build our estimates, we combine harmonized survey microdata with administrative data from tax records and social security registers (based on the re-weighting method put forward by Blanchet, Flores, and Morgan, 2022) before scaling incomes to the national macroeconomic accounts (both the household sector accounts and the total economy sector accounts). Thus, we reconcile all available income data to build inequality estimates that not only adjust for surveys' measurement issues, but also ensure overall macroeconomic consistency. As anticipated in chapter 1, the adjustments end up doubling the total income originally declared in most surveys. Hence, to ensure transparency, we show the impact that each step of our methodology has for the resulting distributions. We distinguish four steps: first, we estimate the distribution of income in the harmonized survey data; second, we adjust for the low representativeness of top incomes in surveys using administrative records; third, we scale the main income components to their matched national accounts aggregates (these are wages, property incomes, mixed income, pensions and imputed rents); and fourth, we impute incomes not flowing to the household sector in the national accounts (corporate retained earnings and other incomes) or incomes that need to be added back (such as net product and production taxes) to reach the net national income of the total economy.

Our main contribution is to revisit the prevailing narrative of falling inequality in Latin America over the first fifteen years of the twenty first century. We attempt to reconcile competing inequality narratives by clarifying issues that affect comparability such as units of analysis, income concepts and the choice of inequality indicators. More importantly, we analyze the contribution of capital incomes and top income groups, which are by all accounts the main missing pieces of household surveys. We document that inequality among the bottom 99% of the total income distribution and among wage earners falls even after making all adjustments. We show how divergent trends in total income inequality are the result of an increasing contribution of capital incomes and top 1% incomes after each adjustment procedure. This is due to both an increasing distance between the top 1% and the bottom 99%, and to increasing inequality within the top 1%. Thus, we do not fully contradict the prevailing narrative; if anything, we confirm it with some qualifications. Fundamentally, we claim that the role played by capital incomes and top 1% incomes reveals the limits of Latin America's much heralded re-distributive effort of the early twenty-first century, even if certain policies appear to be key for a robust inequality decline (such as public spending on health and education).

Part II. Falling inequality in Uruguay

Chapter 3

The second part zeros-in the Uruguayan case, which may be interesting insofar it is, by all accounts, the least unequal country in the region. It is also made of two essays. Chapter 3 is written with Gabriel Burdín, Andrea Vigorito and Joan Vilá. It is an essay that provides a detailed combination of tax, social security and survey data to account for personal income inequality. We investigate whether the recent inequality fall in Uruguay is robust to the use of different data sets and whether it implies modifications of the shares held by top income groups, and, particularly, capital income receivers. Specifically, we analyse primary income inequality, comparing harmonized household surveys and corrected micro-data from tax records. We provide an in-depth analysis of the main factors underlying the evolution of the income distribution in the two data sets, focusing on the upper tail and the evolution of the different income sources. We delve into the characteristics of the top income earners and the firms that they work for or own, which also allows us to account better for capital income' shares. Uruguay is an interesting case study because we are able to exploit a unique data set of matched social security data and personal and firm income tax records at the individual level that covers the period of significant GDP growth and inequality decrease (2009 to 2016).

This research is mainly based on a comprehensive anonymized administrative personal income tax micro-database (*Impuesto a la Renta de las Personas Físicas (IRPF) and Impuesto a la Seguridad Social (IASS)*) matched to the balance sheets that corresponding firms submitted to the tax authorities (*Dirección General Impositiva, DGI*) in 2009-2016. The latter step is necessary to identify completely the capital incomes and characteristics of employers. Since they include information from social security records, these data cover the universe of formal workers (with earnings below or above the minimum taxable income), capital income earners and pensioners, comprising around 75% of the adult population aged 20 and above. At the same time, we use the micro-data from the official household survey *Encuestas Continuas de Hogares, ECH)* gathered by the *Instituto Nacional de Estadística (INE)* and a subsample of 2012/2013 ECH-DGI observations linked at the individual level to compute the underreporting rates in the lower tail of administrative data. The broad coverage of our administrative micro-data and the availability of a unique data set of survey-tax data matched at the individual level for a sub-set of households allow us to depart from the tax records and correct the lower half of the income distribution with household survey information, building on the methodology initially proposed by Atkinson (2007). Specifically, we add labour earnings from informal workers and underreported formal income, creating a corrected tax income variable. We also present several robustness checks by correcting harmonized household survey income with tax data (Alvaredo, 2011; Blanchet et al., 2022). To identify the main characteristics of top income receivers, we carry out a multivariate analysis exploiting the matched individual-firm databases.

Our findings indicate that the synthetic indexes present declining trends in corrected tax income and harmonized survey income and, in both cases, inequality declined at the bottom 99%. However, the driving forces under the inequality reduction are at odds in the two cases. While the equalization process in the harmonized household survey income was lead by a reduction in the concentration of the top 1%, the opposite applies to corrected tax income, in which the redistribution in the bottom 99% outweighed the increasing inequality at the top. In the latter case, the inverted Pareto coefficient has grown steadily since 2012. As a result, the top income shares exhibit a decline in harmonized household surveys and an increase in corrected personal income tax data.

We also show that the evolution of the top income shares in corrected tax income is closely connected to the increased participation and concentration of capital income in the upper tail of the income distribution. Furthermore, we document that the top income holders are closely connected to the increased share of capital income in the top 1% and 0.5% of the income distribution. Most top income holders are men and capital income receivers. Meanwhile, among the subset of top income earners receiving labour income, the most salient group corresponds to health services.

Chapter 4

Chapter 4 is written with Joan Vilá and builds on Chapter 3. 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 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. Chapter 1 showed 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 (Chapter 2). 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, Yagan, Zidar, and Zwick 2019). Adequately accounting for capital incomes therefore requires detailed data on firms and owners (WIL, 2020), which is very rarely available (Fairfield and Jorratt De Luis, 2016; Alstadsæter, Jacob, Kopczuk, and Telle, 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 chapter, 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 build on Chapter 1, which we supplement 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. 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.

Part III. The distribution of wealth

Chapter 5

In the economic literature on the accumulation and distribution of wealth, there is a sharp disparity between the evidence gathered in the past decade for a handful of mostly developed countries and the scarcity of estimates available for the rest of the world. The bar is set high by the literature, which provides wealth distribution estimates based on a variety of methods and sources (e.g. Saez and Zucman (2016, 2022); Alvaredo et al. (2018); Garbinti, Goupille-Lebret, and Piketty (2017)). These estimates are in turn consistent with national wealth estimates (Piketty and Zucman, 2014; Blanco, Bauluz, and Martínez-Toledano, 2021), hence providing a full account of wealth dynamics, both micro- and macro-economically.

Wealth estimates for the developing world are for the most part the result of regression analyses based on countries with available data or household surveys in best case scenarios, with little reference to macro wealth or income aggregates, thus hindering our ability to provide comparable estimates and a credible narrative. I contribute to the closing of this widening gap by estimating the first fully consistent set of estimates of aggregate wealth and its distribution for a Latin American country. A wide array of surveys, detailed personal income tax micro-data, cadastre administrative data, national accounts, owner-decedent micro-data, and firm balance sheets are carefully combined to provide a consistent overview of the level, composition and personal distribution of wealth in Uruguay.

As in most of the developing world and even in many rich countries, there are no national accounts estimates of Uruguay's balance sheet. Therefore, I estimate wealth-to-income ratios based on a wide range of secondary sources, including cadastral administrative data, prices of land and housing properties, firm tax records, and central bank financial data, among others. These estimates, although imperfect, follow the tradition of aggregate wealth estimation by independent scholars of the 17th to early 20th century. Results show that the book value wealth-to-income ratio is around 500%, comparable to what is observed in developed economies, where it is around 500-700% (Piketty and Zucman, 2014). Public net wealth is positive but decreasing, from 50% at the beginning of the century to around 25% by the end of the period under analysis. Gross domestic capital is 30 percentage points higher than net national wealth as a result of a negative net foreign asset position. Approximately half of household wealth takes the form of financial assets (including ownership of the private corporate sector), while housing reaches 100-150% of national income.

The main distributional estimates are based on the capitalization method, which consists of estimating individual net wealth by capitalizing personal capital incomes, using capitalization factors for each type of wealth that are equivalent to the inverse of their macro rate of return, i.e., consistent with macro wealth aggregates. Capital incomes are mainly drawn from high-quality tax record micro-data—which covers 75% of adult population—merged and combined with firm tax records, household survey data, and national accounts. This capital incomes database is the result of the national income distribution series estimated in Chapter 4, and it therefore provides income-wealth consistency, both at the micro and macro levels. Private personal wealth inequality is relatively stable over the period 2009-2016 but at a very high level: over 38-40% of net private wealth is owned by the wealthiest 1%, and the top 10%'s share is around 77-79%. These estimates would locate Uruguay as a relatively high wealth inequality country compared to France, closer to estimates for Spain or the US, but lower than extreme cases such as South Africa. Moreover, I characterize wealth owners in terms of age and gender, showing that men have higher wealth for all age groups (with little evidence of life-cycle accumulation patterns) and represent 70-80% of the top fractiles. I also document a high correspondence between wealth distribution and total income distribution for the top fractiles.

The inequality estimates in this article are triangulated with four other empirical approaches to provide greater certainty about the overall conclusions. First, the main results are compared with the Wealth Household Survey (*Encuesta Financiera de los Hogares Uruguayos*, EFHU) of 2013, which covers similar assets as the ones estimated with the capitalization method. The top 10% is 10 percentage points lower in the survey, with most of this difference being explained, as expected, by the top 1 and top 0.1%. Moreover, the different concentration profiles are explained by wealth composition rather than by the distribution of each asset: contrary to the results of the capitalization method, in the wealth survey, the least unequally distributed asset—housing—represents the bulk of wealth. Second, I compare data from the *Forbes* billionaires list to the very top wealth holders from the capitalization method data; these show very similar net wealth levels, which provides reassurance about the capitalization method's accuracy for individuals at the very top. Third, a novel administrative dataset of decedent real estate owners is constructed, allowing me to estimate urban and rural real estate distribution based on a simplified version of the estate multiplier method (Alvaredo et al., 2018; Berman and Morelli, 2021), which essentially entails weighting the decedent population by the average mortality rate. Results show that the top 1%'s share of real estate wealth reaches around 20-25%, whilst top 0.1% is stable around 10%. When considering urban properties only, it shows higher concentration for the top 10 and 1% than the capitalization method does (10 and 5 percentage points, respectively), but 2 points lower for the top 0.1%. The estimates are, however, remarkably close to wealth survey estimates, which suggests that capitalization method estimates represent a lower bound. Fourth, the top 0.1%'s housing share is calculated based on a wealth tax that covers a small fraction of the population (a little over 0.3%) and targets mainly real estate assets, resulting in a top 0.1% of 4-5%, almost identical to capitalization method estimates.

Chapter 6

Capital has made an astonishing comeback to the empirical distributional research agenda. Based on novel data and a variety of revised and new methodologies, the recent wealth accumulation, inheritance and inequality results are undoubtedly better estimated than ever before. This allows researchers to simultaneously account for variables such as growth, the capital share, the wealth to income ratio, inheritance flows, rate of return and wealth distribution under SNA's framework, endowing the empirical inequality literature with renewed firepower. Consistently estimating all these variables for Latin American countries was the main objective of this thesis. In its second part and particularly in Chapter 5, I estimated all key wealth variables for Uruguay under the Distributional National Accounts framework (WIL, 2020), in an attempt to catch up with the new literature and to provide comparable estimates for a developing country, which are extremely rare. In all these five articles, we showed that both as an estimation challenge as well as an inequality driving force, *capital* was the gravity-center of our empirical endeavors.

These variables and capital itself are not only estimated consistently with National Accounts, but can also be easily linked with standard neoclassical growth models, hence providing the theoretical foundation for the wealth accumulation dynamics (Piketty and Zucman, 2015). While there is no 'unified theory of inequality' (Atkinson and Bourguignon, 2000), for macro-distributional purposes standard neoclassical growth models are the main reference point. These models are admittedly limited, in particular the "one-good, perfect competition model is not a very satisfactory model, to say the least " (Piketty, 2015, p. 81). Nevertheless, they do provide important insights and intuitions on these macro variables and their likely evolution in the future (Piketty, 2014), and have been extensively debated (see e.g. Acemoglu and Robinson (2015); Jones (2015); Piketty (2015)).

Debates over these broad set of models are not new, as the standard neoclassical growth theory is rooted in extremely questioned assumptions and definitions. To begin with, the very definition of *capital* is problematic and has been subject to a significant amount of controversy (Hodgson, 2014). The last of the great controversies that raged during the 1950s until mid 1980s, i.e. the famous Cambridge Capital controversies, was primarily focused on how to measure capital. Yet, that was only the corollary: the main issue at stake was the very essence of what capital was (Harcourt, 2014). Kick-started by the call to arms of Joan Robinson's assault on the existence of a production function (Robinson, 1954), it was shown that neoclassical growth theory was unable to provide convincing explanations for the main driving variables of the capitalist system, especially for the rate of return r (Cohen and Harcourt, 2003). In particular, it was shown that once one leaves the one-commodity-model assumption, it is not longer possible to determine the rate of return to capital, hence turning it impossible to provide an explanation for the macro-distribution of income. Although the exact significance of the overall conclusion was not settled (and nor will I try to do so in this essay), it was indeed admitted even by the neoclassical side that standard growth models were unable to produce an adequate theory of factor prices, i.e. a theory of income distribution (Solow, 2000).

Marxian insights have been highlighted by many authors who were not themselves Marxists as Veblen, Schumpeter or post-Keynesians such as Sraffa (Bellofiore, 2008) or Robinson (Alves, 2022). Indeed, in her *essay on Marxian economics*, Robinson said that if "the orthodox notion of a definite supply price of capital thus disintegrates upon examination, we are left with nothing but Marx's notion that capital is accumulated and maintained because capitalists are forced to accumulate in order to survive" (Robinson, 1942, p. 61). Following this thread, I explore the accounting links between the neoclassical macro distributional theory and Marx's labour theory variables (Marx, 1867), in an attempt to better understand their differences and the implications for empirical research. The aim is simply to understand the conditions under which there is *accounting correspondence* between these two approaches. The starting point is the recognition of the conceptual gaps between similarly named variables. In particular, I contrast the definitions of the rate of return r and Marx's rate of profits P and, more importantly, I discuss the fact that –unlike neoclassical theory– Marx's capital is closer to a flow than to a stock. These two points alone help clarify much of the confusion which often result from these terms. Moreover, I show that under a one-good model assumption, accounting correspondence does exist between available estimates and Marx's labour theory of value, which can in turn be linked with main variables of standard neoclassical growth model and hence to the empirical inequality literature.

After documenting the theoretical differences between the definition of capital and the rate of return (profits), I show that, under a one-good model and closed economy's assumption, it is possible to interpret both labour theory of value and the new empirical literature in a simple unified accounting framework. Moreover, it is possible to establish the accounting links between famous drivers of the capitalist system such as $r > g$ and Marx's falling rate of profits. What I show is that under such a model, $r > g$ and the falling rate of profits yield perfectly consistent results. Specifically, I show that a stable rate of return r larger than the growth rate g, results in both an increase in the wealth to income ratio and the capital share, as well as in falling rate of profits P, which is offset at the beginning by the increase in the capital share but falling in the long run regardless. In the best case scenario, this simple exercise provides a bridge with existing empirical literature and classical political economy, including Marx. The main takeaway is that empirical wealth and income inequality literature needs not to get corseted in a restrictive neoclassical framework. However, I argue that while at the one-commodity-model level there is overall accounting consistency between the different approaches, the shift from a scarcity theory of value to a the classical approach entails relevant implications. One immediate consequence is that to the accumulation of wealth mechanisms one should add exploitation to work and inheritance, and that completely changes the narrative.

Overall contribution

The overall contribution of this dissertation is threefold. First, it is a contribution to the efforts of countless researchers to adequately estimate levels and trends of income and wealth inequality in Latin America. The wide array of methodologies, the massive amount of data and the efforts to dialogue with the current narrative directly aim at this primary and important goal. The second one is the identification of capital, both flow and stock-wise, as the main dimension that allows to understand divergent trends across data sources and methodologies for the study of inequality. By estimating the effect of previously unaccounted for capital incomes, there is not only a methodological point to be made, but also a political one, since it reveals one of the main limitations of the Latin American re-distributional process of the first fifteen years of this century, i.e., the ineffectiveness to target capital incomes and the top 1%. Finally, it is an attempt to contribute to better theoretically understand capital, and especially to provide different lenses through which we can look at these trends, without being trapped by the neoclassical mystification of capital.

PART I. The distribution of income in Latin America

Chapter 1

The inequality (or the growth) we measure: data gaps and the distribution of incomes

Joint work with Facundo Alvaredo, Ignacio Flores and Marc Morgan.

Abstract

There is a large gap between income estimates used in inequality studies and macroeconomic statistics. This makes it hard to assess how economic growth is distributed across the population, and to what extent mainstream distributional statistics are an accurate representation of income flows. We take stock of these discrepancies by confronting estimates of the income distribution from surveys, administrative records and aggregates from the system of national accounts, thoroughly documenting them over the past two decades for ten Latin American countries. We find that surveys only account for around half of the national income and 60% of household income in the region. Measurement gaps account for just over half of the overall gap on average, while the rest is due to conceptual differences across data sets. Measurement gaps have been growing fast for many countries, the bulk being due to non-covered capital income. We also compare the top tails in administrative data and surveys, finding diverging averages –especially for non-wage incomes– and different shapes. We discuss the degree to which inequality levels and trends could be affected.

1. Introduction

The development of economic statistics is a lengthy historical process that involves the views of the dominant doctrine, the construction of a body of conventions, and the limits of available data. The production of such statistics engage governments, central banks, official statistics offices, and research institutions at different stages of the process. The macroeconomic aggregates from the System of National Accounts (SNA), such as Gross Domestic Product (GDP) and National Income (NI), are the most widely used measures of economic activity and are considered as benchmark numbers. In the early years of the SNA, national accountants were also experts in distributional issues, as the inter-linkages between the estimation of national income and its distribution were clearly recognized.[1] In the subsequent decades, this link was lost and the two fields went separate ways. The focus of the SNA has so far been on the aggregates of the main institutional sectors in the economy, distinguishing the household sector, the corporate sector, the government sector and the foreign sector. At the same time, the applied analysis of the distribution of income has mostly relied on household surveys and administrative data, and this has usually been approached quite independently from the SNA.

One of the observed results of such a disconnect has been the development of a large and sometimes increasing gap between aggregates from inequality studies based on microeconomic data, i.e. surveys and administrative records, and the SNA. The discrepancies can be seen in the levels of income, as well as in their growth rates (see, for example Ravallion (2003); Deaton (2005); Bourguignon (2015); Nolan et al. (2019)), and can attain particularly high levels in developing countries.[2] While it may not be surprising that national income is larger than the income concepts traditionally used to study inequality, it has also been growing faster. It has been argued that these discrepancies make it hard to assess how macroeconomic growth is distributed across income groups, and to what extent existing distributional statistics are a proper representation of the income flows in an economy. Recent work has embarked on a process of combining the various available data sources (surveys, administrative records, rich lists) upscaled -with further imputations- to SNA totals, aimed to produce comparable distributional results. These include, among others, WIL (2020), Fixler et al. (2017) and a project coordinated by the OECD (Zwijnenburg, 2019). While the existing gaps have sometimes strengthened the feelings of uncertainty about inequality measurement, these new approaches have taken for granted the numbers provided by the national accounts, a practice

[1] See, among others, Kuznets, Epstein, and Jenks (1941); Kuznets (1953).

[2] Discrepancies can also be observed in wealth and consumption data, but these are beyond the scope of this paper.

that does not always contribute to diminish those feelings, at least in the case of developing countries.

The discrepancies between different sources of income statistics have long been recognized in different parts of the world. None more so than in Latin America. As an important precedent, it is worth citing CONADE (1965), which set out to estimate the distribution of income in Argentina in great detail for the years 1953, 1959 and 1961, making use of surveys, population and industrial censuses, income tax registries, and social security records, and attempting a reconciliation with the national accounts. A few decades later, in a seminal study, Altimir (1987) critically analyzed available tax, social security and census data, as well as a variety of household surveys, systematically comparing the latter with the SNA, and concluding that there was a 15-30% gap with aggregate household income, which could be significantly higher for income sources such as property income. These results were explicitly assumed to be an indicator of the underestimation of each type of income in the surveys, and thus Altimir applied adjustments with notable implications for inequality analysis (e.g. an increase of the Gini index of 10-15%). Altimir's approach was adopted by the United Nations-Economic Commission for Latin America and the Caribbean (ECLAC), but the adjustment had many caveats, and was recently discontinued. This experience clearly illustrates the need (as well as the demand) for a reconciliation between micro and macro datasets – or at least the need to fully understand its potential consequences – and of the significant challenges of such an endeavor.

Data availability is arguably one of the main restrictions to properly study the distributional aggregates that feed the research on income distribution. In Latin America, most of this research has used survey data to analyze the evolution of inequality. The notable finding of this research is that the region experienced a historic decline in income inequality since the twenty-first century, attributed to a mix of vigorous economic growth and redistributive public policies (López-Calva and Lustig, 2010; Cornia, 2014b; Rodríguez-Castelán et al., 2016; Messina and Silva, 2017; Bértola and Williamson, 2017; Gasparini et al., 2018). However, question marks over the reliability of household surveys persist, as evidence on top incomes from tax records accumulates (Alvaredo, 2010; Alvaredo and Londoño Velez, 2014; Alvaredo et al., 2017; Burdín et al., 2022; Cano, 2015; Rossignolo et al., 2016; Morgan and Souza, 2019; Flores et al., 2020; Zuniga-Cordero, 2018). Recent research has also found that capital incomes appear to be remarkably less covered than labor incomes when survey aggregates are compared to SNA aggregates (Törmälehto, 2011; Bourguignon, 2015; Flores, 2021).

Underpinning the recent projects that seek to marry micro data sets and macro aggregates is the conviction that the statistical combination of data from several sources, based on researchers' own judgement and the resulting imputations, would allow for an acceptable correction and mitigation of the problems. Given the discrepancies at stake in a region like Latin America, this undertaking becomes a sensitive issue, with potentially large revisions to currently accepted inequality trends. The point of this paper is to dig into the aggregate differences between survey incomes and the incomes in the SNA, as well as those from administrative sources.

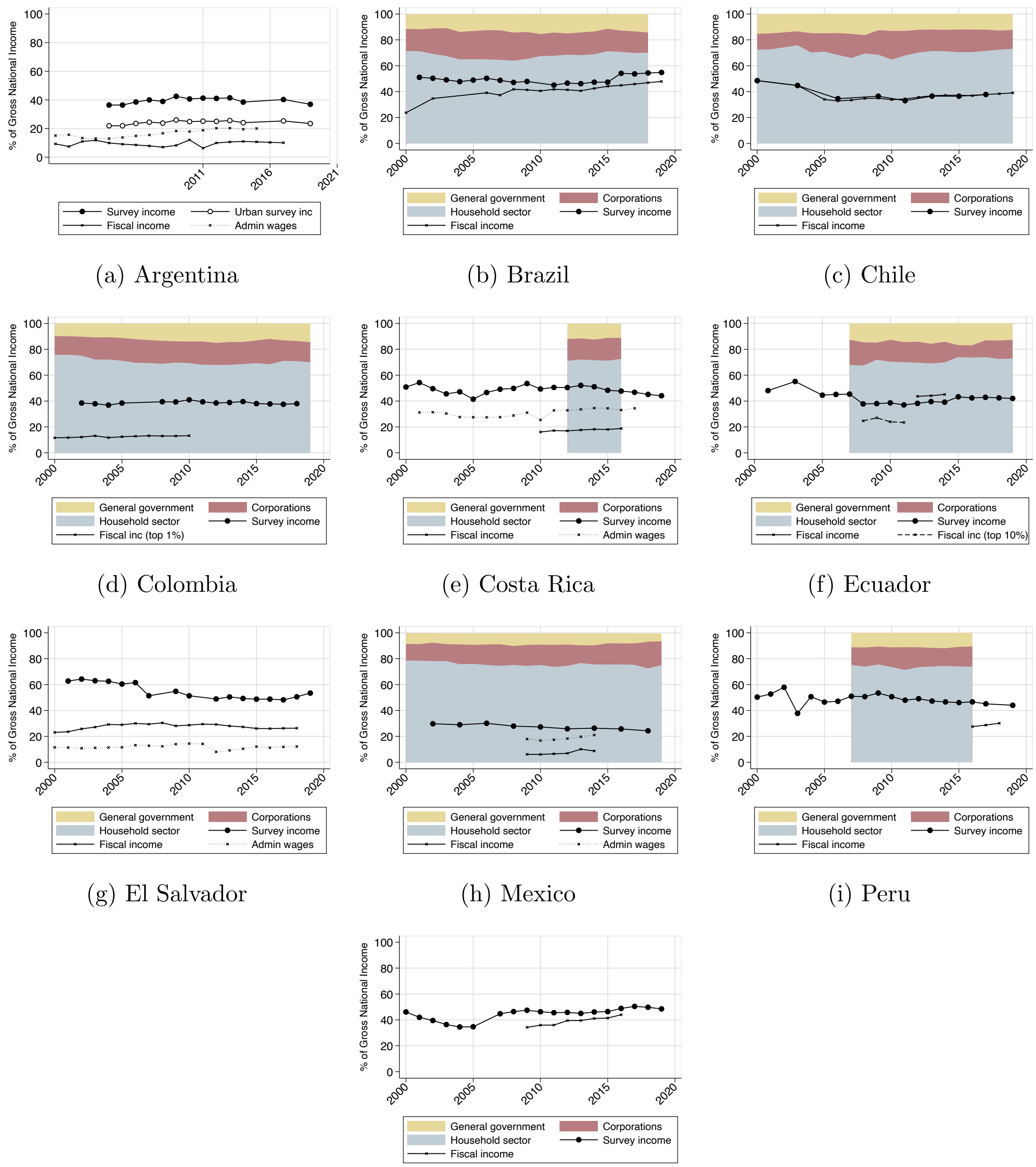
Designing the SNA meant accepting that the standard could not be set at the level of the best: it had to be feasible in less advanced systems. This is a central concern in this paper. The ultimate aim is to make the reader aware of the magnitude of the challenge of reconciling micro-level and macro-level statistics on income in Latin America, and of the consequences of such an enterprise for inequality statistics. For this we need to take a step back, and provide a renewed view of the scenario before the combination of datasets are put forward.

To achieve this we first map the available data sources on income in the region. These include the SNA, household surveys, income tax data and social security records. We then perform a detailed accounting of the discrepancies between macroeconomic and microeconomic aggregates in terms of income coverage, population coverage and distributional statistics for most countries in the region. Finally, we compare the shape and average incomes of top tails in surveys and tax data. This provides a starting point to establish the suitability of approaches that combine these different data sources in order to re-examine inequality trends in the continent, such as those aforementioned projects.

We document for ten countries – covering 80% of the region's population – that official inequality estimates coming from household surveys only account for around half of national income and 60% of household sector's income.[3] Important differences in this ratio exist between some countries, ranging from 50-60% in Brazil, to 25-30% in Mexico, which result in 65-70% and 35-40% of household income coverage respectively. Of particular significance is the fact that in most countries, the total survey income is a declining share of national income over the course of the last two decades. Figure A.3 provides a preview of these results. These are further commented in what follows, where we decompose the gap into two quantifiable components, which we call the "measurement gap" – the gap between measurable household income in surveys and equivalent household income in the SNA – and the "conceptual gap" – the portion of national income that is not directly received by households or measurable

[3] These statistics are "official", as opposed to "experiemental", in that they have been routinely published and cited by government departments, national statistics offices and supranational organizations for decades.

Fig. 1.1. Comparing total income in national accounts, surveys and administrative data



(a) Argentina (b) Brazil (c) Chile

(d) Colombia (e) Costa Rica (f) Ecuador

(g) El Salvador (h) Mexico (i) Peru

(j) Uruguay

Notes. Own elaboration based on UN national accounts data, ECLAC harmonized surveys, and countries' administrative records; for Colombia, the top 1% is taken from Alvaredo and Londoño Velez (2014)), and for Ecuador the top 10% is taken from Cano (2015) and fiscal income comes from Rossignolo et al. (2016). Survey income and fiscal income represent total pretax income in both sources, while admin. wages represents total pretax wage income in administrative wage data. Shaded areas are the balance of primary incomes of the household sector (B.5g, S.14), corporations (B.5g, S.11 + S.12) and the general government (B.5g, S.13).

in survey questionnaires. We find that the measurement gap accounts for roughly 52% of the survey-SNA gap on average, with substantial variation between countries, and a general increasing tendency in recent years. We estimate that overwhelming majority of this gap is due to missing capital income received by households, consistent with other recent literature cited above. Moreover, assuming that administrative data better account for incomes in the right tail of the distribution, we find an increasing undercoverage of top incomes in surveys, especially when non-wage incomes are considered. Additionally, the top tail of the tax data distribution not only depicts higher concentration levels than surveys, but also a higher degree of top income inequality, meaning that the income of individuals at the top in surveys are progressively less covered as one moves up the distribution. These findings have notable implications for analyses of inequality levels and trends within and between countries, which should be given greater attention by the literature.

The remainder of the paper is structured as follows. Section 2 outlines the data we assess and the conceptual framework of each data source, covering the main income variables and the links between each source. We also address the extent to which the SNA is a benchmark for economic indicators. Section 3 presents our findings on aggregate data discrepancies, mapping the evolution of total income across sources, the evolution of gaps by income component and the possible explanations for the observed gaps. Section 4 documents the top income deficit in surveys and the different shape of the top tails of the survey distribution and the tax data distribution. Section 5 concludes with a discussion of the implications of these discrepancies for inequality analysis, asking whether a reconciliation of these data sets is possible, and even desirable.

2. An inventory of data sets

We rely on four main sources to study aggregate income and its distribution: household surveys, income tax registers, social security records, and the national accounts. Yet there are still other sources that could and should be considered: population and economic censuses, banking information, firm-level data, etc; these are beyond the reach of this paper. Table A.1 presents the availability of the microeconomic data sources for the countries in our study. The following subsections elaborate on both the microeconomic and macroeconomic databases we use, highlighting their strengths and weaknesses, as well as assessing their conceptual compatibility.

2.1. *Micro-data: segments of a distribution*

Microdata refers here to datasets for which information on income is collected at the individual level. Unlike macroeconomic data, which comprises aggregate income by institutional sectors in the economy, this approach allows for direct distributional analysis. Microdata includes both household surveys and administrative records (from income tax declarations and wage data from social security contributions). Historically, surveys have been the most widely used source to study the income distribution and its covariates. They mainly rely on randomized sampling and post-stratification techniques to represent the whole population. It is generally accepted that surveys are a reliable representation of a wide segment of the income distribution, but are a less reliable indicator of the tails of the distribution. On the other hand, administrative records do not generally need to rely on sampling because they cover the universe of tax payers and formal wage earners.[4] However, by definition, administrative records mainly focus on the formal sector and are also subject to misreporting. For these reasons, tax data has been typically used to better study the dynamics of top incomes, often allowing to extend the time coverage of estimates far beyond what surveys enable.[5] In section 4, we confront the distributions described by both administrative sources and survey data where they overlap.

2.1.1. *Data from households surveys*

We use the survey micro-data harmonized by the Statistics Division of the UN Economic Commission for Latin America and the Caribbean (ECLAC) for ten countries over the years 2000 to 2019. These countries include Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Mexico, Peru, and Uruguay. ECLAC's harmonization process builds on the original surveys produced by the official statistics institutes of the countries listed in Table A.1. It seeks to create comparable annual income variables across countries, including the decomposition in terms of labor, capital and mixed incomes, pensions, owner-occupied rental income, transfers and other incomes.[6] In all cases but two, post-tax incomes are recorded on an individual basis, the exceptions being Brazil and Costa Rica, where gross (pre-tax) incomes are recorded.[7] Owner-occupied rental income and some capital incomes are

[4]Some public data sets based on administrative records do not provide information of the universe of reference, but of a representative sample, and, in this sense, they also require a sampling strategy. Examples include the Survey of Personal Incomes Public Use Tape, in the UK, and the Longitudinal Sample of Registered Employment (*Muestra Longitudinal de Empleo Registrado*) in Argentina.

[5]Tax data estimates can be brought back to the early years of the 20$^{\text{th}}$ century, when comprehensive income tax systems were created, whereas regular household surveys date commonly from the 1970s onwards.

[6]The only exceptions concerning the frequency of the surveys are Chile and Mexico, which collect data every two to three years.

[7]Gross incomes in Brazil and Costa Rica are before personal income tax and employee social contributions.

collected at the household level, and distributed among the adults (aged 20 years and over) of the household.

The household surveys provided by ECLAC thus represent one of the key data inputs for this study. More broadly, national surveys are an extremely important reference point in their own right in Latin America, since they are the only source publicly available in almost all the countries. Official statistics on inequality, poverty, unemployment, etc., are drawn from them. The countries that remain excluded from this study are mostly from Central America and the Caribbean. They either do not report distributional data at all (Belize, Cuba, Haiti, Jamaica, Suriname and Trinidad and Tobago), do not run household surveys on a regular basis (Bahamas, Nicaragua, Venezuela), or only run surveys but do not have any kind of publicly accessible administrative data (Bolivia, Dominican Republic, Honduras, Panama and Paraguay).[8]

2.1.2. Data from administrative records

Available distributional data from administrative sources in Latin America can be classified into four groups:

(i) microdata covering salaried employees with their wages declared at source by their employers, for social security records (e.g. Argentina, Costa Rica, Uruguay and Mexico);
(ii) microdata covering people with non-wage income sources (e.g. Costa Rica, Uruguay and Mexico);
(iii) grouped data (tabulations) based on the universe of tax payers, or those required to declare incomes, arranged by ranges of income (e.g. Argentina, Brazil, Chile, Peru and El Salvador);
(iv) grouped data (tabulations) based on the universe of formal wage earners, arranged by ranges of wages (e.g. Chile and El Salvador);

We exploit two new administrative sources for countries where, to our knowledge, tax data was never available for public-use purposes. One case is Peru, for which the tax authorities kindly prepared tabulated income statistics for this study. The data covers three years (2016-2018). It excludes entrepreneurial incomes, but includes pre-tax wages, dividends, rents, interests and other incomes. The other case is El Salvador, for which we gained access to two types of income tax tabulations, covering 2000-2018. One of the tables includes pre-tax wage income, while the other only includes individuals reporting income from diverse sources.

[8] For more details on these countries see appendix table A.2.

The rest of the countries in Table A.1 can be divided in two groups. On the one side, those regularly publishing and updating their administrative records (Argentina, Brazil, Chile, Mexico and Uruguay). On the other side, those that gave access to microdata to other researchers at some point, but do not produce distributive information from tax registers on a regular basis (Colombia, and Ecuador). For these cases, we use estimates prepared by the authors of previous studies (Alvaredo and Londoño Velez, 2014; Cano, 2015; Rossignolo et al., 2016), which are restricted to the top fractiles of the distribution. For Costa Rica we avail of grouped data from Zuniga-Cordero (2018), given the restricted access to administrative microdata on wage and independent income. Overall, as Table A.1 reveals, there is a wide range in the proportion of the population covered in the available tax statistics in each country, with less than 5% in Colombia, El Salvador and Mexico (the latter for diverse income), to over 70% in Chile and Uruguay.

2.2. Macro-data: a reference for aggregates

Macroeconomic data refer here to aggregates that follow the UN System of National Accounts (SNA). These are generally used to monitor domestic and national economic activity and are centered around the concept of Gross Domestic Product – or Value Added – which can be defined in three ways, giving rise to three sets of tables in the SNA: the production approach, the expenditure approach, and the income approach. We focus on the latter, which distinguishes flows between five institutional sectors – the foreign sector, financial corporations, non-financial corporations, the government sector and the household sector. Noteworthy items, for our purposes, are the income of salaried workers (recorded as "compensation of employees") and capital incomes (recorded as "property incomes").

The information from the SNA was obtained by scraping the UN Statistics Division database (`http://data.un.org`) and the websites of each country's national statistics office. Although the macro aggregates produced by national accountants are often considered among the most reliable and internationally comparable sources, detailed information on the income approach is scarce in the region. Even in countries that produce this kind of data regularly, statistics offices can update their estimates with two to five years of lag. The level of aggregation also varies across countries. For instance, despite the fact that United Nations (2008b) recommends distinguishing the Operating Surplus of Households (the income produced by owner-occupied housing and rented dwellings) from Mixed Income (the income of the self-employed), three countries – Chile, Ecuador and Bolivia – report both in the same aggregate. Furthermore, we observe large disparities in the level of detail provided for other relevant variables, such as property income and the consumption of fixed capital (capital depreciation). A lower level of detail in the decomposition of aggregates hinders our capacity to accurately match and compare income concepts across data sets and countries. We are thus forced into a trade-off between the precision of our estimates at the individual country level and their comparability at the regional level.

2.3. Matching micro and macro concepts

There are multiple ways to match incomes across data sets, with options ranging from the most aggregated definition of income to the most dis-aggregated. In this paper, our specific choice depends on a trade-off between the level of detail of the income components and the conceptual consistency of the definitions. Our ability to properly compare incomes depends directly on whether national statistics offices provide sufficient detail in their accounts to disentangle income components.

Table 3.1 displays the matching we perform for the empirical estimates presented in the next section. We match five types of income from our harmonized surveys to those in the SNA, in columns 1 and 2, respectively. Since concepts are generally wider in scope in the SNA, column 3 lists the associated income components in the survey, while column 4 lists non-matched or problematic items. In column 4, the items followed by an SNA code (e.g. D61 for social contributions) can be subtracted from the items in column 2 for a better matching (depending on the detail provided by national agencies), while those without an SNA code cannot be separated from the associated aggregates.

Table 1.1: Mapping households' income-concepts across data sets

Survey	**National Accounts**	**Comparable incomes**	**Less comparable incomes**
Salaried work	Compensation of employees (D1)	Wages, salaries (D11)	Social security contributions (D61)
Rental income	Operating surplus (B2)	Imputed rent of owner occupiers	Effective rent of residential buildings
Non-salaried work	Mixed income (B3)	Self-employed income	Effective rent of non-residential buildings
Investment income	Property income (D4)	Interests received (D41r) Dividends (D42)	Interests paid (D41u) Rent of natural resources (D45) Investment income of insurance policy holders (D441) Investment income of pension funds (D442) Investment income of investment funds (D443)
Other incomes	Social transfers (D62) Other transfers (D7)	Pensions Other cash benefits	Unemployment insurance Sick leave Private transfers (remittances)

Note: Based on United Nations (2008b) and OECD (2013). All incomes are gross of capital depreciation. SNA item codes are in brackets.

Different types of incomes have different degrees of conceptual overlapping. Labour income from salaried work, for instance, is among the least problematic. In general, one can easily subtract social security contributions from the compensation of employees in the SNA, so that only wages and salaries are compared with surveys reported net-of-contribution wages.[9] Social benefits are relatively straightforward too. Most countries do not distinguish them by type in their national accounts, so we achieve matching consistency by adding all the social transfers together in surveys (pensions and other cash benefits). Often, however, unemployment insurance may not be adequately captured in surveys.[10] Where it is reported, ECLAC's harmonized household surveys confound it with other incomes from employment, such as sick-leave and other wage-related incomes. However, unemployment insurance and sick-leave are included in social transfers in the SNA, and not in wages and salaries. This creates a minor conceptual inconsistency in the matching of aggregates from both sources.

A slightly more complicated case is the income from non-salaried work, which is included in the definition of mixed income in the SNA. The measurement of this aggregate is riddled with inconsistencies across countries, as well as being subject to the highest degree of misreporting out of all income items (ILO, 2019). A particular issue for us is that the SNA guidelines (United Nations, 2008b) also include effective rents from non-dwelling buildings owned by households, as self-employed units, in the mixed income aggregate. The ECLAC's harmonized surveys report all rental income collectively with other capital incomes so the item is indistinguishable by construction. However, this mismatch is likely to be very small in practice, compared to other comparability issues with measuring the income of the self-employed across countries (e.g. inclusion of some of the self-employed in the corporate sector, inclusion of a part of employer income in compensation of employees, under-reporting of income, especially among informal self-employed workers, etc.).[11] A further complication is that three countries in the region (Chile, Ecuador and Bolivia) report the household sector's operating surplus together with the mixed income aggregate, which limits the analysis to a relatively less precise and more aggregate level.

Something similar occurs when comparing imputed rents to owner-occupiers from surveys to the operating surplus of households from the SNA, which also includes households' actual rental income from leased dwellings. In most cases, we are unable to disentangle what rents are imputed or realized in the SNA. However, thanks to more detailed SNA data from the

[9] The only exceptions are: Argentina for the whole period, and Costa Rica before 2011. In the former, aggregate social security contributions are never reported, we thus compare survey wages directly to the compensation of employees. In the latter they are only available since 2012, we thus assume a constant ratio between contributions and to compensation of employees before that year.

[10] This is the case notably for Brazil before 2016. Unemployment benefits are thus imputed using information on periods of unemployment reported in the survey and statutory payment levels from the ministry of labour, again following Morgan and Souza (2019).

[11] At this stage we cannot properly verify the extent of all these comparability problems and their variation across countries. We thus leave this avenue open for future investigations.

expenditure approach in Brazil, we can estimate that imputed rents account for 93% of the aggregate on average between 2000 and 2015. In some countries, survey questionnaires do not even ask questions on imputed rents, and actual rental income is often reported together with other capital incomes. This creates a mismatch with how the SNA reports this item (in operating surplus rather than property income), yet Brazilian data suggests that the magnitude of the mismatch should be of second order. When rental income from owner-occupiers is not reported in surveys, ECLAC's harmonization process computes its value based on information from similar rented dwellings in the sample for each country. However, the absence of mortgage interest payments is problematic. In the SNA these are included in interests paid on the uses side of the accounts (D41u). A household with a mortgage equal to its imputed rent is not an owner-occupier but an "acquirer", which is functionally equivalent to a renter (Bourguignon, 2015).

The most complex conceptual match is that of investment income from surveys to property incomes from the SNA, which includes many items that are not considered in the survey at all (returns on investment and pension funds, and imputed investment income to insurance policyholders). These items are, in theory, well identified in the SNA, even if they correspond to imputed incomes. However, they usually correspond to a level of dis-aggregation absent in the accounts of most Latin American countries. For those countries where the detail exists at least for a few years (Brazil, Colombia, Chile, Costa Rica, Ecuador and Mexico) we compare capital incomes in a more conceptually consistent way. In appendix A.3, we show that the non-overlapping concepts are lower than 20% of the aggregate (10% on average), which also suggests a second order issue for those countries where we cannot disentangle these concepts properly. In the case of rent form natural resources, which are usually found among uses in the household sector of national accounts, surveys fail to report them all-together.

Another conceptual difference that affects all factor incomes (labour, capital and mixed incomes) is related to taxes. In the SNA, all of them are recorded as pre-tax, while in the survey the situation is less clear. Most incomes are generally assumed to be declared post-tax (except in Brazil and Costa Rica), especially in the case of incomes that pay the personal income tax at source, like formal wages. In order to solve this issue, in what follows, we use effective income tax rates from administrative data to add income taxes paid across the distribution in the survey (for more details see De Rosa, Flores, and Morgan (2020)).

3. Contrasting aggregates

In this section, we quantify the discrepancy between surveys and national accounts. Coverage is highly heterogeneous across income sources. That is, the gap is not the same for all incomes, for some it is almost negligible, while for some others can be substantial. Since different incomes are distributed differently, the impact on measured inequality levels and trends will depend on both the volume of missing incomes and their distribution.

3.1. Micro-data vs Macro-data

From the micro-data perspective, there are a number of reasons why surveys (and administrative data) may underestimate the total income of the household sector. We review the main causes of this phenomenon.

Household survey samples are, in principle, randomly selected from a target population, which is usually meant to be all resident household units. Despite big efforts to enforce randomness, many sources of biases coexist – heterogeneous response rates, non-random misreporting, small samples – and result in both un-representative samples and biased estimates. In order to address these issues, a long tradition of post-sampling adjustments was developed by data producers. The most common techniques use external data, such as population censuses, to re-weight observations in such a way that minimizes the distance between original and adjusted weights, while improving the representativeness of a series of characteristics. However, these are traditionally socio-demographic, such as age and gender, but not income.[12]

It is only recently – and mostly in developed countries – that survey designs started to address income representativeness to improve the coverage of the top tail. We highlight two main techniques. First, oversampling at the top, which basically consists in increasing the sampling size of the targeted group disproportionately, e.g.in places that are known to be wealthier and expected to have lower response rates. Second, the non-anonymous linkage of surveys and the administrative records of their respondents, especially regarding wages.[13] Unsurprisingly, the second approach has shown to be especially effective at enforcing the consistency of macro and micro estimates of income (Törmälehto, 2011; Flores, 2021). Yet

[12] Other relevant issues are top coding and censoring at the top of the distribution. However since the data we use in our empirical assessments are not subject to it, we ignore them in this work.

[13] Several countries participating in the European Union Statistics on Income and Living Conditions project (EU-SILC) have, over the last decade, progressively moved from the standard interview-based collection of information, to a mixed-strategy where incomes are directly obtained from fiscal registers for individuals in the sample. See Atkinson and Marlier (2010); Jantti, Törmälehto, and Marlier (2013), as well as the papers presented at the EUROSTAT Workshop on the Use of Registers in the Context of EU-SILC, Vienna, 2012. However, fiscal registers are not always used in the sampling design.

none of the Latin American countries currently employ either of these techniques. We thus ask whether the lack of mechanisms to counter these biases is likely to affect measures of income inequality not only in levels, but also in trends.

Figure A.3 provides a visual comparison of aggregates across three sources. It shows the decomposition of gross national income from the SNA into the household sector, the general government and the corporate sector. It also plots the total income reported in household surveys, and the total income reported in administrative data (both total fiscal income and wages in the formal sector when available), as a percentage of gross national income. Three countries, Argentina, Uruguay and El Salvador, do not report aggregates from the income approach in the SNA.[14] One result is clear nonetheless: the gap between raw surveys and national accounts is very large, with total survey income covering usually around 50% of national income. Mexico appears as an extreme case, where the gap reaches close to 80%. Comparing across micro-level sources, we find that administrative data cover similar levels of income (across a smaller population) than surveys (especially in Brazil, Chile, Ecuador and Uruguay). In some countries (such as Argentina, Costa Rica and Mexico), administrative wage data give higher values, as they cover the universe of formal employees, and not just a sub-sample required to file an income tax return. Of particular significance is the fact that in most countries, the total income in surveys is a declining share of income in the national accounts over the course of the last two decades.[15]

The crucial point is to know what part of this difference is relevant in the comparison of sources. We can thus go further and decompose the total survey income—national income gap for each country i in each year t as follows:

$$\textit{Total gap}_{it} = \textit{measurement gap}_{it} + \textit{conceptual gap}_{it}$$

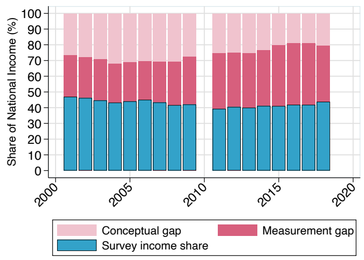
where the measurement gap is the part of the gap associated to matched income items (see table 3.1); and the conceptual gap is the part related to non-matched household incomes (mostly, but not exclusively, those flowing to other institutional sectors). In practice, the latter is calculated as a residual. Figure 1.2 presents this decomposition for years where both surveys and national accounts overlap and for the seven countries that have sufficient detail in their SNA. Blue bars correspond to the share of national income that is covered by surveys, which was previously depicted in Figure A.3.[16] Rose and pink bars divide the difference into

[14] Uruguay recently reported aggregates from the income approach, but only for 2012 and 2016.

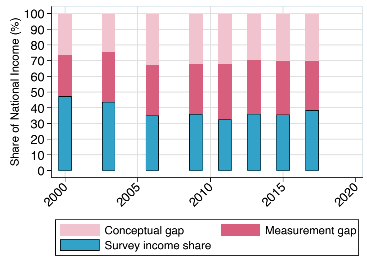
[15] Brazil since 2015 and Uruguay since 2006 are two exceptions, when both country surveys experienced substantial methodological changes in the sampling strategy, and thus the coverage of incomes.

[16] This total is the sum of the incomes in figure A.2.

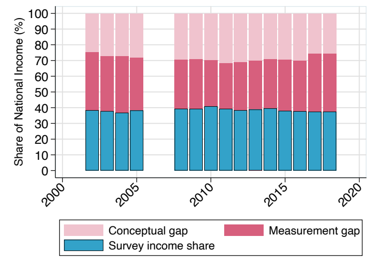
Fig. 1.2. Decomposing the Survey Income—National Income gap



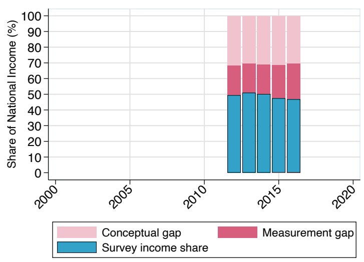
(a) Brazil



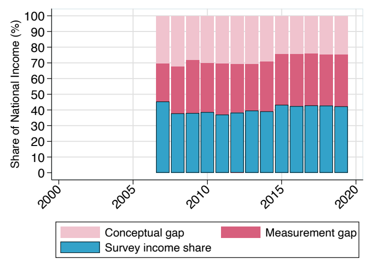
(b) Chile



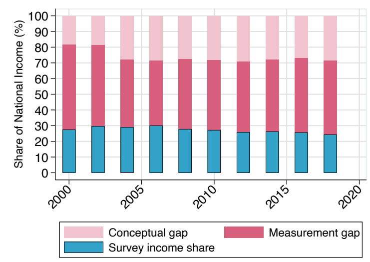
(c) Colombia



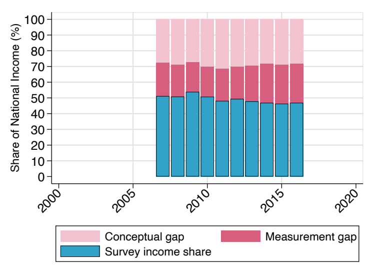
(d) Costa Rica



(e) Ecuador



(f) Mexico



(g) Peru

Notes. Own elaboration based on ECLAC harmonized surveys and the UN national accounts data for countries with sufficient breakdown in the SNA income approach to perform the calculation. Survey income is total pretax income. Measurement gap refers to the part of the gap explained by the under-coverage of household sector incomes for the matched income concepts. Conceptual gap refers to the part of national income not received/reported by households directly.

the measurement gap and the conceptual gap, respectively. With Mexico having the largest gap to cover, it is not all surprising that its measurement gap is the largest out of all the countries. This gap appears to be smallest in Costa Rica and Peru, where it accounts for less than half of the discrepancy. In the remaining countries it is roughly half of the gap. The measurement gap is thus a significant part of the discrepancy between surveys and national accounts, which raises question marks over the survey's capacity to accurately represent the distribution of income. In order to ascertain this we need to break down the discrepancy by income component and estimate the incidence of each component in the distribution of total income. This is what we turn to in the following sections.

3.2. Heterogeneous coverage of income items

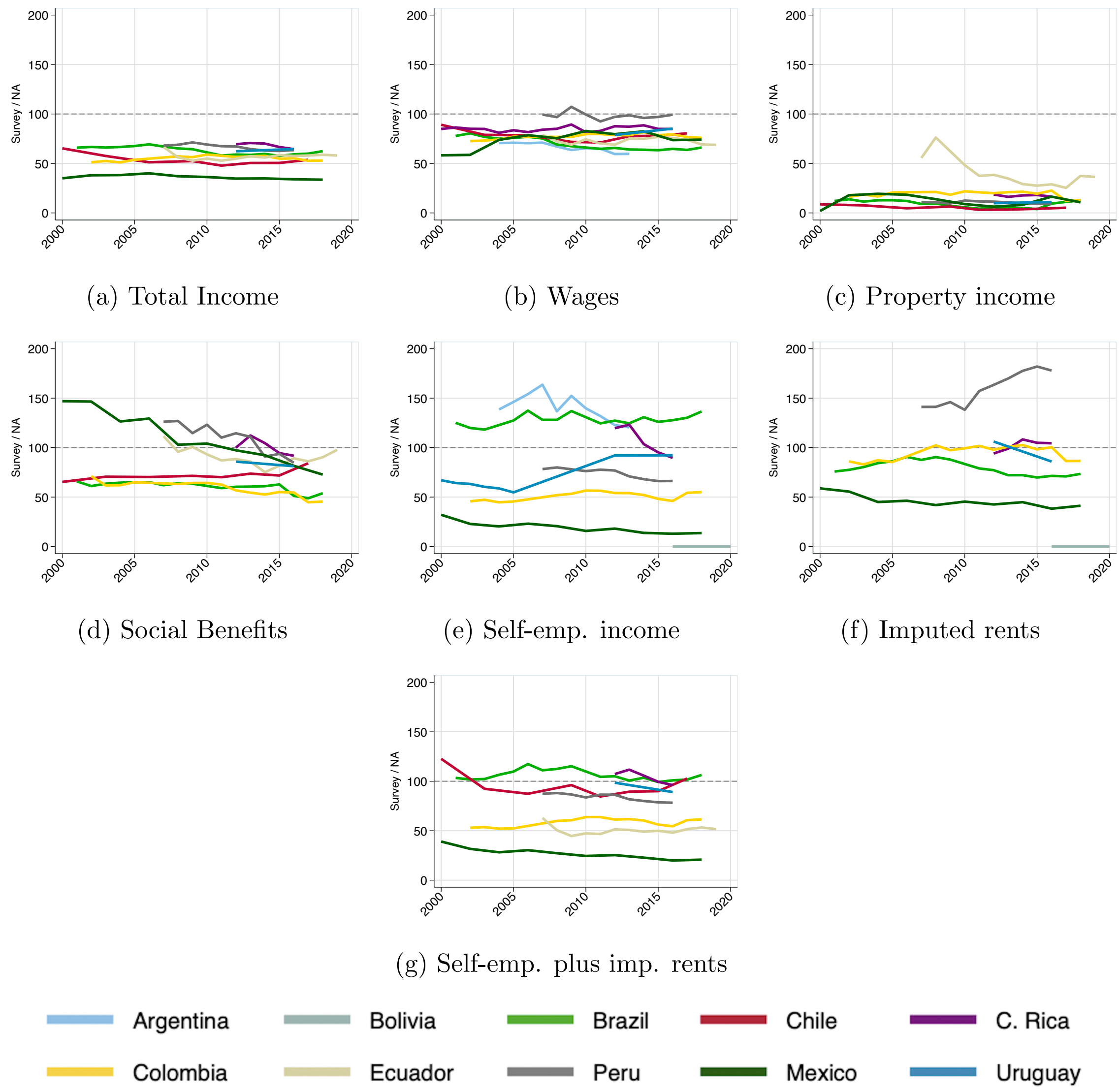
A simple, yet insightful exercise, is to compare the total amount reported, by type of income, in both household surveys and macro data. When definitions are comparable across datasets, the observed measurement gap can be interpreted as an underestimation of income, assuming that national accounts are considered an accurate representation -as discussed, this is a very debatable assumption that we use for the sake of the presentation of numbers.

Figure 1.3 displays coverage of income components in surveys with respect to corresponding items from the national accounts, based on the matching of concepts presented in Table 3.1. As previously commented, the sum of all matching incomes is clearly underestimated in all cases, with coverage rates ranging between less than half, for the case of Mexico, to close to 80% in the case of Brazil (subfigure 1.3(a)). From the rest of the subfigures, one can see the underestimation is not uniform across income items, and that some of them seem to contribute much more than others to the overall underestimation.

The coverage rates of wages and property incomes are polar opposites (subfigures 1.3(b) and 1.3(c)). The former are relatively high, with most countries bunching close to complete coverage, whereas property incomes are severely underestimated in all cases, with the exception of Ecuador, for early years. In the majority of cases, coverage is less than 10%. Other incomes, such as benefits, self-employment income and imputed rents – subfigures 1.3(d), 1.3(e) and 1.3(f) – display relatively more heterogeneous coverage across countries. This includes ratios above one, which suggest that surveys may overestimate certain income components. This, in and of itself, is not wholly unexpected. If certain types of individuals/households with certain types of income are not covered by the survey (due to sparse samples not capturing rare populations and non-response, for example), then the income of certain other individuals/households in the covered sample may well be over-represented. This could affect the populations reporting social benefits or self-employed income. Moreover, the survey reports incomes of a specific reference period, usually a month, or a week that is aggregated to the month of reference, which may not carry over to the entire year. Thus, when annualizing incomes – that is, multiplying declared monthly incomes by twelve – we may be attributing too much income to a certain class of activity whose realized annual income is much more volatile than an assumed persistent monthly earning (e.g. self-employed income).[17] If all types of income were to be then adjusted proportionally, the overall impact would depend on their magnitude and distribution. We turn to the former in the following section.

[17] The 'over-estimation' of imputed rents for some countries is more likely to be due to the methods employed by ECLAC (see section 2.3). We recommend that future revisions of this estimate be calibrated to the national accounts estimate of imputed rent where possible.

Fig. 1.3. Discrepancies by income component in surveys with respect to NA



(a) Total Income

(b) Wages

(c) Property income

(d) Social Benefits

(e) Self-emp. income

(f) Imputed rents

(g) Self-emp. plus imp. rents

Notes. Wage incomes are relatively well represented in surveys, while capital incomes are heavily underestimated. The coverage of other types of income is more heterogeneous, with both under- and over-estimation, depending on the case. Conceptual matching follows the benchmark in table 3.1. For a further decomposition of capital incomes, see appendix A.3. Chile and Ecuador report the corresponding aggregates of self-employment income and imputed rents together in the same item, they are not included here. Own elaboration based on ECLAC harmonized surveys and UN National Accounts.

3.3. The size of 'missing' survey income

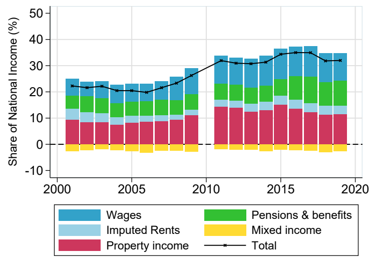
Independently on how to achieve the consistency of micro and macro data, the study of income gaps, and their composition, provides insights into how poor a guide official inequality estimates, and their trends, can potentially be.

Figure 1.4 summarizes these aggregates for the countries that have sufficiently detailed data. It displays the amount of each income item that survey's fail to capture. As we can see, the overall magnitudes are significant, ranging from 10% to 40% of national income across countries. Not surprisingly, the country with the highest discrepancies is Mexico, for which the magnitude of missing household sector incomes amounts to 40% of national income. Between 10% and 20% are due to the underestimation of property incomes. A similar amount is due to the underestimation of self-employment income, while the remaining magnitudes seem to be less significant. In all countries, although wages tend to be the least underestimated item (see figure 1.3), they show a relatively stable, if not increasing, tendency over time. In the cases of Brazil, Chile and Peru, the amount of missing survey income increases considerably during the period. In these cases, especially for Brazil and Chile, property incomes seem to play a major role in this evolution. On the contrary, in Colombia and Ecuador, it is the underestimation of mixed incomes and wages that seems to be driving the overall underestimation of incomes.

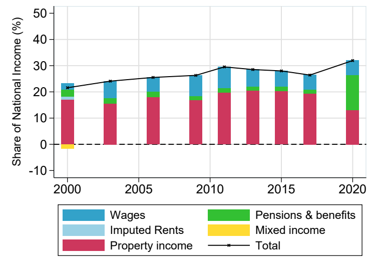
For almost all countries we also observe a minor portion of incomes that are over-represented in surveys. These are incomes whose share of national income are negative, and thus make the total missing income lower than what it would otherwise be. The overall pattern seems to suggest that, consistent with figure 1.3, self-employed income and imputed rents are the most susceptible to being over-represented in surveys, followed by pensions (see the previous section). While these household incomes are much less than the incomes under-represented, they should nevertheless be accounted for in any procedure seeking to make surveys macro-consistent.

To summarize, data gaps affect different income-types differently. Moreover, property incomes, which are always more concentrated at the top than other types of income (see appendix A.4), explain a large part of these gaps. We thus ask to what extent are the gaps presented thus far the result of an underrepresentation of the top tail in household surveys? This is what we turn to in the next section.

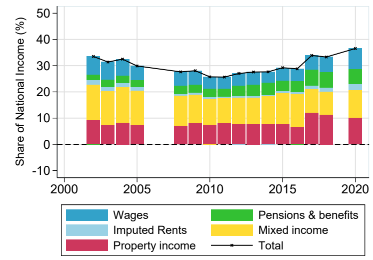
Fig. 1.4. What's missing (or spare) in surveys



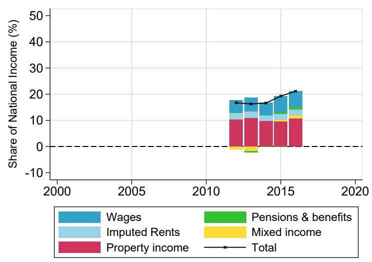
(a) Brazil



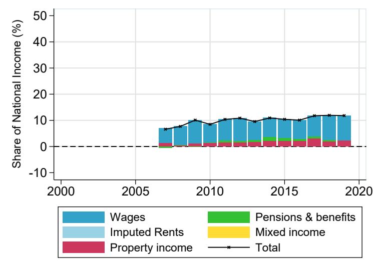
(b) Chile



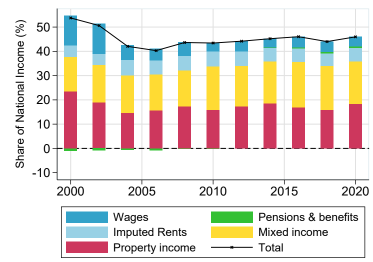
(c) Colombia



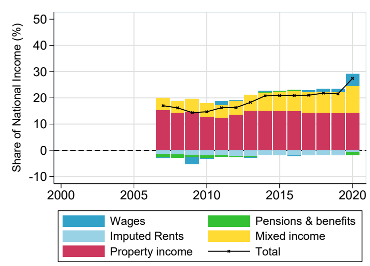
(d) Costa Rica



(e) Ecuador



(f) Mexico



(g) Peru

Note. Own elaboration based on ECLAC harmonized surveys and UN-Data national accounts. GOS stands for gross operating surplus. In the national accounts of Chile and Ecuador mixed income and gross operating surplus are combined in a single aggregate.

4. Distributional implications

In the previous section, the micro-macro gap that emerged from contrasting income aggregates was analyzed both by decomposing it into a measurement gap and a conceptual gap, as well as in terms of income components and their magnitudes. This section studies and compares the distributions described by different micro-data sources where they overlap, which is in the top tail.

4.1. Top income levels in surveys vs tax data

As has been profusely documented in the literature (see Section 1), income differences are particularly important in the top tail of the distribution when survey and administrative data are compared; hence the importance of going beyond the analysis of aggregates. Indeed, one of the key reasons given in the recent applied literature for combining survey and tax data is to properly account for incomes in the top tail of the distribution, which are assumed to be better captured by the latter. Thus, it is important to compare them in order to assess the likely effect of using administrative tax data to provide more accurate inequality estimates. It should be reminded, though, that the absence of the rich from surveys per se does not necessarily imply that measured inequality levels are biased downwards (see, for instance, Deaton, 2005).

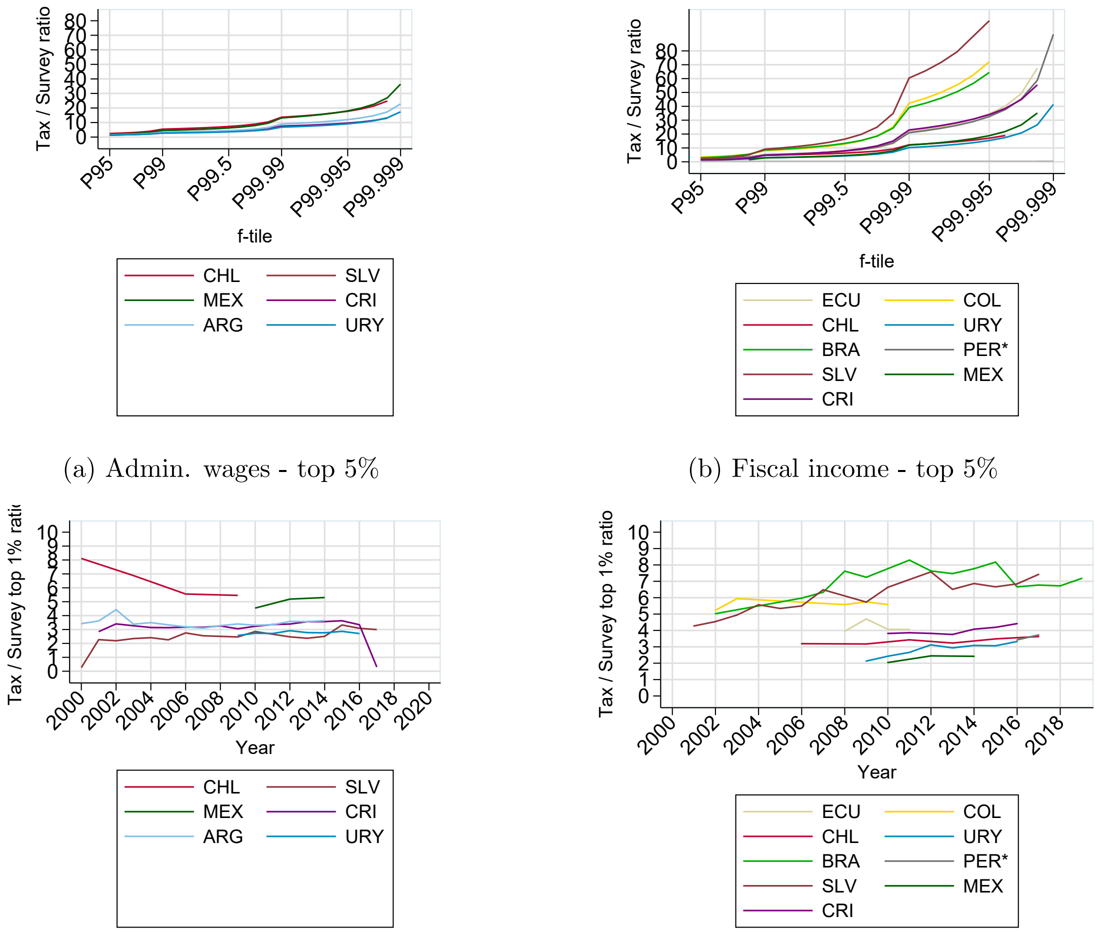
One straightforward way to proceed would be to systematically compare top income shares across data sources, but this requires adjustments to the tax data to account for incomes and population not captured in this sort of administrative data. At least, it would involve considering population and income controls (Atkinson, Piketty, and Saez, 2011), which are usually taken from censuses and national accounts and/or survey data, respectively.

Considering that the objective of this paper is to compare data sources without combining them in any way, and since estimating top shares involves precisely that kind of procedure, one possibility is simply to compare income levels in tax data and surveys for top income groups. This is depicted in Figure 1.5, distinguishing between countries for which tax data only accounts for wages (panel a), and the ones for which all incomes are considered (panel b; see Table A.1).[18] Overall, incomes in tax records are higher for top fractiles. For all countries, pre-tax wages in the tax records surpass household pre-tax wages at some point within top 5% and are considerably larger within the top 1%. Ratios are substantially higher in panel b, which point to the fact that capital incomes are considerably less covered than wages in surveys, in line with previous research (e.g. Burdín et al. (2022); Morgan and Souza (2019).

Perhaps more importantly, tax and survey pre-tax incomes diverge for top fractiles over time. In panels c and d of Figure 1.5 the top 1% ratios are depicted, showing an increasing gap between both data sources, which is clearer when total fiscal incomes are considered. This may suggest that non-wage incomes (especially capital incomes) play an important role in this divergence. If in fact administrative data better capture incomes at the top, this pattern

[18]For this analysis, tax data was interpolated based on a Generalized Pareto distribution to account for all income fractiles in the top 5%. These fractiles are defined relative to an external population control, which is taken from official population projections from country statistics offices.

Fig. 1.5. Tax-survey pre-tax income ratio



(a) Admin. wages - top 5%

(b) Fiscal income - top 5%

(c) Admin. wages - top 1%, 2000-2019

(d) Fiscal income - top 1%, 2000-2019

Note. Own elaboration based on last available tax data point and ECLAC harmonized survey data. In panels a and b, the ratio of average incomes in tax/survey for each fractile is depicted among the top 5%. Ratios for the year 2010, or closest neighboring years are depicted. Panels c and d depict the pre-tax incomes ratio for the top 1%. In panel b, Brazil and El Salvador were caped at P99.995 as they increased exponentially beyond that point. (*) In the case of Peru, tax data on total income excludes entrepreneurial income.

indicates that the poor performance of surveys in top fractiles is indeed worsening, which has significant implications for assessing inequality trends.

Nevertheless, comparisons of this type, even if they could be somewhat informative, are flawed when surveys are considered to be affected not only by under-reporting at the top but also by under-sampling of richer households. See Bourguignon (2018) for a discussion on simple adjustments of observed distributions for missing income and missing people, and Blanchet et al. (2022) for a proposed adjustment using external administrative data.

4.2. The shape of the top tail in surveys vs tax data

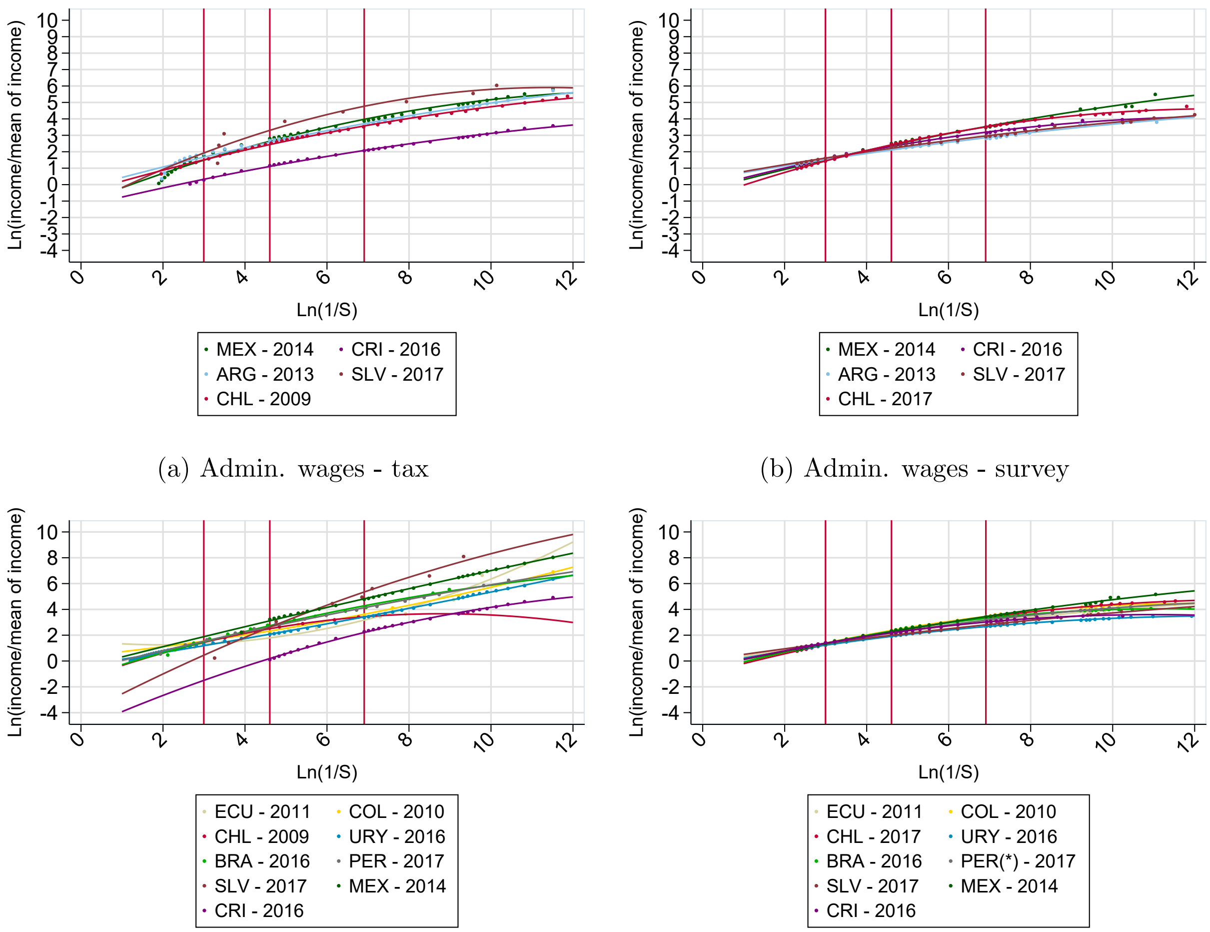
To supplement the preceding analysis, we proceed by comparing the shape of the top tails directly (Atkinson, 2017; Cowell, 2011). In Figure 1.6 we plot – both in tax and survey data – a function of income against a transformation of the survival function S, defined as $1 - F$, i.e. the complement of the cumulative distribution. The y axis depicts the log of income as proportion of mean income, while the x axis depicts the log of $^1/s$.

By construction, if the fitted functions are linear, the steeper the line is the more concentrated is the income (note that the slope of the curves is equivalent to the inverse of the Pareto coefficient). More importantly, when the data points are in a straight line, the distribution is a Pareto-I, while a concave one is the result of what Atkinson (2017) calls a *baronial* top tail, and a convex one is *regal*. The latter two shapes are departures from the familiar Pareto distribution, the first representing a distribution in which top positions tend to be more homogeneous, while the opposite holds in the *regal* shape, i.e. individuals in top positions tend to be further apart from each other. The main advantage of this approach is that it provides a way to directly compare top tails without the need to assume any sort of income control. This allows us to visually inspect not only concentration, but also the tail's shape. When only tabulations are available, tax bracket thresholds are plotted (without interpolation), while in the cases with micro data (both for surveys and for a some countries' tax data) selected data points are depicted. In all cases, a second degree polynomial function was adjusted to more clearly visualize the tails.

Three key features stand out. First, as expected, there is more concentration in tax data than in surveys, which are substantially closer to the x axis. Second, there is more heterogeneity in tax data, which may be the result of different shapes of the top tails across countries, or of the different quality and structure of the data. Third, while in survey data the top tails are in all cases *baronial*, this is not the case in the tax data, in which several countries present a clear *regal* shape, given by the convexity of the survival functions.

To see it more clearly, the quadratic coefficients are presented in Figure 1.7. Despite being quite noisy, several facts emerge. The tax data of Ecuador, Uruguay, Mexico and Colombia present *regal* top tails, while the remaining countries have *baronial*-shaped tails. In contrast, barring a few exceptions, survey data presents *baronial* top tails, with incomes more similar to each other at the top. Overall, these sets of results illustrate incomes are substantially higher at the top of the distribution in administrative data than in surveys, driven especially by non-wage income; and that income dispersion within the top tail is generally greater in administrative data.

Fig. 1.6. Pre-tax income in relation to rank



(a) Admin. wages - tax

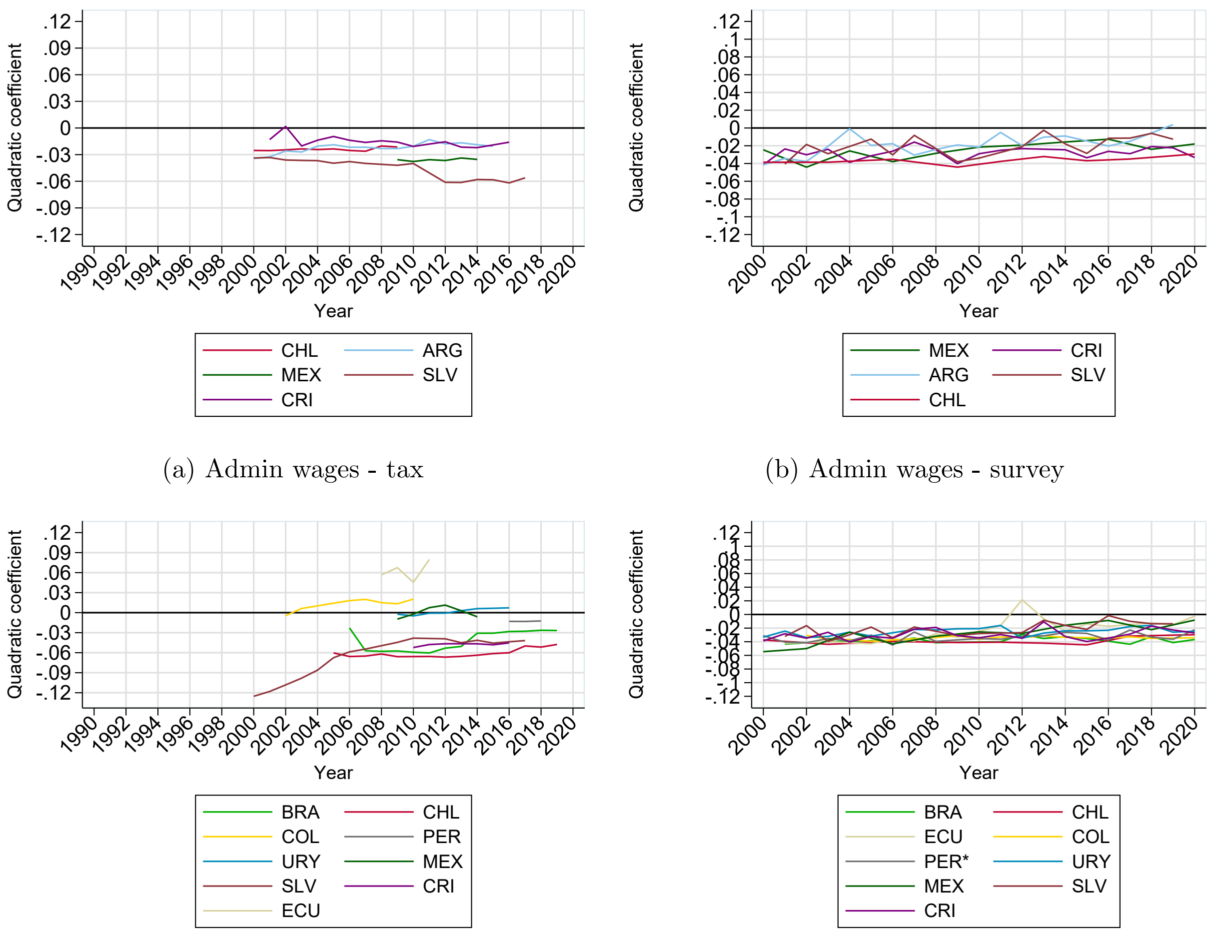
(b) Admin. wages - survey

(c) Fiscal income - tax

(d) Fiscal income - survey

Note. Own elaboration based on country tax data and ECLAC harmonized survey data. For survey data, incomes are equal-split (broad) pre-tax income. Last available year for each country. The y axis depicts the log of income as proportion of mean income, while the x axis depicts the log of $1/S$, with S being the survival function. Vertical lines represent the thresholds for top 5 %, top 1% and top 0.1% incomes, respectively. (*) In the case of Peru, business incomes were not considered as they are not accounted for in tax data.

Fig. 1.7. Fitted survival function's quadratic coefficients



(a) Admin wages - tax

(b) Admin wages - survey

(c) Fiscal income - tax

(d) Fiscal income - survey

Note. Own elaboration based on country tax and ECLAC's harmonized survey data pre-tax incomes. β coefficient of the regression $\log(income/meanincome) = \beta \log(1/S)^2 + \alpha \log(1/S) + \epsilon$. (*) In the case of Peru, business incomes were not considered as they are not accounted for in tax data.

5. Final remarks: implications for inequality research

What lessons can we draw from the preceding analysis for applied inequality research? We found a large distance between the aggregates used in inequality studies and those from the SNA in Latin American countries. We also noted a growing undercoverage of household incomes in surveys vis-à-vis the SNA, especially property incomes. We documented that this overall undercoverage of magnitudes is due to a mixture of measurement and conceptual gaps. The former relates to item and unit misreporting, sparseness of the survey sample, especially in capturing 'rare' populations such as the rich, as well as heterogeneous non-response rates of individuals/households. The latter concerns unaligned income definitions between household surveys and the SNA, i.e., incomes in the SNA that by definition are not covered in surveys. For Latin American countries with detailed enough data we showed that at least half the survey–SNA income gap is due to measurement, and a significant portion of this gap is indeed due to missing capital incomes of households, but not only. Underrepresented items also affect wages, self-employed income, imputed rents, and pensions, to differing degrees across countries. The sheer magnitude of these gaps should not leave anyone indifferent.

To what extent then is income inequality being underestimated in household surveys? To approximate an answer to this question we require to know both the volume and the incidence of misreported income items. Although there is more uncertainty in the latter, we know that an important characteristic of survey measurement error is its heterogeneity across the distribution. Discrepancies thus have distributional implications by construction. We showed, for the Latin American case, how these different income items are distributed in the surveys of the different countries in comparison to the distribution of total income. The general pattern that emerged is that those income items showing the largest magnitude of discrepancy were those whose distribution were the most unequal, namely property incomes. This fact seems to be behind the increasing gap between top incomes in surveys and administrative data that we also document. Other income items showed less concentration, but were relatively less underestimated vis-à-vis the SNA. Some items for some countries seem to be overrepresented in surveys relative to the SNA, such as imputed rent, which is generally the least unequally distributed item across our set of countries. Based on these discrepancies, it can thus be anticipated that income inequality is being underestimated in household surveys in Latin America, even before making any survey adjustments. The question then becomes whether microeconomic data sources are *increasingly* underestimating inequality levels with time so that trends too become inaccurate.

It must be acknowledged that this positioning of surveys in relation to the SNA assumes that the latter are an accurate benchmark for income flows in an economy, including those received by households. From a conceptual viewpoint, and on the basis of international standards, the accuracy of the SNA should be relatively high. However, we are aware that in the case of Latin American countries particularly, but not exclusively, the construction of the national accounts must be a cause of concern. Indeed, the SNA remains somewhat of a black box, in the sense that, faced with a lack of information, accountants need to make judgements and assumptions for the calculation of sectoral incomes (i.e. wages, mixed income, interests, dividends, operating surplus...), even where the 'income approach' of the SNA is available, which is not always the case for the countries we study. Some countries have responded to this lack of transparency by publishing detailed methodological documents on the construction of the SNA across sectors, notably in Brazil (IBGE, 2016). Nevertheless, our findings suggest that macroeconomic income levels, their resulting growth rates, and the capital shares in the developing Latin American countries require a thorough re-examination.

Despite these shortcomings, for roughly thirty years, the UN-ECLAC re-scaled reported incomes in surveys to comparable incomes registered in the SNA, following in the footsteps of Altimir (1975, 1987). This adjustment was done proportionally with the exception of capital incomes, which where imputed only to the top quintile. Yet, it was precisely because of the diverging trends between surveys and the SNA over time and the increasingly costly procedure to reconcile both sources that this adjustment was abandoned by ECLAC. Re-evaluating such an adjustment in light of our findings presents additional challenges. On the one hand, a proportional adjustment is certainly far from perfect, since the true distribution of income components could be very different from the one described by surveys. A better alternative, for example, would require high-quality administrative data, with income decomposed by item, whose availability to researchers is quite rare. From our comparison of surveys and administrative data available in the region, we showed that surveys largely fail to reproduce the shape of the right tail of the distribution in tax data, in some cases more significantly than others, without resorting to incomes in the SNA. The comparison of wages and total incomes in the right tail of tax and survey data suggests that part of the decomposition of income at the top of the distribution is also mis-measured in surveys, in line with previous research. To a certain degree, administrative data of the sort available for Latin American countries could be used to implement a finer survey adjustment, especially at the very top of the distribution.

The application of adjustment methods were not within the scope of this paper, which instead concentrated on the necessary step of critically comparing incomes across different data sources to attest whether a reconciliation of such sources is indeed possible. On the basis of our findings we judge that such a reconciliation can be made (this is obvious from the theoretical point of view) but substantial investment is still needed. Indeed a specific contribution of this paper is to also draw attention to the greater needed investment in the level of detail and transparency of the national accounts by national producers. Additionally, even when an acceptable reconciliation is reached, the discussion about the definition of income that should be considered for welfare evaluation purposes remains. To cite just one example, as Atkinson (2015) reminds us, adopting a comprehensive definition of income requires that the change in asset values be taken into account; although this element is not part of the SNA's definition of national income, it becomes particularly relevant in times of inflation.

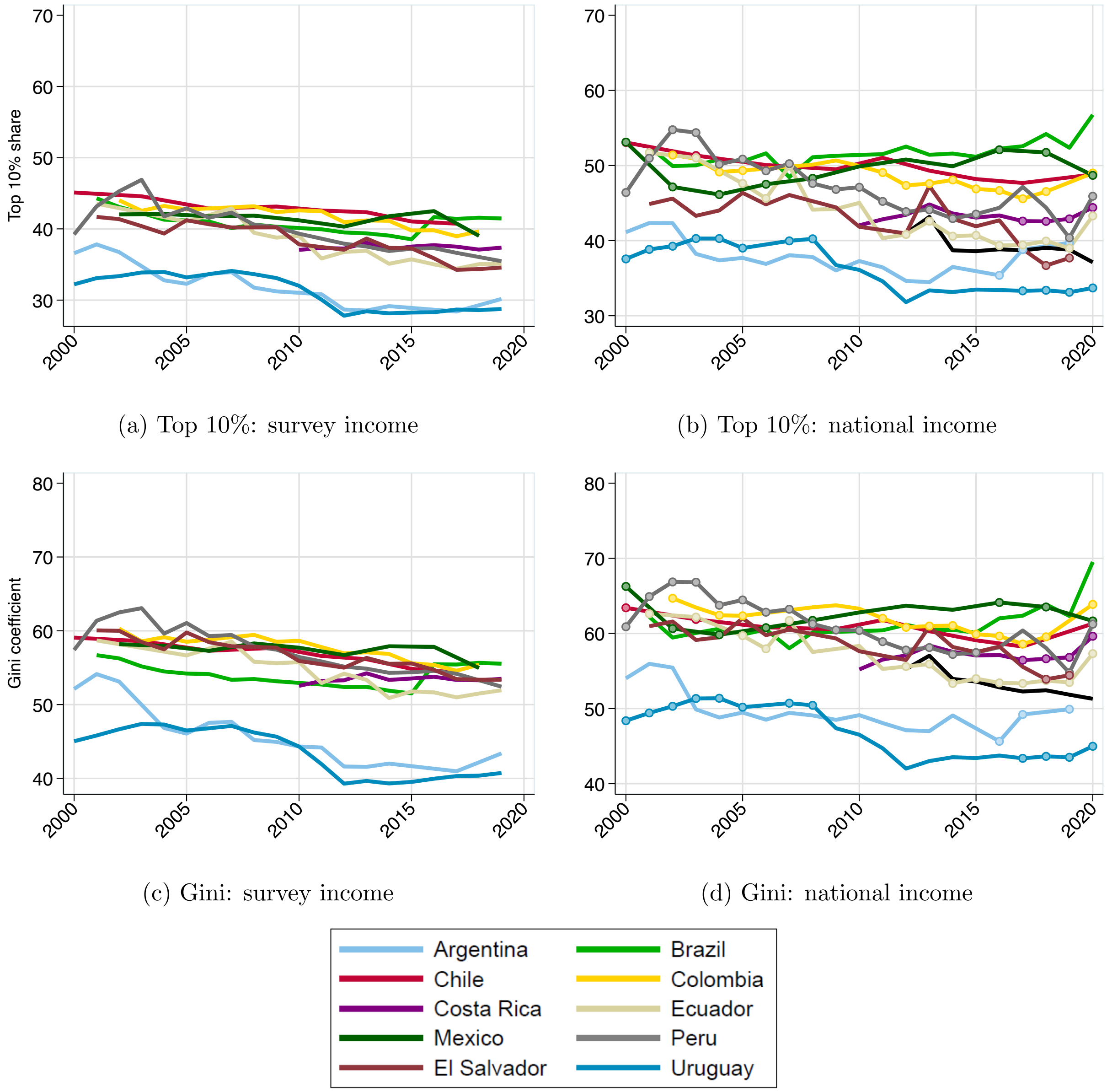
Attempting to fill the a sizeable income gap between official household surveys and official macro statistics may seem like too much of a stretch in the current state of affairs. Yet, it should be pointed out that no data source is perfect, and that in any empirical distributional analysis assumptions must be made, whether they are implicit (e.g. taking surveys to represent the national distribution of household income) or explicit (e.g. imputing the micro-macro gaps according certain allocation rules).

In the meantime, an alternative to this view is the vector approach of inequality statistics to study the evidence provided by different and competing data sources. In this approach a series of internally consistent indicators (synthetic indexes, per capita macro incomes, factor and fractile shares, etc.) are compared to one another to determine the plausible direction of inequality over a certain period of time. This approach places less emphasis on inequality levels and more emphasis on trends, which in itself is cost efficient, as less work is needed to reconcile difference data sources. On the other hand, this approach may suffer from contradictory evidence among the difference series, as has been found in much of the top incomes literature (see the Introduction), leading to inconclusive evidence on inequality trends. Thus, the reconciliation of different income sources would seem to be an inescapable task, if only to understand the diverging trends across these multiple data series. But reconciliation is also merited from the view that the official macroeconomic accounts of a country are themselves the product of a reconciliation of different sources, among them some of the routinely used microdata sources in inequality studies. Something is therefore not adding up, and it is high time that researchers and national data producers try to figure out why.

A. Appendix

A.1. Supplementary figures and tables

Fig. A.1. Top 10% share and Gini coefficient of survey income and national income



(a) Top 10%: survey income

(b) Top 10%: national income

(c) Gini: survey income

(d) Gini: national income

Notes: Income corresponds to pre-tax income of individuals, with the income of couples being split equally. Panels (a) and (c) show the distribution of income as reported in household surveys, while panels (b) and (d) refer to the distribution of national income from the SNA using a combination of surveys, administrative data and aggregate macroeconomic income accounts. Source: De Rosa et al. (2022).

Table A.1: Countries and data sets available for comparison

	Survey microdata		Administrative data			
Country	Source	Availability	Source	Availability	Population (% of total)	Definitions
Argentina	Encuesta Permanente de Hogares (EPH) and EPH-Continua from 2003, Insituto Nacional de Estadística y Censos (INDEC)	2000-2014, 2016-2019	Income tax tabulations, Administración Federal de Ingresos Públicos (AFIP), Employee microdata, Ministerio de Trabajo, Empleo y Seguridad Social	2000-2017, 2000-2015	2% 40%	Survey is representative of urban areas (28-31 cities). Income tax data is total pre-tax fiscal income. Employee microdata includes only private sector wages.
Brazil	Pesquisa Nacional por Amostra de Domicílios (PNAD), Instituto Brasileiro de Geografia e Estatística (IBGE)	2001-2009, 2011-2019	Income tax tabulations, Receita Federal (RFB)	2000, 2002, 2006, 2007-2019	14%	Income tax data is total pre-tax fiscal income.
Chile	Encuesta de Caracterización Socioeconómica Nacional (CASEN), Ministerio de Desarrollo Social	2000-2009 (triannual), 2011-2017 (biannual)	Income tax tabulations, Servicio de Impuestos Internos (SII)	2000-2018	70%	Wages reported separately from other fiscal incomes in 2000-2004.
Colombia	Encuesta continua de hogares (Gran Encuesta Integrada de Hogares from 2008), Departamento Administrativo Nacional de Estadística (DANE)	2002-2005, 2008-2018	Alvaredo and Londoño-Vélez (2013)	2000-2010	1%	Income tax data is total pre-tax fiscal income.
Costa Rica	Encuesta Nacional de Hogares, Instituto Nacional de Estadística y Censos (INEC)	2000-2019	Wage income, Non-wage income Zuñiga-Cordero (2018)	2000-2017 2010-2016	28% 5%	Wage earners from social security records, Independent workers from income tax declarations.
Ecuador	Encuesta Periódica de Empleo y Desempleo (EPED) and Encuesta de Empleo, Desempleo y Subempleo (ENEMDU) from 2003, Insituto Nacional de Estadística y Censo (INEC)	2001, 2003 2005-2019	Cano (2015) Rossignolo et al. (2016)	2008-2011 2012-2014	14% 38%	Distributional data on total fiscal incomes is only available from Cano (2015) for the 10%.
El Salvador	Encuesta de Hogares de Propósitos Múltiples, Dirección General de Estadística y Censos (DIGESTYC)	2000-2007, 2009, 2010, 2012-2019	Tax tabulations (wages), Tax tabulations (diverse income) Dirección General de Impuestos Internos (DGII)	2000-2018	4% (wages) 4% (diverse)	Wages of salaried workers are reported separately from income from diverse sources.
Mexico	Encuesta Nacional de Ingresos y Gastos de los Hogares, Instituto Nacional de Estadística, Geografía e Informática (INEGI)	2002-2018 (biannual)	Income tax microdata, Servicio de Administración Tribuataria (SAT)	2009-2014	20% (wages) 2% (diverse)	Wages of salaried workers are reported separately from income from diverse sources.
Peru	Encuesta Nacional de Hogares - Condiciones de Vida y Pobreza, Instituto Nacional de Estadística e Informática (INEI)	2000-2019	Income tax tabulations, Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT)	2016-2018	25%	Income tax data excludes entrepreneurial incomes.
Uruguay	Encuesta Continua de hogares (ECH), Instituto Nacional de Estadística (INE)	2000-2005, 2007-2019	Income tax microdata, Dirección General Impositiva	2009-2016	75%	Income tax data is total pre-tax fiscal income.

Note: Own elaboration. Population in administrative data is shown as a share of the total population of the country for the latest years.

A.2. Excluded countries

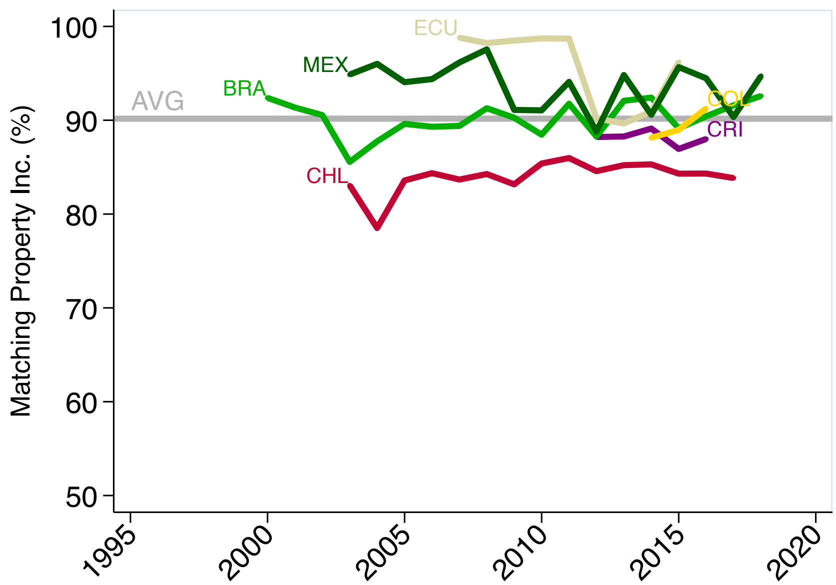
Table A.2: Countries with insufficient data (Excluded)

Country	Survey microdata: Source	Sample size, thousands of individuals	Availability
Bahamas	Bahamas Living Conditions Survey	6	2001
Belize	-	-	-
Bolivia	Encuesta de Empleo,Desempleo y Subempleo, Insituto Nacional de Estadística y Censo (INE)	15 – 40	2000-2019
Cuba	-	-	-
Dominican Republic	Encuesta Nacional de Fuerza de Trabajo (ENFT)	15 – 30	2000-2019
Guatemala	Encuesta Nacional de Condiciones de Vida and Encuesta Nacional de Empleo e Ingresos	10 – 70	2000, 2002-2004, 2006, 2011, 2014
Guyana	-	-	-
Haiti	-	-	-
Honduras	Encuesta Permanente de Hogares de Propósitos Múltiples (EPHPM), Institutio Nacional de Estadisticas (INE)	20 – 100	2001-2018
Jamaica	-	-	-
Nicaragua	Encuesta Nacional de Hogares sobre Medición de Nivel de Vida, Instituto Nacional de EStadística y Censos de Nicaragua	20 – 35	2001, 2005, 2009, 2014
Panama	Encuesta de Hogares, Instituto Nacional de Estadística y Censo (INEC)	40 – 55	2000-2019
Paraguay	Encuesta Integrada de Hogares (EIH) and Encuesta Permanente de Hogares (EPH) from 2002, Dirección General de Estadística, Encuestas y Censos (DGEEC)	15 – 40	2001-2019
Suriname	-	-	-
Trinidad and Tobago	-	-	-
Venezuela	Encuesta de Hogares Por Muestreo (EHM), Oficina Central de Estadística e Informática	80 – 240	2000-2006

Note. Own elaboration.

A.3. Consistency of capital incomes from surveys and national accounts

Fig. A.2. Share of conceptually consistent property incomes



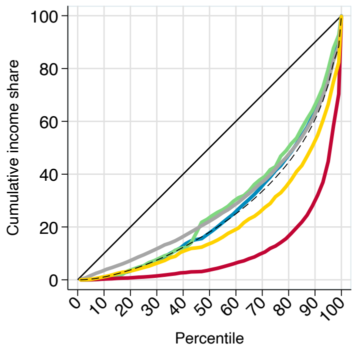
Notes. The share of property incomes from SNA that matches the definition of surveys' capital incomes (dividends and interests) is mostly above 80% of total property income, closer to 90% in most cases. Conceptual differences thus seem to play a minor role in the underestimation of capital incomes displayed in figure 1.3(c). The level of detail that is necessary to observe this is rare in Latin America. Non-matching concepts for the household sector are SNA items D.44, which is composed by D.441, D.442 and D.443 (see table 3.1). Own elaboration based on the public national accounts reported by each country's relevant institutions.

A.4. The composition and distribution of income in household surveys

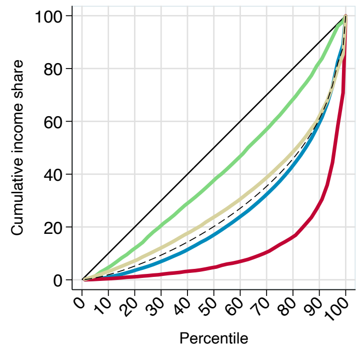
To best understand the distributional impact of the under- or over-estimation of each item on the inequality estimates, figure A.3 displays Lorenz curves of total income (dashed line), along with the cumulative distribution of each income item (colored lines). They all rank individuals by increasing total income. In all cases, we confirm that capital income is the most unequally distributed component. This is especially true in the biggest economies in the region: Brazil, Chile and Mexico. In these cases, the top 10% richest households receives between 70% and 80% of the capital income declared in the survey. Other countries such as Colombia, Costa Rica, Peru, and Uruguay are closer to 60%.

Focusing on the bottom left corner of the subpanels of figure A.3, one can distinguish the income components that are more relevant for poorer households, such as self-employment income and imputed rents. For lower deciles, the curve representing both components is higher than the ones representing other incomes in most cases, but particularly in Colombia, Ecuador, Mexico, and Peru. They appear to be less unequally distributed, since they are found in both poorer and richer households. For instance, the income of both shopkeepers, or street vendors, and doctors is usually included as self-employment income. Similarly, the imputed rent to homeowners includes both housing projects acquired through vouchers and luxury dwellings. In the case of Mexico, for instance, where both types of incomes are greatly underestimated (figures 1.3(e) and 1.3(f)), if we were to adjust the survey –say, by scaling these incomes proportionally– to include the missing part, it would thus have and equalizing impact on overall inequality estimates. Similar conclusions can be reached for Colombia, whereas, in the case of Peru, for which imputed rents are over-estimated, the same adjustment would probably increase inequality. In the cases of Brazil and Costa Rica, self-employment income is overestimated and seems to be more unequally distributed than other incomes (with the exception of capital income). A proportional adjustment of these incomes would thus probably have an equalizing effect. In contrast, adjusting to match higher aggregates of capital income would lead to substantial increases in the level of inequality in all countries.

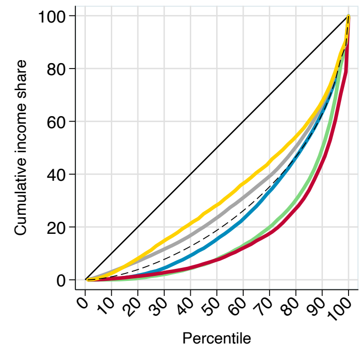
Fig. A.3. Income incidence in latest survey



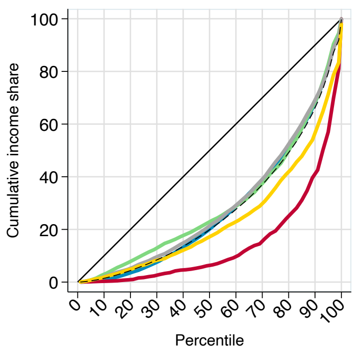
(a) Brazil - 2019



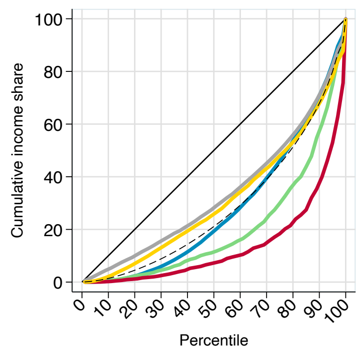
(b) Chile - 2017



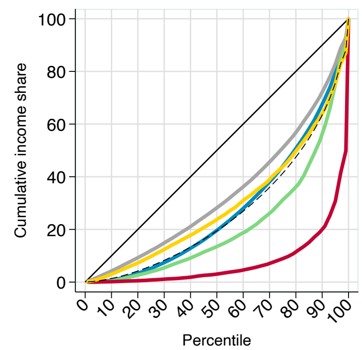
(c) Colombia - 2018



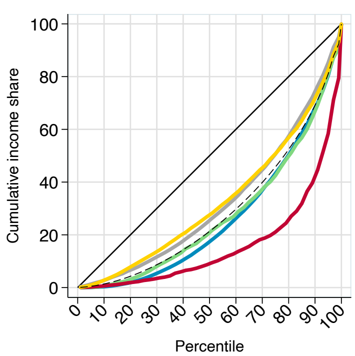
(d) Costa Rica - 2019



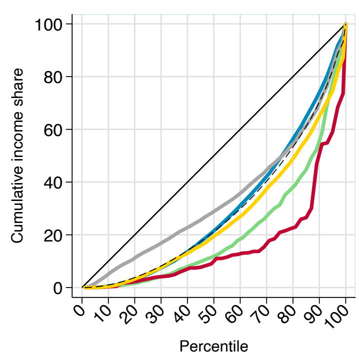
(e) Ecuador - 2019



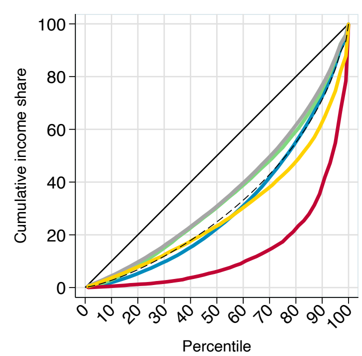
(f) Mexico - 2018



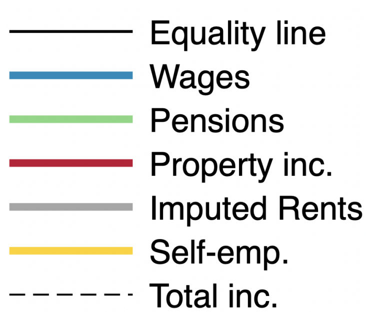
(g) Peru - 2019



(h) El Salvador - 2019

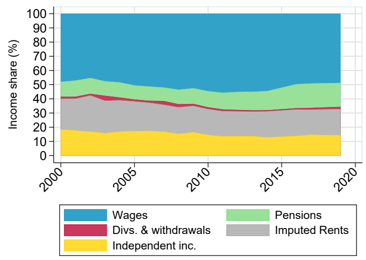


(i) Uruguay - 2019

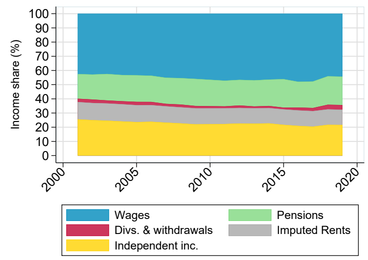


Note. Individuals are ranked by increasing total income. Income is pretax, net of pension contributions. Own elaboration based on ECLAC's harmonized surveys.

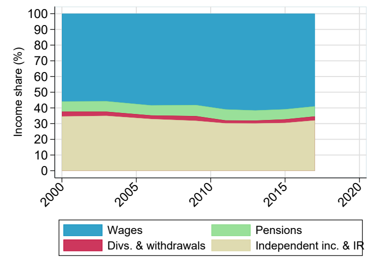
Fig. A.4. Total income composition



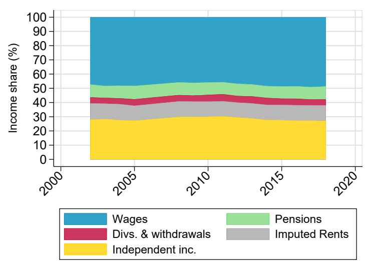
(a) Argentina



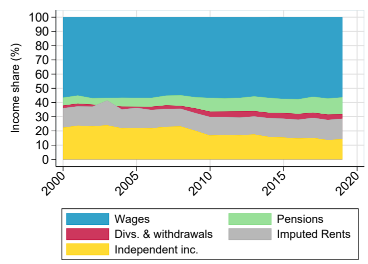
(b) Brazil



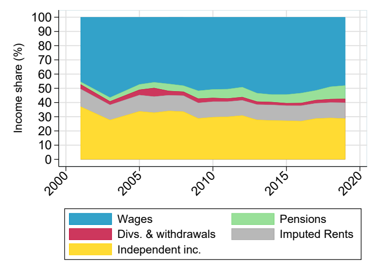
(c) Chile



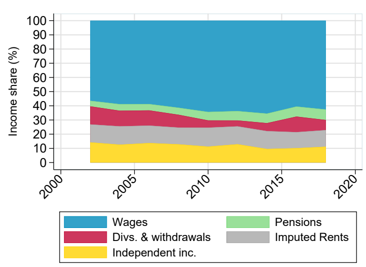
(d) Colombia



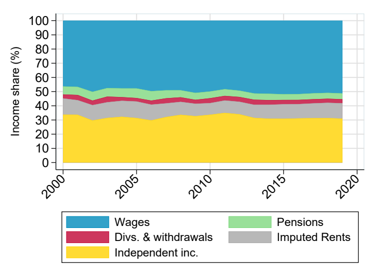
(e) Costa Rica



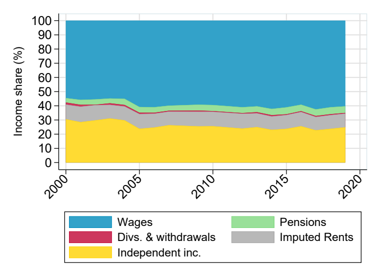
(f) Ecuador



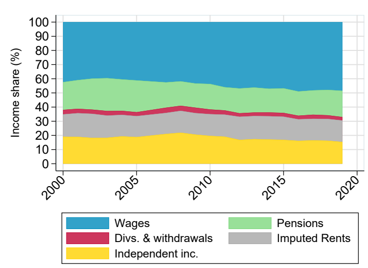
(g) Mexico



(h) Peru



(i) El Salvador



(j) Uruguay

Note. Own elaboration based on ECLAC's harmonized surveys. Income is pretax, net of pension contributions.

Chapter 2

More unequal or not as rich? Revisiting the Latin American exception

Joint work with Ignacio Flores and Marc Morgan.

Abstract

Latin America is often portrayed as a global exception to the rising or consolidating income inequality trends of the early twenty-first century. However, the use of administrative data and macroeconomic aggregates casts doubts on this survey-based narrative. In this paper we revisit the region's exceptionalism by building the most comprehensive data base thus far, which accounts for 80% of the region's population and combines harmonised surveys, social security and tax data, and national accounts. We produce a set of inequality indicators —pre and post-tax, based on alternative units and income definitions— which allows us to track the distributional effects of each methodological step and reconcile divergent trends. The reconciliation of micro and macro data present us with a dilemma: either the region is more unequal or it is not as rich as officially reported. The result of distributing the data gaps is a region much more heterogeneous in its inequality trends. Falling inequality is most visible among the bottom 99%, but the trend flattens or reverses in the largest economies once the top 1% and capital incomes are better accounted for. Post-tax and disposable incomes do not change the picture much, except when in-kind social spending is considered. These results confirm the strengths and highlight the limits of Latin America's redistributive policies during the period, and allow us to reconcile competing inequality narratives.

1. Introduction

Inequality has been on the rise in most countries and regions for the best part of thirty years, spurring academic and political debate worldwide. Latin America seems to have been a notable exception in more recent times. Numerous studies have documented and explained the apparent decline of income inequality taking place throughout Latin American countries during the first decade and a half of the twenty-first century (López-Calva and Lustig, 2010; Lustig et al., 2011; Cornia, 2014b; Alvaredo and Gasparini, 2015; Gasparini et al., 2018). This trend has even been viewed as historically unprecedented in a region characterised by extreme inequality legacies (Bértola and Williamson, 2017). However, its narrative, built on the use of publicly available household survey data, has come to be questioned by the increasing use of administrative data on upper incomes in the region (Alvaredo, 2010; Alvaredo and Londoño Velez, 2014; Morgan, 2018; Souza, 2018; Flores et al., 2020; Burdín et al., 2022), which have shown either milder reductions in top income concentration or more stable, if not increasing, trends in some countries. These doubts are compounded by the large discrepancies between incomes in micro data sources (surveys, tax data) and the national accounts.

Such gaps present us with a distributional conundrum: if they are widest in capital incomes, as has been found historically and more recently, then this would entail significant repercussions for existing inequality indicators (Alvaredo, De Rosa, Flores, and Morgan, 2022). Moreover, if these gaps are subject to changes over time, as also appears to be the case, our assessment of inequality trends would be severely compromised. With all this cumulative information at our disposal, how confident can we be in thinking that Latin America was an exceptional outlier in the global income inequality narrative? The stakes of this question are high, since accepting the survey-based narrative outright, in the context of large and growing micro-macro data gaps, could mean putting into serious question the official macroeconomic growth statistics of countries in the region. Seeing whether the conventional narrative on Latin American income inequality is robust to the reconciling of micro-macro data gaps is the focus of the present paper.

Our first contribution is to use all available data —including several brand new sources— to build novel estimates of macro-consistent inequality in a region with high heterogeneity in data quality. We present distributional results for ten countries (Argentina, Brazil, Chile, Costa Rica, Colombia, Ecuador, Mexico, Peru, El Salvador and Uruguay) over the last two decades, a period when the region as a whole experienced strong economic growth spurred by very favourable terms of trade for the most part (circa 2003-2013) and relative stagnation during the latter years (circa post-2015).[1]

In order to build our estimates, we combine harmonized survey microdata with administrative data from tax records and social security registers (based on the re-weighting method put forward by Blanchet et al., 2022) before scaling incomes to the national macroeconomic accounts (both the household sector accounts and the total economy sector accounts). Thus, we reconcile all available income data to build inequality estimates that not only adjust for surveys' measurement issues, but also ensure overall macroeconomic consistency. As anticipated in Alvaredo et al. (2022), the adjustments end up doubling the total income originally declared in most surveys. Hence, to ensure transparency, we show the impact that each step of our methodology has for the resulting distributions. We distinguish four steps: first, we estimate the distribution of income in the harmonized survey data; second, we adjust for the low representativeness of top incomes in surveys using administrative records; third, we scale the main income components to their matched national accounts aggregates (these are wages, property incomes, mixed income, pensions and imputed rents); and fourth, we impute incomes not flowing to the household sector in the national accounts (corporate retained earnings and other incomes) or incomes that need to be added back (such as net product and production taxes) to reach the net national income of the total economy.

This sequence describes what we call the pre-tax national income series. It includes all gross incomes, including pensions, before taxes, but after social security contributions. We also produce a number of post-tax series which account for taxes, monetary transfers and in-kind public spending. Although we directly observe the incidence of some items, such as the personal income tax in administrative records or social cash benefits in surveys, we use external sources to impute other items. Aggregates and compositions of national taxes and social spending come from OECD and World Bank public databases respectively. We combine these with incidence profiles from the Commitment to Equity (CEQ) database, which are mainly based on budget surveys, to allocate consumption taxes and in-kind spending to individuals. In other cases, we either use micro-simulation techniques or proxies, as described further on.

While the reconciliation of micro and macro estimates of income may seem a relatively new and important topic for the developed world (the next revision of the United Nation's System of National Accounts intends to incorporate it into its guidelines), we recall that this is not a new topic in Latin America. Following the pioneering work by Altimir (1987),

[1]Our methodology, codes and estimates will be made public on a dedicated open-source website that is currently under construction. Users will be able to view and download distributional information at the percentile level for different income definitions and observational units.

macro-adjustments to inequality estimates in Latin American countries were made by the UN's Economic Commission for Latin American and the Caribbean (ECLAC) during more than two decades, before being discontinued for data reasons. Our work seeks to build on this 'lost tradition' in the region.

This exercise allows us to offer more precise answers to basic questions: How is macroeconomic growth distributed within countries? How progressive or regressive are the systems of redistribution in each country if one accounts for all taxes and transfers? To answer these questions, distributional estimates must necessarily be reconciled with macroeconomic aggregates, which follow homogeneous definitions across countries. Despite recent efforts to define benchmark methods to achieve consistency, leading initiatives have mostly focused on a handful of countries with exceptional national statistics so far, overlooking problems that are particular to a majority of countries, including both developed and developing ones (WIL, 2020; Zwijnenburg, Bournot, Giovanelly, and Giovannelli, 2017; Zwijnenburg, 2019). For instance, two pioneering studies in the United States and France heavily rely on detailed tax micro-data to portray income distributions, only using surveys to describe small sections at the very bottom of the distribution (Piketty, Saez, and Zucman, 2018; Garbinti, Goupille-Lebret, and Piketty, 2018). The same approach would be poorly adapted for countries where tax coverage and compliance are much lower, which is the case of most countries in the world. In such a setting, tax data can only be trusted to portray top income groups relatively accurately, while household surveys can better inform on middle and bottom incomes, which generally have higher informality rates and higher shares of un-taxed incomes.

Our second contribution is to revisit the prevailing narrative of falling inequality in Latin America over the first fifteen years of the twenty first century. Regarding the *level* of inequality, we are faced with a mutually-exclusive dilemma. If we assume that the national accounts are an accurate benchmark for aggregate incomes, and proceed to distribute the macroeconomic income of the household sector or the total economy, our conclusion is that inequality is in fact much higher than previously thought. After adjusting surveys based on administrative data and scaling income components to the national accounts, inequality levels increase significantly —the Gini coefficients in our sample increase by about 10 points, with notable heterogeneity across countries. If, on the other hand, we assume that official surveys are closer to the benchmark for household incomes, our results are consistent with the current consensus. However, one would also need to accept that Latin American households are considerably poorer than what is reported by official macroeconomic statistics.

The analysis of inequality *trends* is not as clear cut. The adjustments we make to the survey distribution are enough to cancel out the pre-tax inequality decline in countries where it was present —Brazil, Chile, and Mexico— or to increase inequality where it was stable —Costa Rica. In the remaining countries (Argentina, Colombia, Ecuador, Peru, El Salvador, and Uruguay) the falling inequality trends persist after the three sets of adjustments, although in a milder fashion. In some cases, such as Brazil or Mexico, a trend reversal is visible before ensuring macro consistency (i.e. at the adjusted survey level), so there is room to believe that both statements may be true: inequality did not fall as the prevalent narrative says it did, even if countries are not as rich as what is estimated by national accountants. In other countries, however, changes in trends are more clearly visible when scaling up to household incomes or national income, and so the answer may again be mutually exclusive. Furthermore, although our estimates confirm the regressive distributive effect of national taxes and cash transfers (mainly due to consumption taxes), the progressive impact of in-kind social spending (in health and education) allows for the falling inequality narrative to emerge with greater clarity. Mexico is our sole exception to this trend: inequality in the unadjusted and adjusted surveys do not mirror each other in any of the definitions.

In light of these findings, we attempt to reconcile competing inequality narratives by clarifying issues that affect comparability such as units of analysis, income concepts and the choice of inequality indicators. More importantly, we analyze the contribution of capital incomes and top income groups, which are by all accounts the main missing pieces of household surveys. We document that inequality among the bottom 99% of the total income distribution and among wage earners falls even after making all adjustments. We show how divergent trends in total income inequality are the result of an increasing contribution of capital incomes and top 1% incomes after each adjustment procedure. This is due to both an increasing distance between the top 1% and the bottom 99%, and to increasing inequality within the top 1%. Thus, we do not fully contradict the prevailing narrative; if anything, we confirm it with some qualifications. Fundamentally, we claim that the role played by capital incomes and top 1% incomes reveals the limits of Latin America's much heralded re-distributive effort of the early twenty-first century, even if certain policies appear to be key for a robust inequality decline (such as public spending on health and education).

Given the scale of the data deficiencies we are dealing with, we stress caution in proclaiming definitive statements for the region. Our goal is to contrast competing inequality narratives and provide broad insights on the driving forces of divergent trends. In this sense, country-specific studies are usually better equipped to discuss details about the specific evolution of different series for each country. However, by systematically applying the same set of methodological decisions to the whole region, we are able to provide a bird-eye's view of the evolution of inequality among its six hundred million inhabitants in the only part of the world in which it just might have fallen.

We stress from the outset that this procedure is experimental, intended to answer a specific research question. Although it can also have a broader interest for policymakers and the general public, it is by no means a gold standard. The implication of our work is to highlight the deficiencies in the myriad of current statistics on incomes, which cloud our understanding of the crucial issue of economic growth and its distribution. If anything, it is a call to data producers in the region, and the world at large, to provide better, more integrated and coherent statistics on the incomes of their populations.

The remainder of the paper is structured as follows. Section 2 describes the data and methodologies used to construct the series. Section 3 presents the pre-tax inequality estimates, while section 4 discusses the redistributive effects of taxation and spending. In section 5, we attempt to reconcile the competing narratives that emerge from alternative inequality series, before ending with concluding remarks in section 6.

2. Building macro-consistent inequality estimates

This section summarises the challenge of reconciling micro and macro estimates of income, before assessing results in the following sections. A more detailed description of the data and methods used to build the estimates presented in the rest of the paper is available in Appendix A and B.

2.1. Statistical inconsistency as a rule

There is a longstanding gap between the statistics used to study the distributions of income, wealth and consumption at the micro level and macroeconomic aggregates in the system of national accounts (SNA). A wide body of work shows, in many different contexts, that major discrepancies are found when studying aggregate levels, as well as in their observed growth rates (Ravallion, 2003; Deaton, 2005; Bourguignon, 2015; Nolan et al., 2019). A noteworthy finding is that national income, which is measured by the SNA, is larger and has been growing faster than other income concepts traditionally used to study inequality. Whenever survey aggregates are compared to SNA aggregates, capital incomes appear to be remarkably less covered than labor incomes (Törmälehto, 2011; Bourguignon, 2015; Flores, 2021; Alvaredo et al., 2022). Such gaps make it hard to assess how macroeconomic growth is distributed among the population, and to what extent existing distributional statistics (based both on surveys and tax records) are an accurate representation of material living standards.

An approach taken in the literature on global inequality to address these gaps has been to assume that the discrepancy between total survey income and national income, or Gross Domestic Product (GDP), from the national accounts is entirely due to an underrepresented top tail, usually the top 10% or top 1%. The entire gap is thus imputed to this income group to adjust global estimates of Gini indices (Lakner and Milanovic, 2016; Anand and Segal, 2017). The issue with this type of adjustment is that it is arbitrary and restrictive, in the sense that it attributes the entire difference between two aggregates to a top group, without assessing the decomposition of the aggregate gap across income types and thus population groups in the micro-level statistics.

Recent work in this field has now embarked on a process of combining data sources (surveys, national accounts, administrative registries, rich lists, etc.) through the development of two large-scale projects aiming to ensure the macroeconomic consistency of inequality estimates. On one side, following recommendations by the Canberra Group (2001) and Stiglitz, Sen, and Fitoussi (2009), the Organization for Economic Co-operation and Development (OECD) started hosting periodic Expert Group meetings on Disparities in a National Accounts Framework (EG-DNA), focusing exclusively on the income, consumption and savings of the household sector (Fesseau and Mattonetti, 2013; Zwijnenburg et al., 2017; OECD, 2020).[2] On the other side, the World Inequality Lab at the Paris School of Economics started publishing its own Distributional National Accounts guidelines (WIL, 2020); alongside numerous country-case studies.[3]. The main difference with respect to the OECD's approach is that DINA aims to distribute the national income of the total economy as opposed to just the household sector (for an in depth comparison of these projects see Zwijnenburg, 2019).

In Latin America there is an old tradition of aligning micro and macro data for distributional analysis, largely following the work of Altimir (1987). This seminal study critically analyzed available tax, social security and census data, as well as variety of household surveys, systematically comparing the latter with the national accounts. The author concluded that there was a 15-30% gap in aggregate household income, which could be significantly higher for certain income sources such as property income. These results were explicitly assumed to be an indicator of underestimation of each type of survey-based income. Hence, the United Nations' Economic Commission for Latin America and the Caribbean (ECLAC) proceeded

[2] See https://stats.oecd.org/Index.aspx?DataSetCode=EGDNA_PUBLIC for experimental statistics based on the output of this project.

[3] See Piketty et al. (2018); Garbinti et al. (2018) for pioneering applications of the methodology and https://wid.world for further applications

to correspondingly adjust survey-based incomes, with significant implications for inequality analysis —the Gini coefficients increased by 10-15%). Despite its positive intentions, this methodology was shown to have many caveats (Bourguignon, 2015), and was progressively abandoned by ECLAC in recent years for reasons that are not entirely clear. The rise and fall of this experience are the result of both the need for a reconciliation of data sources —or at least of the need to fully understand its potential consequences— and of the significant challenges of such an endeavour. Our goal in this paper is precisely to recover this critical comparative tradition with the latest data and methods presently available. We turn to these in the following sections.

2.2. Data inputs

Our estimates rely on four main data sources: households surveys, income tax records, social security records, and national accounts.

We use survey micro-data harmonised by the Statistics Division of ECLAC for ten countries over the 2000-2020 period. These countries are: Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Mexico, Peru, and Uruguay. ECLAC's harmonisation process builds on the original surveys produced on a yearly basis by the official statistics institutes of each country. It seeks to create comparable income variables across countries in terms of labour, capital and mixed incomes, pensions, owner occupier rental income, transfers and other incomes.[4] In all cases but one, post-tax incomes are recorded on an individual basis (where "post-tax" also refers to after social contributions), the exceptions being Brazil and Cost Rica, where pre-tax incomes are recorded by surveyors. Part of the harmonisation process involves the imputation of rental income of owner-occupiers, which is absent in the surveys. This calculation is based on an internal estimation model that matches data from similar rented dwellings. Figure A.1 shows the inequality trends computed from ECLAC's survey database mirrors that from the World Bank's database, with differences in levels for the same unit of analysis being minor.[5] Table A.1 describes the original surveys and administrative data available for each country that we study in this paper.[6]

[4] The only exceptions concerning the frequency of the surveys are Chile and Mexico, which collect data every two to three years.

[5] This should be of no surprise given that the underlying surveys are the same in both databases, with differences arising from interventions both both institutions. For a representation of the composition of survey income among five categories in the ECLAC database see Figure A.2.

[6] Excluded countries with some household survey data are presented in Table A.2. These countries do not form part of our analysis due to inefficiencies in survey data and/or unavailable administrative data to complement them.

Available distributional data from administrative sources in Latin America can be classified in four groups.[7] First, microdata covering the population required to submit a tax return on their income (e.g. Mexico and Uruguay). Second, grouped (tabulated) data organised by ranges of total income (e.g. Argentina, Brazil and Chile). Third, distributional data covering income tax payers with wage income only, either in microdata format (e.g. Argentina, Costa Rica), or in tabulated form (e.g. Brazil). And Fourth, in an increasing number of countries, information on the distribution of wages is made available from the social security administration, either in micro-data or grouped-data format. Naturally, this is restricted to the formal sector, and, depending on each country institutional arrangements, this may include the universe of formal workers, or only those in the main social security regime. We use social security records in the case of Costa Rica. The ten countries can be divided in two groups. On the one side, those regularly publishing and updating their administrative records (Brazil, Chile, Argentina, Mexico and Uruguay). On the other side, those that gave external researchers access to microdata at some point, but do not produce distributive information from tax registers on a regular basis (Colombia, Costa Rica and Ecuador). For these cases, we use estimates prepared by the authors of previous studies (Alvaredo and Londoño Velez, 2014; Cano, 2015; Zuniga-Cordero, 2018, 2022; Rossignolo et al., 2016), which are restricted to the top percentile of the distribution only. In section A.2 we report the use of income tax data in the literature on top incomes for eight of our ten countries. In the remaining two countries (Peru and El Salvador), we obtained access to new tabulated data on incomes from the respective country tax offices for the purposes of this project.[8]

The information from the System of National Accounts (SNA) was obtained by scrapping the United Nations Statistics Division database (`http://data.un.org`), which gathers a variety of series produced by national statistical offices or central banks. We complement this source with country-specific data on National Accounts published by either Central Banks or National Statistical Institutes, which are sometimes more up-to-date. We also use data from the World Inequality Database on undistributed corporate profits, the OECD on taxation, and the World Bank on social transfers in kind. Although the macro aggregates produced by national accountants are often considered among the most reliable and internationally comparable data sources (e.g. to rank countries according to their total output, per capita GDP, etc.), detailed information on the income approach, which is the one we need for our purpose, is scarce in the region, to say the least. Even in countries that produce this kind of

[7] See Table A.1 for further details. In this paper we use the terms "administrative", "register" and "'tax" data interchangeably to describe data from personal income tax declarations or social security records on wages and salaries.

[8] At the time of writing, the authorities of the Dominican Republic have made income tax data exclusively available to us for the purpose of a separate study, which is currently under embargo.

data regularly, statistical agencies can update their estimates with three to five years of lag. The level of aggregation also varies across countries.

Figure A.3 (from Alvaredo et al., 2022), provides a visual comparison of the main aggregates across the different sources we exploit. It shows the decomposition of gross national income (GNI) into the household sector, the government sector and the corporate sector. It also presents the aggregate income informed by surveys, before any adjustment, as percentage of GNI, as well as the reported income in administrative data. Three countries, Argentina, Uruguay and El Salvador, do not report aggregates from the income approach in their SNA. For the other countries that do so, the time coverage is rather short, and usually below that of surveys. However, one result is clear: the gap between micro-distributional statistics and the national accounts is very large, usually above 40-50%.

2.3. Methods

2.3.1. Pre-tax distribution

Our estimation procedure is based on four steps. First, we estimate inequality indicators from the harmonised survey microdata of the full population of each country. We compute income shares and Gini coefficients of total income and of wage income. Second, we adjust these surveys to improve their coverage of top incomes using administrative data. Third, we scale the different income components of these top-corrected surveys to match equivalent aggregates from the national accounts. Fourth, we impute the remaining items (corporate undistributed profits and other pretax incomes) that make up national income. We briefly describe each of these steps in turn.

Given the nature of the harmonised household survey database, the estimation of distributional indicators is relatively straightforward. From the microdata, we rank the population by total income (or total wages) and subsequently compute shares of total income or Gini coefficients. Total income is in this stage the sum of net-of-social-contribution wages, pensions, self-employment income and capital income. The reason for including pensions in this definition of income is that wages in the surveys are reported net of social contributions in all countries (except for Brazil), without information on the amounts paid per person. This makes it unfeasible to leave pensions and their contributions for the redistributive analysis. The second step consists in combining household surveys and distributive information from administrative sources, mainly to improve the coverage of top income groups, which are often badly captured (especially when register data is not used in the surveying process, which is the case in all countries in the region).

In practice, we use the method described in Blanchet, Flores, and Morgan (2022), which uses the ratio of survey observations to administrative observations by income percentile beyond a cut-off point (or "merging point") to adjust survey weights. Figure B.3 displays the intuition behind this re-weighting process. The density ratio described can be interpreted as a rate of response, which is generally lower than one for top incomes. For surveys where administrative records do not exist, we assume within-country stability for these coefficients to make the adjustment. Prior to reweighting the survey we deduct tax paid from the declared income in tax data for all countries where the survey reports post-tax income (i.e. all countries except Brazil). Appendix B.2 explains this procedure in more detail. This ensures that we are adjusting the survey using a comparable income definition.

The second step consists in scaling the adjusted survey incomes to equivalent aggregates in the household sector account of the system of national accounts (SNA). Before doing so, we add back the effective income tax paid by percentile group to the distribution so that we compare pre-tax micro-level incomes with pre-tax macro-level incomes. Table 3.1 summarises the matching we perform between incomes in surveys and the SNA. Since the income decomposition of the SNA is not available for every country and every year, we assume within-country stability of these coefficients. For countries where this decomposition is never reported (Argentina, El Salvador and Uruguay), we use the period's regional average to scale each type of income. Figure B.5 reports the "scaling factors" we use for each component in each country, that is, we multiply each survey component by 1/scaling factor, taking comparable household incomes from the national accounts as benchmarks. As found in Alvaredo et al. (2022), capital income in the survey is systematically under-covered in all countries by the largest margin, implying that we multiply the income component by a factor of 5-10 in most cases. Some income aggregates need to be deflated to arrive at the SNA benchmark, typically either imputed rents, mixed income or social benefits. These gaps are a function of comparability issues (outlined in Table 3.1) and complications with annualizing composite survey income variables from reference periods (see Alvaredo et al., 2022 for more details).

The final step of our procedure is to impute the remaining incomes included in the net national income of the total economy. By definition, these do not match any of the income variables that are present in the distributive data we use. Government sector's net primary income (mainly taxes net of subsidies on products), which represents on median 10% of net national income, lacking of a better alternative which does not rely on heavy assumptions, is imputed proportionally. Thus essentially, this stage boils down to the imputation of corporate undistributed profits to households. Figure B.7 shows that aggregate undistributed profits from the SNA (sourced from https://wid.world/) are usually in the order of 10-20% of total survey income, which is a significant amount. In order to distribute this aggregate amount to individuals, we need a proxy for corporate ownership. Since wealth surveys are mostly absent from the region, we use variables from our income surveys as proxies. One option would be impute them to dividends. However, too few people declare dividends in our surveys. Our benchmark allocation is to impute them proportionally to the sum of dividends and employer income, where an employer's income refers to the total income of individuals that declare being an employer in surveys when asked about their occupation. Since the amount of undistributed profits is not available for every country and every year, we proceed similarly to what was done for scaling factors, i.e. we assume within country stability of these coefficients and use regional averages for countries with no data. Figure B.8 documents the incidence of total corporate retained earnings across the distribution, showing that almost all of the amount is allocated to the top decile, especially the top 5% and 1%, as one would expect.

2.3.2. Post-tax distribution

After estimating the pre-tax national income series, we produce a number of post-tax series which account for taxes, monetary transfers and in-kind spending. Although we directly observe the incidence of some items, such as the personal income tax in administrative records or social benefits in surveys, we use external sources to impute other items. We use incidence profiles from the Commitment to Equity (CEQ) database, which are mainly based on family budget surveys, to allocate consumption taxes and in-kind spending to individuals.[9] Macroeconomic aggregates on each tax and social spending category are taken from OECD and World Bank databases, respectively.

From the pre-tax national income distribution we estimate three varieties of post-tax distributions. In the first variety, we deduct all direct taxes on personal and corporate income and add all social assistance transfers in cash. The amount of direct taxes to impute are taken from the OECD/ECLAC/CIAT/IDB (2022) database, which correspond to taxes present in the national accounts.[10] We impute personal income taxes using the profile presented in income tax declarations for countries (depicted in Figure B.1). For corporate income taxes we impute them proportional to the distribution of retained corporate earnings from the pre-tax distribution, that is, the joint distribution of dividends and employer income in

[9] See https://commitmentoequity.org/.

[10] The reason for using this database over UN or country-level national accounts is that it presents a more detailed breakdown of the tax categories to be imputed to the distribution. Figure C.9 describes these categories for Latin American countries as well as the average for the whole region.

surveys. To add social assistance transfers in cash we simply impute the aggregate present in the national accounts (D623 in SNA 2008) proportional to the micro-distribution of these transfers observed in the household surveys (that is, social transfers excluding pensions and other contributory social insurance transfers, D621 and D622). We label this the "post-tax spendable" distribution.

In the second variety, which we label the "post-tax disposable" distribution, we deduct all indirect taxes on production and consumption. The amount of these taxes are taken from the OECD/ECLAC/CIAT/IDB (2022) database (see Figure C.9), while their distribution is imputed proportional to the incidence estimated by the studies in Lustig (2018) and updated in CEQ (2021).

In the third variety, which we label the "post-tax national" distribution, we add social transfers in kind received by households (D63) and all other remaining incomes. Social transfers in kind correspond to government spending on services like education, healthcare and other collective expenditures (defence, roads, administration, etc.). Their aggregate amount by category is taken from the World Bank database.[11] To impute their distribution, we distinguish between education and health expenditures on the one hand, and all remaining expenditures on the other. The reason is that for the former we avail of their estimated incidence from the studies in Lustig (2018) and CEQ (2021), which attribute spending on education and healthcare services to household members according to their use of the services. For the remaining collective expenditures we impute them proportional to the post-tax disposable income distribution due to a lack of reliably justifiable estimates of their incidence. The remaining incomes that make up post-tax national income are imputed proportionally to the disposable income distribution, including other current transfers between households for which we don't avail of a reliable breakdown in the surveys.[12]

Alongside these three varieties of post-tax distributions, which build on the pre-tax national income distribution combining all the sources previously described, we also estimate a distribution of post-tax spendable income just based on the surveys, which we label the "post-tax raw" distribution. This series is the common one used in the inequality literature in the region, and we use it to compare to the series we estimate based on the combination of survey, register and national accounts data. The following sections present our results and

[11] See https://data.worldbank.org/.

[12] The separation of private transfers from other (social) transfers in CEPAL's harmonized surveys is a line of future work that is being explored by the institution. This is important as these private transfers include transfers between households in the country and also between domestic and foreign households, the latter of which (i.e. remittances) can be especially large for smaller economies in the region.

discussion of our findings.

3. Growing richer and less unequal?

The new millennium brought an exceptional growth cycle to Latin America, mainly led by a global increase in commodity prices, which inflated exports from the region and is often cited among the main causes in the falling-inequality narrative that derives from survey-based statistics (Ocampo, 2017; Cornia, 2014b; Sánchez-Ancochea, 2021). This section analyses the evolution of pre-tax income inequality over a period encompassing such event. The first subsection reveals the impact of each step of our macro-adjustment procedure over that period, while the second, investigates on who actually benefited from the commodity boom (2003-2013) if one accounts for the distribution of macroeconomic income.

3.1. Reassessing pre-tax income inequality

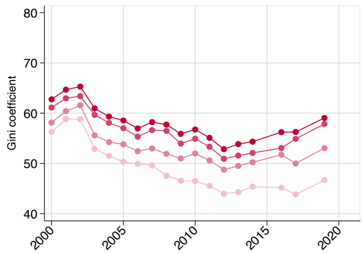
Figure 3.1 depicts the evolution of the Gini coefficient of the four distributions described in the previous section (see section 2.3.1).[13] In all cases, the distributional estimates refer to per-capita household income for comparability reasons with survey-based series (more on this in section 5).

Three key comments regarding the evolution of the Gini coefficients are worth highlighting. First, inequality estimates increase after each of the adjustments to the raw surveys for all countries and years considered. The adjustment of surveys based on available tax data increases overall income inequality as a result of increasing the weight of higher income individuals. The subsequent scaling of household incomes to national accounts increases inequality, as incomes that are scaled up by higher factors are precisely those that are more concentrated in the top tail (especially property incomes, see Figure B.5). The final adjustment to national income increases inequality as the result of the allocation of undistributed corporate profits, which represent a large share of aggregate income (Figure B.7), and are imputed mostly to top percentiles, given the hypothesised structure of business ownership (Figure B.8).

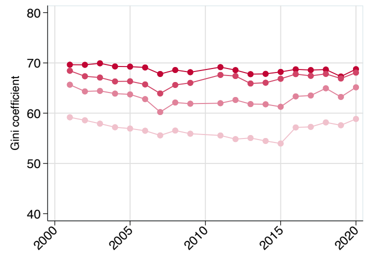
As far as the level of inequality goes, we are left with quite an unambiguous result (or dilemma): the region is either more unequal than previously thought or not as rich as what is reported by official macroeconomic statistics. How should we interpret this finding? Unlike the pioneering efforts by Altimir (1987) or the current agenda of Distributional National

[13]For details on income shares, see Figure C.1, Figure C.2, Figure C.3, and Figure C.4, which depict the shares of the Bottom 50%, Middle 40%, Top 10% and Top 1% for each of the four distributions respectively.

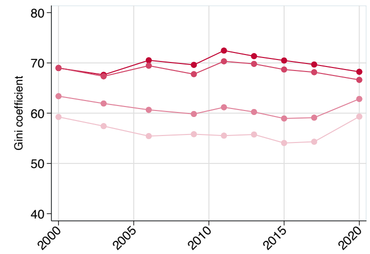
Fig. 3.1. Gini coefficients in four distributions



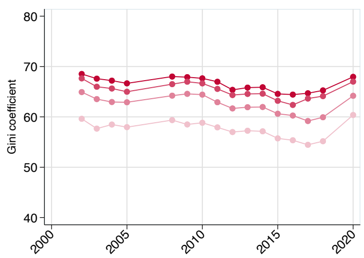
(a) Argentina



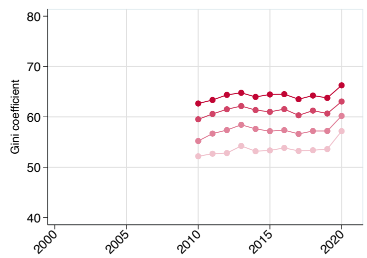
(b) Brazil



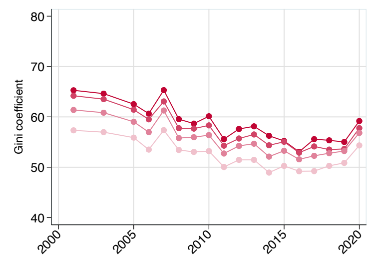
(c) Chile



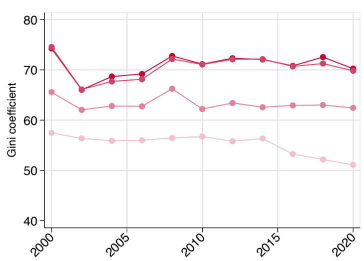
(d) Colombia



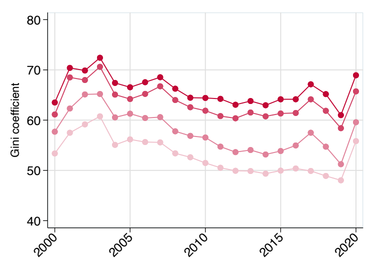
(e) Costa Rica



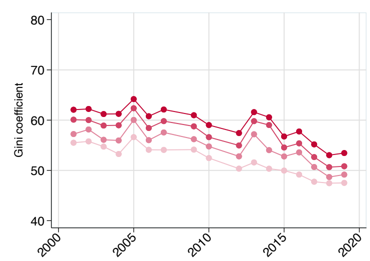
(f) Ecuador



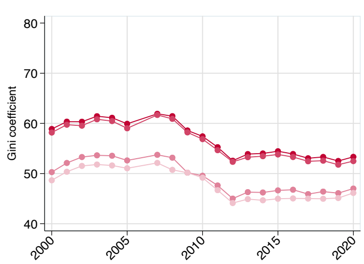
(g) Mexico



(h) Peru



(i) El Salvador



(j) Uruguay

Note. Authors' elaboration. The figures depict four distributions: the household survey-based distribution and the three augmented distributions based on three adjustment steps to the survey. The first step uses administrative data (income tax data or social security wage data) to reweight the raw survey; the second step scales the income totals in the tax-adjusted survey to their equivalent household-level aggregates in the national accounts; the third step imputes missing incomes needed to reach national income. The distributions are of pre-tax household per capita income (including pensions and after social contributions).

Accounts (WIL, 2020), we do not claim that the national accounts are without question the benchmark source for measuring incomes, at least not in the Latin American case, precisely because of the major shortcomings and opacity of national accounting in the region. What we do claim is that if we take all data sources seriously, there is a large micro-macro gap with significant effects on inequality. As already noted in Alvaredo et al. (2022), not all of the survey income–national income gap is the result of measurement issues (only about half). A significant share is explained by conceptual differences, most notably those related to undistributed profits, which are incomes attributed to the financial or non-financial corporate sector in the national accounts, and thus not to households. Imputing these incomes to households ends up adding 5-10 points to the Gini, as the step from the scaled household income distribution to the pretax national income distribution in Figure 3.1 shows. A case can certainly be made for their exclusion from any household income inequality indicator on a conceptual basis (indeed, even national accountants keep them separate from the household sector). Yet, taking data sources seriously also means recognising the purpose of their construction. The national accounts are built around the concept of production, and the distribution of produced value-added between aggregate production units at the institutional level. Although a share of total corporate profits may remain in corporate accounts as retained earnings for future use, a strong argument can be made to impute these earnings to the owners of such businesses, which after excluding foreign shareholders and government involvement, are ultimately household individuals (participating shareholders and working directors) who have property and commend over such incomes. Thus, including them in inequality measures incorporates actually produced incomes as well as the concept of power into the analysis. Moreover, it is a way of assuring that tax-based incentives to distribute or withhold corporate profits do not affect estimates of inequality over time.

Secondly, as far as inequality trends go, in some cases the broad downward trajectory from the beginning of the period to the end holds after each of the three adjustment steps. This is the case for Argentina, Colombia, Ecuador, El Salvador, and Uruguay (at least prior to 2020).[14] For other countries —such as Brazil, Chile, Mexico and Peru— we observe trends that gradually flatten or even increase with each step. In the cases of Brazil or Mexico, stability is already visible after the first step, while for others the trend stability is more visible after scaling incomes to the household sector account, e.g. Chile, and Mexico. Furthermore, around 2015 it appears that the falling inequality trend comes to a halt and even reverses in several countries, detectable already in the raw survey.

[14]In the case of Costa Rica, the trends are also consistent among the four distributions, but in the opposite direction.

Third, as shown by previous literature (see section A.2), falling aggregate inequality may coexist with stable or even growing shares going to the top 1% (Figure C.4). In all cases where the survey-based top 1% share was stable or slightly decreasing, after the top-income and macro-income adjustments it increases (most dramatically in Mexico). Figure C.5 shows that even in the presence of a stable or increasing top 1% share, the dynamics between the top 10% and the bottom 50% and middle 40% shares can still produce falling inequality for a number of countries, such as Uruguay, Argentina or Ecuador. Thus, a much more heterogeneous and complex picture emerges from the anatomy of macro-consistent inequality than that coming from the survey-based narrative.

3.2. Who benefited from the commodity boom?

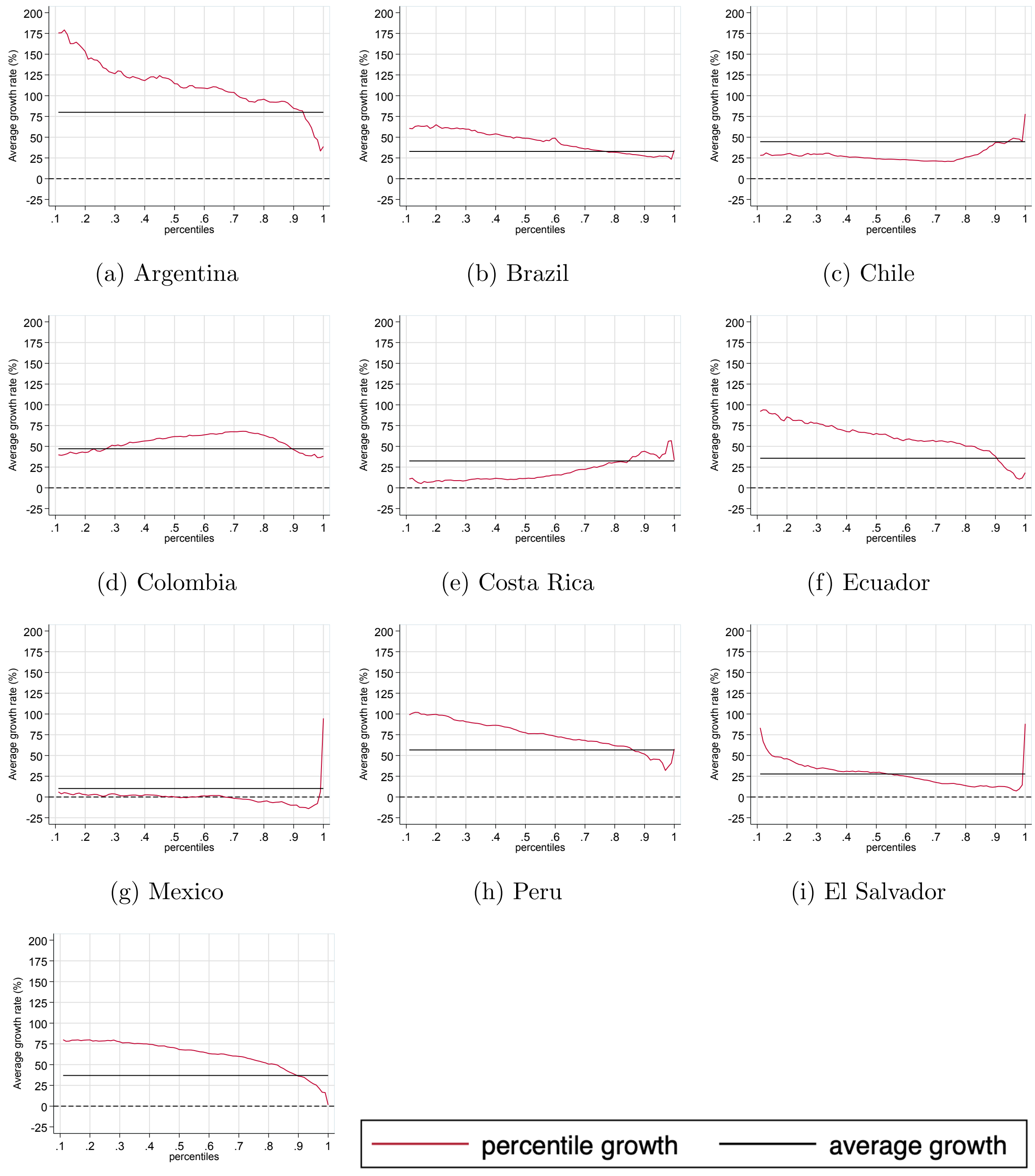
Even though the national income inequality series is not necessarily the benchmark, it does represent, by construction, the only series out of the four that may be used to analyse officially reported economic growth and inequality consistently. In particular, studying the evolution of inequality in surveys together with GDP, although informative, is miss-leading since they each refer to widely different and often divergent aggregates (Nolan et al., 2019; Alvaredo et al., 2022). This makes it quite difficult to directly answer important questions, such as who benefited from the commodity boom in region. By all accounts, the commodity boom that took place roughly between 2003 and 2013, which brought very favourable terms of trade and significant export-led growth (Ocampo, 2017), played a substantive role in influencing the direction of inequality in the region, at least according to inferred evidence from surveys (Cornia, 2014b; Sánchez-Ancochea, 2021). This is precisely the type of event that should be analysed under a micro-macro consistent framework.[15]

Figure 3.2 presents growth incidence curves of pre-tax national income for the commodity boom period (broadly 2003-2013). Upper incomes did not outperform lower incomes in all cases. In fact, only in Chile, Mexico and El Salvador did the top 1% outperform the average.[16] In most cases the next 9% (that is, the top 10-1%) experienced lower growth than the average. Overall, the commodity boom seems to have benefited lower income groups relatively more, except in Chile, Colombia, or Costa Rica, where we find either a neutral or regressive pattern. However, in Brazil, Chile and Colombia, the bottom 5% experienced substantially lower growth rates than the average. This can be explained by the fact that households at the very bottom of the distribution depend overwhelmingly more on social assistance transfers

[15] Income growth for the entire period (including the crisis of the early 2000s and the end of the commodity boom) are depicted in Figures C.6 and C.7 for pre-tax national income and C.8 for post-tax national income.

[16] This is consistent with the picture provided by Figure C.4. The top 1% income share will be stable across this period if its average income grew at the same rate as the average growth rate.

Fig. 3.2. Growth incidence curves during the commodity boom



(a) Argentina (b) Brazil (c) Chile (d) Colombia (e) Costa Rica (f) Ecuador (g) Mexico (h) Peru (i) El Salvador (j) Uruguay

Note. Authors' elaboration. Income is household per-capita pre-tax national income. Baseline year is 2003 for every country except Mexico, Costa Rica and Peru (2004), while the final year is 2013 for all except Mexico (2014). Growth rates represent average annual growth of incomes. They are caped at 3 to facilitate visual analysis (it only surpassed in rare cases at the very bottom, where incomes are extremely low and erratic and thus growth rates are artificially high).

(as we show in section 4), which are not included in pre-tax national income. Thus, one has to distinguish between primary effects (on market income) and secondary effects (on public transfers and spending) of the commodity boom. Interestingly, Argentina is the only case where the distribution of growth was unambiguously progressive, with lower incomes benefiting from higher growth rates than higher groups right across all percentiles.

4. Redistribution: taxation, transfers and spending

The redistributive effect of public policies has been extensively analysed in the region. In general terms, previous research has found that direct taxes and cash transfers have a very limited redistributive effect compared to richer countries (Hanni, Martner Fanta, and Podestá, 2015; Goñi, López, and Servén, 2011). Moreover, the overall redistributive effect is neutralised by consumption taxes, while only when social spending is considered does the re-distributive effect emerge (Lustig, Pessino, and Scott, 2014; Clifton, Díaz-Fuentes, and Revuelta, 2020). We revisit this analysis by considering not only the totality of national income —as opposed to the income reported in household surveys—, but also all national taxes and national social expenditures. Note that national taxes include corporate taxes, which are seldom brought into consideration and most likely affect higher income individuals, potentially increasing overall progressivity (see e.g. Saez and Zucman, 2020). This allows us to provide a supplementary view of redistribution in the region.

Figure C.9 depicts the composition of national taxes in the region since 1990. It suggests that there are diverse patterns in the region: while some countries have high and growing tax receipts as a percentage of GDP throughout the period (Argentina, Brazil and Uruguay), others have a share of around 20% and only slightly increasing or even stable (Colombia, Chile and Peru). Countries such as Mexico have low and stable shares (except for the end of the period), while the rest (e.g. El Salvador and most notably Ecuador) present large increases from low starting points. For the region as a whole, consumption and production taxes make up more than half of the total take, a trend replicated in most countries. Personal income taxes represent a comparatively small share, as expected from the size of the taxable income and the effective rates levied (as shown in Figure B.1). Social security contributions (SSC) vary considerably more by country, with the most important shares in countries where the overall tax take is highest, like Argentina, Brazil, Costa Rica, and Uruguay. These countries are those where pensions represent a higher share of total income (see Figure A.2). Property and corporate income taxes represent about a quarter of total taxes, with corporate income taxes representing the bulk of these receipts.

The distributive effect of these trends depends on both the level of taxes as well as their incidence throughout the income distribution. Figure 4.1 shows effective incidence rates by type of tax, as well as monetary benefits, across the distribution of total pre-tax national income. Among progressive taxes (those whose effective tax rate increases with percentiles), personal income taxation is broadly redistributive in every country except for Peru.[17] This progressive profile is largely because such a small share of the population have positive effective tax rates (as shown in Figure B.2, where only at the very summit of the distribution do effective rates fall). All types of wealth/property taxation and corporate income taxes are also progressive, while taxes on goods and services as well as the residual category "other taxes" are clearly regressive. Given the larger share of the latter two, the overall result is a regressive pattern in the region, which has a steeper gradient in countries such as Argentina, Chile, Mexico or Costa Rica, and more neutral in Colombia, Uruguay or Ecuador. In all countries, corporate income taxation plays a key role in taxing incomes at the top of the distribution, given that its incidence falls on corporate owners (employers and shareholders). Monetary benefits have a clearly progressive profile across all countries, with a higher amount of transfers below the median of the distribution.

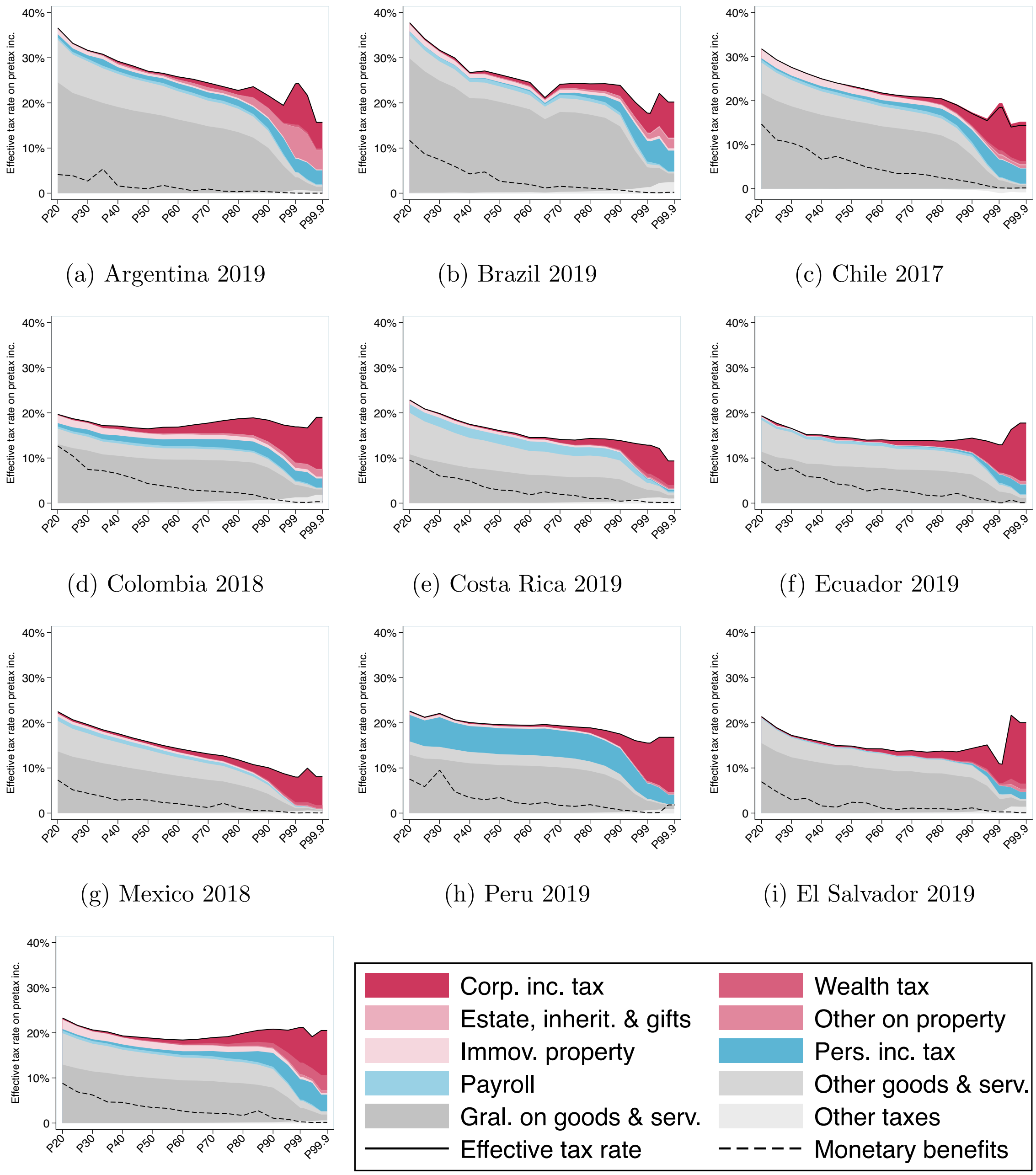
The effect of these taxes and transfers on the income distribution is presented in Figure 4.2. It depicts the post-tax disposable income distribution, which is the result of applying all taxes and monetary transfers of Figure 4.1 to the the pre-tax national income (from Figure 3.1, which is plotted again for comparative purposes). The net effect of taxes and transfers is in general terms slightly regressive, or neutral in the best case scenarios (e.g. Colombia after 2010 or Uruguay after 2009). Most of the regressiveness is given by value-added taxes: when removed, the post-tax spendable income distribution results in a significantly lower inequality throughout the region. The redistributive effect of the remaining taxes and transfers is mild, and close to negligible in countries such as Mexico, Colombia or Costa Rica. More importantly, these taxes and transfers do not seem to be powerful drivers of reducing inequality, since trends do not visibly change, except in Brazil around 2004-2005 or Uruguay in 2007. Thus, changes to the income distribution are substantially driven by pre-tax incomes, stressing the importance of pre-tax inequality as documented for France and the United States (Piketty, Bozio, Garbinti, Goupille-Lebret, and Guillot, 2020).

When social spending in-kind is incorporated, particularly the two categories that affect the distribution —health and education—, trends *do* change.[18] The falling inequality narrative

[17] This outlier could be due to the fact that the personal income tax statistics sent to us by the Peruvian tax office excludes income from foreign sources as well as entrepreneurial incomes.

[18] As stated before, all other social expenditures in kind are imputed proportionally to the disposable income distribution.

Fig. 4.1. Incidence of taxes and transfers



(a) Argentina 2019

(b) Brazil 2019

(c) Chile 2017

(d) Colombia 2018

(e) Costa Rica 2019

(f) Ecuador 2019

(g) Mexico 2018

(h) Peru 2019

(i) El Salvador 2019

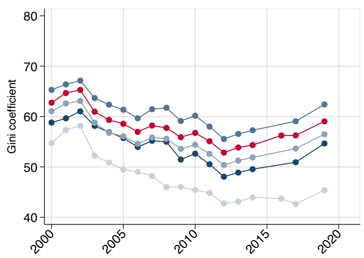
(j) Uruguay 2019

Note. Authors' elaboration. The Pre-tax per capita household income.

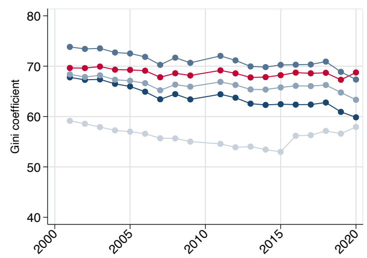
re-emerges even for countries where the sequential process of adjusting and scaling the raw survey results in stable pre-tax inequality trends. The clearest exception is Mexico, for which inequality continues to rise even after in-kind transfers are accounted for. This is because health and education spending in Mexico has remained pretty stagnant over the last twenty years (Figure C.10), despite progressive (or slightly progressive) spending profiles estimated for both categories by the CEQ studies (see Figures C.11 and C.12). For all other countries the mix of growing health and education expenditures and progressive incidence suffices to produce falling inequality across the board.

At this point it is worth recalling that the literature on income inequality in Latin America seldom considers in-kind social spending, due to debatable assumptions about how to impute these expenditures to households. Thus, the conventional narrative is largely built on a disposable income definition, which unlike the national accounts definition does not include consumption taxes. As specified in section 2.3, we construct a post-tax spendable income distribution to compare with the common definition behind the conventional narrative, a survey-based definition of income which we label "postax raw" in Figure 4.2. The comparison of these two series is consistent, allowing us to scrutinise the conventional narrative of falling inequality in the region after accounting for missing top and household incomes. In at least three countries (Argentina, Chile, Mexico) the downward post-tax raw trend is not replicated in the post-tax spendable series. In all other countries the two series track each other pretty well, suggesting that the conventional narrative holds up to scrutiny on its own terms. However, it is worth reiterating that its definition of income does not fully account for the entirety of the tax and transfer system, which on cash terms produces regressive disposable income profiles, due to the weight and regressiveness of consumption taxes, as Figures 4.1 and 4.2 reveal.

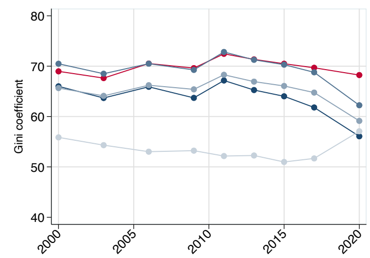
Fig. 4.2. Gini coefficients: pretax vs post-tax series



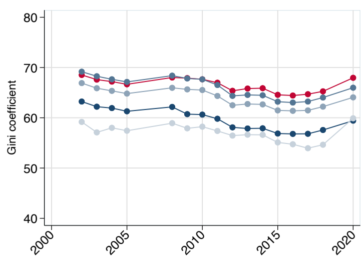
(a) Argentina



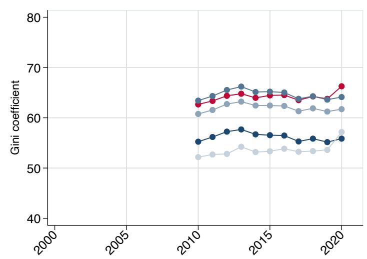
(b) Brazil



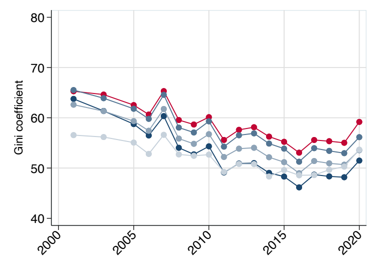
(c) Chile



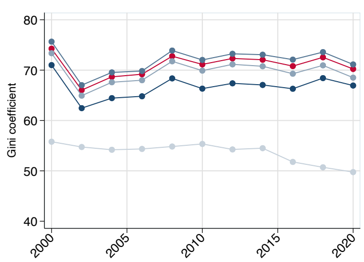
(d) Colombia



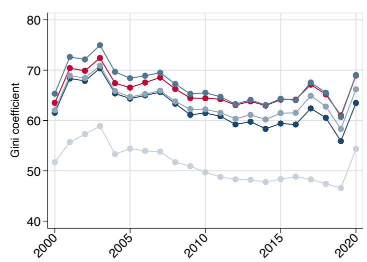
(e) Costa Rica



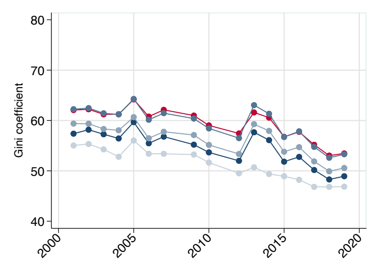
(f) Ecuador



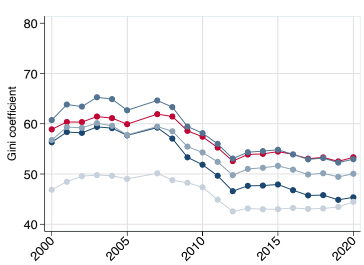
(g) Mexico



(h) Peru



(i) El Salvador



(j) Uruguay

Note. Authors' elaboration. The figures depict the pretax national income distribution and four post-tax distributions: the raw survey series (after taxes and cash transfers as reported in surveys), the spendable series (the surveys combined with administrative data and national accounts, after taxes and cash transfers except consumption taxes), the disposable series (after all taxes and cash transfers), and the national series (after all taxes, cash transfers, and in-kind spending). The distributions are of household per capita income.

5. Reconciling competing narratives

5.1. The conventional narrative and its limits

Research based on household surveys has consistently shown a downward trend in per-capita household income inequality in Latin America between 2000 and 2015, fostered by the improvements in international economic conditions, terms of trade, and a new social policy model in most of the region (Gasparini et al., 2018). These estimates, with relatively minor variations, represent the core of the series shown by all major inequality databases of the region, i.e. the World Bank, SEDLAC, CEQ and ECLAC.[19] Despite its multiple causes, it is wage inequality which has been found to be the main driver of falling inequality in the region (López-Calva and Lustig, 2010; Messina and Silva, 2017).

Whereas the rise in inequality in the 1980s and early 1990s is typically explained by skill-biased technological change (after the liberalisation of international trade flows), the decline is explained by demographic factors and, more importantly, by the reduction in labour income inequality. For the latter, the educational upgrading of the labour force played a major role. Cornia (2014b) documents that the average regional decline in the Gini index was 5.5 points from 2002 to 2010, after two decades of systematic increases. After noting that conventional data sources are not able to properly account for capital incomes or labour incomes of the "working rich", he shows that the evolution in 1990-2010 was driven by wage income inequality, matched by skill premium shifts benefiting the bottom of the distribution. The increase in social assistance also played a role, but its contribution was relatively less important than changes to the labour income inequality. Rodríguez-Castelán, López-Calva, Lustig, and Valderrama (2022) find that the decline in wage inequality was driven by an increase in real hourly earnings among the bottom of the distribution, which in turn was associated to a fall in education and experience premiums, as well as to a reduction in the wage dispersion among workers with the same observable attributes. Amarante (2016) argues based on factor component analysis that it was mainly informal wages which pushed inequality downwards (while the opposite happened with the formal sector).

Tax data have seldom been integrated into the picture nor have findings from this literature been reconciled with the "conventional wisdom" in a systematic way. Where it has been attempted, the conclusion reached is that the conventional wisdom regarding inequality trends

[19] See Bourguignon (2015) for a systematic comparison of SEDLAC and ECLAC's data. See also `https://data.worldbank.org/`, `https://www.cedlas.econo.unlp.edu.ar/wp/estadisticas/sedlac/`, `https://commitmentoequity.org/datacenter` and `https://statistics.cepal.org/portal/cepalstat/index.html?lang=es`, respectively.

remains solid. For example, De la Torre, Yeyati, Beylis, Didier, Castelan, and Schmukler (2014, p.35) ask "does the pro-poor growth story still hold once we incorporate the missing top earners to the distribution?" Complementing survey data with information on top earners from tax data and comparing with survey-based results for Argentina (1998-2003) and Colombia (2002-2010), the authors indeed find that inequality levels are corrected upwards. However, they also find that inequality dynamics are more prone to diverge between both scenarios during times of economic crisis than during smooth business-cycle periods. They nonetheless conclude that while "extending this exercise to the rest of the region could shed more light on the determinants of income distribution over time, we feel confident that the trends in income inequality unveiled by the household survey data are a good approximation to the real Gini for much of LAC" (De la Torre et al., 2014, p. 36). This conclusion is supported by Winkelried and Escobar (2020), who reveal for the case of Peru that a range of simulated adjustments to the top tail of the survey—using Pareto models and top income shares from other countries—still produce declining Gini coefficients. Burdín et al. (2022) study pre-tax adults inequality for the Uruguayan case between 2009 and 2016. They find that synthetic inequality indices fall according to survey data and administrative data (the latter supplemented by survey data for the unaccounted population and incomes), but find divergent trends for top income shares. While the top 1% share decreases in the survey, it remains stable first and then grows based on administrative data. This divergence is the result of increasing inequality in the right tail of the distribution of the administrative data, driven in turn by an increasing share of reported dividends.

While Amarante and Jiménez (2015) do not extrapolate lessons from a small sample of countries to the whole region, they recognize the similar evolution of the standard Gini and tax-adjusted Gini for the countries with survey and administrative data at their disposal (Argentina, Colombia, Uruguay). Even with the acknowledged problems of tax data (evasion, avoidance, exemptions, threshold changes), the authors think that tax data can add value to the study of inequality in the region, particularly from the perspective of top income concentration.

Recent top incomes literature based on tax data has shown a more persistent pattern of inequality —particularly from the perspective of top income concentration— in this period (Alvaredo, 2010; Alvaredo and Londoño Velez, 2014; Morgan, 2018; Souza, 2018; Burdín et al., 2022; Flores et al., 2020).[20] Making sense of divergent trends is not straightforward, since there are differences in units (adults, households), income definitions (pre or post-tax) and more importantly, the differences in the way top income groups are accounted for in different

[20] For more details see Appendix A.2.

sources and the coverage of capital incomes, not only in surveys but also in top-corrected surveys compared with the national accounts (Alvaredo et al., 2022). We turn to these issues in the following section.

5.2. Making sense of divergent trends

Do estimates discussed in sections 3 and 4 of this paper, alongside the tax-based literature contradict the conventional inequality narrative for Latin America? Addressing this question is not straightforward, since competing narratives are riddled with comparability issues that need to be cleared beforehand. Thus, in order to reconcile different inequality narratives that emerge from alternative sets of estimates, we need to distinguish between conceptual and measured sources of divergence. Among the former, it is necessary to clarify: (i) the unit of analysis; (ii) the inequality indicator; and (iii) the income definition, since differences in one or more of these may render estimates incomparable from a purely conceptual point of view. Among the latter, we analyse the role of the two main sources if divergence in each estimation step, i.e. capital incomes and top income groups.

Unit of analysis. Most estimates for Latin America (e.g. World Bank official estimates) are based on per-capita household income, while tax-based studies usually use adults or equal-split adults (where the income of couples is equally split among the individuals) (WIL, 2020). Thus, to avoid comparability problems in this area throughout this study we use per-capita household income as the reference unit of analysis. Note that this is in turn only possible because we depart from surveys and adjust them using administrative data not the other way around as is the case in most studies on developed and underdeveloped countries (Piketty et al., 2018; Garbinti et al., 2018; Burdín et al., 2022; Flores, 2021).[21] Nevertheless, in order to clarify the effect of this issue on trends, in Figure A.1 we show the evolution of the Gini coefficient in the World Bank's database and in the surveys we use from ECLAC's database according to alternative unit definitions.[22] To dismiss differences in harmonisation, we reproduce the World Bank's downward trend in the inequality of per-capita household income based on ECLAC's surveys. Individual-adult and equal-split-adult inequality series are also depicted, showing that trends are not altered, except for Mexico and Chile to a lesser degree. Thus, at least for these two countries one should expect that considering individuals

[21]All series based on alternative unit definitions are available upon request and will be available in a dedicated website that will be made public in November 2022.

[22]Note that the World Bank database on Latin American household surveys is the Socio-Economic Database for Latin America and the Caribbean (SEDLAS), which is produced in collaboration with the CEDLAS institute (*Centro de Estudios Distributivos, Laborales y Sociales*) of the Universidad Nacional de la Plata. As with ECLAC's database, it is based on official household surveys run by country statistical offices or central banks.

instead of per-capita household income should mechanically change trends, even before even changing the data source or incomes considered, while this does not seem to be an issue for the remaining countries.

Inequality indicator. As Figure C.5 shows, even in cases where inequality does fall according to indicators like the Gini index or the top 10% share, this can coincide with stability or even an increase in top 1% income shares. This is not surprising, and has been found for countries such as Brazil, Colombia or Uruguay (Morgan, 2018; Souza, 2018; Alvaredo and Londoño Velez, 2014; Burdín et al., 2022). This should be kept in mind when comparing competing narratives. To what extent one should prioritise one approach or the other depends on normative considerations, such as privileging individuals at the bottom of the distribution in the Rawlsian sense, or favouring a limitarianist approach (see for example Robeyns, 2019). Mechanically, it is to be expected that the income dynamics of a very small group in the population like the top 1%, will not necessarily impact a synthetic indicator like the Gini coefficient in the same direction, when these incomes are included into the distribution. This is largely because the Gini coefficient, still the most widely used summary measure of inequality, weights all income groups equally, and by its construction from the Lorenz curve is more sensitive to changes in the middle of the distribution than its tails. These points have been recognised by the literature (see for example Leigh, 2006, and Atkinson, 2007). But as shown by Alvaredo (2011), if top incomes not covered by surveys experience a large enough increase relative to lower incomes, then trends in the Gini coefficient and top income shares can diverge. What we observe, therefore, is that increases in top incomes from administrative data are not large enough relative to the growth of lower incomes to reverse the downward tendency of the Gini coefficient. What arguably does make more of a difference are dynamics in micro-macro income gaps as we've shown.

Income definition. There are two dimensions to consider regarding the definition of income. The first one is that the conventional narrative is based on household surveys which generally report disposable incomes, that is after tax incomes and including social transfers with the exceptions of Brazil and Costa Rica, where gross wages are reported. On the other hand, tax-based studies rely on pre-tax income excluding social transfers, except where they are taxable (like pensions). Naturally, the redistributive effect of taxes and transfers changes income distribution as shown in section 4 and hence hinders comparability. In Figure 4.2, we depict the Gini index based on post-tax definitions. We include a series for post-tax disposable income in the raw survey, without being subject to any adjustments for top incomes or macro incomes. Together with our post-tax spendable series (post-tax disposable income without subtracting value-added taxes (VAT)), they both represent income inequality after taxes and transfers and before VAT. Thus, both are comparable and importantly both show downward trends. The post-tax spendable income series shows higher inequality than the post-tax national income series since it does not consider heath and education spending (but lower than the post-tax disposable, as a result of ignoring the regressive effect of VAT). In most cases, with the notable exception of Mexico, significant periods of downward inequality trends are observed in the series, mirroring what happens with the raw survey.

The second dimension is the aggregate income concept the series refer to, which was discussed in sections 3 and 4. Expanding the income reported in the raw surveys to include missing top incomes and especially absent aggregate capital incomes softened the downward inequality trends in most countries, and was enough to stabilise or reverse the trend in at least three large countries of our sample. The crux of the debate thus lies in the contributions of the top 1% and of capital incomes to the narrative. This is what we turn to in the remainder of this section.

Bottom 99% vs top 1%. The contribution of top incomes to overall inequality is the result of their distance from the rest of the distribution and to the distance between themselves. These between-group and within-group dimensions can be decomposed in the Theil index of inequality. Figure C.13 shows the contribution of the inequality between the bottom 99% and the top 1% for three pre-tax distributions and the post-tax spendable distribution. With each step of the adjustment the top 1% and the rest grow apart and hence between group inequality contributes more to overall inequality. Note that in the pre-tax national income series, it explains 40-50% of total inequality, while it was around 30% or less in the raw survey for most countries (results are the same with for post-tax spendable series). Moreover, in some cases such as Mexico, and to a lesser degree Chile, the contribution of the distance between groups increases in time for the national income series relative to other series (or as in the case of Peru it flattens while remaining series are falling).

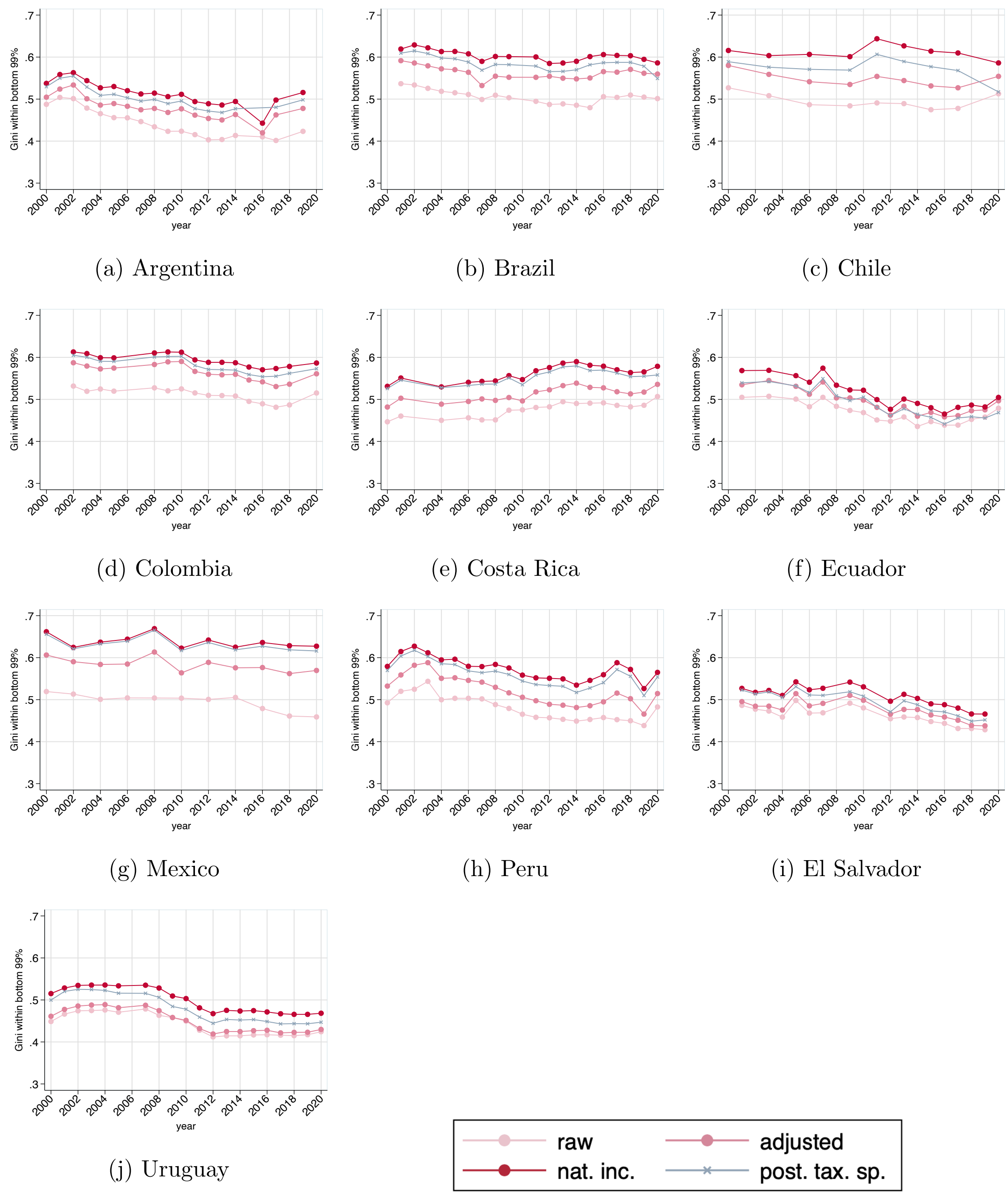
Overall within-group inequality is the sum of inequality within the top 1% and with the rest of the distribution. Figure C.14 shows that not only is the top 1% more distant from the rest after each adjustment, but also that is more internally unequal, fostering overall inequality. Moreover, in most countries across a significant portion of the period there is an upward trend in the national income series, which means that the adjustments push inequality up through time via inequality within the top 1%. Thus, each step of the sequence from raw survey to national income increases the distance between groups and inequality within the top 1%, and in some cases this gap widens over time, contributing to diverging inequality trends between the raw survey series and the augmented series. Within-group inequality among the bottom 99% (Figure 5.1), on the other hand, shows a decreasing trend for most countries, even in the pre-tax national income series. The only clear exception to this pattern is Costa Rica, for which inequality among the bottom 99% increases regardless of the series.

Capital incomes vs wages. To assess the effect of capital incomes on divergent trends, we first look at the distribution of each of the income sources for each aggregate definition of income, depicted in Figures C.15 to C.18. The first thing to note is that the Gini index of of wages decreases for most countries between the early 2000s and the mid 2010s. In fact, as shown in Figure C.19, the trend in wage inequality after adjusting the survey using administrative and macroeconomic data, respectively, mirrors the raw survey. This is an important result given that wage inequality is one of the driving forces behind the falling inequality narrative, and it remains in most countries after incomes have been adjusted using administrative data on top incomes and macro data for different income sources.

Secondly, capital incomes are as expected extremely concentrated, which is already a feature of raw surveys. To better understand the role capital incomes, Figure 5.2 depicts their contribution to overall inequality based on Lerman and Yitzhaki (1985)'s Gini decomposition, which is the result of within source inequality and their share in total incomes. As can be seen it is the scaling up to household sector and national income that significantly increases the contribution of capital income to overall inequality, increasing by a factor of 3 in many cases.

To sum up, competing narratives are sometimes affected by comparison problems affecting the unit of analysis, the income definition or the choice of inequality indicators. In some cases these can lead to divergent results even if the same data sources were used. However, a reconciliation of various micro and macro data sources on income can produce diverging trends relative to raw surveys when the contribution of ignored top income groups and aggregate capital incomes to overall inequality increase over time, even in the presence of the decreasing concentration of wages. Thus, our results confirm the conventional narrative of falling Latin American inequality within the bottom 99% of the post-tax income distribution and especially related to earnings, but they also suggest that these trends change for some countries once top income groups and capital incomes are better accounted for. The "debate" among the research community over Latin American inequality largely boils down to one about trust in micro and macro data sources in region where all suffer from glaring imperfections.

Fig. 5.1. Within-group Gini coefficient (bottom 99%)



(a) Argentina

(b) Brazil

(c) Chile

(d) Colombia

(e) Costa Rica

(f) Ecuador

(g) Mexico

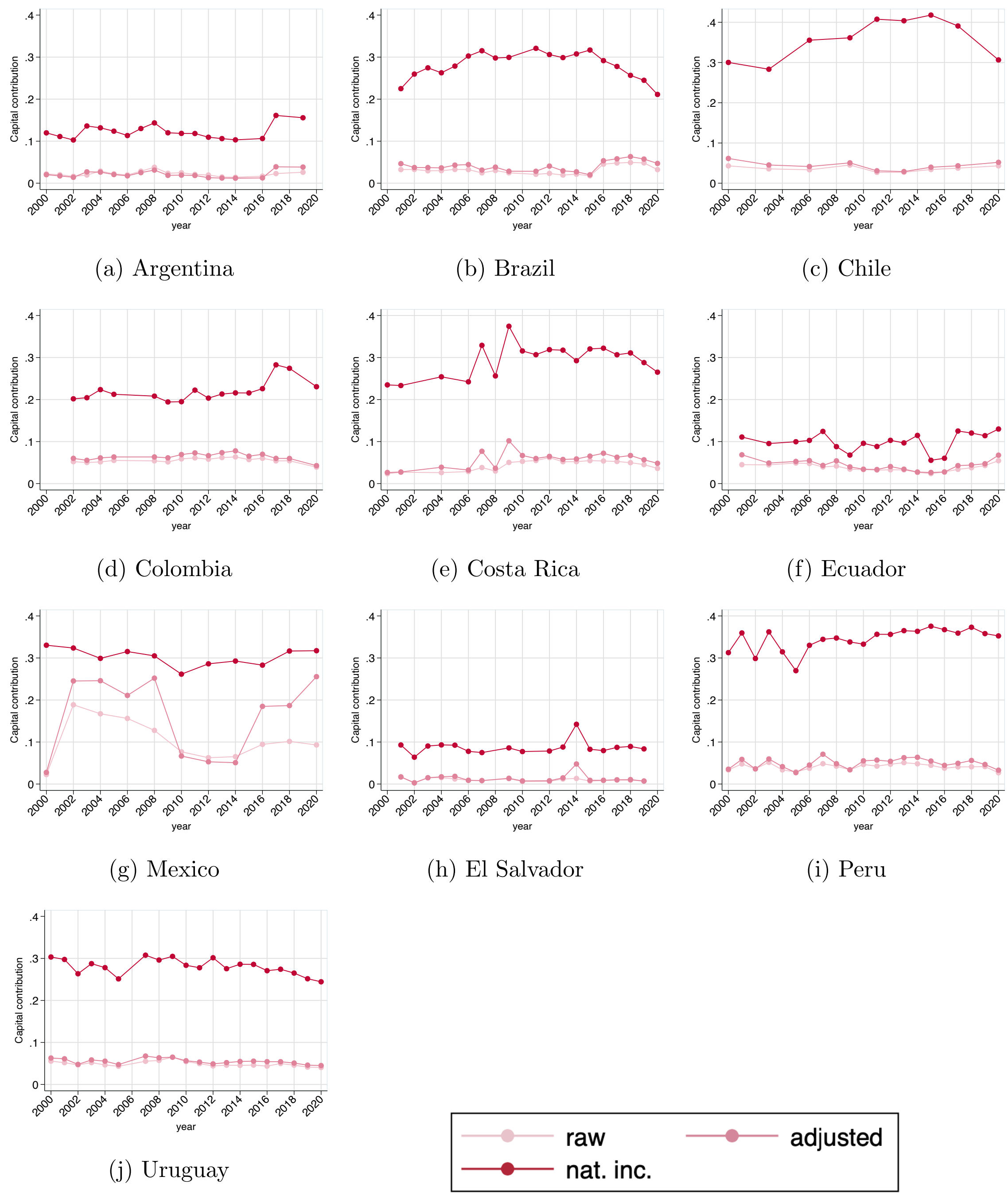
(h) Peru

(i) El Salvador

(j) Uruguay

Note. Authors' elaboration. Gini coefficient of per-capita household income for the bottom 99% in the raw pre-tax survey, the adjusted pre-tax survey with tax data, the pre-tax national income series and the post-tax spendable income series (i.e. disposable income without excluding value-added taxes).

Fig. 5.2. Capital income contribution to inequality



(a) Argentina (b) Brazil (c) Chile

(d) Colombia (e) Costa Rica (f) Ecuador

(g) Mexico (h) El Salvador (i) Peru

(j) Uruguay

Note. Authors' elaboration. Capital income contribution to the Gini coefficient of per-capita household income, based on Lerman and Yitzhaki (1985) in the raw pre-tax survey, the adjusted pre-tax survey with tax data, and the pre-tax national income series.

6. Concluding remarks

Trust in data sources is at the heart of the dilemma we pose in this paper when revisiting the Latin America inequality story. If it is accepted that the region is as rich as macroeconomic data report, then it is also significantly more unequal. A rejection of this conclusion implies accepting that Latin America grew less rich, but remained less unequal throughout. The former outcome may be easier to digest if it were just about levels. What the debate is really about, however, is trends, and here we showed that the region is more inequality-heterogeneous than previously understood. In at least three countries of our sample of ten countries (Brazil, Chile and Mexico) inequality trends during the high-growth years (2003-2013) change after the survey's reported income is augmented to include ignored top incomes from administrative data and macroeconomic incomes of the household sector and total economy from the national accounts. This holds even for the same income concepts and units of analysis commonly used in the literature. In all cases the declining inequality trend of the high growth years softens with each of the adjustments made to the raw survey. Moreover, during the low-growth years at the end of our period of analysis (post-2015), inequality has increased faster in the augmented series than in the raw series.

Was Latin America exceptional after all? It turns out to be a matter of degree. Taken at face value, our results suggest that the region's exceptionalism is no longer uniformly shared across all countries. Broadly speaking, we showed that while inequality did fall for the bottom 99% and for wages across the region, this is not the case for every country once top income groups and capital incomes from extra-survey sources are accounted for. Even if only a part of this were true, on account of the many weaknesses of both the region's administrative data and national accounts, it does reveal certain limits of the Latin America's redistributive experience of the early twenty first century. While it was widely successful in increasing the incomes of the poor and reducing overall inequality, it was relatively unsuccessful in redistributing income from the rich and from capital in particular. Interestingly, we find that the falling inequality narrative emerges with most strength once in-kind social spending is considered, which highlights an important feature of the redistributive process that deserves greater attention in future research.

It is worth stressing once again that this exercise relies on imperfect and heterogeneous data alongside numerous necessary assumptions to bridge them all together. However, it is also true that it represents a unique attempt to make use of such a wide array of data sources in a coherent manner to provide conceptually consistent inequality estimates. Moreover, we see it as an effort to build a bridge between different inequality approaches and narratives. In this sense, this work should be regarded as a contribution to open a debate on an important topic and not to close it.

Following the path laid out by Alvaredo et al. (2022), the large gap between the micro distribution and macro distribution of household incomes we estimate shows that the seminal findings by Altimir (1987) are still essentially true. The credibility of the scaling of survey incomes to the national accounts obviously depends on our confidence in macroeconomic statistics, as well as the way in which we view incomes that households do not directly receive on an annual basis. From our perspective, regardless of the accounting convention on whether to allocate corporate retained earnings to firms or to their owners, it is evident that they are resources controlled by individuals and they should be accounted for in any meaningful inequality analysis, if only to avoid cross-country biases affecting the distribution of profits.

Naturally, the above conclusions are highly dependent on the particular assumptions made. Considering that surveys miss about half of national income, we are perfectly aware that many other distributions can theoretically be estimated with a different set of assumptions. Yet, we find it difficult to plausibly settle on alternative assumptions given the data at our disposal. Having said this, as we pointed out in the introduction, our procedure should not be taken as a gold standard going forward. Further research is still needed at the country-level —exploiting the rich country data lost in our generalised approach and the local knowledge of data producers and researchers— to provide greater clarity on data gaps and their implications for inequality analysis. A host of public policies lie in the balance of such an approach, especially if policymakers wish to adequately tailor them to the distribution of actually measured economic growth.

A. Appendix: Data

We rely on four main data sources: households surveys, income tax records, social security records, and national accounts. Table A.1 schematically presents the data sources for countries included in this study, together with the years covered by each source, while table 2 displays data availability for countries that remain excluded from our analysis. The following subsections elaborate on the databases used.

A.1. Households surveys

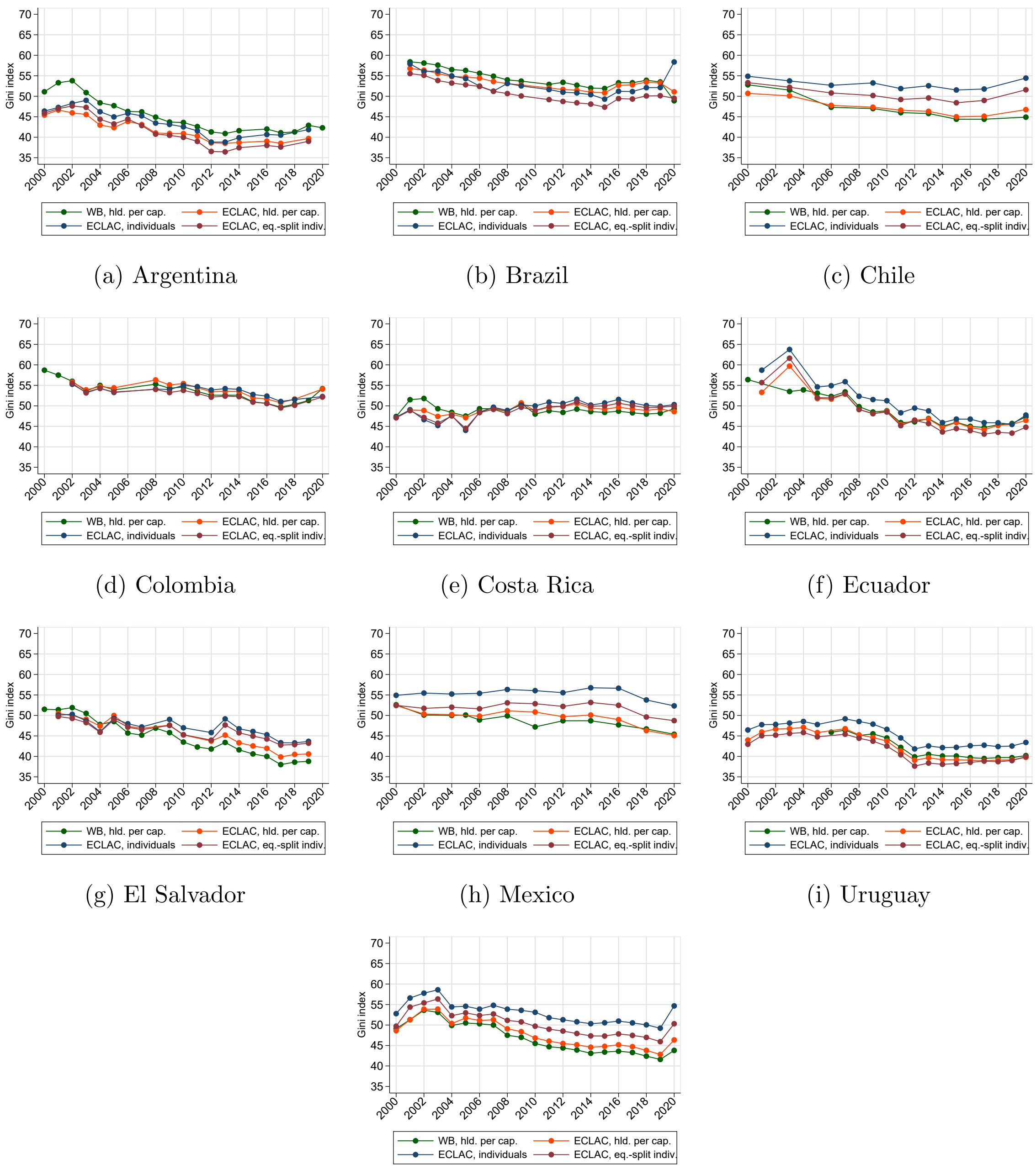
Household surveys provided by ECLAC represent one of the key data inputs for this study. More broadly, national surveys are an extremely important reference point in their own right in Latin America, since they are the only source available in almost all the countries. Official statistics on inequality, poverty, unemployment, etc., are drawn from them. Based on ECLAC data, we are able to reproduce country level inequality estimates by the World Bank (WB), as depicted in Figure A.1. This points to the fact that, even if the two harmonization processes (ECLAC-WB) are independent, they produce very similar results in terms of income distribution.[23]

Table A.2 makes it explicit that many of the countries that remain excluded, mostly from Central America and the Caribbean, either do not report distributive data at all (Belize, Cuba, Haiti, Jamaica, Suriname and Trinidad and Tobago), do not run household surveys on a regular basis (Bahamas, Nicaragua, Venezuela), or only run surveys but do not have any kind of publicly accessible administrative data (Bolivia, Dominican Republic, Honduras, Panama and Paraguay).

Figure A.2 shows the decomposition of income in surveys, before any adjustment, in terms of wages, pensions, capital income, self-employment income, and imputed rents. Wages and self-employment income represent 60-90% of total household incomes, while capital incomes are much lower.

[23]El Salvador up to 2010 is the clearest exception, since World Bank estimates are considerably higher and falling very rapidly. The surprisingly large inequality decrease of over 10 points in the Gini index, casts doubts on this trend, while the one resulting from ECLAC's harmonized surveys seems more reasonable. Both the Wold Bank and ECLAC use the same underlying household surveys prior to harmonization. The World Bank database is the Socio-Economic Database for Latin America and the Caribbean (SEDLAC), produced in cooperation with the Center for Distributive, Labor and Social Studies (CEDLAS) of the Universidad Nacional de La Plata in Argentina.

Fig. A.1. Survey-based Gini indexes by source and income definition



(a) Argentina (b) Brazil (c) Chile

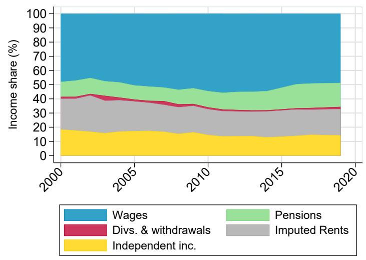
(d) Colombia (e) Costa Rica (f) Ecuador

(g) El Salvador (h) Mexico (i) Uruguay

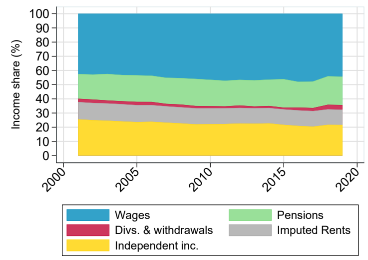
(j) Peru

Note. Authors' elaboration based on World Bank data (https://data.worldbank.org/) and ECLAC's harmonized surveys. World Bank (WB) and ECLAC's household per capita income series ("hld. per cap.") show identical trends and very similar levels. Personal income Gini indices for adult population (20 and more years) based on ECLAC's harmonized surveys are also depicted along two dimensions – individual earners and equal-split individuals (where the total income of couples is divided by two).

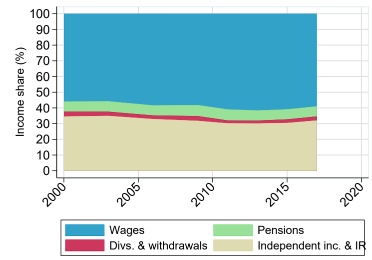
Fig. A.2. Income composition - raw surveys



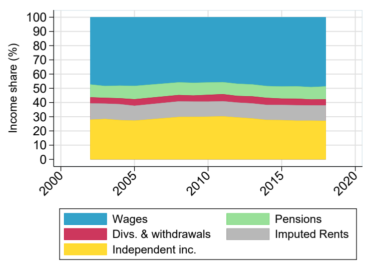
(a) Argentina



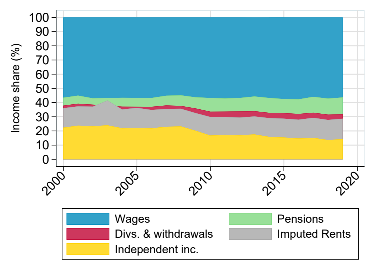
(b) Brazil



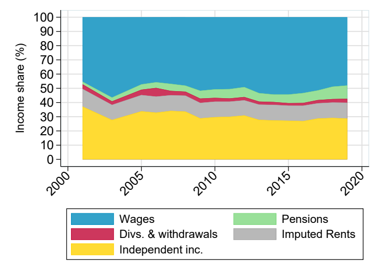
(c) Chile



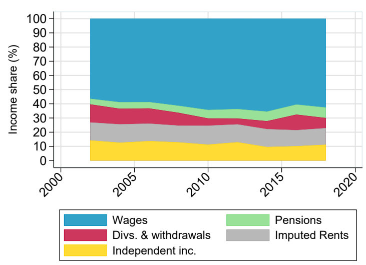
(d) Colombia



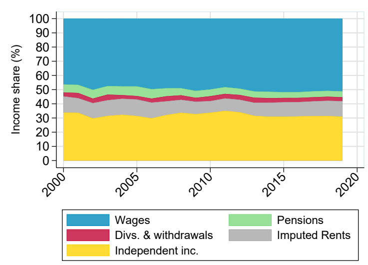
(e) Costa Rica



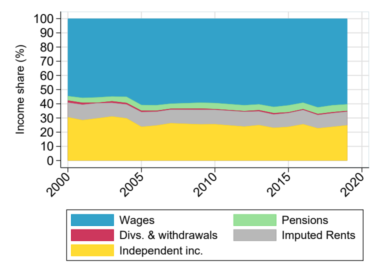
(f) Ecuador



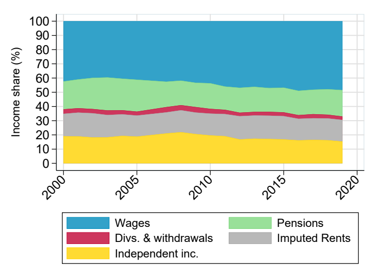
(g) Mexico



(h) Peru



(i) El Salvador



(j) Uruguay

Note. Authors' elaboration based on ECLAC's harmonized surveys. Income is pretax, net of pension contributions.

Table A.1: Updated countries

Country	Survey microdata: Source	Survey microdata: Availability	Administrative data: Source	Administrative data: Availability	Population (% of total)	
Argentina	Encuesta Permanente de Hogares (EPH) and EPH-Continua from 2003, Insituto Nacional de Estadística y Censos (INDEC)	2000-2014, 2016-2019	Income tax tabulations, Administración Federal de Ingresos Públicos (AFIP), Employee microdata, Ministerio de Trabajo, Empleo y Seguridad Social	2000-2017, 2000-2015	2% 40%	Survey is of urban Income t Employe sector wa
Brazil	Pesquisa Nacional por Amostra de Domicílios (PNAD), Instituto Brasileiro de Geografia e Estatística (IBGE)	2001-2009, 2011-2019	Income tax tabulations, Receita Federal (RFB)	2000, 2002, 2006, 2007-2019	14%	Income t
Chile	Encuesta de Caracterización Socioeconómica Nacional (CASEN), Ministerio de Desarrollo Social	2000-2009 (triannual), 2011-2017 (biannual)	Income tax tabulations, Servicio de Impuestos Internos (SII)	2000-2018	70%	Wages re fiscal inc
Colombia	Encuesta continua de hogares (Gran Encuesta Integrada de Hogares from 2008), Departamento Administrativo Nacional de Estadística (DANE)	2002-2005, 2008-2018	Alvaredo and Londoño-Vélez (2013)	2000-2010	1%	Income t
Costa Rica	Encuesta Nacional de Hogares, Instituto Nacional de Estadística y Censos (INEC)	2000-2019	Wage income, Non-wage income Zuñiga-Cordero (2018)	2000-2017 2010-2016	28% 5%	Wage ea Indepenc
Ecuador	Encuesta Periódica de Empleo y Desempleo (EPED) and Encuesta de Empleo, Desempleo y Subempleo (ENEMDU) from 2003, Insituto Nacional de Estadística y Censo (INEC)	2001, 2003 2005-2019	Cano (2015) Rossignolo et al. (2016)	2008-2011 2012-2014	14% 38%	Distribut only avai
El Salvador	Encuesta de Hogares de Propósitos Múltiples, Dirección General de Estadística y Censos (DIGESTYC)	2000-2007, 2009, 2010, 2012-2019	Tax tabulations (wages), Tax tabulations (diverse income) Dirección General de Impuestos Internos (DGII)	2000-2018	4% (wages) 4% (diverse)	Wages o separatel
Mexico	Encuesta Nacional de Ingresos y Gastos de los Hogares, Instituto Nacional de Estadística, Geografía e Informática (INEGI)	2002-2018 (biannual)	Income tax microdata, Servicio de Administración Tribuataria (SAT)	2009-2014	20% (wages) 2% (diverse)	Wages o separatel
Peru	Encuesta Nacional de Hogares - Condiciones de Vida y Pobreza, Instituto Nacional de Estadística e Informática (INEI)	2000-2019	Income tax tabulations, Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT)	2016-2018	25%	Income t
Uruguay	Encuesta Continua de hogares (ECH), Instituto Nacional de Estadística (INE)	2000-2005, 2007-2019	Income tax microdata, Dirección General Impositiva	2009-2016	75%	Income t

Note. Authors' elaboration.

Table A.2: Excluded countries

Country	Survey microdata: Source	Sample size, thousands of individuals	Availability
Bahamas	Bahamas Living Conditions Survey	6	2001
Belize	-	-	-
Bolivia	Encuesta de Empleo,Desempleo y Subempleo, Insituto Nacional de Estadística y Censo (INE)	15 – 40	2000-2019
Cuba	-	-	-
Dominican Republic	Encuesta Nacional de Fuerza de Trabajo (ENFT)	15 – 30	2000-2019
Guatemala	Encuesta Nacional de Condiciones de Vida and Encuesta Nacional de Empleo e Ingresos	10 – 70	2000, 2002-2004, 2006, 2011, 2014
Guyana	-	-	-
Haiti	-	-	-
Honduras	Encuesta Permanente de Hogares de Propósitos Múltiples (EPHPM), Institutio Nacional de Estadisticas (INE)	20 – 100	2001-2018
Jamaica	-	-	-
Nicaragua	Encuesta Nacional de Hogares sobre Medición de Nivel de Vida, Instituto Nacional de EStadística y Censos de Nicaragua	20 – 35	2001, 2005, 2009, 2014
Panama	Encuesta de Hogares, Instituto Nacional de Estadística y Censo (INEC)	40 – 55	2000-2019
Paraguay	Encuesta Integrada de Hogares (EIH) and Encuesta Permanente de Hogares (EPH) from 2002, Dirección General de Estadística, Encuestas y Censos (DGEEC)	15 – 40	2001-2019
Suriname	-	-	-
Trinidad and Tobago	-	-	-
Venezuela	Encuesta de Hogares Por Muestreo (EHM), Oficina Central de Estadística e Informática	80 – 240	2000-2006

Note. Authors' elaboration.

A.2. Literature on top incomes using administrative data

Argentina. Alvaredo (2010), covering the period 1932-2004, is the seminal reference on the topic, with no precedent to our knowledge. This line of work was recently picked up again by Jiménez and Rossignolo (2019), who similarly use tax registries alongside updated national accounts statistics, for the period 2004-2015. The latter emphasize certain caveats regarding the use of statistical information, which they deem to be "scarce, incomplete, inconsistent or still nonexistent."

Brazil. Mortara (1949) was the first scholar to use personal income tax records in Brazil, applying the Pareto interpolation to tabulated data to study income inequality. His contribution did not spur further studies until the 1970s, when scholars with ties to the military dictatorship, such as Kingston and Kingston (1972) and Langoni (1973), also relied on income tax data to try to push for more benign views of the rise in inequality in the 1960s. The use of tax records to study top incomes would not re-surface until the 2010s when newly-released income tax tabulations became available to researchers. Not only did this data show that surveys exaggerated the fall in inequality in the 2000s (Medeiros, Souza, and Castro, 2015; Morgan, 2017a), it was also used to measure distributional effects of taxation (e.g. Castro and Bugarin (2017); Gobetti and Orair (2017); Fernandes, Diniz, and Silveira (2018)). Coupled with archival data on historical income tax tabulations, this new data was used by Souza and Medeiros (2015), Morgan (2015) and Souza (2016, 2018) to estimate top income shares in the long-run for the first time. While the combination of survey and tax data into a single measure of inequality was attempted by Medeiros et al. (2015); Souza (2016); Medeiros, de Castro Galvão, and de Azevedo Nazareno (2018), their reconciliation with national income statistics over the 2000s was studied by Morgan (2017a) and by Morgan (2018) over the long run.

Chile. The earliest attempt to study top income trends did not come from the use of administrative tax data but from surveys (Sanhueza and Mayer, 2011). López, Figueroa, and Gutiérrez (2013) were the first scholars to employ personal income tax tabulations to study top incomes over the 2000s. Administrative microdata of tax declarations were used by Fairfield and Jorratt De Luis (2016) to better study top incomes in the context of an institutional set-up tailored for the retention of a large amount of corporate profits not included in income tax returns for two individual years, refining the similar estimates made by López et al. (2013). Flores et al. (2020) has been to date the most comprehensive study on top incomes, combining features from previous attempts – long run estimates from income tax tabulations (1964-2017) with imputations of retained earnings from national accounts.

Colombia. Londoño-Vélez (2012) was the first work to incorporate income tax databases, which were used in Alvaredo and Londoño Velez (2014) for the study of top incomes and their composition between 1993 and 2010. The latter reconciled the results with survey-based measures using Gini-adjustment methods from Atkinson (2007) and Alvaredo (2011).

Costa Rica. Zuniga-Cordero (2018) is the first study to use multiple administrative sources of income (social security records, income tax data, national accounts) to study inequality, alongside household surveys, for Costa Rica, for the 2000-2017 period. Zuniga-Cordero (2022) revised these numbers and updated the series until 2020.

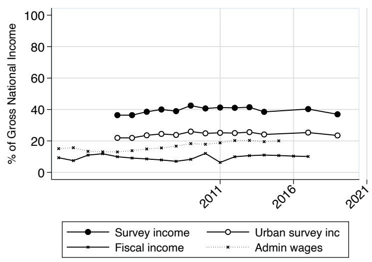
Ecuador. Few studies exist for the analysis of top incomes, with Cano (2015) initiating the trend based on microdata from tax registries over the period 2004-2010. This attempt was followed by Rossignolo et al. (2016), updating the previous series to 2014.

Mexico. Alvaredo et al. (2017) is the only study that used income tax data on universe of personal income taxpayers from the Mexican Tax Administration Service (SAT) and formal wage data from the universe of employer-reported information in the Declaración Informativa Múltiple (DIM) from 2009 to 2014. The authors perform a comparative analysis of incomes declared in these administrative datasets with those reported in the household surveys (ENIGH) for the same years in order to explore a potential reconciliation.

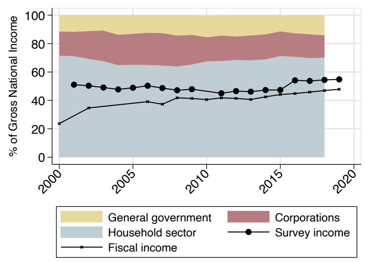
Uruguay. The decrease in income inequality shown in household surveys (e.g. Cornia (2014b)) has been confirmed by the use of income tax records (Burdín et al., 2022) for the 2009-2016 period, although milder and with stability in top income groups. Capital incomes are the key drivers of divergent trends between survey and administrative records. Falling inequality also emerged from Distributional National Accounts (DINA) estimations (De Rosa and Vilá, 2022), which is found to be more pronounced than in the fiscal incomes series given the decreasing share of undistributed profits. In all cases, unlike this study, the departure point is the administrative dataset, which is supplemented with household surveys and national accounts, as opposed to survey correction.

A.3. National Accounts

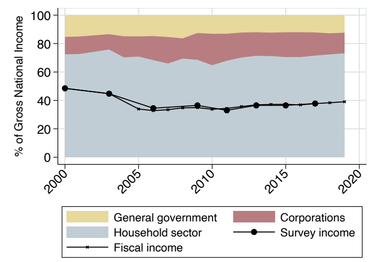
Fig. A.3. From Household Surveys to National Income



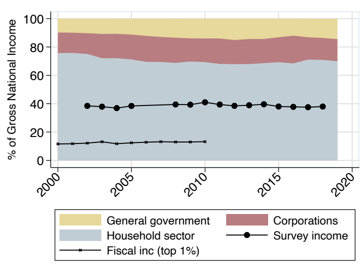
(a) Argentina



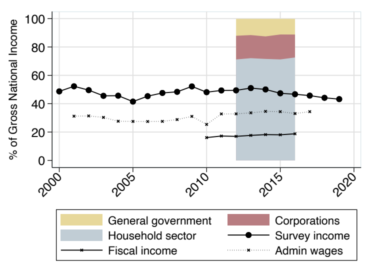
(b) Brazil



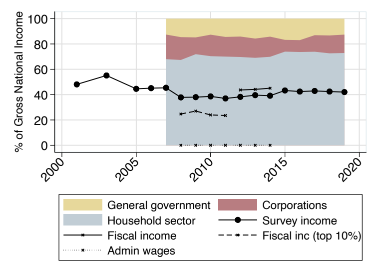
(c) Chile



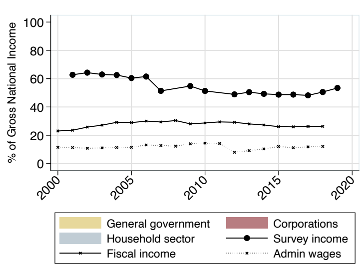
(d) Colombia



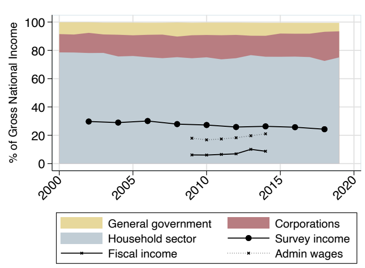
(e) Costa Rica



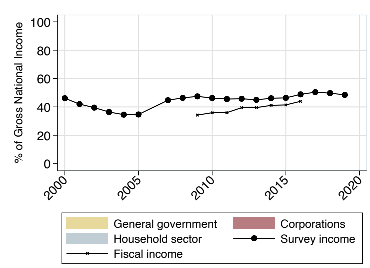
(f) Ecuador



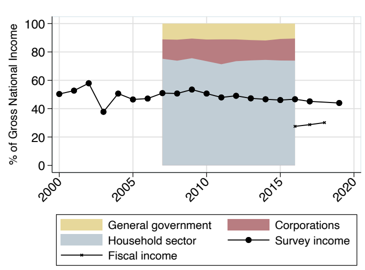
(g) El Salvador



(h) Mexico



(i) Uruguay



(j) Peru

Note. The survey series are for total pretax income. Shaded areas are the balance of primary incomes of the household sector (B.5g, S.14), corporations (B.5g, S.11 + S.12) and general government (B.5g, S.13). Source: Alvaredo et al. (2022).

B. Appendix: Estimation Methods

Our estimation procedure is based on four stages. We first estimate a survey-based distribution of income. The transition from this distribution to the distribution of national income as measured in the national accounts is accomplished in three subsequent steps. In the first step, we adjust household surveys to include distributive information from administrative records; in the second step, we proportionally scale the different income components to match aggregates from the national accounts; finally, in the third step, we impute corporate undistributed profits (retained earnings) and remaining missing incomes. In this section we provide a brief summary of these adjustment steps.[24]

B.1. Estimation of pre-tax incomes in surveys

The inequality estimates we present in this paper concern pre-tax incomes. However, the main data source on which our estimates are based are harmonized household surveys, which account for post-tax incomes in Latin America.[25] In order to scale incomes to their pre-tax aggregates in the national accounts it is necessary to calculate pre-tax incomes in surveys.

As data on direct taxes paid by households is not collected in surveys we tax data to estimate pre-tax incomes. Broadly speaking, we compute effective tax rates by income fractile in the tax data, and use these tax rates to calculate pre-tax incomes in the survey, based on the income fractiles to which individuals belong to.[26] Effective tax rates by income fractile are computed for the years for which we have access to income tax data, and the average effective tax rate by fractile is used to calculate pre-tax incomes when this data is not available.[27] Tax data quality and coverage, however, varies significantly across countries and so specific procedures and assumptions have to be made for each country. In Table B.1, the main characteristics of the data and estimation procedure by country are shown.

In the cases where data comes from tax tabulations, effective rates are computed for observed points (e.g. the average of a given income bracket) and linearly interpolated. For Colombia

[24] For a more detailed description of the general procedure we employ in this paper see WIL (2020).

[25] The only exceptions are Brazil and Costa Rica, whose survey accounts for pre-tax incomes.

[26] We consider, whenever possible, 127 income fractiles, which account for the whole income distribution (the first 99 percentiles) and a very detailed break-down of the top 1%, where tax rates may experience significant changes.

[27] This assumption is potentially problematic in the cases for which the absence of tax data reflects the absence of progressive income taxation (e.g. Uruguay prior to 2009), or when the availability of data followed a large tax reform.

Table B.1: Effective tax rates estimation by country

Country	Period	Pop. Cov.	Data	Method	Ref. income	Rates
Mexico	2009-2014	Top 2%	Microdata	Directly computed	Gross income	Tax rate
Argentina	2002-2017	Universe	Tabulations	Interpolated	Gross income	Tax rate
Brazil	2008-2016	Universe	Microdata	Directly computed	Net income	Tax rate
Colombia	2006-2010	Top 1%	Tabulations	Interpolated*	Gross income	Tax & SS rate
Chile	2005-2017	Universe	Tabulations	Interpolated	Net income	Tax rate
El Salvador	2000-2017	Universe	Tabulations	Interpolated	Gross income	Tax rate
Uruguay	2009-2016	Universe	Microdata	Directly computed	Gross income	Tax & SS rate
Peru	2016-2017	Universe	Microsim.	Interpolated	Net income	Tax rate
Ecuador	2008-2011	Top 10%	Tabulations	Interpolated*	Gross income	Tax & SS rate
Costa Rica	2010-2016	Universe	Tabulations	Interpolated	Gross income	Tax rate

Note. Authors' elaboration.

and Ecuador, effective tax rates are taken directly from the same studies we use to extract top income information – Londoño-Vélez (2012) for Colombia or Cano (2015); Rossignolo et al. (2016) for Ecuador. In the case of Peru, effective rates were microsimulated based on the statutory tax schedule. Finally, for countries in which we have tax micro-data or very detailed tabulations, the effective tax rates were computed directly (e.g. Mexico and Uruguay).

Taxes are progressive, but effective rates decrease significantly in the very right tail of the distribution for most countries. In countries where this is not the case (e.g Argentina and Chile), we cannot observe the very high income fractiles in the data without extrapolating. When social security contributions are observed (Colombia, Uruguay and Ecuador), they are a lot more regressive than the income tax, especially for top fractiles, where it converges to zero as a result of truncated schedules (i.e. schedules were a maximum income is defined for contributions). The absence of information on social contributions is not problematic, given that the income definition we use in our estimates includes social security transfers, net of social contributions.

B.2. Surveys adjusted with administrative data

The use of administrative data refers to both personal income tax declarations and social security records. These sources are mainly used to improve the coverage of top income groups in the survey, which are often badly captured; especially when register data is not used in the surveying process, which is the case in all countries in the region.

In general, administrative records not only include individuals that are richer than the richest survey respondents, but also report larger numbers of moderately high incomes. Therefore, when we compare the income distributions described in both sources, we usually find that the densities reported by administrative records tend to be higher for top incomes relative to surveys. Given that income tax declarations are made by real people, who might under-declare their income but are unlikely to over-declare, it seems natural to consider the distribution in register data as a lower bound that the survey should aim to match, at least when tax-data densities are higher.

In order to adjust the surveys we use the method described in Blanchet, Flores, and Morgan (2022), which mainly uses the ratio of survey to tax data densities to adjust survey weights. Although the method includes a "replacing" option, which allows users to impute incomes above the maximum income observed in surveys, we only use re-weighting without replacing for practical reasons (it makes the extrapolation of years without tax data clearer). The impact of not using the replacing option does not seem to affect inequality estimates in any meaningful way. Figure B.3 displays the intuition behind this re-weighting process, while Figure B.4 depicts the theta coefficients of the adjustment, i.e. the ratio of the survey density to the administrative density by income fractile. Results indicate that the median of the ratio between density functions is equal to one within the top 10% in all cases, although relatively closer to top percentiles in cases such as El Salvador or Costa Rica. Morevover, it may be seen that this ratio decreases during the period (clearly in the cases of Brazil, Colombia, Argentina and Uruguay), indicating that administrative incomes are higher than survey income after a decreasing threshold, as noted in (Alvaredo et al., 2022).

B.3. Scaling to incomes from national accounts

Figure B.5 displays the adjustment factors used to scale five types of income (wages, capital incomes, mixed incomes, imputed rents, and social benefits) to corresponding aggregates from the national accounts. This is done proportionally to survey incomes after adjustment with administrative data. Since the relevant macro aggregates are reported before income tax in the national accounts we add effective income tax paid across the adjusted survey distribution for the nine countries with post-tax survey incomes. Appendix B.1 explains how these tax rates are computed.

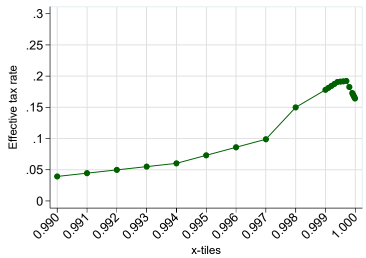
Table 3.1 summarizes our benchmark matching of income concepts. For labor incomes, we subtract social security contributions from the compensation of employees before computing scaling factors. Since most countries' national accounts report pensions along with other benefits, we scale total benefits to that aggregate, assuming the joint distribution of pensions and other benefits is accurately described by the survey. The level of detail that is necessary to split the part of property incomes related to investment income disbursements (D44) – which includes investment income from insurance funds (D441), pension funds (D442), and collective investment funds (D443) imputed to households – in the national accounts is not available in most countries in the region. For the countries where the detail exists at least for a few years (Brazil, Colombia, Chile, Costa Rica, Ecuador and Mexico) we estimate that investment income disbursements represent a relatively stable 10% of total property income of households on average. Therefore, we scale total capital income in the surveys to 90% of total property income (D4) in the national accounts for each country to match the incomes actually received by households (i.e. interests and dividends).

Table B.2: Conceptual relation between incomes in surveys and national accounts

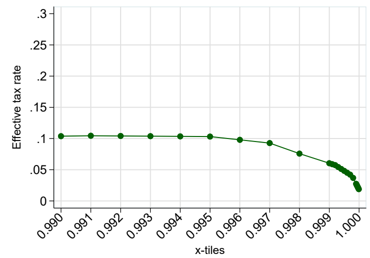
Survey	**National Accounts**	**Comparable incomes**	**Less comparable incomes**
Salaried work	Compensation of employees (D1)	Wages, salaries (D11)	Social security contributions (D61)
Rental income	Operating surplus (B2)	Imputed rent of owner occupiers	Effective rent of residential buildings
Non-salaried work	Mixed income (B3)	Self-employed income	Effective rent of non-residential buildings
Investment income	Property income (D4)	Interests received (D41r) Dividends (D42)	Interests paid (D41u) Rent of natural resources (D45) Investment income of insurance policy holders (D441) Investment income of pension funds (D442) Investment income of investment funds (D443)
Other incomes	Social transfers (D62) Other transfers (D7)	Pensions Other cash benefits	Unemployment insurance Sick leave Private transfers (remittances)

Notes: Table taken from Alvaredo et al. (2022), based on United Nations (2008b) and OECD (2013). In column 4, the items with a code next to them can be subtracted for a better matching (depending on the detail provided by national agencies), while those without a code cannot be separated from the aggregates in column 2. Listed items are pre-tax in SNA, while most of them are post-tax in surveys. Operating surplus and mixed income are gross of depreciation in surveys and in the SNA.

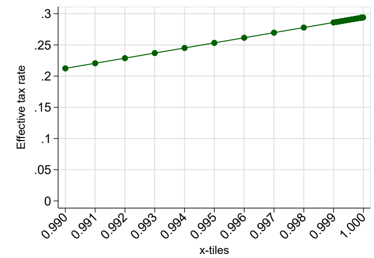
Fig. B.1. Effective tax and social security rates - Top 1% - Latest year



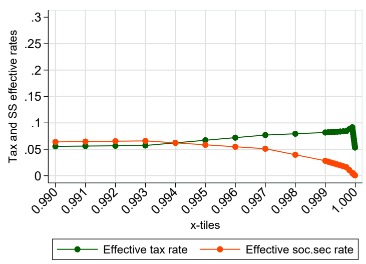
(a) Argentina 2017



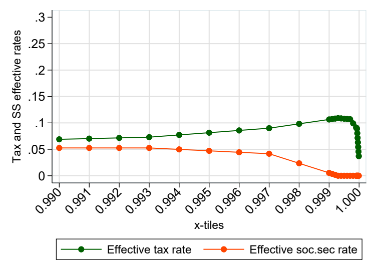
(b) Brazil 2016



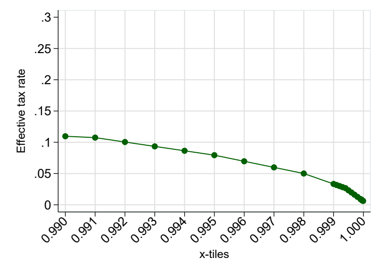
(c) Chile 2017



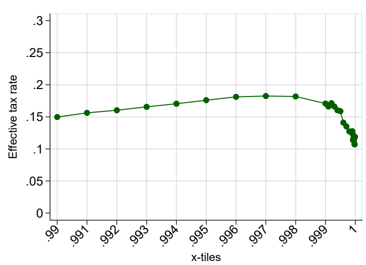
(d) Colombia 2010



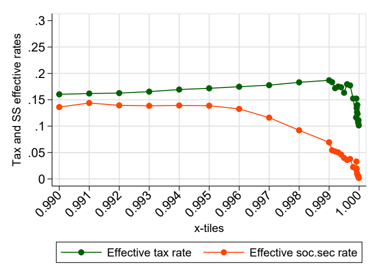
(e) Ecuador 2011



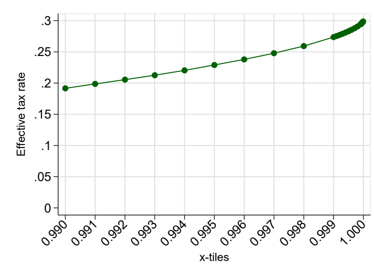
(f) El Salvador 2017



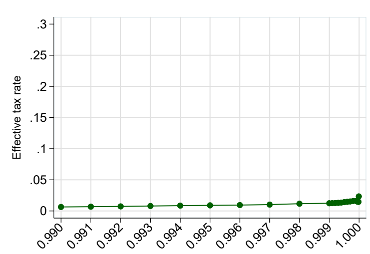
(g) Mexico 2014



(h) Uruguay 2016



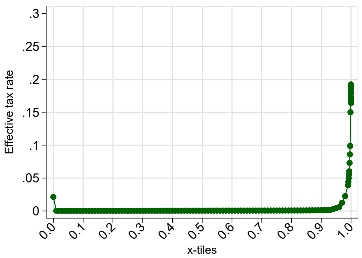
(i) Peru 2017



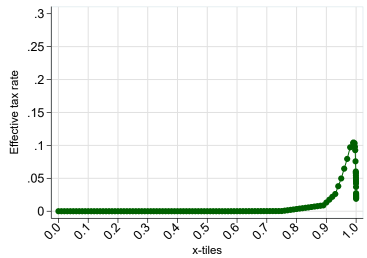
(j) Costa Rica 2016

Note. Authors' elaboration.

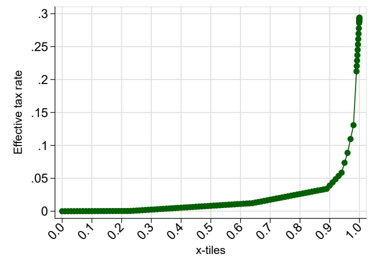
Fig. B.2. Effective tax and social security rates - Latest year



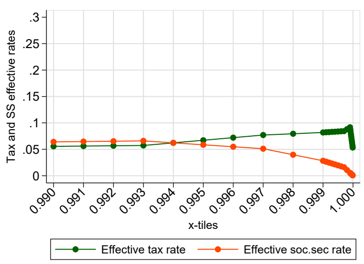
(a) Argentina 2017



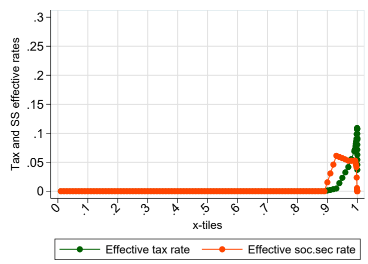
(b) Brazil 2016



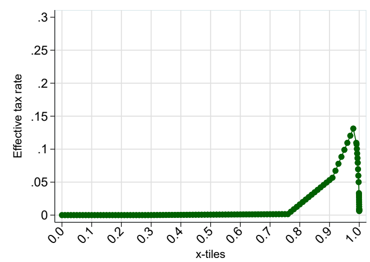
(c) Chile 2017



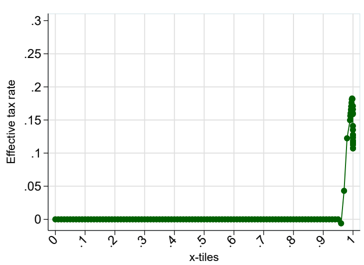
(d) Colombia 2010



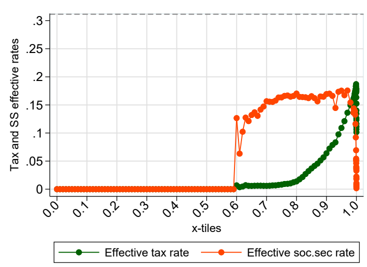
(e) Ecuador 2011



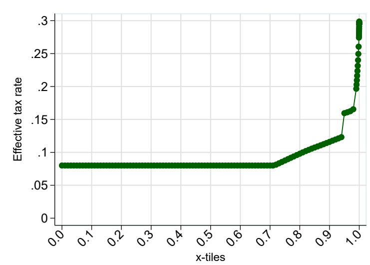
(f) El Salvador 2017



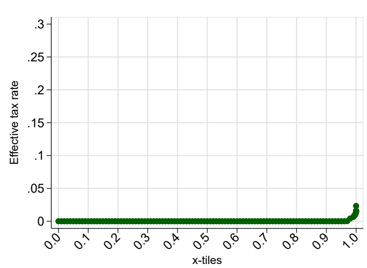
(g) Mexico 2014



(h) Uruguay 2016



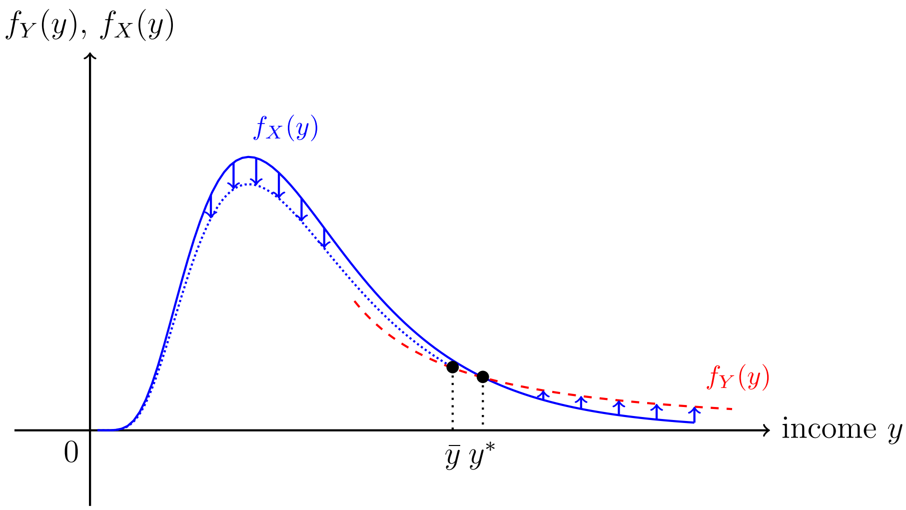
(i) Peru 2016



(j) Costa Rica 2016

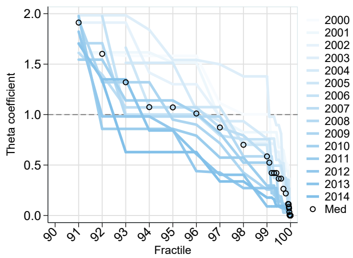
Note. Authors' elaboration.

Fig. B.3. The intuition behind reweighting

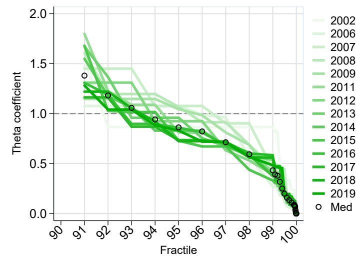


Source. Blanchet, Flores, and Morgan (2022). The solid blue line represents the survey density f_X. The dashed red line represents the tax data density f_Y. Above the merging point $\bar{y}$, the reweighted survey data have the same distribution as the tax data (dashed red line). Below the merging point, the density has been uniformly lowered so that it still integrates to one, creating the dotted blue line.

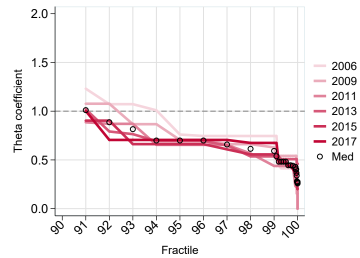
Fig. B.4. Theta coefficients, by country and year



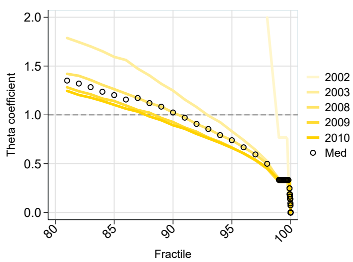
(a) Argentina



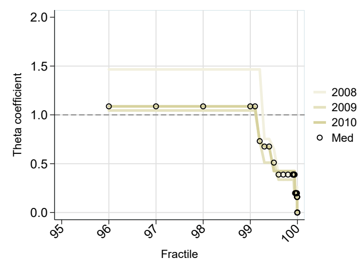
(b) Brazil



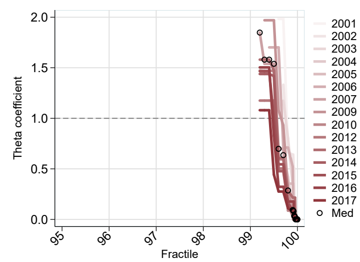
(c) Chile



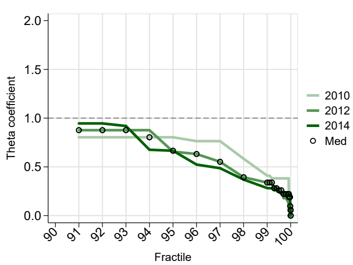
(d) Colombia



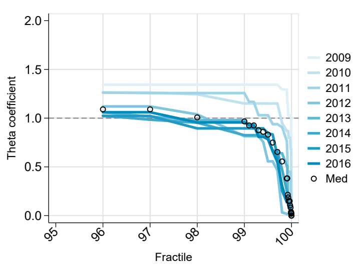
(e) Ecuador



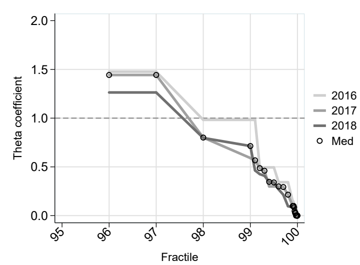
(f) El Salvador



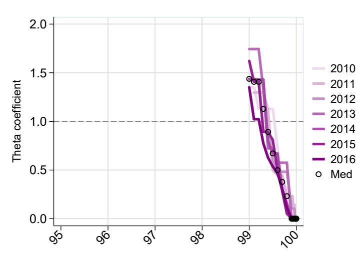
(g) Mexico



(h) Uruguay



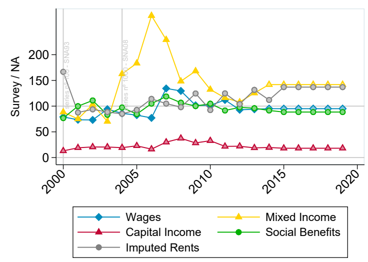
(i) Peru



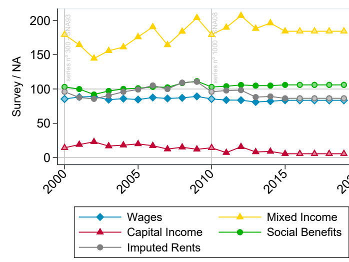
(j) Costa Rica

Note. Authors' elaboration based on Blanchet et al. (2022)

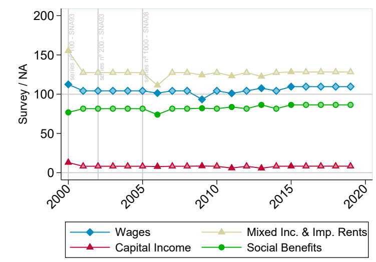
Fig. B.5. Scaling factors for re-weighted surveys



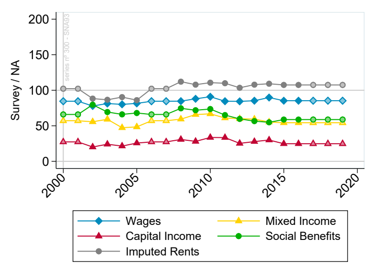
(a) Argentina



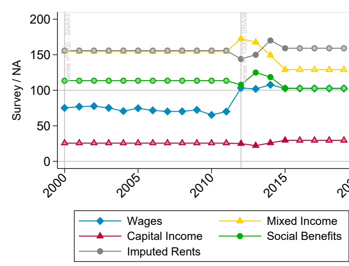
(b) Brazil



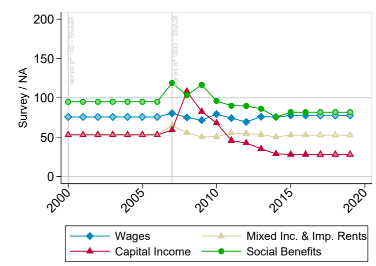
(c) Chile



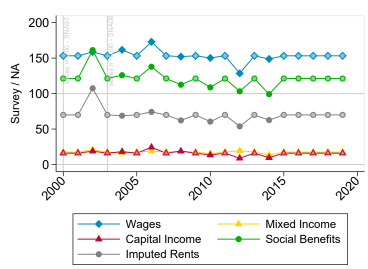
(d) Colombia



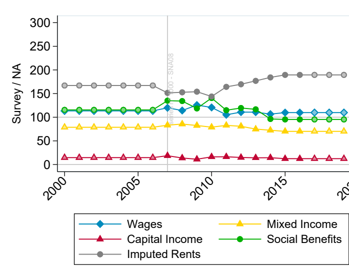
(e) Costa Rica



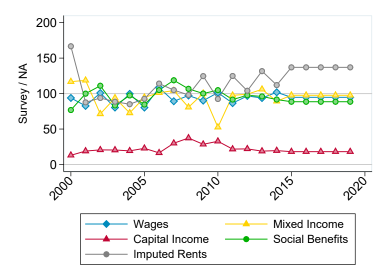
(f) Ecuador



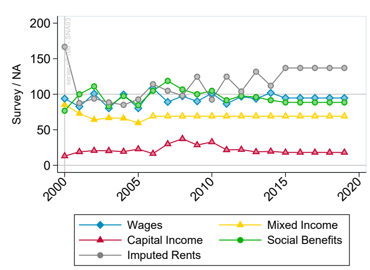
(g) Mexico



(h) Peru



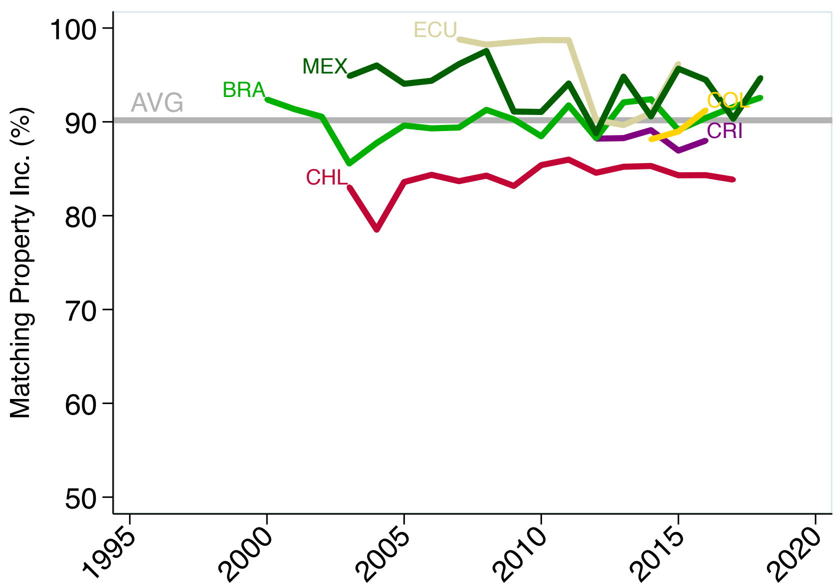
(i) El Salvador



(j) Uruguay

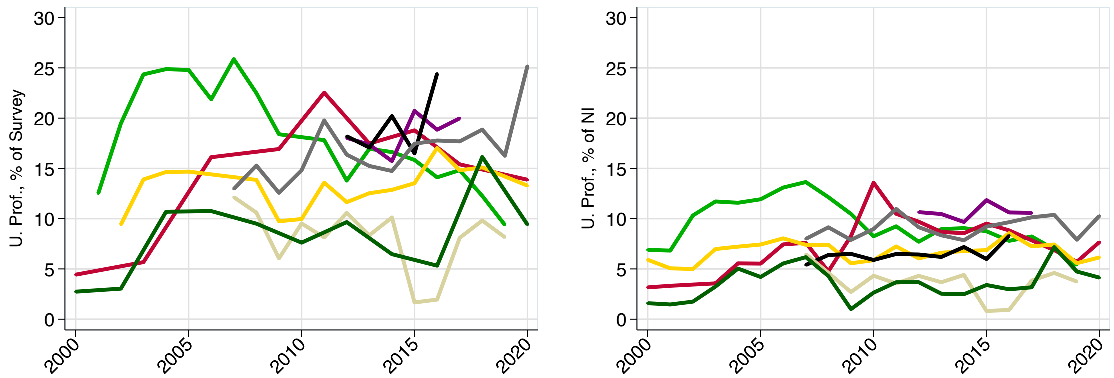
Note. Authors' elaboration using surveys, administrative data and national accounts. Each series is the ratio of survey income (adjusted using administrative data) to national accounts income for each component. Brighter points indicate imputed scaling factors due to missing information in National Accounts. Each survey income component is multiplied by the scaling factor (1/ratio) for components where coverage is less than 100%, and divided by the factor for components where coverage is greater than 100%.

Fig. B.6. Share of conceptually consistent property incomes

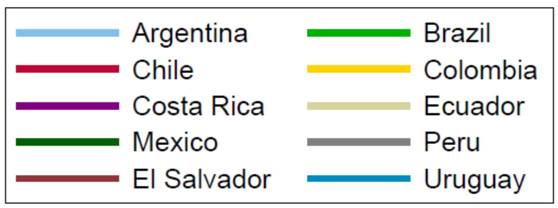


Notes. The share of property incomes from SNA that matches the definition of surveys' capital incomes (i.e. dividends and interests) is mostly above 80% of total property income, closer to 90% in most cases. Conceptual differences thus seem to play a minor role in the underestimation of capital incomes displayed in figure 1.3(c). The level of detail that is necessary to observe this is rare in Latin America. Non-matching concepts are SNA codes D.43 and D.44 (see table 3.1). Authors' elaboration based on the public national accounts reported by each country's relevant institutions.

Fig. B.7. Undistributed Profits as % of Aggregate Incomes

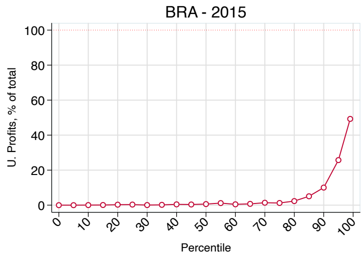


(a) % of Survey Income (b) % of National Income

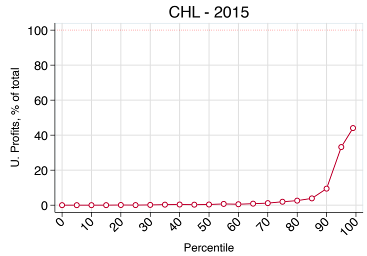


Note. Authors' elaboration using data from the World Inequality Database on undistributed profits, UN data or country-level data on national income and ECLAC on household surveys.

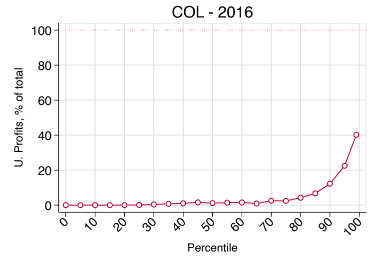
Fig. B.8. Share of total undistributed profits imputed to each fractile



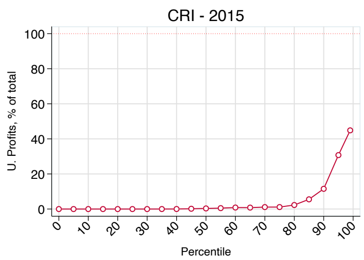
(a) Brazil 2015



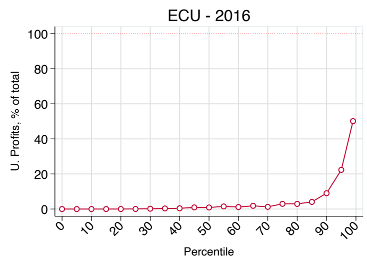
(b) Chile 2015



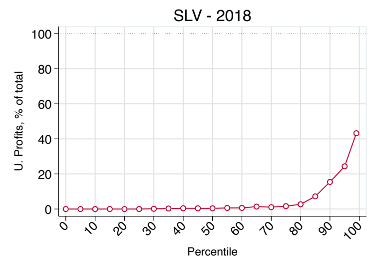
(c) Colombia 2016



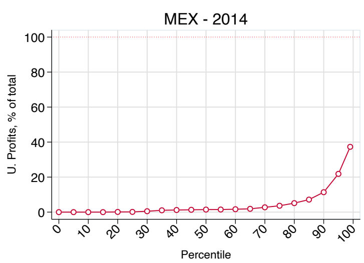
(d) Costa Rica 2015



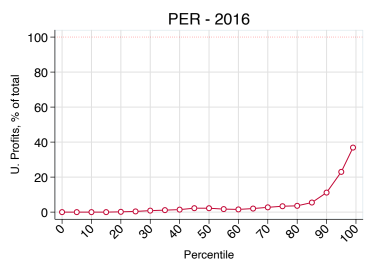
(e) Ecuador 2016



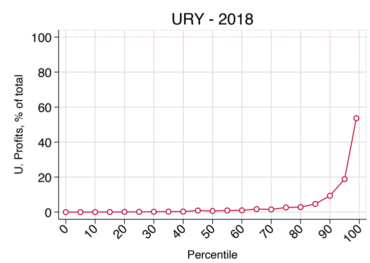
(f) El Salvador 2018



(g) Mexico 2014



(h) Peru 2016

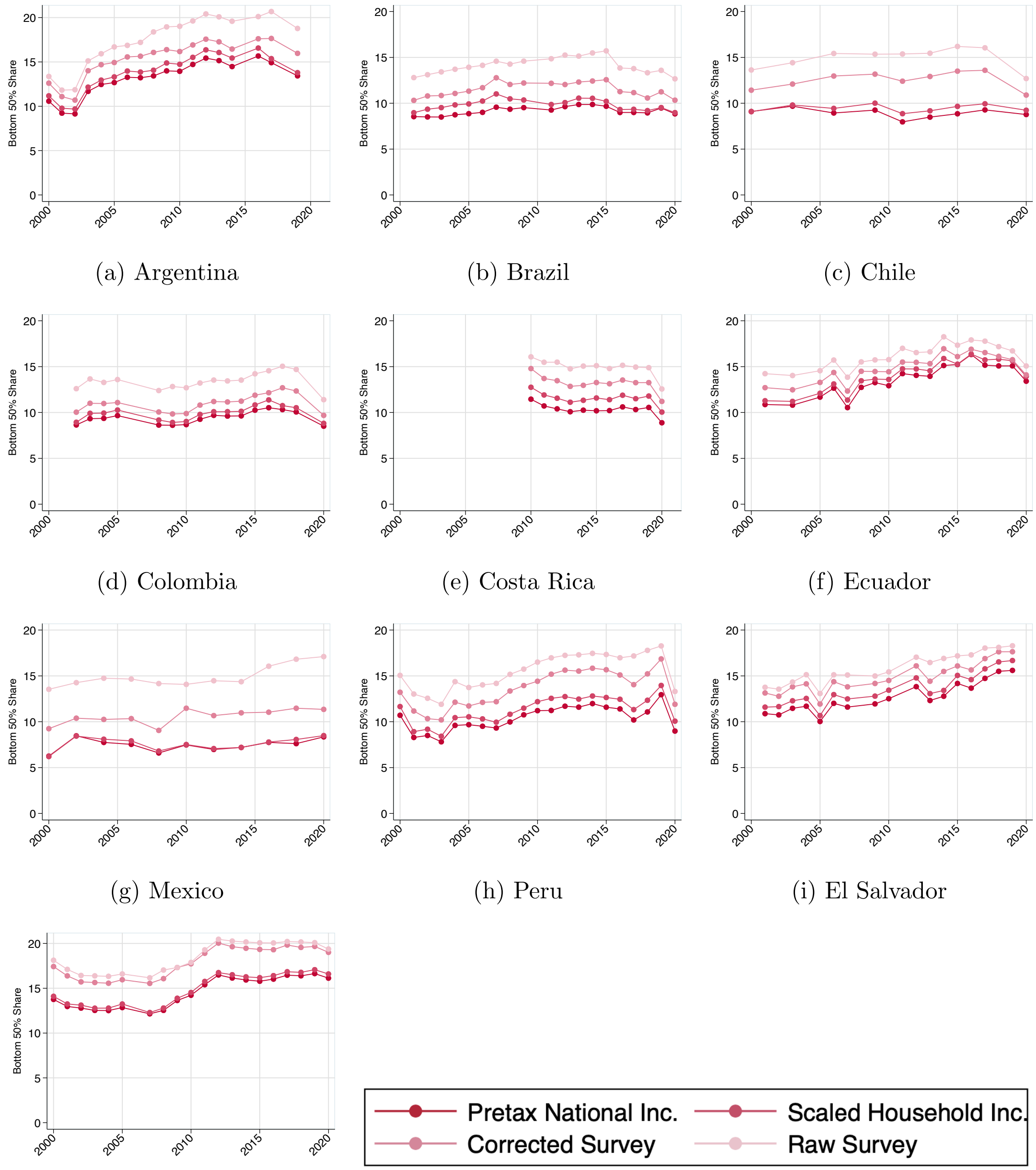


(i) Uruguay 2018

Note. Authors' elaboration using distributional data from surveys on dividends and employer income and aggregate data on undistributed profits from the World Inequality Database (https://wid.world/).

C. Supplementary Figures

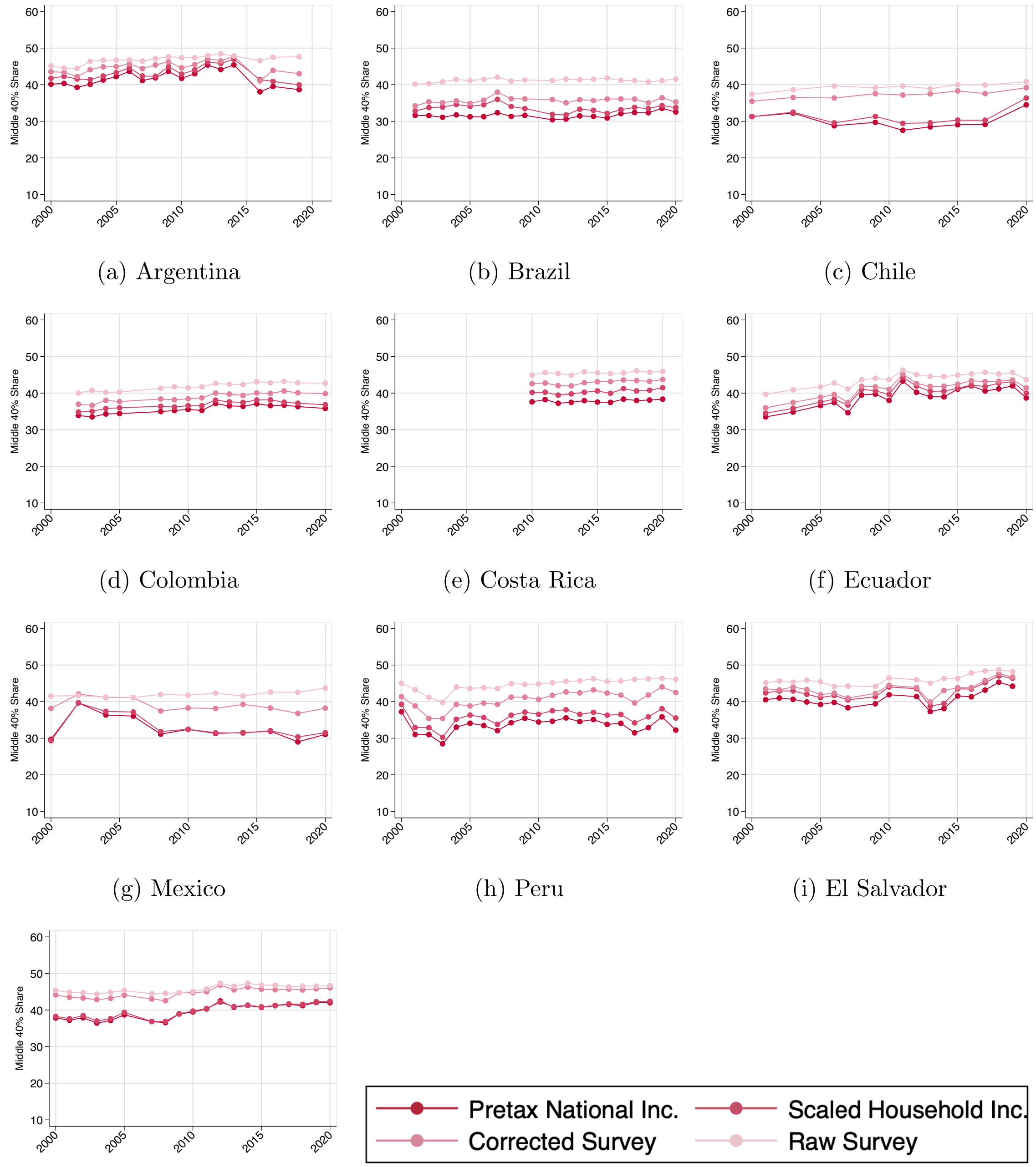
Fig. C.1. Bottom 50% Share in four distributions



(a) Argentina (b) Brazil (c) Chile (d) Colombia (e) Costa Rica (f) Ecuador (g) Mexico (h) Peru (i) El Salvador (j) Uruguay

Note. Authors' elaboration. The figures depict four distributions: the household survey-based distribution and the three augmented distributions based on three adjustment steps to the survey. The first step uses tax data to reweight the raw survey; the second step scales the income totals in the tax-adjusted survey to their equivalent household-level aggregates in the national accounts; the third step imputes missing incomes needed to reach national income. Brighter points indicate that at least part of the data necessary for the adjustment step was imputed based on remaining country/year averages. The distributions are of pre-tax household per capita income (including pensions and after social contributions).

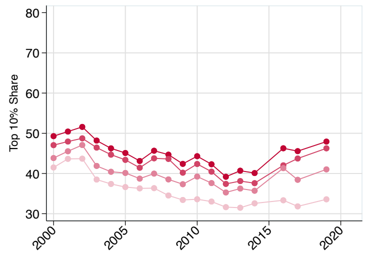
Fig. C.2. Middle 40% Share in four distributions



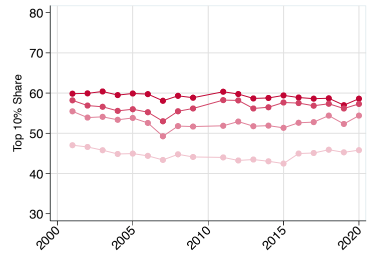
(a) Argentina (b) Brazil (c) Chile (d) Colombia (e) Costa Rica (f) Ecuador (g) Mexico (h) Peru (i) El Salvador (j) Uruguay

Note. Authors' elaboration. The figures depict four distributions: the household survey-based distribution and the three augmented distributions based on three adjustment steps to the survey. The first step uses tax data to reweight the raw survey; the second step scales the income totals in the tax-adjusted survey to their equivalent household-level aggregates in the national accounts; the third step imputes missing incomes needed to reach national income. Brighter points indicate that at least part of the data necessary for the adjustment step was imputed based on remaining country/year averages. The distributions are of pre-tax household per capita income (including pensions and after social contributions).

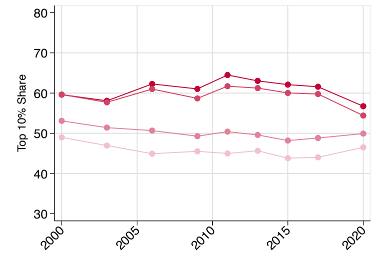
Fig. C.3. Top 10% Share in four distributions



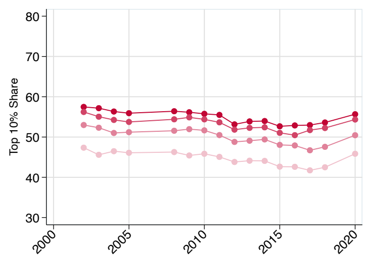
(a) Argentina



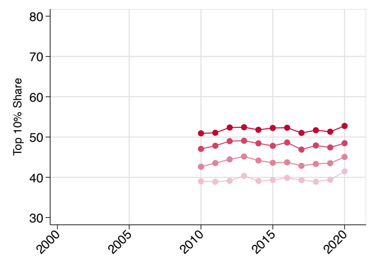
(b) Brazil



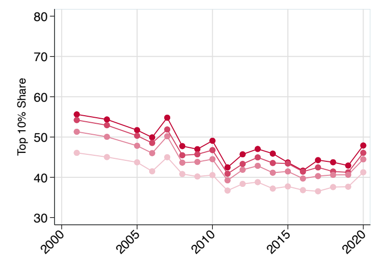
(c) Chile



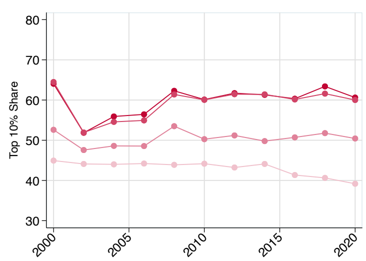
(d) Colombia



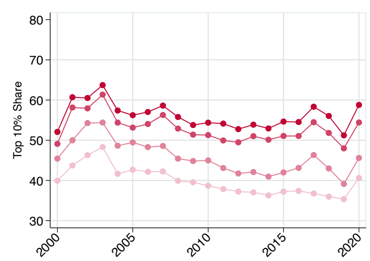
(e) Costa Rica



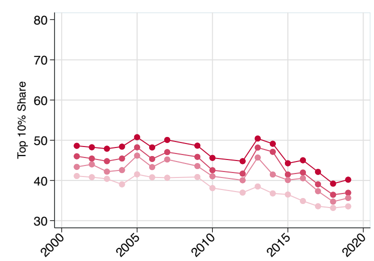
(f) Ecuador



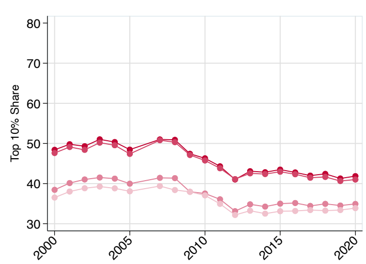
(g) Mexico

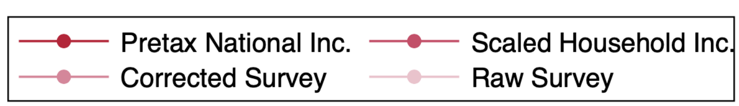


(h) Peru



(i) El Salvador

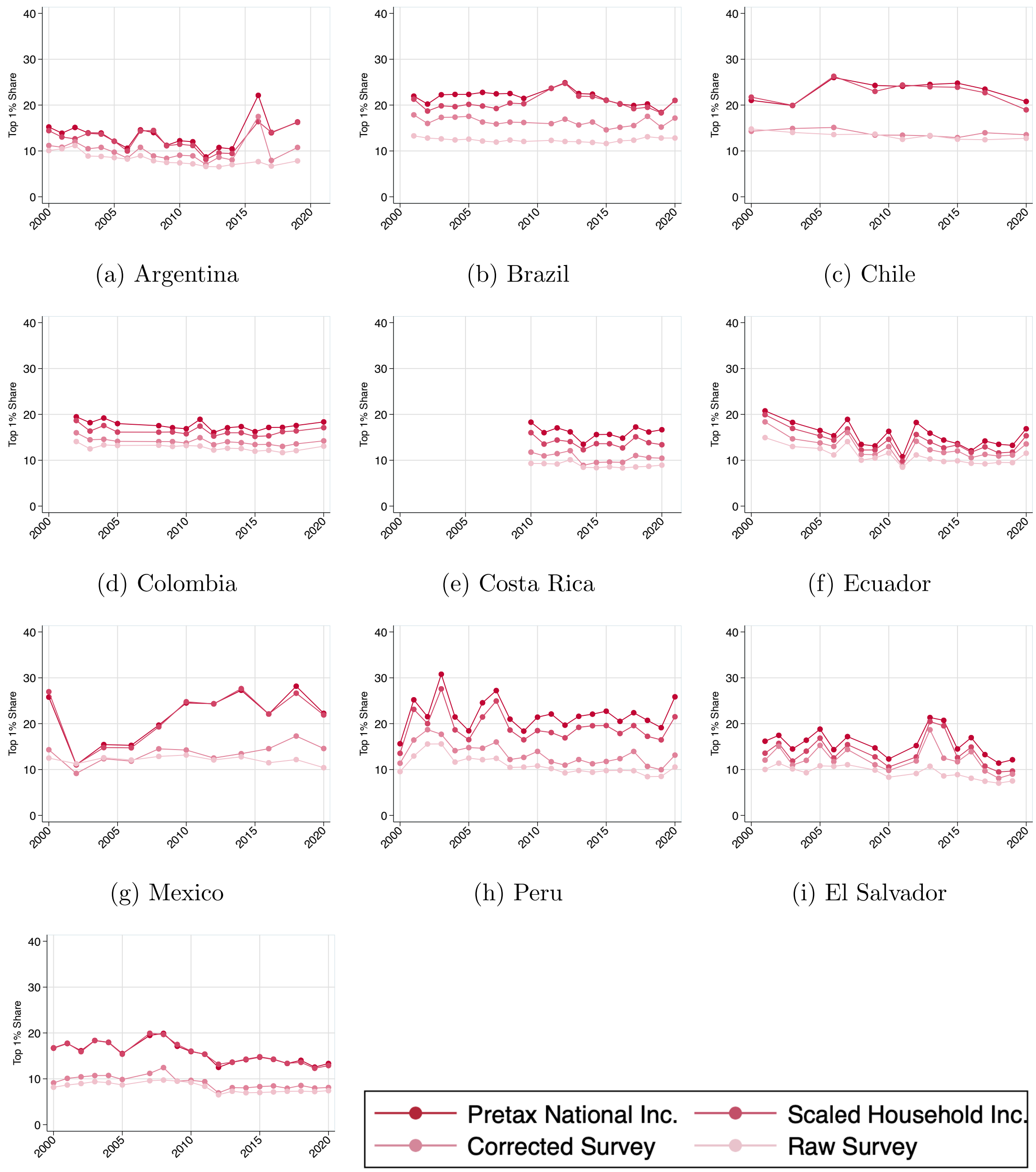




(j) Uruguay

Note. Authors' elaboration. The figures depict four distributions: the household survey-based distribution and the three augmented distributions based on three adjustment steps to the survey. The first step uses tax data to reweight the raw survey; the second step scales the income totals in the tax-adjusted survey to their equivalent household-level aggregates in the national accounts; the third step imputes missing incomes needed to reach national income. Brighter points indicate that at least part of the data necessary for the adjustment step was imputed based on remaining country/year averages. The distributions are of pre-tax household per capita income (including pensions and after social contributions).

Fig. C.4. Top 1% Share in four distributions



(a) Argentina

(b) Brazil

(c) Chile

(d) Colombia

(e) Costa Rica

(f) Ecuador

(g) Mexico

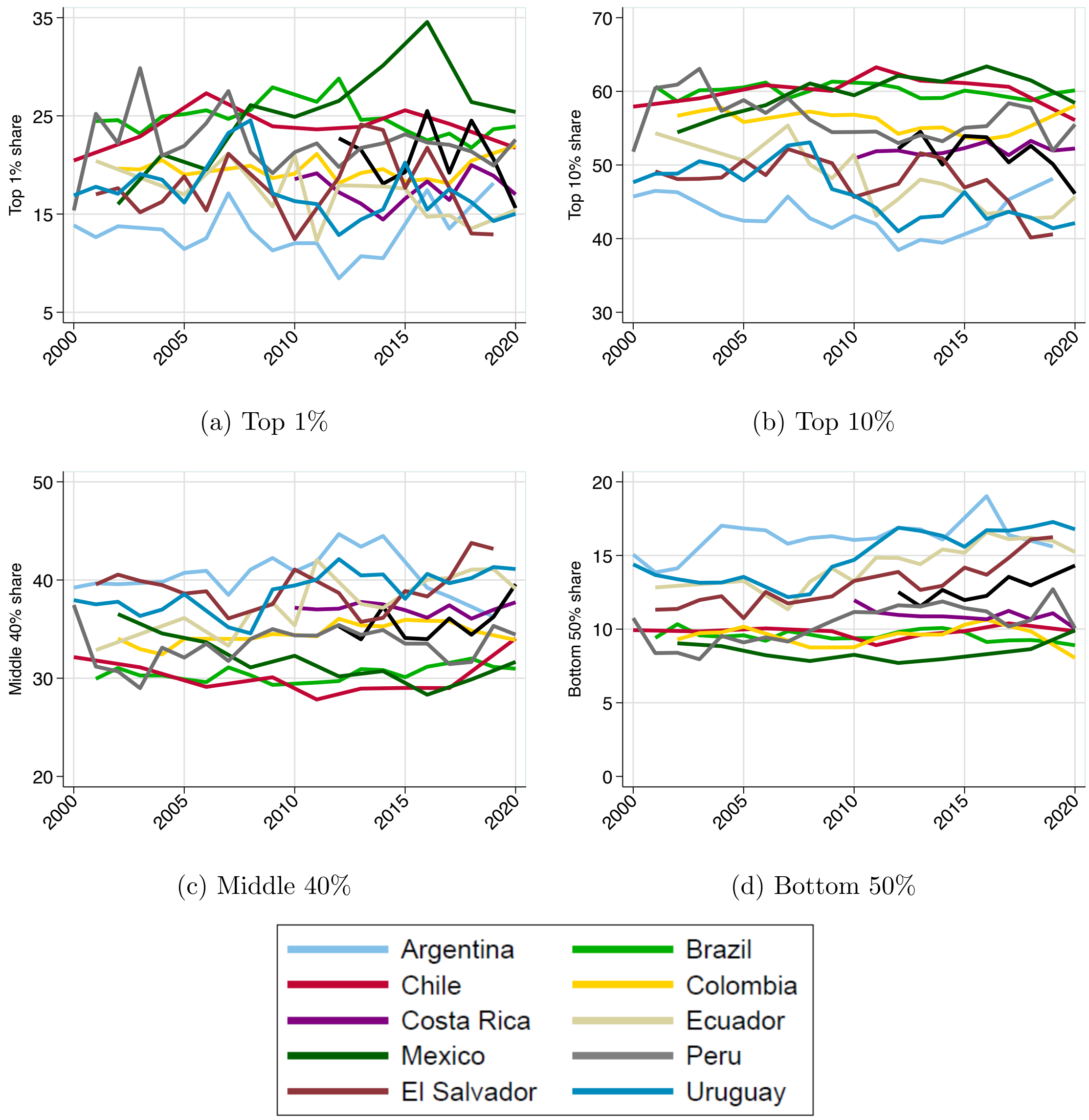
(h) Peru

(i) El Salvador

(j) Uruguay

Note. Authors' elaboration. The figures depict four distributions: the household survey-based distribution and the three augmented distributions based on three adjustment steps to the survey. The first step uses tax data to reweight the raw survey; the second step scales the income totals in the tax-adjusted survey to their equivalent household-level aggregates in the national accounts; the third step imputes missing incomes needed to reach national income. Brighter points indicate that at least part of the data necessary for the adjustment step was imputed based on remaining country/year averages. The distributions are of pre-tax income (including pensions and after social contributions).

Fig. C.5. Pre-tax national income shares

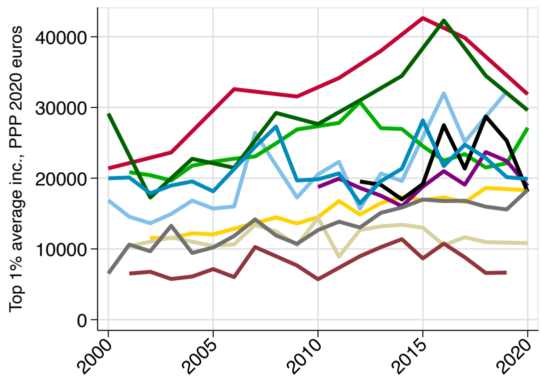


(a) Top 1% (b) Top 10%

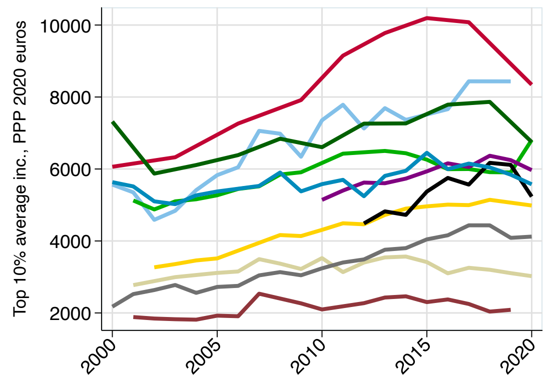
(c) Middle 40% (d) Bottom 50%

Note. Authors' elaboration based on the combination of household surveys, administrative data and national accounts.

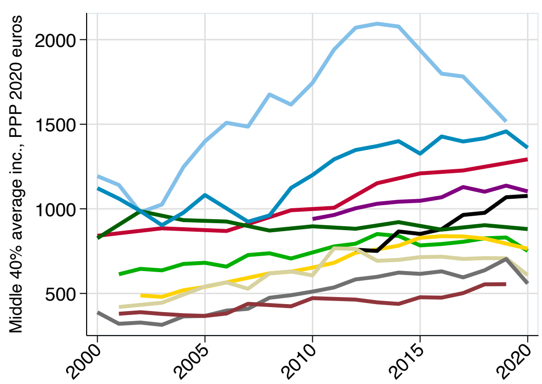
Fig. C.6. Pretax average national incomes by group



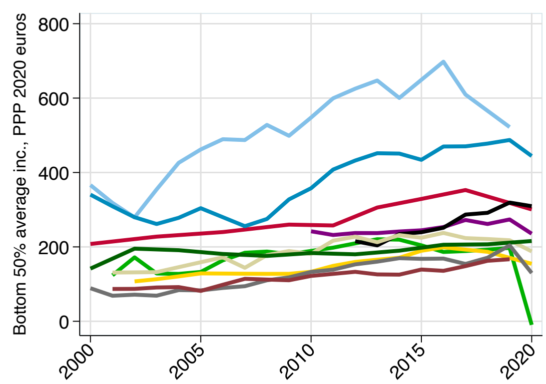
(a) Top 1%



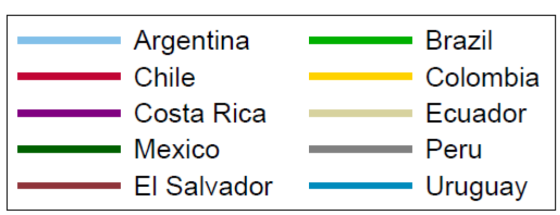
(b) Top 10%



(c) Middle 40%

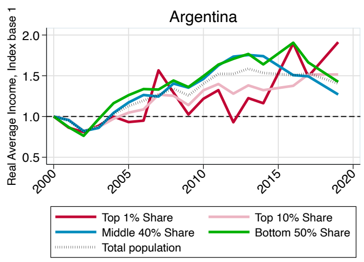


(d) Bottom 50%

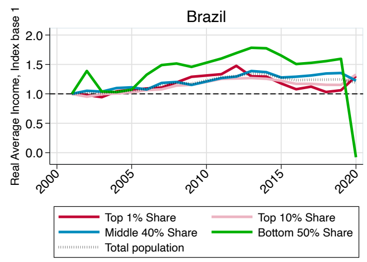


Note. Authors' elaboration based on the combination of household surveys, administrative data and national accounts.

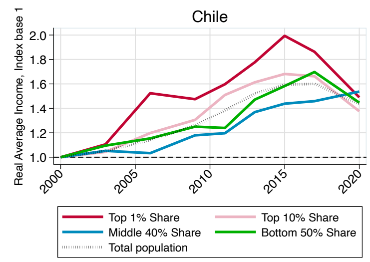
Fig. C.7. The distribution of pretax income growth across groups



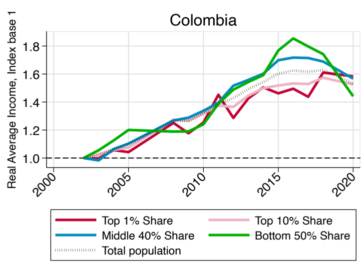
(a) Argentina



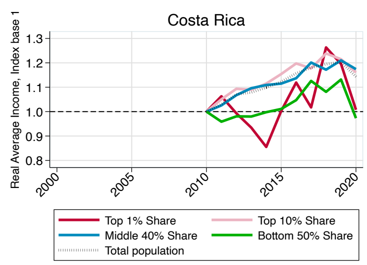
(b) Brazil



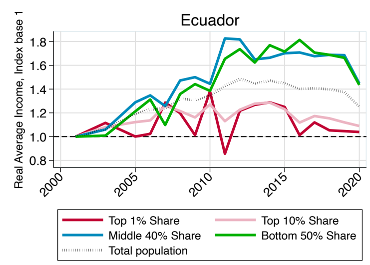
(c) Chile



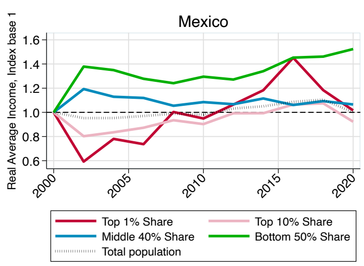
(d) Colombia



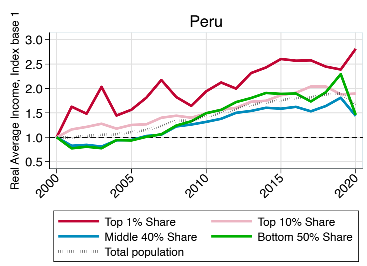
(e) Costa Rica



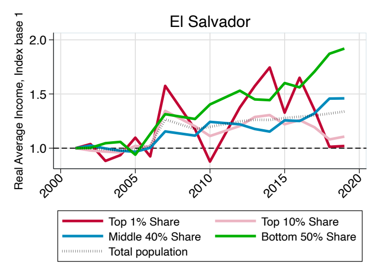
(f) Ecuador



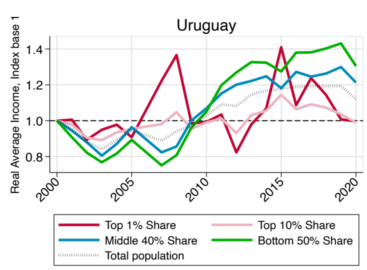
(g) Mexico



(h) Peru



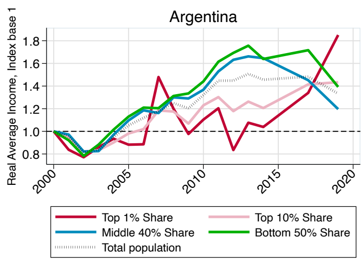
(i) El Salvador



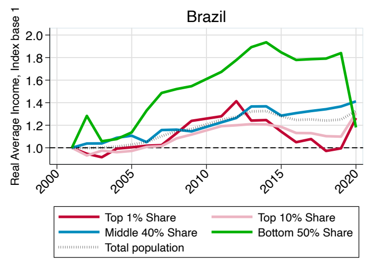
(j) Uruguay

Note. Authors' elaboration. Income is pre-tax national household per capita income (surveys, tax data and national accounts, before all taxes, transfers and public spending, including pensions and deducting social contributions).

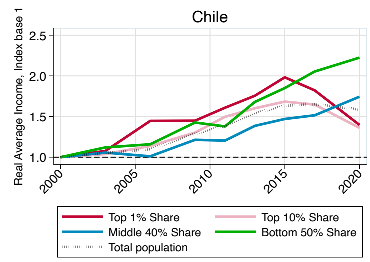
Fig. C.8. The distribution of post-tax income growth across groups



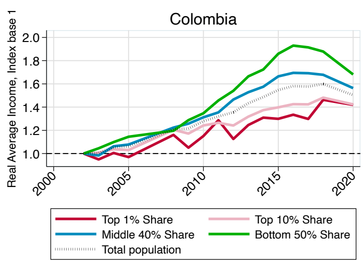
(a) Argentina



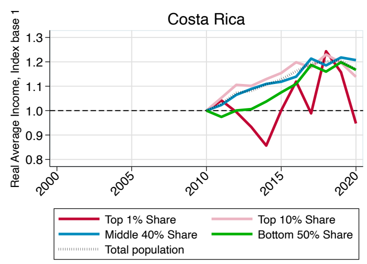
(b) Brazil



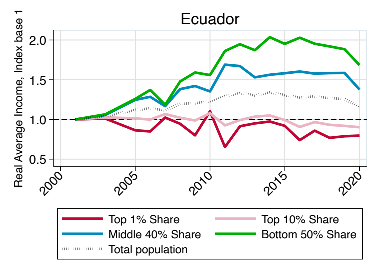
(c) Chile



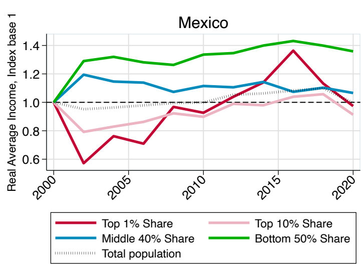
(d) Colombia



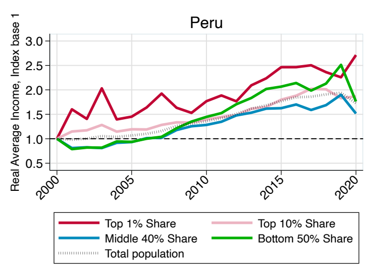
(e) Costa Rica



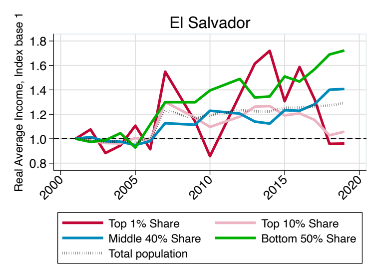
(f) Ecuador



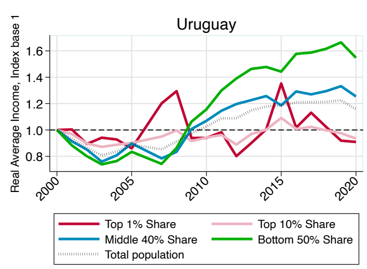
(g) Mexico



(h) Peru



(i) El Salvador



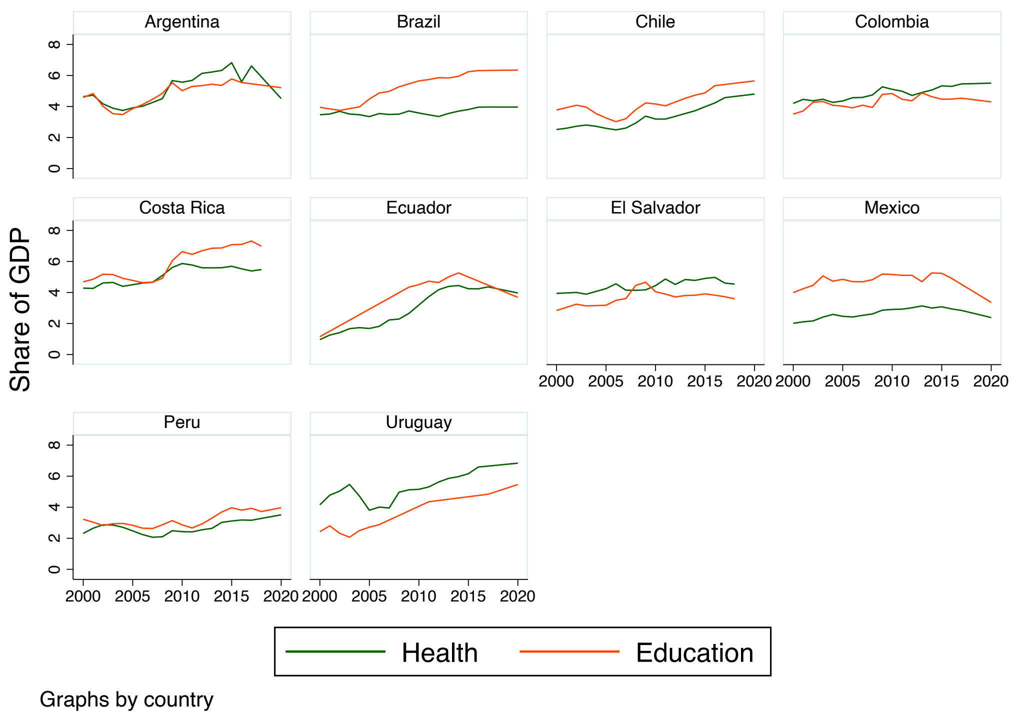
(j) Uruguay

Note: Authors' elaboration. Income is post-tax national household per capita income (surveys, tax data and national accounts, after all taxes, transfers and public spending).

Fig. C.9. The composition of national taxes

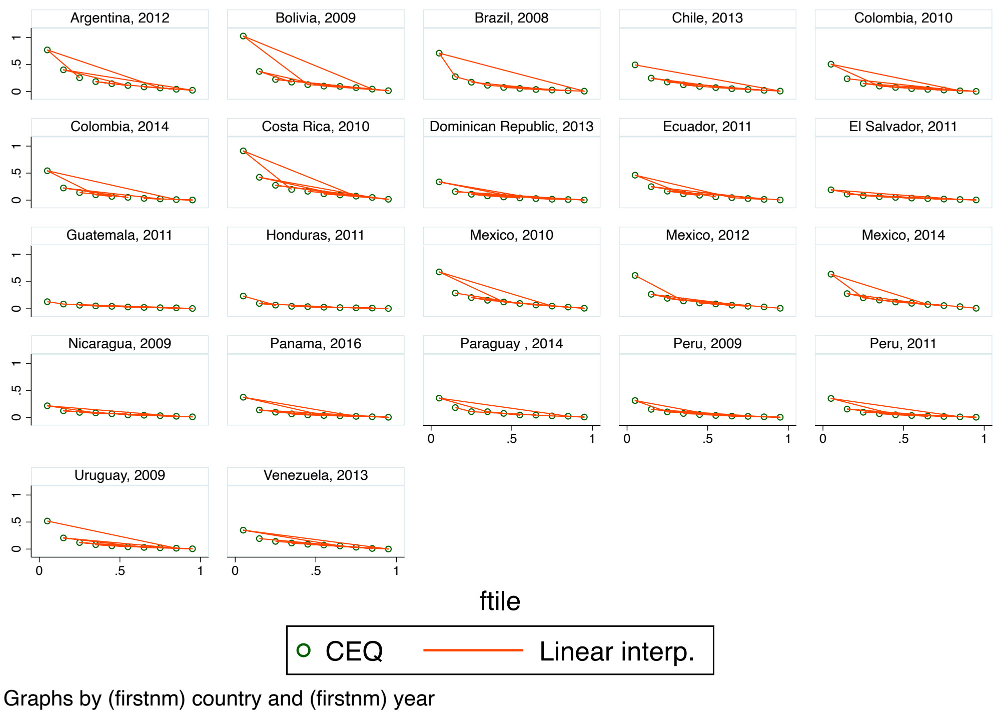
Source: OECD, CIAT and CEPAL (2021).

Fig. C.10. The evolution of in-kind social expenditures



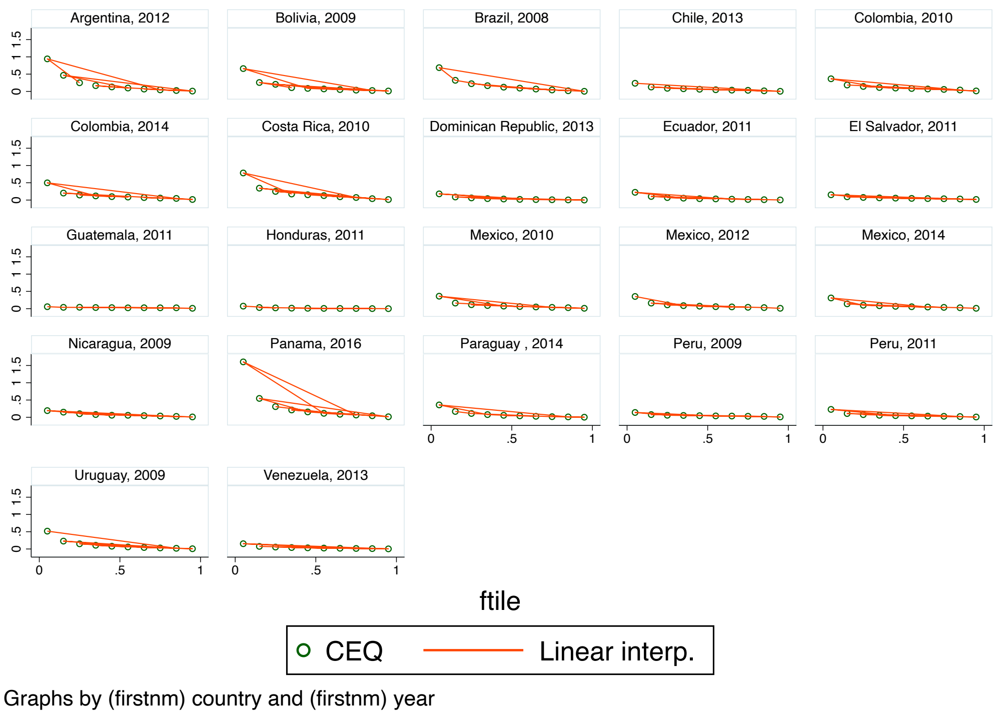
Note: The graphs show the evolution of government expenditures on health and education as a share of GDP. Source World Bank (https://data.worldbank.org/).

Fig. C.11. The incidence of education spending



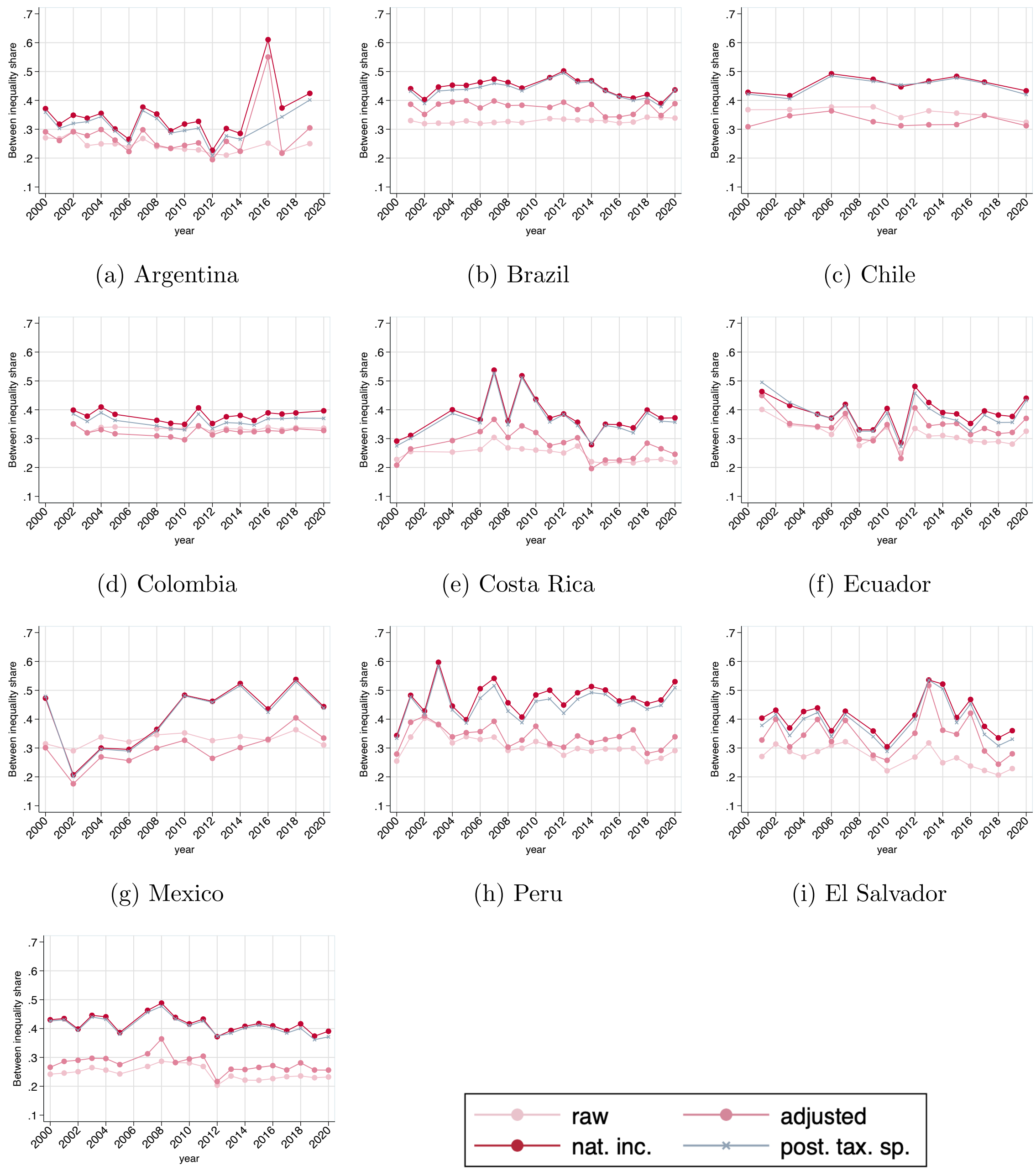
Note The graphs show the share of public education spending attributed to each fractile in the distribution. Source: CEQ (2021).

Fig. C.12. The incidence of health spending



Note: The graphs show the share of public health spending attributed to each fractile in the distribution. Source: CEQ (2021).

Fig. C.13. Contribution of between-group inequality (bottom 99% and top 1%)



(a) Argentina (b) Brazil (c) Chile

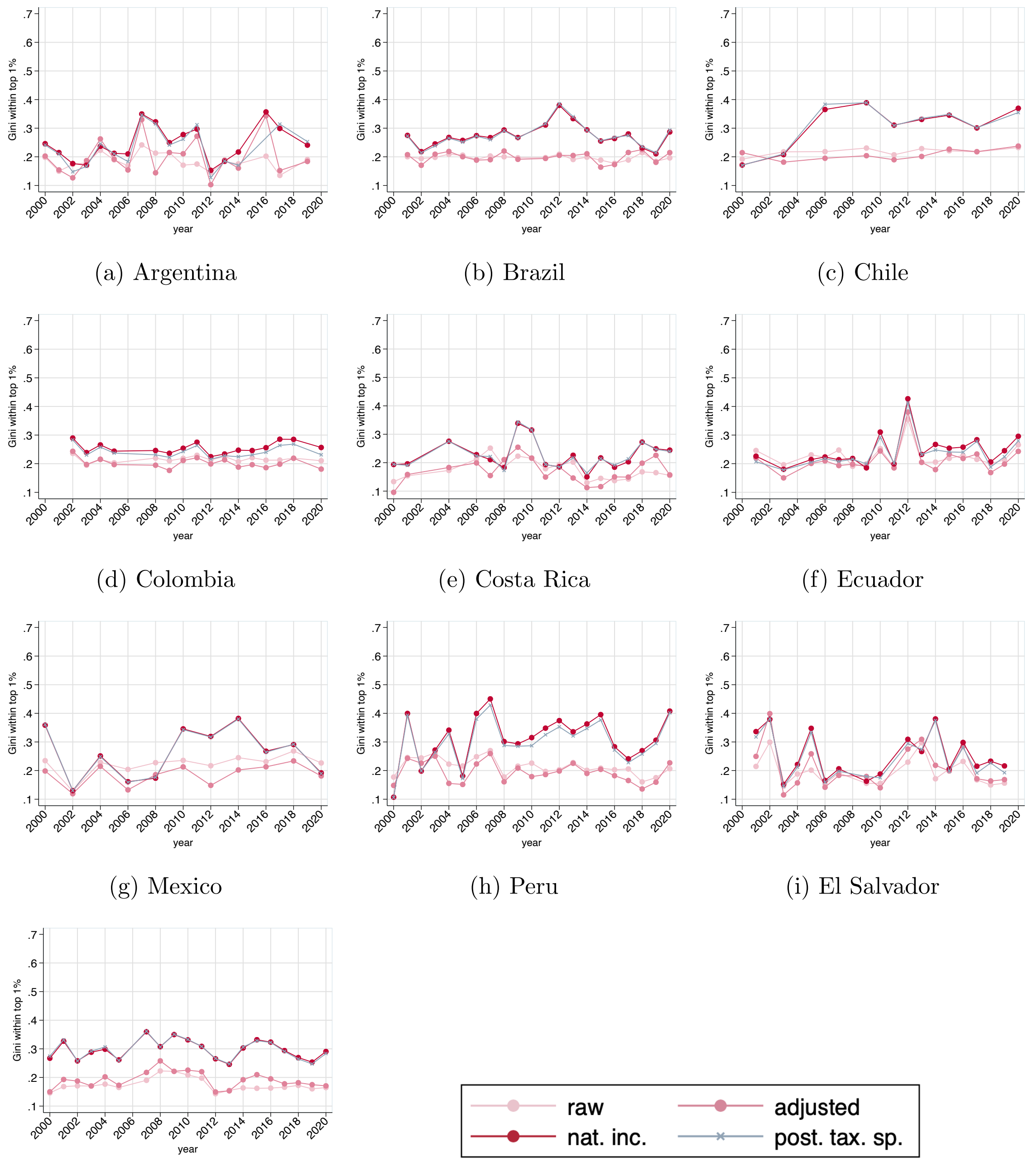
(d) Colombia (e) Costa Rica (f) Ecuador

(g) Mexico (h) Peru (i) El Salvador

(j) Uruguay

Note. Authors' elaboration. The graph shows the contribution of between-group inequality (between the bottom 99% and the top 1%) to total inequality of per capita household inequality using the Theil index decomposition.

Fig. C.14. Within-group inequality (top 1%)



(a) Argentina (b) Brazil (c) Chile

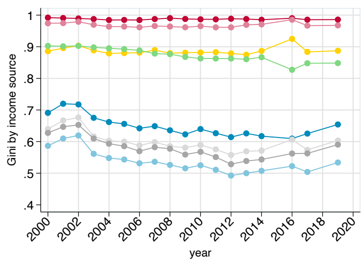
(d) Colombia (e) Costa Rica (f) Ecuador

(g) Mexico (h) Peru (i) El Salvador

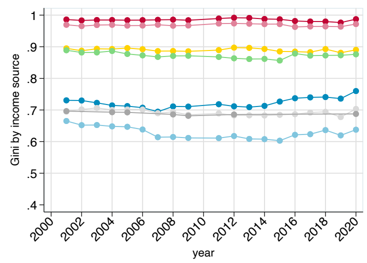
(j) Uruguay

Note. Authors' elaboration. The graph within group inequality of per capita household income among the top 1% using the Gini coefficient.

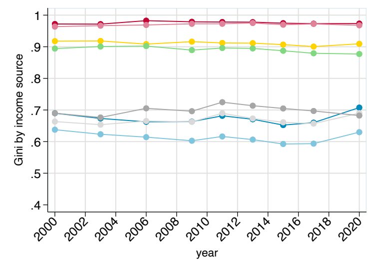
Fig. C.15. Inequality by income source (pre-tax national income)



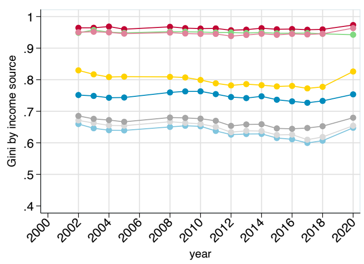
(a) Argentina



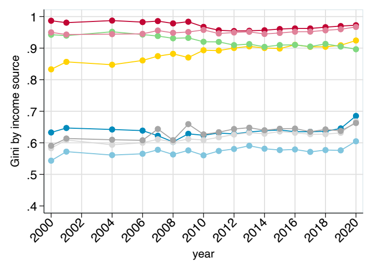
(b) Brazil



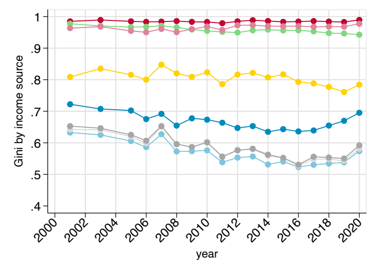
(c) Chile



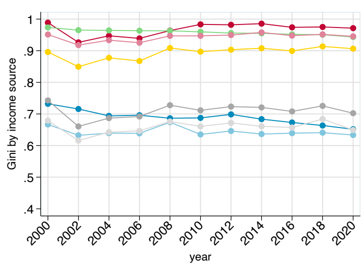
(d) Colombia



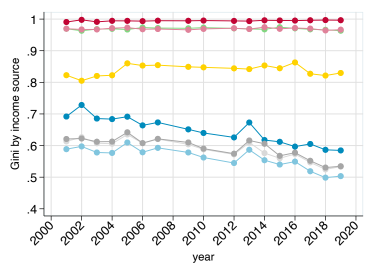
(e) Costa Rica



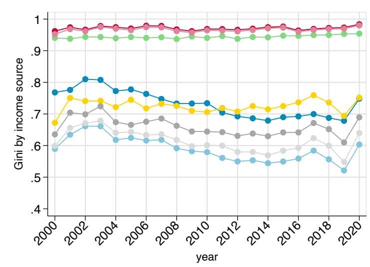
(f) Ecuador



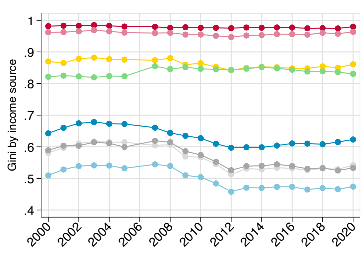
(g) Mexico

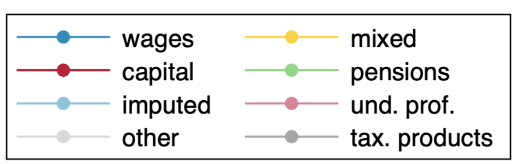


(h) El Salvador



(i) Peru

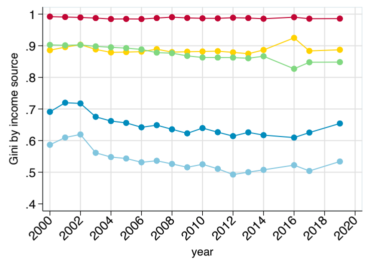




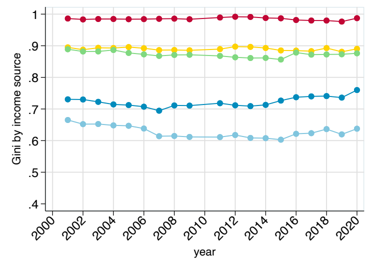
(j) Uruguay

Note. Authors' elaboration. The graphs show the Gini index by source of national income for household per capita units based on Lerman and Yitzhaki (1985).

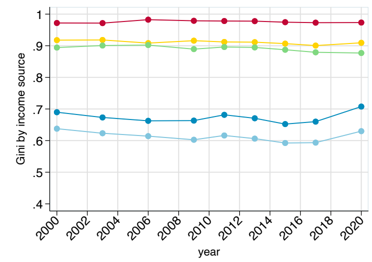
Fig. C.16. Inequality by income source (household sector income)



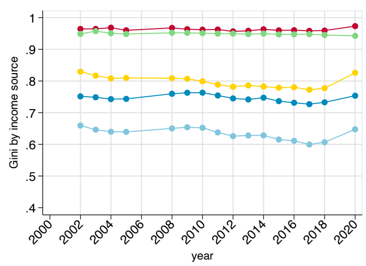
(a) Argentina



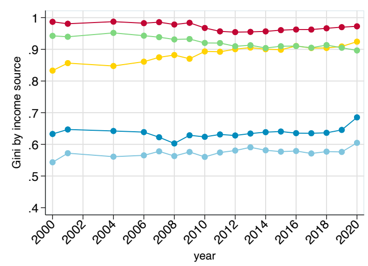
(b) Brazil



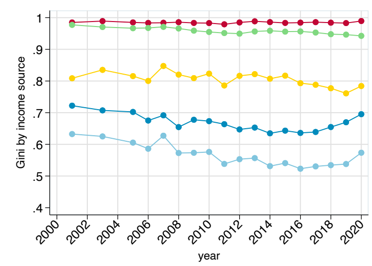
(c) Chile



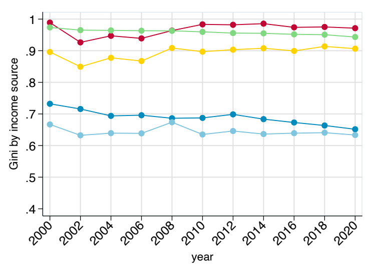
(d) Colombia



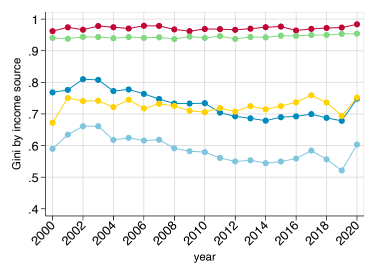
(e) Costa Rica



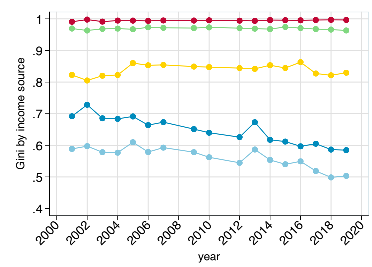
(f) Ecuador



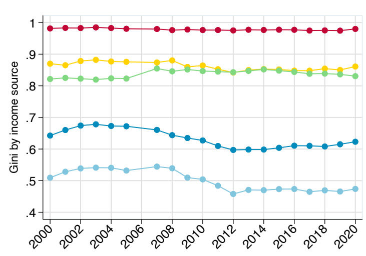
(g) Mexico

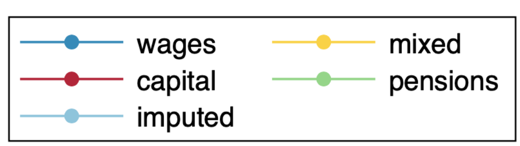


(h) Peru



(i) El Salvador

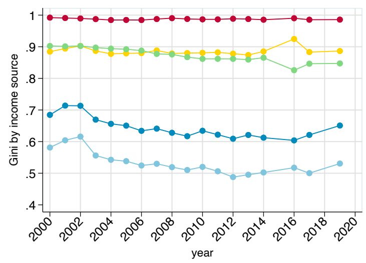




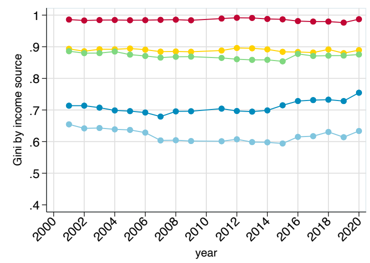
(j) Uruguay

Note. Authors' elaboration. The graphs show the Gini index by source of household sector income for household per capita units based on Lerman and Yitzhaki (1985).

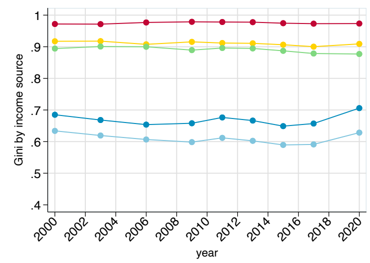
Fig. C.17. Inequality by income source (top-corrected survey)



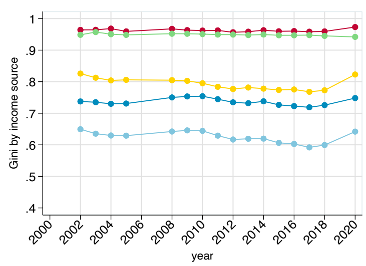
(a) Argentina



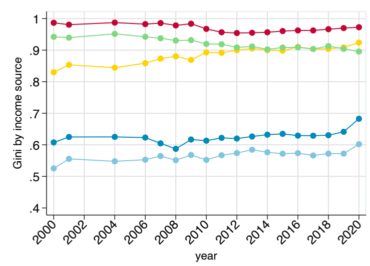
(b) Brazil



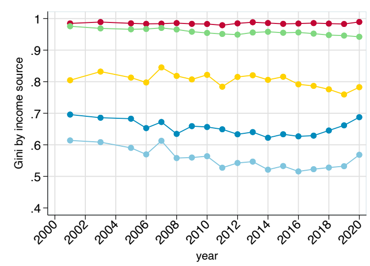
(c) Chile



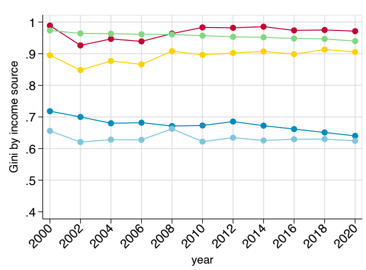
(d) Colombia



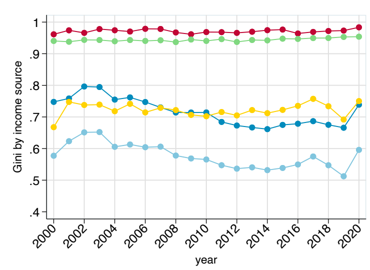
(e) Costa Rica



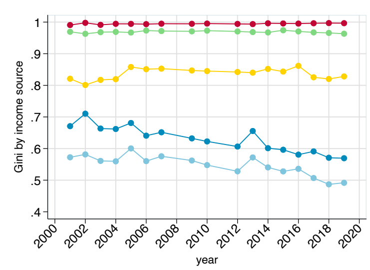
(f) Ecuador



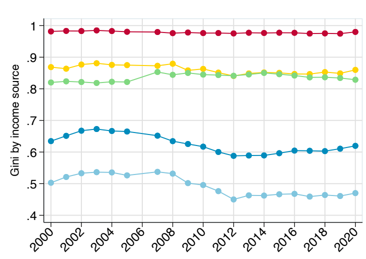
(g) Mexico



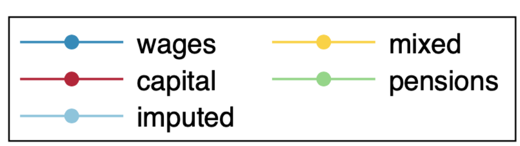
(h) Peru



(i) El Salvador

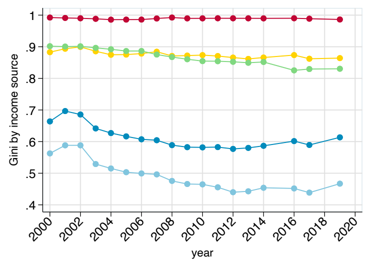


(j) Uruguay

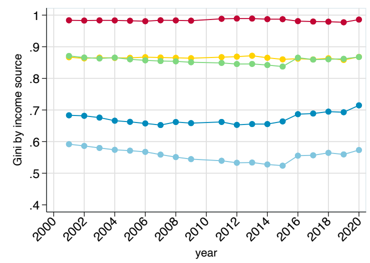


Note. Authors' elaboration. The graphs show the Gini index by source in the top corrected surveys for household per capita units based on Lerman and Yitzhaki (1985).

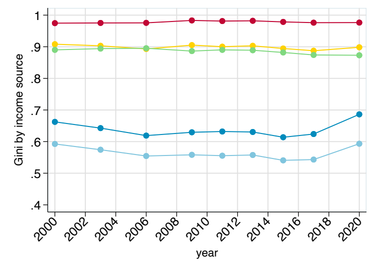
Fig. C.18. Inequality by income source, (raw survey)



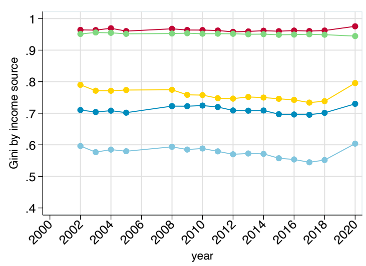
(a) Argentina



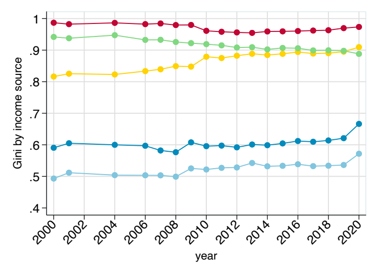
(b) Brazil



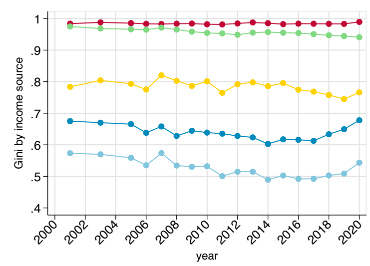
(c) Chile



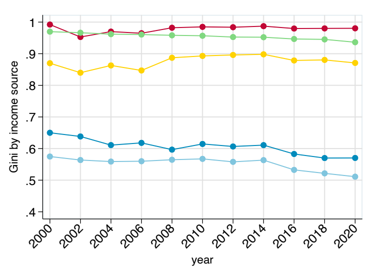
(d) Colombia



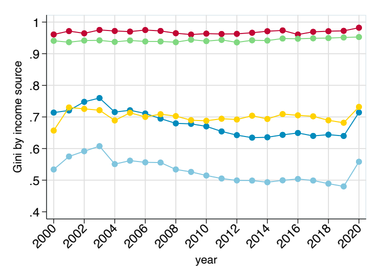
(e) Costa Rica



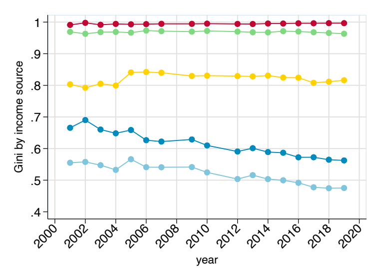
(f) Ecuador



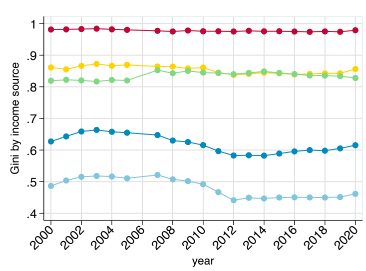
(g) Mexico

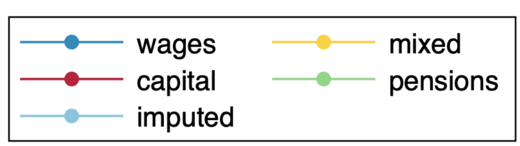


(h) Peru



(i) El Salvador

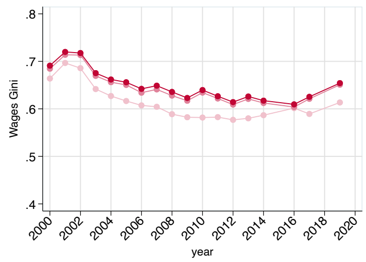




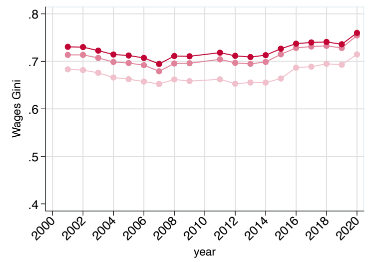
(j) Uruguay

Note. Authors' elaboration. The graphs show the Gini index by source in the raw surveys for household per capita units based on Lerman and Yitzhaki (1985).

Fig. C.19. Gini index of wages



(a) Argentina



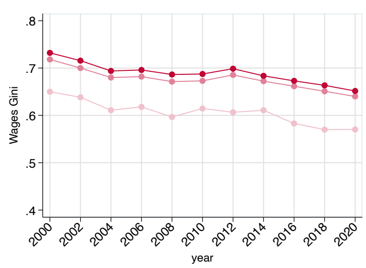
(b) Brazil

(c) Chile

(d) Colombia

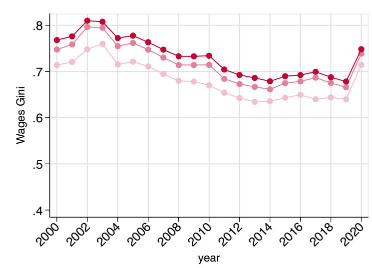
(e) Costa Rica

(f) Ecuador

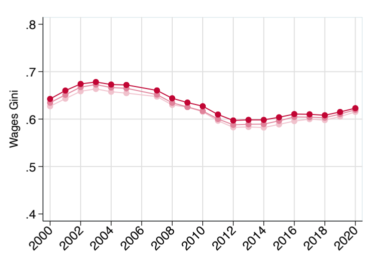


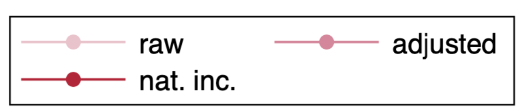
(g) Mexico

(h) El Salvador



(i) Peru





(j) Uruguay

Note. Authors' elaboration. The figure shows the Gini index of the wage distribution in household per capita units.

PART II. Falling inequality in Uruguay

Chapter 3

Falling inequality and the growing capital income share: reconciling divergent trends in survey and tax data

Joint work with Gabriel Burdín, Joan Vilá and Andrea Vigorito.

Abstract

In contrast to the remaining regions of the world, the available evidence from household surveys indicates that most Latin American countries experienced substantial reductions in monetary poverty and personal income inequality in the first 15 years of the 21^{st} century. However, it is still unclear whether these trends are robust to the inequality index and database. Based on a unique array of matched social security and personal and firm income tax records, and household survey microdata, we provide detailed evidence on inequality trends for the period of survey-based inequality reduction in Uruguay (2009-2016), focusing on the top income groups and the evolution of the capital income share. We correct administrative data to account for informality and social security/income tax underreporting. Trends are sensitive to the data source and inequality measure. Synthetic indices decreased in both datasets and the top income shares diverged. This results from increasing inequality in the upper tail of administrative data, mainly driven by a growing share of capital income, and particularly dividends. The probability of reaching top income positions is higher for men, liberal professionals, capital income receivers, and occupations associated to medical services. In contrast to evidence for developed countries, the financial and tech sectors are less represented. These findings have strong implications for the design of public policies aimed to reduce persistent inequalities in developing countries.

1. Introduction

In contrast to the remaining regions of the world, the available evidence from household surveys indicates that most Latin American countries experienced substantial reductions in monetary poverty and personal income inequality in the first 15 years of the 21st century (López-Calva and Lustig, 2010; Cornia, 2014b; Alvaredo and Gasparini, 2015). While this decline was very fast in 2000-2010, it continued at a milder pace in the subsequent 5 years and, in most cases, ended by 2015 in a context of economic slowdown (ECLAC, 2019; SEDLAC, 2019).

However, the findings of the top incomes research based on tax returns, both worldwide (Piketty, 2003; Atkinson et al., 2011) and in Latin America (Alvaredo and Londoño Velez, 2014; Flores et al., 2020; Morgan, 2017b) have reinvigorated the discussion on the validity of household survey data in providing accurate inequality estimates. It is well known that household surveys correctly capture income information of the low and middle strata as well as pension and labour earnings but that they are subject to underreporting and undercoverage at the top end of the distribution and underestimate capital income (Altimir, 1987; Székely and Hilgert, 1999; Cowell and Flachaire, 2015; Bourguignon, 2015; Lustig et al., 2019).

The available literature for developed countries has shown that these draw-backs are particularly important when appraising inequality trends (Piketty, 2003; Atkinson et al., 2011; Jenkins, 2017). Moreover, correctly assessing the evolution of capital income is particularly relevant in a period of rapid economic growth such as the one experienced recently by Latin American countries. If capital income levels or/and shares increased, this phenomenon itself might erode the capacity of household surveys to capture income at the upper tail and could provide a more optimistic view of inequality trends in a region that has been characterized historically by a high concentration of income and wealth (Alvaredo and Gasparini, 2015). Furthermore, the undercoverage of richer strata can lead to wrong evaluations of the redistributive effects of public policies and, in general, of what can successfully reduce inequality. Since persistent inequalities are a major challenge for public policies design, this problem is particularly relevant in the context of developing countries.

Comparisons among household surveys and tax record-based inequality measures are not straightforward due to differences in income definitions and population coverage. Because tax units are individuals in many schemes, top income studies are not able to reconstruct per capita household income, leaving aside homogamy, fertility differentials and other relevant features that affect household conformation and might amplify or mitigate primary income inequality. At the same time, in most cases, administrative data lack information from non-taxable income sources, such as non-contributory cash transfers and other public benefits. Thus, reconciling these two strands of the literature requires access to micro-data from household surveys and tax records to carry out a careful harmonization process (Burkhauser, Feng, Jenkins, and Larrimore, 2012; Burkhauser, Hérault, Jenkins, and Wilkins, 2018).

In this article, we investigate whether the recent inequality fall in Uruguay is robust to the use of different data sets and whether it implies modifications of the shares held by top income groups, and, particularly, capital income receivers. Specifically, we analyse primary income inequality, comparing harmonized household surveys and corrected micro-data from tax records. We provide an in-depth analysis of the main factors underlying the evolution of the income distribution in the two data sets, focusing on the upper tail and the evolution of the different income sources. We delve into the characteristics of the top income earners and the firms that they work for or own, which also allows us to account better for capital income' shares. Uruguay is an interesting case study because we are able to exploit a unique data set of matched social security data and personal and firm income tax records at the individual level that covers the period of significant GDP growth and inequality decrease (2009 to 2016) (Figure A.1).[1]

This research is mainly based on a comprehensive anonymized administrative personal income tax micro-database (*Impuesto a la Renta de las Personas Físicas (IRPF) and Impuesto a la Seguridad Social (IASS)*) matched to the balance sheets that corresponding firms submitted to the tax authorities (*Dirección General Impositiva, DGI*) in 2009-2016. The latter step is necessary to identify completely the capital incomes and characteristics of employers. Since they include information from social security records, these data cover the universe of formal workers (with earnings below or above the minimum taxable income), capital income earners and pensioners, comprising around 75% of the adult population aged 20 and above. At the same time, we use the micro-data from the official household survey *Encuestas Continuas de Hogares, ECH)* gathered by the *Instituto Nacional de Estadística (INE)* and a subsample of 2012/2013 ECH-DGI observations linked at the individual level to compute the underreporting rates in the lower tail of administrative data. The broad coverage of our administrative micro-data and the availability of a unique data set of survey-tax data matched at the individual level for a sub-set of households allow us to depart from the tax records and correct the lower half of the income distribution with household survey

[1] Household survey information reveals that inequality was constant from 1986 to 1997, started to increase in 1998, peaked with the severe economic crisis in 2002 and remained steady from 2003 to 2008 (Amarante, Colafranceschi, and Vigorito, 2014).

information, building on the methodology initially proposed by Atkinson (2007). Specifically, we add labour earnings from informal workers and underreported formal income, creating a corrected tax income variable. We also present several robustness checks by correcting harmonized household survey income with tax data (Alvaredo, 2011; Blanchet et al., 2022). To identify the main characteristics of top income receivers, we carry out a multivariate analysis exploiting the matched individual-firm databases.

Our findings indicate that the synthetic indexes present declining trends in corrected tax income and harmonized survey income and, in both cases, inequality declined at the bottom 99%. However, the driving forces under the inequality reduction are at odds in the two cases. While the equalization process in the harmonized household survey income was lead by a reduction in the concentration of the top 1%, the opposite applies to corrected tax income, in which the redistribution in the bottom 99% outweighed the increasing inequality at the top. In the latter case, the inverted Pareto coefficient has grown steadily since 2012. As a result, the top income shares exhibit a decline in harmonized household surveys and an increase in corrected personal income tax data.

We also show that the evolution of the top income shares in corrected tax income is closely connected to the increased participation and concentration of capital income in the upper tail of the income distribution. Furthermore, we document that the top income holders are closely connected to the increased share of capital income in the top 1% and 0.5% of the income distribution. Most top income holders are men and capital income receivers. Meanwhile, among the subset of top income earners receiving labour income, the most salient group corresponds to health services.

This study contributes to three main avenues of the existing literature. First, we provide further evidence on the evolution of primary income inequality for a Latin American country. The available top incomes studies for Argentina, Brazil, Chile, Colombia and Uruguay cast doubts on the magnitude of the recent inequality reduction and, in some cases, even on its trend (Alvaredo, 2010; Alvaredo and Londoño Velez, 2014; Flores et al., 2020; Morgan, 2017b).[2] Compared with previous studies, we undertake a broader reconciliation exercise. To our knowledge, this is the first study to provide a detailed account of the differences in the evolution of inequality and top incomes in Latin America observed in household surveys and tax records, correcting the lower tail of administrative micro-data to account for

[2] In the case of Uruguay, previous studies for a shorter time span have also concluded that income inequality estimates based on tax and survey data, although not showing opposing trends, did not fully coincide (Burdín, Esponda, and Vigorito, 2014; Burdín et al., 2022). Even though the conclusions are qualitatively similar overall to the ones reached in the present article, the time span was shorter and the data were less comprehensive.

underreporting and informal income.

Even though synthetic inequality indices show similar trends in the two data sets, top incomes in the corrected tax income series remained almost steady and slightly grew at the end of the period under analysis. These findings suggest that the Uruguayan redistribution process occurred in the lower and middle strata and coexisted with increasing share and concentration of capital income at the top of the distribution.

Second, we show that household surveys indicate a reduced capacity to reach the top of the distribution, which might be connected to the increasing participation of capital income and the subsequent concentration observed in the upper tail. Although we cannot generalize our results to other Latin American countries, our exercise illustrates the limits of the recent redistributive process and casts doubts on the validity of assessments that rely only on household survey data.

Third, for the first time, we provide evidence of the characteristics of top income earners in a developing country. The scarce representation of women among the top income holders is in line with previous studies on developed countries (Aaberge and Mogstad, 2015; Hansen, Harmenberg, Öberg, and Sievertsen, 2021). However, different from the findings reported by Bertrand and Mullainathan (2001), Bivens and Mishel (2013), Kopczuk and Zwick (2020) and Smith et al. (2019), top income holders are mainly capital income receivers and the growing share of capital income (and particularly dividends) is the driving force underlying the increase in top income shares. The predominance of capital income in the upper tail of primary income distribution is in line with previous work by Alvaredo, Atkinson, Piketty, and Saez (2013) for Colombia, suggesting that rentiers rather than CEOs hold the top income positions in Latin American countries. Our multivariate analysis shows that the probability of reaching top income positions is higher for men, liberal professionals, capital income receivers, and occupations associated to health activities. In contrast to the findings by Lemieux and Riddell (2015) for Canada, the financial and tech sectors are scarcely represented at the top.

The remainder of this article is organized as follows. Section 2 reviews the previous research on inequality and top incomes shares in Latin America and Uruguay. Section 4 describes the data sources and methods used in this study. Section 4 presents the main inequality estimates across income definitions and data sources, while section 5 attempts to reconcile the divergent trends. Section 6 documents the growing share of capital incomes at the top, and presents some distinctive features of the top income groups, and finally section 7 concludes.

Additional information can be found in Appendices.[3]

2. Inequality and top incomes shares in Latin America: recent evidence from survey and tax records data

To overcome the caveats of household surveys' ability to capture top incomes, in the last decades, distributional studies have revivied the tradition of analysing personal income tax records (Feenberg and Poterba, 1993; Piketty, 2003; Atkinson, 2007; Atkinson et al., 2011; Alvaredo et al., 2013). These studies have shown that, even when high income groups by definition represent a very small fraction of the population, not only can the top income share levels and trends be different but also synthetic inequality measures, such as the Gini index, have proved to be sensitive to misreporting and survey undercoverage at the upper tail of the income distribution (Leigh, 2007; Alvaredo, 2011).[4]

However, tax records also present many caveats that have been acknowledged in the related literature. Due to informational constraints, most assessments based on administrative data can only analyse primary income inequality among individuals.[5] At the same time, administrative data are subject to tax evasion and avoidance, as well as behavioural responses to changes in tax rates (Atkinson et al., 2011; Feenberg and Poterba, 1993).[6] The challenges are even larger in developing countries, where informal workers represent a large proportion of the labour force and personal tax systems are not fully developed. Thus, recent studies have moved in two main directions: i) creating harmonized income variables to carry out accurate comparisons among different data sources to assess inequality trends correctly and ii) developing methodologies to combine survey and tax data properly.

Regarding i), Burkhauser et al. (2012) analysed the inequality trends in household surveys and personal income tax data for the United States in 1967-2006, previously harmonizing the Current Population Survey to make it consistent with the administrative data. They

[3] Appendix 2 is an online supplement that mainly contains additional information on the databases used in this study.

[4] In spite of this, Leigh (2007) argued that the top 1% estimates are a good proxy for Gini index rankings across countries.

[5] Depending on the tax regime and the definition of taxable income, in most cases this information does not allow us to reconstruct households (which might be the relevant unit for many assessments and, particularly, for public policy design) and leaves aside non taxable income sources, such as cash and in-kind transfers.

[6] For instance, Feenberg and Poterba (1993) assessed the participation of top income groups in the United States based on personal income tax information between 1951 and 1990 and showed that the rise in top income shares was partly driven by a substantial reduction in the top marginal tax rates from 70% to 28% implemented in 1986, which affected the evasion rates at the top.

found that, once income and tax units are defined consistently across data sources, the differences decrease, even though modifications to the tax system and survey design may explain differential trends in some periods. A limited number of earnings validation studies, relying on survey-tax linked data at the individual level, have identified a mean reversion pattern in reported income, with survey information yielding higher incomes at the bottom of the income distribution and lower values in the upper tail (Abowd and Stinson, 2013; Adriaans, Valet, and Liebig, 2020). This reporting pattern has been associated with cognitive difficulties, social desirability behaviours, off-the-book payments and informality (particularly at the bottom of the distribution).

The recent literature addressing ii) has been progressing in providing a common ground by developing new methods that combine household survey and tax data to ensure that the upper tail is captured properly(Jenkins, 2015; WIL, 2020; Piketty, Yang, and Zucman, 2019; Anand and Segal, 2017). However, to date, there is no consensus on the "true" distribution, which largely depends on researchers' priors (Abowd and Stinson, 2013), and there is an ongoing discussion on the appropriate correction methods. While some studies have departed from tax data and supplemented them with household survey information, other studies, relying on reweighting and replacing methods, have corrected the upper tail of household survey data with information from tax data and, in some cases, fitted a parametric distribution at the top (see, for instance, Jenkins (2017), Blanchet et al. (2022) and Lustig et al. (2019)).

In Latin America, the first attempts to correct household survey income underreporting can be traced to Altimir (1987)'s adjustment to national accounts, which was included in the official inequality estimations provided by the Economic Commission for Latin America (ECLAC). However, this methodology has proven to have many caveats (mainly concerning the quality and paucity of national accounts information), and ECLAC discontinued this procedure in 2019.

Despite the longstanding Latin American tradition in distributional studies, research focusing on the top income groups has been less frequent, partly due to scarce data availability and the weaknesses of personal income taxation in the region. To date, there is evidence for Argentina (Alvaredo, 2010), Colombia (Alvaredo and Londoño Velez, 2014), Brazil (Souza and Medeiros, 2015; Morgan, 2017b), Chile (López, Figueroa, Gutiérrez, et al., 2013; Fairfield and Jorratt De Luis, 2016; Flores et al., 2020) and Uruguay (Burdín et al., 2014; De Rosa and Vilá, 2022). However, most of these studies covered a shorter period than the scholarship on top incomes for developed countries and either relied on tax data tabulations or were based on micro-data that covered tax-payers only or the upper income strata.

In a recent study, De Rosa et al. (2020) provided inequality estimates for ten Latin American countries by correcting household survey information with tax data (before scaling up to national income components), based on the reweighting methodology developed by Blanchet et al. (2022). They found mixed evidence regarding the recent inequality decline. Specifically, in the cases of Argentina, Colombia, Ecuador and Uruguay, the results are robust to the correction, whereas, in the case of Brazil, the findings are similar those presented by Morgan (2017b) (see below).

In-depth studies on specific countries, comparing survey and tax data, have concluded that inequality trends are sensitive to the data source and inequality measure (Table 2.1). For instance, Alvaredo and Londoño Velez (2014) found that the top income shares in Colombia remained steady (at around 20%) in the period in which household survey-based Gini indices fell (2006-2010), even after correcting for underreporting in the upper tail. In turn, Flores et al. (2020) identified opposite trends for Chile, with an increase in tax-based top income shares since 2000 and a decline in household surveys. Souza and Medeiros (2015) analysed the case of Brazil during the period 2006-2012 and concluded that the inequality indices remained steady, with the top income shares representing around 25% of the total income throughout the whole period. However, more striking results came from Morgan (2017b), who, using the Blanchet et al. (2022) correction, analysed a longer span and found an increasing trend or, at best, a steady income concentration level in Brazil, contradicting most of the previous research based on household survey data, which unanimously identified a consistent and long period of rapid inequality decline (Lustig et al., 2011; Barros, Foguel, and Ulyssea, 2006). It is noteworthy that this study also reported a decline in labour income inequality, which is consistent with the previous literature and with the income sources mainly captured by household surveys. Since previous studies on Latin American countries were not able to exploit micro-data for a significant fraction of the population, the corresponding comparisons used the Alvaredo (2011) correction and did not include tax record-based synthetic inequality indices. In sum, the existing evidence on the robustness of the recent decrease in inequality in Latin America is not conclusive.

Table 2.1: Top income shares and Gini indices in Latin American countries: circa 2000 and 2015

Country	Year	Top 1% share (primary income)	Source	Gini coefficient
Argentina	2001/06	14.3 / 16.8%	Alvaredo (2010)	0.504 / 0.493
Brazil	2001/15 2005/12	26.3/ 27.5% 22.7 / 26.4%	Morgan (2017b) Souza and Medeiros (2015)	0.583 / 0.513 0.556 / 0.526
Chile	2000/15	20.2 / 23.7%	Flores et al. (2020)	0.526 / 0.448
Colombia	2007/10	20.7 / 20.4%	Alvaredo and Londoño Velez (2014)	0.59 / 0.554

Note. The sources for the top income share's estimations (primary income) are Alvaredo (2010); Morgan (2017b); Souza and Medeiros (2015); Flores et al. (2020); Alvaredo and Londoño Velez (2014). Income shares are calculated according to fiscal income. Gini indices based on household surveys are available from SEDLAC (2019) and correspond to per capita household income.

3. Data and methodology

In this section we first describe the main features of the data-bases used in this research (3.1) and then present the methods implemented to estimate top incomes shares and the remaining inequality measures (3.2).

3.1. Data

To account for the Uruguayan population aged 20 years and more, we combine personal and firm income tax with household surveys micro-data. Table A.1 summarizes the population coverage and income definition for each data source.

3.1.1. Income tax micro-data

The Uruguayan personal income tax is based on a dual scheme that consists of two separate progressive tax schedules for labour income and pensions (*Impuesto a la Renta de las Personas Físicas (IRPF) cat. II* and *Impuesto de Asistencia a la Seguridad Social*, (IASS)), and a flat tax rate on capital income (IRPF cat. I).[7] There is also a separate corporate income

[7]Personal income tax was originally established in 1961 but, jointly with inheritance taxation, was abolished in 1974 by the *de facto* regime that ruled Uruguay during 1973-1985. Framed in an overarching tax reform, it was restored in 2006. Although pensions were originally included in IRPF, soon after the reform this component was declared unconstitutional. As a result, a new progressive tax on pensions with similar characteristics was passed in July 2008 (IASS).

tax scheme that taxes dividends and profits at a 25% flat rate (*Impuesto a la Renta de las Actividades Económicas*, IRAE). The tax schedule remained unchanged throughout the period 2009-2016, except for a relatively small tax increase for the top income brackets in 2012 (the tax rates can be found in Tables B.1, B.2 and B.3).[8]

In most cases, labour taxes are withheld by employers, who transfer the corresponding payments to the Social Security Institute (*Banco de Previsión Social, (BPS)*). Only the self-employed or those workers with more than one occupation (and an annual income above 16,000 USD) have to file a tax return. Self-employed workers contribute for their full (non salaried) labour income and are entitled to deduct up to 30% of their income. Although tax units are individuals, married couples can fill a joint labour income tax return however, in practice, only 1.8% of taxpayers choose this regime.

DGI created anonymized databases for research purposes that put together two administrative data sources: (a) the universe of IRPF and IASS tax payers for 2009-2016, which contained detailed information on capital, pension and labour income for each occupation, tax burden and deductions (Table B.4); (b) the universe of monthly labour income and pensions payments from social security records (provided to the DGI by the BPS) corresponding to formal workers and pensioners.[9] As the BPS withholds income tax payments for workers and pensioners, DGI information comprises pensioners and the universe of workers contributing to social security, regardless of whether they are net tax-payers. At the same time, each record contains information on sex, age, industry and type of employer (salaried or self-employed). Additionally, DGI provided a supplementary database with information on income and taxes corresponding to the personal services societies that chose to pay corporate income tax (IRAE) instead of IRPF (see the IRAE row in Table B.4). This option is available for liberal professionals and, thus, these earnings can be assimilated either to mixed or to income. The resulting micro-data covers 75% of the population aged 20 years and above.[10]

We group capital income into the following categories: profits and dividends, real estate rents, interest from bank deposits and other concepts (sports persons royalties, authors royalties and everlasting rents). Like most top incomes studies, we exclude capital gains from our analysis. Due to the Bank Secrecy Act and to previous regulations that allowed firms to

[8]Recent evidence has suggested that this change did not result in a major reduction of reported income after the reform, and, therefore, did not affect the top income shares estimations, although it may have had a minor impact on the income composition for some groups of taxpayers, (Bergolo, Burdin, De Rosa, Giaccobasso, Leites, and Rueda, 2019).

[9]The Uruguayan fiscal year corresponds to the calendar year.

[10]The remaining 25% corresponds to informal workers (38.9%), and people who are unemployed (10.9%) or out of the labour force, who are not receiving pensions or capital income (50.2%).

issue bearer shares, we do not have access to micro-data on interests from bank deposits and non-nominative dividends.[11] Table B.5 shows that while interest is not a relevant concern, non-nominative dividends account for half of the total dividends.[12] Since we lack information on the characteristics of non-nominative profit receivers, to minimize the potential reranking among capital earners, we distribute the total amount among individuals in the tax record micro-data proportionally to the total capital income held by each individual.[13]

It is worth pointing out that the analysis presented in this article excludes dividends accrued by non residents. From Table B.6, it is apparent that, as many firms are owned by international corporations and non-residents, a significant fraction of the profits generated in Uruguay are taxed according to a different scheme, the *Impuesto a la Renta de no Residentes* (IRNR).[14] Notice that throughout the period, assuming an IRNR tax rate of 12%, dividends remitted abroad represented between 1.3 and 4 times those held by residents. Compared with the full amount of capital income, these shares varied between 57% and 80%. These figures suggest that a substantial proportion of the capital income generated in Uruguay does not remain in the country.

Even if tax records are available, identifying capital income correctly can be difficult due to the design of the tax systems and particularly the interplay between firm and personal income taxation.[15] It is noteworthy that in Uruguay, firms were allowed to keep undistributed profits that were not reinvested without any time limit until 2017. Thus, to avoid filing a personal income tax return declaring distributed profits or dividends (taxed at a 7% rate additional to the 25% rate on corporate income), many firm owners took cash advances. As these withdrawals have to be singled out on balance sheets as a separate concept, *advance*

[11] Non nominative dividends are profits distributed by firms of which the owners are anonymous, and, thus, it is not possible to identify the receiver in DGI data-base. The DGI provided the total amount of dividends that fall into this category.

[12] In recent years, to comply with the international regulations set by the Basel Agreement, Uruguay has restricted the issuance of bearer shares. In spite of this policy change, the share of non-nominative dividends remained steady in the period under analysis. Thus, potential trespassing from non-nominative to nominative profits does not seem to be a relevant concern here.

[13] As shown by De Rosa, Sinisclachi, Vilá, Vigorito, and Willebald (2018), very few firms declare distributed profits. Therefore, imputing non-nominative profits only to nominative profit receivers, is likely to overestimate the concentration of capital income. By distributing it in proportion to the total capital income, the capital income distribution remains unchanged.

[14] In 2008, the annual influx of foreign direct investment was around 5.5% of the GDP (Bittencourt, Carracelas, Doneschi, and Reig Lorenzi, 2009; Chudnovsky and López, 2007). In the time span covered in this study, at least 13% of the firms were owned by non-residents (Peluffo, 2015).

[15] For instance, in their study on Chile, Fairfield and Jorratt De Luis (2016) and Flores et al. (2020) used information from individuals and firms tax returns and imputed accrued profits and accumulated undistributed profits to taxpayers using ownership shares that were directly estimated from businesses tax-return forms. These studies indicated that although the inequality levels are extremely sensitive to this procedure, trends do not vary.

payments, we are able to partially reconstruct the actual distribution of capital income had these payments in advance been declared as distributed profits.[16] Unsurprisingly, our estimations convey a low number of profit withdrawals per year (fewer than 10% of the firms distributed benefits). Nevertheless, throughout the whole period, the total amount of profit withdrawals in DGI is considerably higher than the amount that we obtain in ECH. As shown in B.8, in 2009 and 2016, individuals receiving in advance payments respectively represented 188% and 146% relative to distributed profits.

As in most tax record based research, in Uruguay tax units are individuals and we cannot reconstruct households. Because they are not included in the taxable income definition, we also do not consider relevant income sources such as the value of owner-occupied housing and private and non-contributory public transfers.[17]

3.1.2. The Uruguayan household surveys

The National Statistical Office (INE) gathers household survey (*Encuestas Continuas de Hogares*, (ECH)) since 1968. At present, ECHs are nationally representative and are carried out throughout the whole year. They collect information in detail on household composition, labour force status and employment characteristics, socioeconomic variables and personal income by source. The sample design and further methodological details can be found in Instituto Nacional de Estadística (2021).[18]

After-tax labour income is gathered for each household member aged 14 years or above, including cash and in-kind payments for salaried workers, self-employed workers and business owners (separately recording the main occupation and the remaining ones). The survey also collects information on the contributory status of employed workers in each occupation. After tax pensions are collected separately for each individual.

The questionnaire also collects interest, dividends, rents, benefits and the imputed value of owner occupied housing. Except for profit withdrawals reported by self-employed workers and business owners, capital income is captured in the household questionnaire, which implies

[16] However, corporate tax declarations and balances are available only for the sub-set of firms with revenues above US$40.000 per month (around 60% of registered firms).

[17] Many studies indicate that both factors are relevant in Latin America. Besides, the increased coverage of cash transfers contributed to the recent reduction of inequality (López-Calva and Lustig, 2010; Cornia, 2014b; Alvaredo and Gasparini, 2015). Moreover, in the case of Uruguay, household survey based studies conclude that the static contribution of child benefits and other cash transfers is similar to the equalizing effect of the personal income tax (Bucheli, Lustig, Rossi, and Amábile, 2013; Amarante et al., 2014).

[18] Sample size was 46,550 households and 120,781 individuals in 2009 and 46,669 households and 128,204 in 2016.

that each item is added up for the whole household and attributed to the household head.

As in other regions, the accuracy of household surveys in capturing incomes has been the subject of a longstanding discussion in Latin America (Altimir, 1987; Székely and Hilgert, 1999). In the same vein, during the 1990s, several studies analysed the accuracy of ECH in capturing household income by source compared with the national accounts and expenditure surveys (Groskoff, 1992; Mendive and Fuentes, 1996; Amarante and Carella, 1997). More recently, Amarante, Arim, and Salas (2007) found that ECH captures 39.7% and 23% of the total amount of housing rents and interest on bank deposits. Based on an ECH subsample of households with children aged 0 to 3 that gathered ID numbers and was merged with tax records, Higgins, Lustig, Vigorito, et al. (2018) and Flachaire, Lustig, and Vigorito (2022) harmonized household survey formal income to make it comparable with tax records, and identified the expected misreporting pattern (Abowd and Stinson, 2013): underreporting in DGI income below the median and underreporting in ECH income thereafter. For the top 1%, ECH captures around 56% of DGI income.

Thus, if we only correct DGI income to account for informal income, we are still losing misreported formal income at the bottom of the distribution and we could overestimate inequality. To account for this problem, we also use information from the Nutrition, Child Development, and Health Survey (*Encuesta de Nutrición, Desarrollo Infantil, y Salud*, ENDIS; (Instituto Nacional de Estadística, 2013, 2021)). ENDIS follows households with children aged 0 to 3 that were originally included in ECH between February 2012 and August 2013 and gathered information on the unique national identification number (*cedula*) of the respondents, and, in this way, INE and DGI were able to merge all adults from the 2012-13 ECH that were also in ENDIS, to tax records and provided an anonymized data-set for research purposes. 1,471 individuals have positive harmonized formal income in the two datasets and are the ones we use to compute differences in labour earnings from formal occupations (see Flachaire et al. (2022) for details).

To harmonize ECH information with the income tax micro-data, we compute formal and informal labour earnings, pensions and capital income on an individual basis and restrict income sources to the ones captured by DGI micro-data according to the definition of taxable income (see Appendix B for details). Additionally, we use two ancillary tables created using ECH data. The first one is computed on the basis of the ECH-ENDIS linked tax data sub-sample and contains misreporting ratios by tax income percentile and available for 2012/2013 only. The second one identifies the extent of overlapping among formal and informal income in ECH by computing informal/harmonized formal income in ECH ratios using DGI percentile tax thresholds for each year.

3.2. *Variables of interest: corrected income and population control*

As we are particularly concerned with reconciling inequality trends in household surveys and tax data, and the previous literature has pointed out that the differences rely heavily on undercoverage of the upper tail, we depart from DGI data and supplement it with ECH information to account better for informal income and misreporting in the lower tail. This option is feasible because of the wide population coverage of DGI data. Furthermore, as mentioned in previous sections, evidence from Uruguayan linked data suggests that since underreporting starts in the median of the income distribution, the advantages of departing from the household survey are not clear as we are not attempting to reconstruct households, use ECH covariates or assess the impact of redistributive policies targeting the lower tail of the distribution.

Thus, adapting the methodology to estimate the top income shares based on tax records developed by Atkinson (2007), we depart from tax data and add survey information to create full income distributions that allow us to compute income and population control totals, quantile shares and synthetic inequality measures. We also carry out two robustness checks by correcting survey data with tax information to account for underreporting in the upper tail, implementing the corrections proposed by Alvaredo (2011) and Blanchet et al. (2022).

3.2.1. *Population control*

Since tax micro-data represent formal workers, capital income earners and pensioners, computing of income shares (and inequality measures in general) requires the definition of a reference population. The standard practice in top incomes research is to consider the population projections of individuals aged 15 to 20 years and above. Since most top income studies on Latin America consider the latter, we follow this practice. Besides, the number of observations in DGI micro-data in the age interval 15-19 is extremely low.

Uruguayan tax records account for around 75% of the population aged 20 and above (Table 3.1).[19] As we show in detail in section 3.2.2, we carry out a set of adjustments to account for the total number of income earners and adults in labour force.

[19]One of the facts explaining the broad coverage of the adult population of the data base used in this study derives from the fact that informality rates in Uruguay are lower than in most Latin American countries. In 2009 social security coverage rates were 67.8% of total workers and 80.6% among salaried workers, in 2016 these figures rose to 74.7% and 87.9% respectively.

Table 3.1: Population control

		Population control							
		2009	2010	2011	2012	2013	2014	2015	2016
1	Total population (¿ 20)	2,348,300	2,370,788	2,390,888	2,410,258	2,430,379	2,451,739	2,474,284	2,497,361
2	Tax unadjusted	1,840,111	1,842,057	1,917,702	1,914,829	1,973,759	2,003,804	2,017,146	2,019,465
3	Survey unadjusted	760,713	743,279	697,776	687,517	686,487	676,524	692,600	710,096
4	Informal	369,224	368,758	338,103	323,440	317,494	313,705	314,273	327,252
5	Inactive	391,489	374,521	359,673	364,077	368,993	362,819	378,327	382,844
6	Total unadjusted (tax + survey)	2,600,824	2,585,336	2,615,478	2,602,346	2,660,246	2,680,328	2,709,746	2,729,561
7	Excess of population (%)	10.8%	9.0%	9.4%	8.0%	9.5%	9.3%	9.5%	9.3%
8	Tax unadjusted	1,840,111	1,842,057	1,917,702	1,914,829	1,973,759	2,003,804	2,017,146	2,019,465
9	Survey adjusted	508,315	528,857	472,301	495,431	456,739	448,163	458,216	477,885
10	Informal	116,826	154,336	112,628	131,354	87,746	85,344	79,889	95,041
11	Inactive	391,489	374,521	359,673	364,077	368,993	362,819	378,327	382,844
12	Survey population adj.	-33%	-29%	-32%	-28%	-33%	-34%	-34%	-33%
13	Informal population adj.	-68%	-58%	-67%	-59%	-72%	-73%	-75%	-71%
14	Tax adj. (months w/income)	1,649,109	1,662,313	1,729,522	1,741,108	1,796,395	1,947,126	-	-
15	Survey adjusted	706,912	715,121	667,823	678,038	644,099	572,252	-	-
16	Informal	315,423	340,600	308,150	313,961	275,106	209,433	-	-
17	Inactive	391,489	374,521	359,673	364,077	368,993	362,819	-	-
18	Survey population adj.	-7%	-4%	-4%	-1%	-6%	-15%	-	-
19	Informal population adj.	-15%	-8%	-9%	-3%	-13%	-33%	-	-

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

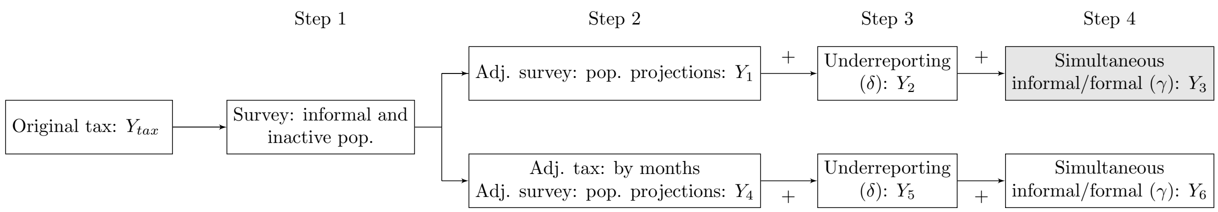
3.2.2. Income variables

Atkinson et al. (2011) proposed two main methods to estimate top incomes shares when tax data are available. Departing from the population control, most top income studies used the first variant and estimated the total income held by a certain quantile according to tax records and compared it to National Accounts System (SNA) information on income totals. However, in Uruguay national income estimations by institutional sector were discontinued from 1997 to 2012. In addition, we are able to work with social security records matched with personal income tax records combined with firms micro-data. Thus, our preferred option is the second procedure proposed by Atkinson (2007), that can be used when administrative data have a large coverage of the population control, as in the case of the Netherlands. This method combines tax and survey micro-data (henceforth Method 1). To check the robustness of our results, we also use the limited SNA information for the sub-period 2012-2016 (henceforth Method 2).

Based on corrected DGI micro-data, we computed the pre- and post-tax top income shares, the synthetic inequality indices (Gini and Theil) and the corresponding between group and income source decompositions (Shorrocks, 1981; Lerman and Yitzhaki, 1985; Shorrocks, 1999; Boschini, Gunnarsson, and Roine, 2020). We also include confidence intervals, calculated by boostrapping the corresponding inequality measures.

Figure 3.1 presents a general overview of the steps that we follow to create the set of corrected income variables and aggregates used in this study. The main purpose of our correction is to adjust the lower tail of the tax records distribution, in order to account for informal and simultaneous formal/informal income. Thus, we depart from the tax records database (Y_{tax}) that includes the universe of individuals receiving formal labor, capital and pension income and add up ECH observations corresponding to purely informal workers and non income receivers with their respective survey weights and expansion factors (Step 1). However, as a proportion of individuals might switch from informal to formal work, the total number of individuals we get is larger than the population control. Hence, to fit the total number of observations to the actual value of the population projections, we perform two alternative adjustments to assess the sensitivity of our results (Step 2). In the first option, we only downsize the number of purely informal individuals that were added from the survey (Y_1), while, in the second alternative, we also adjust each DGI individuals by the number of months of formal labour income received (Y_4). Next, using the misreporting ratios obtained upon the linked data, we inflate DGI earnings to account for formal labour income underreporting in the lower tail of the tax record distribution obtaining Y_2 and Y_5 (Step 3). Up to this point, we included pure informal individuals and corrected formal labour income but we still do not account for individuals that jointly receive formal and informal labour income. Thus, in a final step, based on the proportion of informal to formal labour income reported by ECH respondents, we add a second imputation to the corrected labour earnings vector, creating Y_3 and Y_6 (Step 4). In the remaining of this subsection, we describe each step in detail.

Fig. 3.1. Overview of Method 1



Note. Own elaboration.

In Step 1, we depart from the tax records' income variable, Y_{tax}, and include ECH observations corresponding to individuals aged 20 or above who have zero harmonized income in ECH (Y_{survey}) -that is, they are not contributing to the social security system (informal labour income) and are not receiving pensions or capital income- with their survey weights. As can be checked in Table 3.2, the added ECH informal income represents around 6% to 9% of DGI income and, as expected, it is heavily concentrated in the lower tail of the income distribution (Figure B.2 panel (a)).

However, as pointed out before this procedure yields a number of observations that exceeds the population total by approximately 10% (Table 3.1). Thus, in a first variant we compress the survey weights to achieve consistency with the population projections (Step 2). This excess number of observations arises from the fact that this correction implicitly assumes that workers are either formal or informal and do not switch from inactivity or informality to formal work, or combine formal and informal earnings, a salient feature of developing countries. Thus, to match the actual population total, we need to include an additional reweighting factor to downsize ECH observations (Y_1). To compute this factor, we assume that the inactive population is estimated accurately in ECH and reweight the number of informal workers to match the corresponding total (Table 3.1, lines 4 and 10). In this case, the added ECH informal income falls to 2 to 3% of DGI income.

In a second variant of Step 2, to account better for inflows and outflows to and from formal work and the joint reception of formal and informal labor earnings we exploit the information (available for 2009-2014 only) on the number of months for which a certain worker has been recorded in the labour earnings database (Y_4). In this way, we are able to weight those individuals with positive labour income in DGI by the number of months they received formal labour income ($\lambda_{it} = \sum_{n=1}^{12} m_{it}$ if $Y_{DGIit} > 0$), in each year. Following this procedure, the population total that we obtain is very close to the actual one and, thus, the residual ECH adjustment factor is negligible (line 18, Table 3.1). Notice that, as the sum of the earnings reported by the informal population in ECH are very low, the income control falls by approximately 5% and the additional informal income from ECH represents 4% to 8% of DGI total income (Table 3.2). Unfortunately, since we lack this monthly information for 2015-2016, we discard this option.

In Step 3, we incorporate information from the linked sub-sample to account for the fact that income from formal occupations reported in ECH is under-captured in tax records. We include this correction because this potential underestimation of the lower tail might yield overestimated inequality measures in tax records. To overcome this problem, we use an ancillary table containing DGI/ECH harmonized formal labour income ratios for each DGI labour income centile (p) and adjust Y_{tax} as follows (Step 3, income Y_2 and Y_5):

$$\delta_q(Y_{tax}) = Y_{survey}/Y_{tax} \text{ if } Y_{survey} > 0 \text{ and } Y_{tax} > 0$$

Under this adjustment, we inflate DGI total income by 7.9% to 8.7%, depending on the year (Table 3.2). Figure 3.2 shows that in this case, the adjustment mainly affects the centiles in the middle 40% of the distribution. Nonetheless, in the previous steps we did not account for the fact that formal workers might be receiving formal and informal income simultaneously. Hence, to introduce the corresponding correction, we compute the total labour (Y_{survey}) to harmonized labour ECH income ratios by DGI percentile thresholds in ECH micro-data (Step 4). Multiplying the DGI labour earnings by this factor, we obtain an approximation to total labour income ($\gamma_{qit} = Y_{surveyqit}/Y_{formalqit}$ if $Y_{formalqit} > 0$). In this case, we add a 4% increase to the original DGI income (Table 3.2).

Table 3.2: Income control

	2009	2010	2011	2012	2013	2014	2015	2016	Income
Tax unadjusted	309,532	353,322	412,898	488,090	567,955	659,210	740,858	842,939	Y_{tax}
Survey unadjusted	27,923	30,130	31,795	33,570	36,697	39,513	43,342	48,780	Y_{survey}
% of original tax	9.0%	8.5%	7.7%	6.9%	6.5%	6.0%	5.9%	5.8%	-
Tax unadjusted	309,532	353,322	412,898	488,090	567,955	659,210	740,858	842,939	
Survey adjusted (total pop.)	8,832	12,624	10,589	13,617	10,138	10,744	11,023	14,167	Y_1
% of original tax	2.9%	3.6%	2.6%	2.8%	1.8%	1.6%	1.5%	1.7%	
Tax + informal + under.	348,080	394,894	474,495	555,252	670,657	744,342	846,704	951,598	
% of original tax	112.5%	111.8%	114.9%	113.8%	118.1%	112.9%	114.3%	112.9%	Y_3
Tax adj. (months w/income)	302,344	344,953	403,298	478,309	556,630	640,172	-	-	
% of original tax	97.7%	97.6%	97.7%	98.0%	98.0%	97.1%	-	-	
Survey adjusted	23,852	27,829	28,974	32,556	31,799	26,383	-	-	Y_4
% of original tax	7.7%	7.9%	7.0%	6.7%	5.6%	4.0%	-	-	
Tax adj. + informal + under.	331,677	377,115	447,488	530,169	637,712	720,605			
% of original tax	107.2%	106.7%	108.4%	108.6%	112.3%	109.3%	-	-	Y_6

Note. Own calculation based on tax records (DGI), household surveys (INE) and population projections. Total income in millions of Uruguayan pesos (1 US dollar=30 Uruguayan pesos).

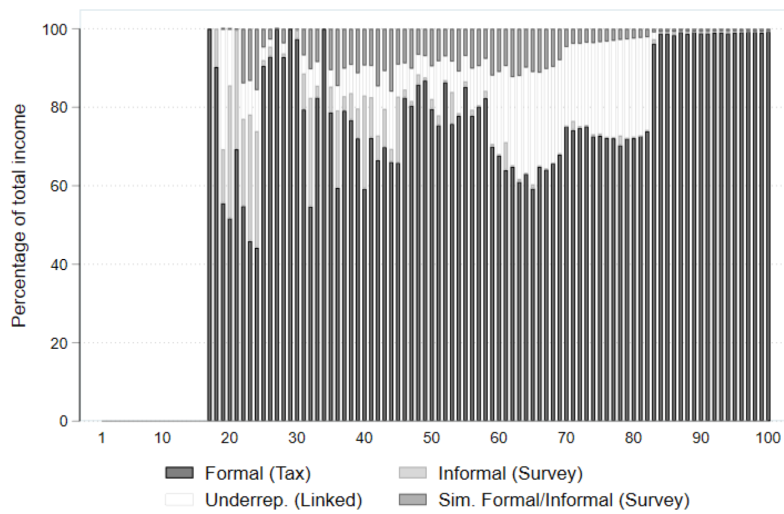
As a whole, we are inflating the original DGI income by approximately 15%. It can be noticed that the additional ECH income variables are mainly placed in the lower tail and middle of the income distribution. Table 3.2 and Figures 3.2, 3.3 and B.2 summarize the full correction process. Due to space constraints, the table does not include Y_2 and Y_5, but this information is available from the authors on request.

Following the previous steps, we create two adjusted tax income variables (Y_3) and Y_6). As stated, even re-weighting DGI observations by the number of months in formal work (Y_6) might reflect the dynamics of formal and informal employment more accurately since we lack this information for the whole period, Y_3 is our preferred option:

$$Y_{3it} = \begin{cases} Ysurvey_{it} \text{ if } Ysurvey_{formal} = 0 \\ Ylabour_{tax,it,q} * \gamma_{q,it} * \delta_q + Ypensions_{it} + Ycapital_{it} \text{ if } Ytax_{it} > 0 \end{cases}$$

In this way, we account for income underreporting at formal occupations, informal income in the lower tail and simultaneous reception of formal and informal income. In the next section, we refer to Y_3 as corrected tax income. Figure 3.2 shows the contribution of each data source to the composition of this variable by percentile. It can be noticed that the first 17 centiles correspond to the population aged 20 or more with zero income. For all quantiles with positive earnings, income is mostly composed from information from tax records and the corrections are concentrated at the bottom 90%. As expected, pure informal income is concentrated at the bottom 50% of the distribution, whereas income underreporting from formal occupations and simultaneous reception of formal and informal income affect the lower and middle strata.

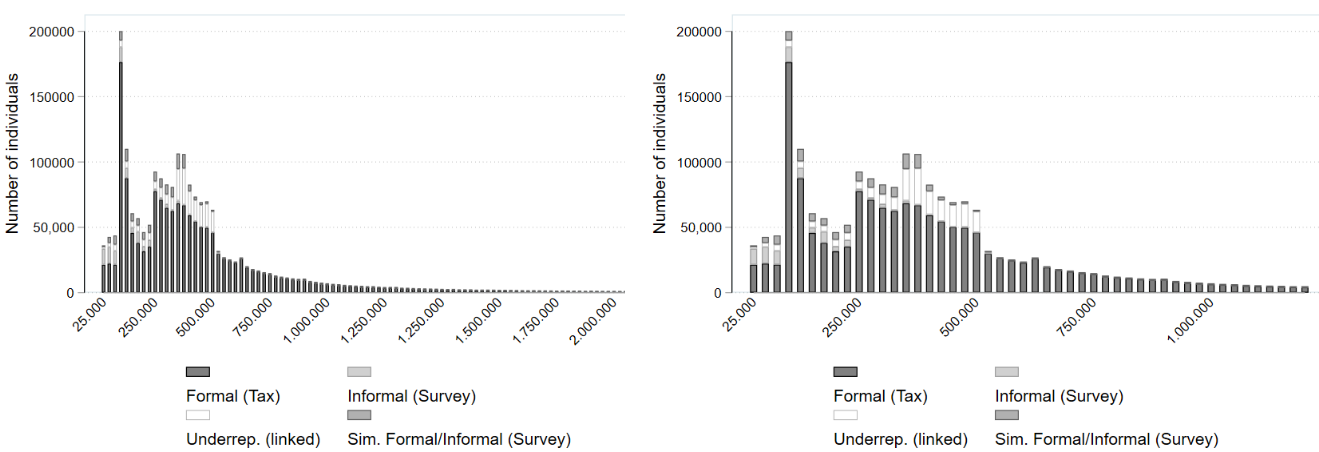
Fig. 3.2. Income composition by percentile of total income (Y_3)



Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Fig. 3.3. Composition of the corrected tax income distribution by data source

(a) Total Distribution (b) Bottom 99%



Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

3.2.3. *Robustness checks*

As a first robustness check, in the top income shares estimation we also computed alternative income totals for 2012 to 2016 (Method 2) based on SNA information (Y_7). In this case, we use our preferred corrected income variable (Y_3) as the numerator but we use the 80% of the households income account as income control.

In the second place, to assess the sensitivity of our results, we compute the Alvaredo (2011) correction departing from harmonized ECH income and adding the top 1% share calculated for corrected tax income (Y_3).[20] Additionally, we implement the reweighting methodology developed by Blanchet et al. (2022) and create an additional income vector (Y_8).[21] This method identifies a merging threshold at the maximum point at which the survey-tax quantile ratio equalizes the survey-tax densities ratio. To carry out the correction, researchers need to define the minimum percentile at which the tax data are reliable, which we set at p50 due to the considerations presented previously. The endogenous merging point varies around percentiles 50 and 70, depending on the year. Additionally, to check the sensitivity of our results, we imposed merging points at quantiles 50 and 70 and obtain similar results.

4. The recent evolution of primary income inequality in Uruguay

In this section, we analyse income inequality, focusing on the evolution of top income shares and synthetic indices for corrected tax income and harmonized survey income. Unless specified in the text, from this point onwards, corrected tax income refers to pre-tax $Y3$ and harmonized survey income refers to pre-tax individual earnings from formal and informal occupations plus pensions and capital income computed using ECH information.

4.1. *Income shares*

At first glance, the distribution of corrected tax income did not experience significant modifications throughout the period under analysis (Table 4.1). The share of the bottom 50% exhibits a mild increase, whereas the middle 40% remained almost unchanged. It is noteworthy that the top 1% holds a larger proportion of the total income than the bottom 50%, although this gap has reduced slightly over the years. A similar comment applies to the

[20] According to Alvaredo (2011), the corrected Gini Index can be approximated by: $G = G*(1-S)+S$, where $G*$ is the Gini Coefficient for the bottom 99% of the distribution, and S is the share held by the top 1%.

[21] To implement this method we resort to the stata code (bfmcorr) provided by the authors.

middle 40% with respect to the top 10%, although the gap widened in this case and, by 2016, the proportion of the total income accrued by the latter was smaller. In the harmonized survey income, the lower strata increased their participation and, conversely, the top shares decreased. Notice that, in 2009, the income distribution was not very different in the two income variables considered, but diverged over the years.

Table 4.1: Pre-tax income shares, 2009-2016

Corrected tax income								
Inc. groups	2009	2010	2011	2012	2013	2014	2015	2016
Bottom 50%	10.8%	11.0%	11.9%	12.2%	13.0%	12.8%	13.2%	12.4%
Middle 40%	45.4%	45.5%	46.1%	45.9%	46.0%	46.1%	46.2%	45.7%
Top 10%	43.8%	43.5%	42.0%	42.0%	41.0%	41.1%	40.6%	41.9%
Top 5%	31.0%	30.8%	29.9%	29.8%	28.9%	29.2%	29.0%	30.3%
Top 1%	13.5%	13.5%	13.5%	13.2%	12.7%	13.2%	13.5%	14.6%
Top 0.1%	4.6%	4.7%	5.0%	4.7%	4.8%	4.8%	5.2%	5.8%
Harmonized survey								
Inc. groups	2009	2010	2011	2012	2013	2014	2015	2016
Bottom 50%	8.7%	9.6%	10.5%	11.1%	11.2%	11.5%	11.3%	11.2%
Middle 40%	47.5%	48.2%	49.4%	51.6%	50.9%	51.0%	50.9%	51.3%
Top 10%	43.9%	42.3%	40.1%	37.3%	37.9%	37.5%	37.8%	37.5%
Top 5%	30.0%	28.5%	26.6%	23.8%	24.6%	24.4%	24.7%	24.4%
Top 1%	11.9%	10.6%	9.6%	7.7%	8.4%	8.4%	8.7%	8.4%
Top 0.1%	3.1%	2.4%	2.1%	1.3%	1.7%	1.7%	1.8%	1.6%

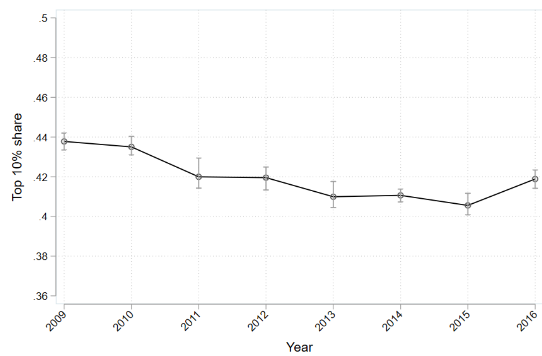
Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). Income thresholds of corrected tax income in Table B.7

Figure 4.1 depicts the evolution of the top 10%, 1% and 0.1% corrected tax income shares and the corresponding confidence intervals. In line with previous inequality studies for Uruguay, the participation of the higher decile exhibits a statistically significant decline. We are not able to assess whether this point estimate is indicating a reversion of the previous trend. However, the top 1% and 0.1% shares remained almost unchanged in 2009-2013 and exhibit a slight increasing trend since 2014, which is statistically significant in the first case and imprecise in the latter.

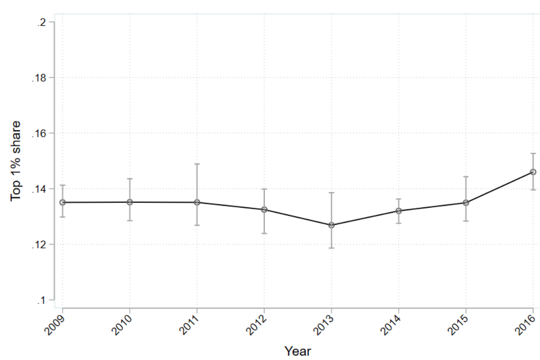
Considering the whole period, the point estimate of the top 1% share rose from 13.5% to 14.6%. These values place Uruguay among the countries with the highest concentration at the top among the group of countries for which tax record-based top income estimates are available, only appearing below the remaining Latin American countries, South Africa and the United States (see Atkinson (2007)).

Fig. 4.1. Pre-tax top income shares, 2009-2016. Corrected tax income.

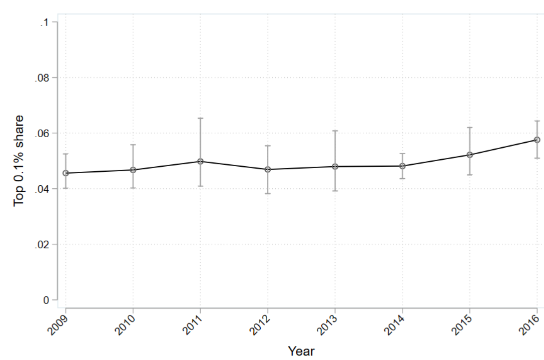
(a) Top 10%



(b) Top 1%



(c) Top 0.1%



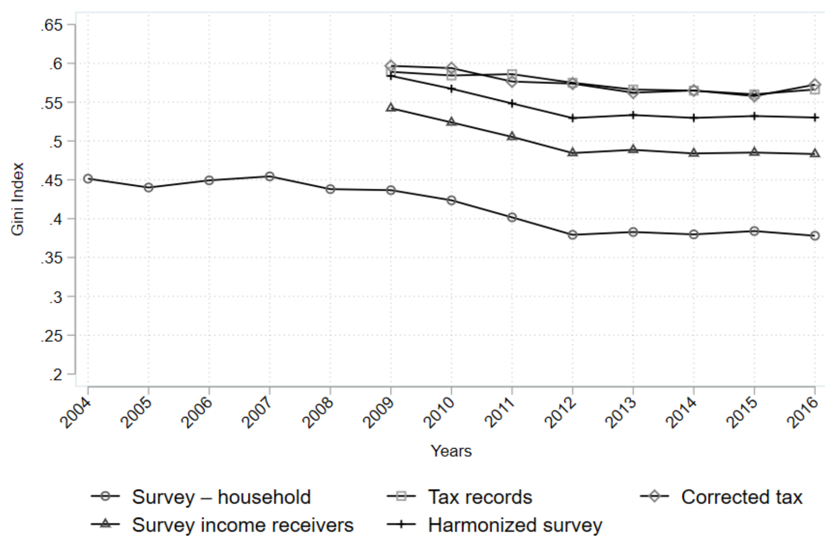
Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). See the point estimates in Table A.3. Bootstraps with 100 repetitions, with confidence intervals at the 5% level.

The slight increase in the top income shares in the corrected tax income is in sharp contrast to the declining trends observed in the harmonized survey income (Table 4.1). The corrected tax income to harmonized survey income ratio of the top 1% shares was 0.88 in 2009, falling to 0.57 in 2016. At the same time, the top 1% thresholds ratio fell from nearly 0.95 to 0.74 (Table A.2). The evolution of these two ratios suggests that the ability of the household survey to capture incomes in the upper tail was eroded in these years. In fact, the 10% thresholds ratio is very close to 1, although it exhibits a mild decline (from 1 to 0.92) throughout the whole period.

4.2. Synthetic inequality indices

Figure 4.2 depicts the synthetic Gini indices computed on the basis of different survey and tax income variables. The longest line corresponds to the survey per capita household income, the income aggregate mostly used in personal income inequality studies. As stated in the introduction, its evolution indicates a sharp decline between 2008 and 2013 and stability thereafter. Although the levels are higher, inequality among income receivers in the survey mimics the path of household income distribution, considering either the full set of income sources or the more restrictive harmonized survey variable used in this study. The 2009-2013 and 2009-2016 Gini and Theil reductions are statistically significant in all cases.[22]

Fig. 4.2. Inequality trends by income definition and source, pre-tax income Gini index, 2004-2016



Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Figure 4.2 also depicts the original pure administrative information and the corrected tax income variable. The two lines indicate a mild decline, with the inequality indices converging after 2012 and slightly increasing by 2016. Again, the 2009-2016 and 2009-2013 differences are statistically significant.[23]

[22] See confidence intervals in Table A.3. If we restrict the corrected tax income and harmonized survey income to the subset of observations with positive income, the results are similar in the former case, whereas we find a larger fall (12.6%) in the latter one (Figure B.1).

[23] It is noteworthy that these results also hold when considering only the original DGI data without undistributed and non nominative profits imputations. The corresponding tables are available from the authors on request.

Thus, the full set of income variables conveys an inequality reduction from 2009 to 2016, which mainly occurred in the first five years. This finding suggests that the equalization trend is robust to the data base and harmonization criteria, even when the levels and slopes are different. Considering the whole period, the harmonized survey income indicates an 8.6% inequality reduction. Since the corrected tax income only experienced a 2% decrease, the gap has widened in the last years. [24]

4.3. Robustness checks

As mentioned in section 3.2, to validate our main conclusions, we carry out a set of robustness checks. First, we compute the inequality measures presented in the previous subsections for the seven income variables that we created following Methods 1 and 2. As Figure A.2 shows, the levels vary within a relatively bounded interval, particularly regarding the top 1%. However, the trends resemble the ones presented in the previous subsections: stability or an increase in the top 1% and 0.1% income shares and a statistically significant decline in the Gini and Theil indices. Again, the top 10% share falls steadily until 2015 and rises in 2016.

Second, we take the opposite approach and correct the harmonized ECH data with the tax record information (Figure A.3). In the first place, we implement the correction proposed by Alvaredo (2011). Thus, we compute the Gini coefficient for the bottom 99% with harmonized survey income and carry out the corresponding decomposition using the corrected tax income's (Y_3) top 1% share. As shown in Figure A.2, although the levels are lower, inequality also decreased in this case. If we use the uncorrected tax data ($Y_t ax$) instead, we obtain similar results.

In the second place, we implement the reweighting procedure proposed by Blanchet et al. (2022). The endogenous merging point varies over the years, but is always found between the median and the 70th percentile, which implies that the correction starts in a lower quantile than the one usually considered in the empirical implementation of Alvaredo (2011) used here. As Figure A.3 shows, the absolute value of the Gini index is very similar to the one we obtain with Method 1; hence the trend is similar. This conclusion also holds for the different fractiles' levels and trends. In sum, our robustness checks validate the conclusions presented in the previous sections.

[24]Table B.9 confirms that these results also hold in the case of Theil's indices.

5. Reconciling the inequality trends in tax and household survey data

The increasing divergence in the top 1% thresholds in harmonized survey and the corrected tax income might be consistent with the larger reduction in inequality in the former case vis á vis the latter. To dig further into these differences, we first present the Gini and Theil indices decompositions by income subgroups, to isolate the movements and the contribution to inequality of the top 1%. After that, we analyse the evolution of the densities and inequality indices at the top, singling out the intervals in which the tax and survey overlap and those that are beyond the survey maximum. Finally, we compare the composition of income by source (pensions, labour earnings and capital income) in the harmonized survey and corrected tax income.

5.1. Inequality decompositions by income groups

We decompose the Gini and Theil indices by income groups, considering the bottom 99% and the top 1% (Table 5.1 and Table B.9).[25] In both in corrected tax income (Y_3) and the original tax income variable (Y_{tax}), the proportion of between groups inequality remained steady and grew slightly in the last two years, indicating an increased distance in the two groups' average income. Meanwhile, the harmonized survey income exhibits the opposite pattern, with a substantial decline in the between group inequality fraction over the years. The results for the Theil's index decomposition are similar, with an slightly increase in the between group fraction both in pure tax and in Y_3 income (from 30% to 40% for the latter).

The last two rows in the panels depicted in Tables 5.1 and B.9 present the inequality indices for the two income subgroups. The two DGI based income variables indicate a sharp contrast between the equalizing trend of the bottom 99% (-6%) and increased concentration at the top 1% (20%).[26] Nevertheless, in harmonized survey micro-data the two income groups experienced a substantial inequality decline. Moreover, the reduction is larger for the top 1% (11% and 35%, respectively).[27]. The two subgroups present the same patterns as the Theil index decompositions.

[25] Since we are using income quantiles, we can obtain exact population subgroups decompositions for the Gini and Theil indices (Cowell, 2011).

[26] These results also hold for all the DGI income variants, either considering the original uncorrected tax data (without adding bank deposits, non nominative profits and undistributed profits), or in the case of the remaining corrected income variables.

[27] The results are similar for the lower 99% in the three subgroups.

These results strengthen the hypothesis that the equalizing trends observed in the synthetic indices in the harmonized survey and tax based variables stem from very different movements throughout the income distribution. The between group inequality shares indicate that the subgroup's average income diverged in the tax records and converged in harmonized survey income. This finding is consistent with the falling survey/tax top 1% threshold ratio presented in Table A.2. At the same time, the mild inequality reduction observed in the tax data results from an offsetting fall in the concentration of the bottom 99% against the increased inequality at the top. Conversely, in the harmonized survey income, inequality fell in all the income groups, although the reduction was considerably larger at the top. It is worth noticing that even when the fall was steeper (11% versus 7%), inequality trends for the lower 99% were relatively similar in the harmonized survey income and in the tax data.

Table 5.1: Pre-tax Inequality decomposition between two income groups, 2009-2016.

	2009	2010	2011	2012	2013	2014	2015	2016
				Corrected tax income				
Gini index	0.597	0.594	0.577	0.574	0.562	0.565	0.558	0.573
% between	21.0%	21.1%	21.7%	21.3%	20.8%	21.6%	22.4%	23.8%
% within	79.0%	78.9%	78.3%	78.7%	79.2%	78.4%	77.6%	76.2%
% overlap	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gini bottom 99	0.550	0.547	0.527	0.525	0.515	0.515	0.505	0.516
Gini top 1	0.347	0.356	0.380	0.365	0.390	0.380	0.402	0.417
				Harmonized survey				
Gini index	0.584	0.567	0.548	0.530	0.533	0.530	0.532	0.530
% between	17.9%	16.3%	15.0%	11.8%	13.4%	13.3%	13.9%	13.5%
% within	82.1%	83.7%	85.0%	88.2%	86.6%	86.7%	86.1%	86.5%
% overlap	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gini bottom 99	0.547	0.535	0.519	0.508	0.508	0.505	0.505	0.505
Gini top 1	0.261	0.221	0.205	0.133	0.185	0.175	0.192	0.177
				Tax records				
Gini index	0.589	0.584	0.586	0.575	0.566	0.565	0.560	0.566
% between	17.5%	15.8%	14.7%	11.7%	13.1%	13.0%	13.6%	13.1%
% within	82.5%	84.2%	85.3%	88.3%	86.9%	87.0%	86.4%	86.9%
% overlap	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gini bottom 99	0.540	0.535	0.533	0.523	0.513	0.511	0.503	0.505
Gini top 1	0.355	0.364	0.389	0.373	0.399	0.385	0.408	0.422

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). The table is divided in three panels, presenting the corrected income in harmonized surveys and tax records respectively. By construction, both micro-data bases refer to the same individuals and the same incomes (pre-tax and total formal income). In each panel, the Gini index is decomposed into *between* and *within* components, among the groups defined (bottom 99% and top 1%). Within group inequality is shown in the last two rows of each panel.

5.2. Movements in the upper tail of the income distribution

In the preceding sections, the top 1% thresholds were endogenously defined for each data source. However, as shown in the previous subsections, the harmonized survey/corrected tax income thresholds ratio decreased monotonically. Hence, the top 1% share of corrected tax income is defined with an increasingly larger absolute income value than the one in the harmonized survey. To test whether the conflicting trends in relative income and within group inequality at the top might result from these differences, we compute the proportion of observations beyond the harmonized survey top 1% threshold and inequality measures in corrected tax income, separately considering: 1) observations with income above the 1% threshold in the harmonized survey and below the survey's maximum; and 2) observations with income above survey's maximum (see Figure A.4).

In 2009 and 2016 the proportion of corrected tax income observations belonging to each group (group 1: 1.3% and 2.0%; group 2: 0.15 and 0.25%) indicates that most of the observations used to compute the top 1% lie in the common support. Thus, the problem is not only reaching the rich who are above the survey maximum but representing correctly those individuals located in the common support. Both subgroups, but particularly group 1, present an increasing share, again reflecting the divergence between the two data sources. Lowering the threshold (beyond the survey threshold) to compute the Gini index of the corrected tax income does not affect the inequality trends at the top of corrected tax income.

Figure B.3 (panel a) depicts kernel density functions for those observations pertaining to the top 2% of the corrected tax income in selected years. The vertical red line represents the maximum of the harmonized survey income (or the limit between group 1 and group 2). Two features are noteworthy: an inequality increase in group 1 and an augmented fraction of income received by the top 1% and 2%. Thus, the observed differences inthe top incomes shares and top 1% inequality indices are noticeable in the common support and are not only driven by the corrected tax income capturing richer individuals but seems to result from an increasingly lowered density in the common support. Notice that, in both groups, the gap increases in 2012, close to the end of the inequality reduction period.

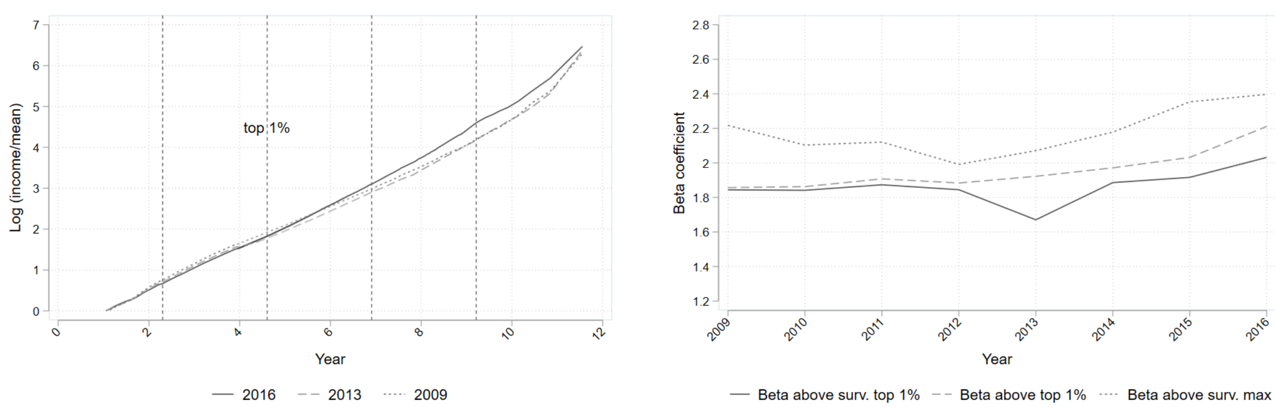
To conclude, we present a brief parametric analysis of the evolution of inequality at the top-end, based on the Pareto I distribution.[28] Figure 5.1 shows the survival function (Cowell,

[28]The purpose of this exercise is not to analyse in depth the parametric function that best fits the Uruguayan data, but to inspect briefly the shape of the upper tail, that is, the income differences at the top-end of the distribution. As Jenkins (2017) and Charpentier and Flachaire (2019) show, Pareto I estimates are very sensitive to the threshold. To overcome this potential draw-back, we also consider the three thresholds analysed previously (the top 1% in the harmonized survey and the corrected tax income (Y_3) respectively,

2011; Atkinson, Casarico, and Voitchovsky, 2018). First, in all cases the survival function is concave at the top, indicating that the Pareto parameter (α) decreases with income. Atkinson et al. (2018) labeled this shape as "regal" to indicate the large distances between the different observations at the top, opposing it to the "baronial" pattern in which the distances among observations at the top are smaller. Second, the slope of the 2016 survival function is less steep than those for 2009 and 2013, indicating an inequality increase in the upper tail throughout the years. In turn, the evolution of the beta clearly shows an increasing differentiation of incomes at the top-end, despite the income threshold (see panel b) of Figure 5.1). 2012-2013 again seems to be a watershed regarding inequality trends. Third, the β coefficients ($^{\alpha}/_{\alpha-1}$) indicate an increasing differentiation of incomes at the top, despite the threshold.

Fig. 5.1. Inequality at the top tail of corrected tax income, 2009-2016

(a) Income in relation to rank (b) Inverted Pareto coefficients (β)



Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). In panel a, the y axis depicts the log of income as a proportion of the mean income, while the x axis depicts the log of $\frac{1}{S}$, with S being the survival function. Vertical lines respectively represent top 10, 1, 0.1 and 0.01% thresholds. All the incomes are annual and at 2016 prices. In panel b, the top 1% threshold refers to the total income distribution in corrected DGI data.

These findings suggest that differences in inequality trends might result from diverging concentration patterns at the upper tail in ECH and DGI data. Considering the short period under analysis, a 32% reduction in the harmonized survey income Gini coefficient for the top 1% seems extremely high compared with previous evidence on inequality reduction trends at the top. On the side of administrative data, two main features might create an artificial inequality increase: reduced informality with the subsequent entry of low-salaried workers in the data-base and a greater ability of the tax authority to enforce tax-payments. Furthermore,

and the maximum at ECH) and the results are similar.

the evolution of inequality in the bottom 50% rules out the possibility of corrected tax income trends being driven by the formalization process. Although the available data do not allow us to solve this puzzle, in the next subsection we dig a little further into these differences, focusing on the capital income share in both distributions.

6. Income sources and characteristics at the top

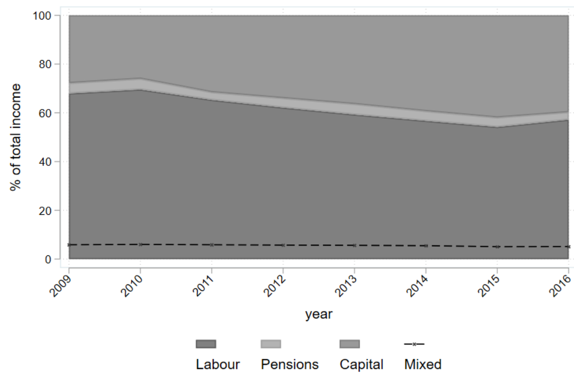
6.1. The growing share of capital income

The previous section findings suggest that the differences in inequality trends among administrative and survey data result from divergent trends at the top of the income distribution. Thus, the ability of household surveys and administrative data to capture the different income sources can contribute to shedding light on these discrepancies, particularly, if during this period, capital income earnings increased as this income source is associated with higher underreporting rates in ECH. To explore this point further, we first analyse the composition of income by source (Figures 6.1 and A.6) and present the Lerman and Yitzhaki (1985) inequality decomposition by income source.

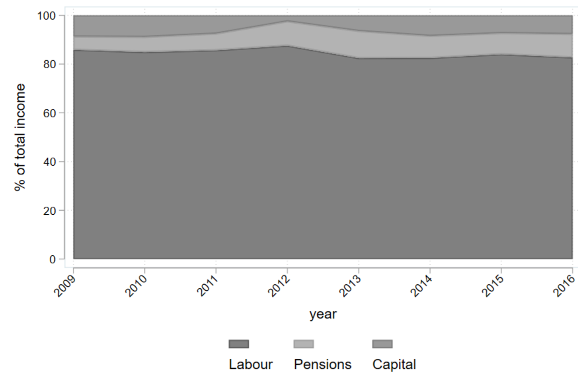
These results uncover the expected pattern: labour earnings account for around 75% of the total income in harmonized survey income and fall to 66% in corrected tax income. Since the share of pensions is similar in the two data-sources, the whole difference is due to the capital income share, which is around three or four times larger at the tax records database and increases throughout the period, whereas it falls in household survey data. Again, this pattern is consistent with the different trends in the top incomes shares observed in the two data-sets. The available SNA data on the capital income share in the households account show a slight increase from 10.9% in 2012 to 12.8% in 2016. These figures are closer to the ones computed using the corrected tax income, ruling out the possibility that the corrected tax income trajectory has been lead by the increased capacity of the tax authority to reach the rich.

In the corrected tax income estimations there is a substantial increase in the participation of capital income at the top throughout the whole period, which is not mirrored in the harmonized survey income. In fact, our estimations indicate that whereas the top 1% receives 37% of total capital income in the harmonized survey, this figure rises to 62% in the corrected tax income. The increasing share of capital income at the top might be the driving force explaining the divergent trends at the top. It is worth noticing that in 2016, the capital income and mixed income equalize the share of labour earnings for the top 1% and surpassed it for the top 0.1% at corrected tax income.[29]

Fig. 6.1. Pre-tax income composition by source, 2009-2016



(a) Corrected tax income - Top 1%



(b) Harmonized survey income - Top 1%

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). In tax records, mixed incomes are depicted as a share of the labour income for comparison purposes.

Table A.5 presents the results of the Gini coefficient decomposition by income source for the corrected tax income and the harmonized survey income. As expected, capital income and mixed income are the most unequally distributed income components, followed by pensions (probably related to the number of individuals who are not pensioners). In both cases, labour earnings make the greatest contribution to overall inequality, with a larger share in harmonized survey income. In spite of its diminishing share in ECH, the contribution of capital income to overall inequality increased over the years, in both data sources. Again, the decomposition yields to different patterns in the two data sources, with a larger contribution of labour income to inequality in the ECH data. Conversely, in the corrected DGI income, the contribution of capital income and pensions is substantially larger.

To investigate further the interplay between the evolution of the relative participation of the different income sources and the concentration at the top of the corrected tax income distribution, we decompose the evolution of the top 1% income share in two factors (Boschini et al. (2020)): the total share of the different income sources and the variation in the share of the different income sources held by the top 1%. We group the share of labour income and pensions (q) on one side, and the capital share ($1-q$) on the other. The share of the top 1% in the joint labour earnings and pensions distribution is a and b is the corresponding top 1% share in the total capital income.

[29]Due to the number of cases these estimations cannot be carried out with ECH micro-data.

$$\Delta s = s_{t+1} - s_t = (a_{t+1} - a_t)q_t + (b_{t+1} - b_t)(1 - q_t) + (a_{t+1} - b_{t+1})(q_{t+1} - q_t)$$

The first term represents the contribution from changes in non-capital income, the second one reflects the contribution of changes on capital income, and the third one corresponds to the contribution of the changes between sources.

As it can be noticed from Table 6.1, the 1.1% percentage points increase in the top 1% share results from a 41% increase in $1 - q$ coupled with a 10% increase in b that is not outweighed by the equalizing trend in the distribution of labour and pensions earnings. Meanwhile, in the harmonized survey income $1 - q$ was constant and exhibits a smaller share, and the 46% reduction of the top 1% share (4 percentage points) results from a 50.5% reduction of its share in capital income and a 41.6% decrease in labor income. Notice that the top 1% share in labour income declined in the corrected tax income and harmonized survey, although the reduction was larger in the latter case. In sum, the diverging trends of the top 1% share in survey and tax data are closely related to the evolution of capital income inequality, which in the tax corrected income seems to outweigh the equalizing trend of labour income and pensions.

In their study on the United States, Kopczuk and Zwick (2020) pointed out that correctly identifying the different income sources at the top of the income distribution is not an easy task. Personal income taxation as well as the mechanisms used by firms and, particularly, big corporations to set payments to managerial personnel (partially driven by the specific features of the income tax schedule) clearly shape the definition of income sources. At the same time, in case if it is possible to observe it, it will be necessary to determine whether the annual distribution of profits to liberal professionals can be considered to be capital income. Thus, the limits between labor and capital income can be an unintended result of the personal income tax schedule and corporate decisions.

It is worth pointing out that personal income taxation did not offset the increased share of capital income and the top fractiles. Although personal income taxation in Uruguay is progressive, it has modest redistributive effects. It approximately reduces the top 10% and 1% shares by 12-14% and 5-6% (2.5 and 2 percentage points respectively), with a subsequent increase in the middle 40% and the bottom 50% (Table A.4). In addition, the IRPF became less redistributive in the period under analysis. This effect is probably related to the dual nature of the Uruguayan taxation scheme, coupled with the increased share of entrepreneurial profits and dividends at the top; these are taxed at a lower rate than the remaining capital income sources (7 versus 12%, Table B.1). As a result, tax rates effectively paid by the top

1% are lower than the ones for lower neighbouring fractiles and the same pattern holds for the top 0.5 and 0.1% (A.8). This regressive capital income taxation scheme is reflected in the total effective rates. Even when they exhibit a progressive pattern for the first 99 percentiles, they fall from 11.5% for the top 1% to 9.5% for the top 0.1% (see Figure A.8).

To conclude this subsection, we assess the share of the different capital income concepts for the corrected tax income quantiles. As previously shown, capital income is disproportionately concentrated at the top of the income distribution.[30] Property rents exhibit a larger share for centiles 90-99, whereas dividends account for around 45% of the capital income at the top-end (see Figure A.7). Dividends are clearly the most unequally distributed capital income sub-component. The predominance of capital income and, specifically, dividends in the richest strata has been highlighted by the top incomes literature as a distinctive feature of developing countries, since in the developed world, executives compensations and high salaried workers predominate (Alvaredo and Londoño Velez, 2014).

In subsection 5.2 we have shown that there is an increasingly lowered density at the common support in ECH and a 32% fall in the concentration of the top 1% in the survey. The reduced capacity of household surveys to capture high incomes is consistent with the fact that the increase at the top is mainly caused by capital income growing to a larger extent than labour income. In fact, our decomposition exercises indicates that the increased top 1% share is explained by capital income inequality and that personal income taxation does not morigerate this evolution. As mentioned at the beginning of this subsection, the SNA information indicates that the capital income share rose from 10.9% in 2012 to 12.8% in 2016. The information presented in this section shows that in ECH, the capital share remained almost steady and exhibited a considerably lower share (4.7 to 5.1%), whereas in Y_3 it grew from 10 to 15.3%, which is consistent with SNA information. At the same time, from Table B.8, it can be noticed that the participation of dividends within capital income rose from 13.4 to 29.6%. These findings suggest that the evolution of dividends played a key row in the growth of the capital income share, an income concept considerably undercaptured in household surveys.

6.2. Top income holders: a brief characterization

In this section we examine the main characteristics of the individuals belonging to the different income fractiles, focusing on the top of the corrected tax income distribution. Since in the previous section we show that the upper tail is misrepresented in the harmonized survey

[30] Recall that since individuals own occupied housing is not included in the Uruguayan personal income tax scheme, our results might be biased as we exclude the most widespread form of capital income from our calculations.

Table 6.1: Inequality decomposition by income source, 2009 - 2016. Pre-tax corrected income and harmonized survey income

Panel A: Capital and non-capital incomes shares by source (Y3 and harmonized survey)					
Corrected tax income					
	Top 1% share	Labor + pensions (q)	Capital (1-q)	Labor + pensions top 1 (a)	Capital top 1 (b)
2009	13.5%	94.1%	5.9%	10.4%	63.1%
2016	14.6%	91.7%	8.3%	9.6%	70.1%
Hamonized survey income					
	Top 1% share	Labor + pensions (q)	Capital (1-q)	Labor + pensions top 1 (a)	Capital top 1 (b)
2009	11.9%	96.1%	4.0%	13.8%	31.9%
2016	8.4%	96.2%	3.8%	10.1%	21.2%
Panel B: Contribution of each source to the change in the top 1% share					
Corrected tax income					
	Top 1%	Labor + pension	Capital	Change between sources	
Change 2009-2016	1.1%	-0.7%	0.4%	1.4%	
Contribution to change	100%	-66.3%	37.7%	128.7%	
Hamonized survey income					
	Top 1%	Labor + pension	Capital	Change between sources	
Change 2009-2016	-4.0%	-3.5%	-0.4%	0.0%	
Contribution to change	100%	89.0%	10.6%	0.4%	

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

income, this exercise can only be carried out with tax records information.[31] Furthermore, we exploit the matched firm-worker/owner data-base. We present evidence on gender differentials and carry out a multivariate analysis.

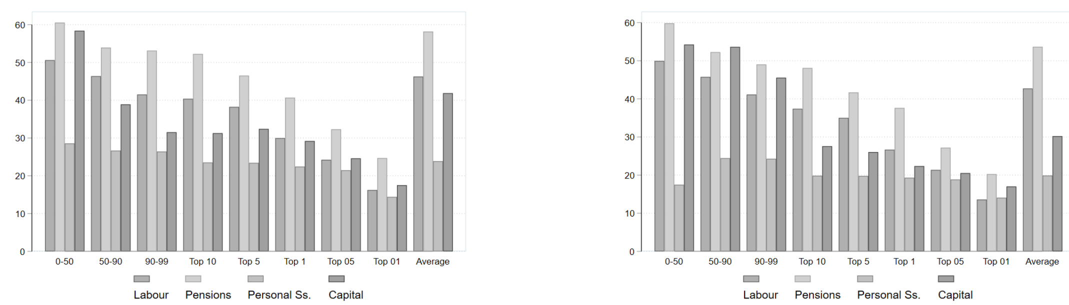
In line with previous studies on wage differentials, our estimations show that the proportion of women in the total and labour income decreases with the quantile (Figure 6.2, panel a), ranging from more than 50% below the median to 25% at the highest percentile. The estimations reported by Atkinson et al. (2018) for eight high income countries yielded to very similar results. Due to differences in life expectancy patterns coupled with the wide coverage of the Uruguayan pensions system, the presence of women is larger among pensioners. Even though the differences are smaller in this case, the presence of women declines with income (60% and 40% respectively). Conversely, women are severely underrepresented among liberal professionals and capital income receivers. Considering the distribution of income instead (panel b), the results are very similar, although women's share is even smaller in most cases, probably reflecting the earnings gap within these categories. In sum, capital income and earnings from liberal professionals mirror and widen the gender gap documented for labor income in previous studies on Uruguay (Amarante, Arim, and Yapor (2016); CEPAL (2020);

[31]The data used in this section are representative of formal occupations, pensioners and capital owners, leaving aside informal workers, who represent approximately 20% of the Uruguayan labour force and are by large self-employed. Unfortunately social security and tax data lack information on schooling

Espino, Isabella, Leites, and Machado (2017); Domínguez-Amorós, Batthyány, and Scavino (2021).

In their study covering five decades in Sweden, Boschini et al. (2020) reported that the participation of women evolved from 6% to 19% in the top 1% and from 5 to 15% in the top 0.1%. This trend was lead by their increased participation at the top of the labour earnings distribution. However, men increased their share at the top of the capital income distribution. Despite the short time span considered in this article, a similar pattern can be identified here for labour earnings and pensions (see Figure A.5, panel a). Meanwhile, the participation of women in capital income as a whole remained relatively constant, with an increase in housing rents and stability in business income (Figure A.5, panel b).

Fig. 6.2. Participation of women in total income and receivers (by income source and income group, 2016



(a) Proportion of female income receivers

(b) Proportion of pre-tax income held by women

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

To further deepen into the characteristics of top income earners, we estimate two different probit models on the probability of being in the top 1% against the remaining 99% or the top 10%, for the total population and opening by gender (marginal effects can be checked in Table 6.2).[32] Among the covariables, we include individual characteristics (sex and age), type of employment (liberal professional, salaried worker, self-employed, and multi-employment), a set of binary variables reflecting the different income sources received by the individual (pensions, labour earnings, capital income, dividends, housing rents and other capital income) and firms characteristics (size, type of business and industry). Industries are opened at the section level (ISIC, rev. 4).[33] Additionally, as our descriptive analysis indicates that workers at financial activities (K) and human health services (Q) are over-represented at the top of

[32]Additional estimations restricted to individuals with positive labor earnings are available on request to the authors.

[33]We grouped agriculture, forestry and fishing, and mining and quarrying (A and B, the omitted category);

the labor earnings distribution, we further disaggregate these divisions.[34] Regarding section Q, we incorporate three binary variables reflecting the classes that are overrepresented at the top 1%: hospital activities (8610), (medical and mental activities (8620) and other human health activities (8690). In the case of section K, we include the three financial activities divisions (64, 65 and 66).

Probit estimates for 2016 show that, relative to the bottom 99%, individuals belonging to the top 1% are more likely to be men and liberal professionals. At the same time, they exhibit a higher probability of receiving capital income, dividends and, to a lesser extent, labor income. Conversely, pensioners are less likely to belong to this group. There are also gender differences in the marginal effect of receiving labor income, which is positive but very low for the total population and men, while it is not statistically significant in the case of women.

Regarding the estimates within the top 10%, it is worth pointing out that the same differences hold but, again in line with the descriptive findings, the gender gap is thirteen times larger than the one corresponding to the top 1% versus the whole population. Differences by income source also hold within the top 10%, but in this case, the marginal effects of receiving dividends are considerably larger than the capital income ones. Liberal professionals exhibit a high probability of reaching the the top 1% relative to remaining in the top 10% . In contrast, the marginal effects of receiving labor earnings is negative for the average and women, while they turn to be not statistically significant in the case of men.

With respect to the differences by industry, higher income positions are associated to the manufacturing, financial, wholesale and retail, public administration, health activities and financial services sectors. These results hold both for men and women. The first three columns of Table 6.2, indicate that the magnitudes of the marginal effects of the health related activities, and particularly hospitals, are considerably larger than the financial sector ones. The last three columns of the table (differentiation within the top 10%), exhibit considerably larger marginal effects for the health services classes (4 times for the total), particularly in the case of hospitals. In this case, the differences with the financial services classes marginal effects are substantial. These findings are consistent with the descriptive

transportation and communications (H and J); and other services activities, activities of households as employers and activities of extraterritorial organizations (S,T and U).

[34](Table A.7) shows that almost half of the top 1% of the labour earnings distribution is concentrated in three sectors: liberal professional and health services (29%), financial and business services (11.9%) and other liberal professional services and public administration (6.2%). In sharp contrast, no sector predominates in the capital income distribution. The share of the health sector decreases to 6% and the financial sector shrinks considerably to 3.6%, which can be explained by the significant share of the public sector and foreign firms in the banking sector.

information presented in Table A.7 for 2016, that shows that approximately 1 out of 4 top income holders receiving labour income are occupied in health services, whereas this figure declines to 8% for the financial sector.

In their characterization of Canada's top 1% earners, based on Census data for a larger time span (1981 and 2006), Lemieux and Riddell (2015) identified the leading force under the increasing share of the top 1% as executives compensations and financial and business services, whereas the medical sector has lost relative relevance. It is hard to determine whether the different pattern found in this study is an Uruguayan feature or whether it holds in other Latin American countries, since previous top incomes studies for Argentina, Brazil, Chile and Colombia do not provide similar disaggregations.

Table 6.2: Probability of belonging to the top 1% (by gender, versus bottom 99% or centiles 90-99, 2016. Probit estimates. Marginal effects)

	Top 1 vs bottom 99%			Top 1 vs remaining Top 10%		
	Total	Female	Male	Total	Female	Male
Male	0.001***			0.0153***		
	(0.000)			(0.00130)		
Age	0.000***	3.80e-05***	0.000966***	0.000488***	0.000298***	0.00529***
	(0.000)	(1.48e-06)	(4.67e-05)	(1.70e-05)	(1.65e-05)	(0.000360)
Age2	-0.000***	-1.54e-08***	-4.28e-06***	-1.89e-07***	-1.01e-07***	-2.13e-05***
	(0.000)	(7.50e-10)	(4.50e-07)	(9.39e-09)	(8.65e-09)	(3.50e-06)
Liberal professional	0.103***	0.0208***	0.0350***	0.332***	0.142***	0.189***
	(0.004)	(0.000744)	(0.000767)	(0.00994)	(0.00605)	(0.00497)
Capital inc. recipient	0.022***	0.0166***	0.0237***	0.0555***	0.0971***	-0.0244***
	(0.000)	(0.000254)	(0.000386)	(0.00140)	(0.00214)	(0.00426)
Dividends	0.030***	0.0174***	0.0396***	0.173***	0.106***	0.203***
	(0.000)	(0.000496)	(0.000632)	(0.00289)	(0.00419)	(0.00401)
Property rents	0.011***	0.00530***	0.0142***	0.0684***	0.0353***	0.0776***
	(0.000)	(0.000250)	(0.000360)	(0.00160)	(0.00203)	(0.00236)
Others	0.006***	0.00135***	0.00856***	0.0557***	0.0189***	0.0697***
	(0.000)	(0.000510)	(0.000571)	(0.00301)	(0.00457)	(0.00407)
Labour inc. recipient	0.002***	0.000342	0.00691***	-0.0133**	-0.0177**	0.00212
	(0.000)	(0.000610)	(0.000863)	(0.00554)	(0.00754)	(0.00684)
Pensioners	-0.000	-0.00292***	-0.00477***	0.00662***	-0.0195***	-0.0179***
	(0.000)	(0.000248)	(0.000389)	(0.00179)	(0.00221)	(0.00282)
Multi-job - Dependent	0.018***	0.0114***	0.0213***	0.0803***	0.0670***	0.105***
	(0.001)	(0.000230)	(0.000284)	(0.00208)	(0.00228)	(0.00193)
Self-employed	0.004***	0.00502***	0.0103***	0.0422***	0.0483***	0.0789***
	(0.000)	(0.000594)	(0.000718)	(0.00385)	(0.00573)	(0.00522)
Dependent/Self-employed	0.031***	0.0158***	0.0268***	0.140***	0.103***	0.146***
	(0.001)	(0.000336)	(0.000449)	(0.00428)	(0.00305)	(0.00303)
Manufacturing	0.005***	0.00347***	0.00571***	0.0133***	0.0293***	-0.00849*
	(0.000)	(0.000576)	(0.000638)	(0.00381)	(0.00641)	(0.00508)
Electricity, gas, air	0.003***	0.000750	0.00531***	-0.00424	-0.00727	0.000557
	(0.001)	(0.00117)	(0.00102)	(0.00535)	(0.00998)	(0.00686)
Construction	-0.002***	-0.000281	-0.00261***	-0.0245***	-0.00192	-0.0265***
	(0.001)	(0.000858)	(0.000773)	(0.00464)	(0.00912)	(0.00586)
Wholesale and retail trade	0.004***	0.00232***	0.00606***	0.0172***	0.0310***	0.0156***
	(0.000)	(0.000554)	(0.000642)	(0.00380)	(0.00625)	(0.00494)
Transportation, Information and communication	0.000	-0.00235***	0.00187***	-0.0183***	-0.0285***	-0.0131***
	(0.000)	(0.000624)	(0.000655)	(0.00387)	(0.00647)	(0.00499)
Accommodation and food service	-0.002**	-0.00605***	0.00182	0.00225	-0.0381***	0.0149
	(0.001)	(0.000994)	(0.00122)	(0.00717)	(0.0106)	(0.00960)
Real estate activities	-0.002***	-0.00284***	-0.00295***	-0.0163**	-0.0193*	-0.0183**
	(0.001)	(0.000957)	(0.00111)	(0.00661)	(0.0104)	(0.00865)
Professional and technical activities	0.002***	-0.00112*	0.00620***	0.00381	-0.0150**	0.0173***
	(0.000)	(0.000611)	(0.000767)	(0.00426)	(0.00646)	(0.00575)
Administrative and support service	-0.001	-0.000142	-0.00197**	-0.00267	0.0137*	-0.000449
	(0.001)	(0.000643)	(0.000922)	(0.00540)	(0.00812)	(0.00719)
Public administration and defence	0.003***	0.00193***	0.00394***	0.0166***	0.0116*	0.00932
	(0.001)	(0.000623)	(0.000799)	(0.00427)	(0.00641)	(0.00580)
Education	-0.006***	-0.00712***	-0.00492***	-0.0511***	-0.0577***	-0.0432***
	(0.001)	(0.000653)	(0.000903)	(0.00457)	(0.00663)	(0.00649)
Social work activities	-0.010***	-0.00852***	-0.0121***	-0.0742***	-0.0644***	-0.0776***
	(0.001)	(0.00122)	(0.00212)	(0.00936)	(0.0112)	(0.0146)
Arts, entertainment	-0.006***	-0.00845***	-0.00405***	-0.0291***	-0.0525***	-0.0122
	(0.001)	(0.00107)	(0.00127)	(0.00712)	(0.0104)	(0.00960)
Other service activities	-0.004***	-0.00767***	-0.00177	-0.0319***	-0.0659***	-0.0199**
	(0.001)	(0.00104)	(0.00120)	(0.00638)	(0.00993)	(0.00845)
Hospital activities	0.071***	0.0151***	0.0343***	0.274***	0.115***	0.193***
	(0.007)	(0.00112)	(0.00206)	(0.0201)	(0.00972)	(0.0139)
Medical and dentral activities	0.026***	0.00862***	0.0195***	0.119***	0.0616***	0.102***
	(0.004)	(0.00127)	(0.00227)	(0.0175)	(0.0110)	(0.0154)
Other health activities	0.024***	0.00881***	0.0171***	0.125***	0.0704***	0.0924***
	(0.003)	(0.00118)	(0.00217)	(0.0168)	(0.0103)	(0.0148)
Financial service activities	0.023***	0.00376***	0.0201***	0.0913***	0.0169***	0.0911***
	(0.001)	(0.000584)	(0.000738)	(0.00640)	(0.00613)	(0.00532)
Insurance	0.007***	-0.000810	0.0121***	0.0522***	0.00829	0.0577***
	(0.001)	(0.000834)	(0.00115)	(0.00797)	(0.00805)	(0.00780)
Auxiliary activities to finanical ss.	0.012***	0.00366***	0.0128***	0.0474***	0.0169**	0.0564***
	(0.001)	(0.000823)	(0.00114)	(0.00794)	(0.00824)	(0.00818)
Stock corporation	0.007***	0.00698***	0.00986***	0.0175***	0.0118	0.0271***
	(0.000)	(0.000754)	(0.000989)	(0.00599)	(0.00929)	(0.00840)
Public sector	0.006***	0.00606***	0.00775***	0.00389	-0.00487	0.0132
	(0.000)	(0.000791)	(0.00110)	(0.00625)	(0.00949)	(0.00901)
Observations	1,952,876	986,420	966,377	240,044	102,470	137,561

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

The set of dummy variables reflecting type of business shows that individuals that work or withdraw dividends from corporations receive larger earnings, and these effects substantially increase when observing differences within the top 10%. Approximately 60% of the individuals at the top 1% are occupied in corporations. Being a public employee is also positively associated to belonging to the top 1% versus the rest, but when we compare within the top 10%, these coefficients fall and lose statistical significance.[35]

The comparison of the estimations obtained for 2009, 2013 and 2016 (Table A.8), indicates a that the occupations associated to health activities increased their probability of being a the top of the income distribution and widened their distance with the classes pertaining to the financial sector. At the same time, the association among being a liberal professional or receiving dividends and belonging to the top 1% increased, and the same happened (but at lower absolute levels) with property income. This pattern is in sharp contrast with the coefficients reflecting the reception of labour income, whose magnitude fell and even changed sign in 2016.

7. Final remarks

As in most Latin American countries, previous studies based on household survey micro-data have shown that Uruguay experienced a substantial decrease in inequality in the period 2008-2013, which resulted from high economic growth rates that fostered the demand for unskilled workers, coupled with a package of reforms that included the restoration of centralized wage-setting mechanisms, the inception of a progressive personal income taxation scheme and the expansion of non contributory cash transfers (Amarante et al., 2014). To determine whether this trend resulted from household surveys draw-backs in capturing the upper tail of the income distribution, in this article we analysed primary income inequality among the adult population aged 20 and above, creating a corrected tax records income variable and comparing it with harmonized household survey micro-data. Differently from previous studies for other Latin American countries, we had access to a unique data-set that covers a substantial fraction of the adult population; this allowed us to include informal income and correct underreporting from formal occupations in the the personal income tax records

[35]We also estimated a set of quantile regressions valuated at the median, the top 10, 1 and 0.1% that are available on request to the authors. The results are consistent with the ones obtained from the probit models estimations. The magnitudes of the coefficients associated to receiving capital income and its components substantially increase with the quantile. Conversely, receiving labor earnings is more relevant in the median and the magnitudes of the coefficients are considerably lowered at top points of the distribution. Again, being a pensioner yields a negative sign along all the quantiles considered. The patterns regarding industry and business type are similar to the ones obtained in the probit estimations.

distribution, to compute both synthetic indices and top income shares, and to investigate the characteristics of the top income holders.

We found that, in both databases, synthetic indices experienced a statistically significant reduction (although milder for corrected tax income) in 2009-2013, which remained unchanged afterwards. The top 10% share in our corrected tax income variable mimicked this evolution, although in 2016 experienced a statistically significant increase. It is still soon to understand whether this is reflecting a new trend or it is a point variation. At the same time, the income share accrued by the top 1% was stable and grew slightly in our corrected tax micro-data income variable in the last years, whereas it fell significantly in the harmonized household survey income throughout the whole period. We carried out a wide set of robustness checks that strengthened these findings. Our study contributes further evidence to that already provided by Alvaredo and Londoño Velez (2014), Flores et al. (2020) and Morgan (2017b) for Colombia, Chile and Brazil on the divergence between household survey inequality measures and top income shares based on tax data.

Whereas the inequality indices within the bottom 99% present a declining trend in both datasets, the different trajectories of the top 1% explain the diverging trend in top income shares. In the harmonized household survey data, inequality within the the top 1% experienced a 35% reduction that contributed substantially to the overall equalization observed in 2009-2013. Meanwhile, for the corrected tax income, the top 1% experienced an increasing concentration trend over the years, which we document in several ways. After 2012, the inequality reduction at the bottom 99% could not offset the concentration at the top.

The significant inequality reduction experienced by the harmonized household survey income in the top 1% and the income redistribution observed for the bottom 50% of the tax-records distribution convey the idea that these differences are driven by the eroded ability of ECH to capture the upper tail of the distribution, rather than by the formalization process or an improved capacity of the tax authorities to reach the rich. Moreover, the increased inequality at the tax records top-end is mainly explained by the increasing share of capital income, which can be associated with a higher misreporting in household survey data. Our decomposition exercise shows that increased participation of capital income, along with the augmented inequality within this income source and the rise in the participation of dividends, accounts for the increase in the proportion of income held by the top 1%. These findings also highlight the relevance of monitoring and renewing the ways in which household surveys gather information and the need to articulate this information with other valuable data-sources, such as information from tax records.

Our study suggests that the recent fall in inequality in Uruguay was driven by equalization at the bottom and middle of the distribution, whereas the top remained unchanged. The meagre effect of personal income taxation provide further evidence on the weaknesses of redistributive policies and dual tax schemes in reaching the top-end of the distribution. The Uruguayan effective rates are relatively low when compared with those of the OECD countries, although they are double the available ones for Colombia (Alvaredo and Londoño Velez, 2014).

We also document that the Uruguayan top income holders are mainly male and obtain a significant proportion of their earnings from capital income and, specifically, dividends. Different from the available information for developed countries, labour earnings at the top are highly concentrated in the health and professional services' sectors. Broader issues such as analysing the socio-economic stratification on the basis of a wider scope of variables need to be investigated further.

Although our results indicate that the dividends obtained by top income holders are generated in a wide set of industries, it is worth mentioning that this empirical exercise assessed national income and, thus considered approximately 15% to 33% of the total amount of dividends generated in Uruguay. Consequently it lacked information on non resident owners of domestic assets. The consideration of dividends that are remitted abroad might lead to a very different characterization of the top of the distribution. A similar point holds for income obtained abroad by Uruguayan residents, as the recent literature on tax havens has suggested (Zucman, 2013, 2014). These specific features of small open economies need to be studied in further research.

The apparent contradiction between the stability of the top income shares and the evolution of the Gini and Theil indices in our tax based income variables calls into discussion several issues related to the kind of inequality reduction is sought. Furthermore, it contributes to the appraisal of the relationship between economic growth and redistribution as well as the extent of the equalizing effect and limitations of the menu of redistributive policies launched in Latin America and in Uruguay in the last two decades. As Lemieux and Riddell (2015) argue, most of these interventions affect the low, middle and upper-middle sectors, rather than the top incomes.

A. Appendix: Figures and Tables

Table A.1: Characteristics of the data sources used in this study

Data source	Unit	Population coverage (*)	Income variable used in this article	Time coverage
Tax records	Individuals	Formal earners (potential income tax payers receiving labour, capital or pensions income)	Pre and post tax income by income source. It does not include non taxable income (e.g. cash transfers, imputed owners housing value)	2009-2016
Household survey	Households/ Individuals	All income earners (formal and informal income from all sources).	i) Subset of individuals aged 20 or more with 0 income or being informal earners only; ii) Ratios of informal to formal income for individuals simultaneously receiving both types of income	2009-2016
Linked hh survey - tax records	Households/ Individuals	Sub-sample of the household survey with children aged 0 to 3 in 2012/13 with positive income in tax records and household survey	Ratios of tax records to household survey harmonized income for the subset of linked observations	2012/2013
Firms balance sheets	Firms	Firms required to provide annual balance sheets to the tax authorities (annual income above 40000UI)	Withdrawals from firm owners that had not been distributed as profits in next year	2009-2016
Population projections	Individuals	Uruguayan population aged 20 years or more	-	2009-2016

Note. (*)We restrict the population to individuals aged 20 or more.

Table A.2: Top fractiles thresholds by data source, 2009-2016

	2009	2010	2011	2012	2013	2014	2015	2016
Top 1 - threshold	1,036,537	1,157,498	1,302,751	1,526,879	1,656,311	1,912,940	2,100,272	2,404,508
Top 1 - threshold (survey)	980,025	1,048,896	1,112,222	1,121,837	1,316,246	1,499,245	1,650,291	1,792,000
Survey/Tax	95%	91%	85%	73%	79%	78%	79%	75%
Top 10 - threshold	320,095.5	361,134.7	408,598.2	475,083.4	563,590.3	612,658.6	669,908.3	751,771.8
Top 10 - threshold (survey)	334,079.9	361,940.4	411,079.1	458,437.8	520,729.2	579,817.1	644,707.0	701,523.9
Survey/Tax	104%	100%	101%	96%	92%	95%	96%	93%

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). The first block depicts the top 1%'s share in the tax records and harmonized survey.

Table A.3: Inequality measures- bootstrap confidence intervals (95%). Selected indicators, 2009-2016

	Gini index			Top 1%			Top 10%			Top 0,1%		
Year	Point est.	Lower b.	Upper b.	Point est.	Lower b.	Upper b.	Point est.	Lower b.	Upper b.	Point est.	Lower b.	Upper b.
2009	0,500	0,497	0,504	13,5%	13,0%	14,1%	43,8%	43,4%	44,2%	50,0%	49,7%	50,4%
2010	0,503	0,499	0,507	13,5%	13,0%	14,3%	43,5%	43,1%	44,0%	50,3%	49,9%	50,7%
2011	0,477	0,472	0,485	13,5%	12,7%	14,8%	42,0%	41,5%	42,8%	47,7%	47,2%	48,5%
2012	0,484	0,480	0,489	13,2%	12,6%	14,1%	42,0%	41,5%	42,6%	48,4%	48,0%	48,9%
2013	0,469	0,464	0,476	12,7%	11,8%	13,7%	41,0%	40,4%	41,7%	46,9%	46,4%	47,6%
2014	0,476	0,473	0,479	13,2%	12,7%	13,6%	41,1%	40,7%	41,4%	47,6%	47,3%	47,9%
2015	0,468	0,463	0,473	13,5%	12,8%	14,4%	40,6%	40,0%	41,2%	46,8%	46,3%	47,3%
2016	0,486	0,482	0,490	14,6%	14,1%	15,4%	41,9%	41,5%	42,4%	48,6%	48,2%	49,0%

Note. Own elaboration based on DGI and ECH. Bootstraps with 100 repetitions.

Table A.4: Redistributive effect of direct taxation. Pre and post-tax corrected tax income, 2009-2016

	2009	2010	2011	2012	2013	2014	2015	2016
Bottom 50	4.38%	4.71%	4.68%	4.82%	4.79%	5.26%	5.02%	5.22%
50 - 90	3.81%	3.94%	3.79%	3.69%	3.50%	3.65%	3.30%	3.34%
90 - 99	-3.58%	-3.64%	-4.05%	-4.16%	-4.14%	-4.61%	-4.58%	-4.49%
Top 10	-5.04%	-5.31%	-5.48%	-5.44%	-5.45%	-5.75%	-5.40%	-5.19%
Top 5	-6.68%	-6.99%	-6.98%	-6.80%	-6.84%	-7.05%	-6.47%	-6.15%
Top 10	-8.31%	-9.03%	-8.50%	-8.20%	-8.36%	-8.14%	-7.03%	-6.50%
Top 0.5	-8.42%	-9.39%	-8.53%	-8.35%	-8.19%	-8.00%	-6.75%	-6.21%
Top 0.1	-7.14%	-8.96%	-7.12%	-6.92%	-5.99%	-5.92%	-4.75%	-4.40%
Gini Index	-0.015	-0.016	-0.017	-0.017	-0.017	-0.018	-0.017	-0.017
Theil Index	-0.074	-0.075	-0.072	-0.098	-0.073	-0.07	-0.052	-0.071

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Table A.5: Inequality decompositions by income source. 2009 - 2016. Corrected tax income (Y_3) and harmonized survey income

		2009	2010	2011	2012	2013	2014	2015	2016
Corrected tax income - Y3									
Gk	Labour	0.620	0.624	0.585	0.601	0.583	0.590	0.580	0.597
	Pensions	0.819	0.813	0.823	0.810	0.813	0.813	0.812	0.810
	Capital	0.989	0.990	0.991	0.984	0.985	0.986	0.990	0.990
	Mixed	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
Share	Labour	0.774	0.782	0.803	0.767	0.783	0.758	0.750	0.754
	Pensions	0.101	0.098	0.057	0.083	0.069	0.079	0.077	0.075
	Capital	0.106	0.101	0.120	0.130	0.129	0.144	0.155	0.153
	Mixed	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009
Harmonized survey income									
		2009	2010	2011	2012	2013	2014	2015	2016
Gk	Labour	0.650	0.641	0.612	0.597	0.596	0.588	0.594	0.592
	Pensions	0.827	0.820	0.830	0.826	0.825	0.829	0.825	0.819
	Capital	0.967	0.967	0.965	0.961	0.968	0.968	0.967	0.967
Share	Labour	0.852	0.848	0.878	0.893	0.864	0.874	0.868	0.868
	Pensions	0.100	0.107	0.085	0.085	0.088	0.074	0.083	0.081
	Capital	0.047	0.045	0.036	0.022	0.048	0.053	0.049	0.051

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Table A.6: Gini index above different income thresholds

Year	Top 1% (Corrected survey income)	Top 1% (Corrected tax income)	Max. survey
2009	0.342	0.347	0.459
2010	0.348	0.356	0.448
2011	0.366	0.38	0.477
2012	0.344	0.365	0.435
2013	0.35	0.39	0.474
2014	0.361	0.38	0.443
2015	0.381	0.402	0.47
2016	0.398	0.417	0.444

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Table A.7: Industries ranking according to their share in top income earners income by income source (ranked by top 1% of corrected tax income - 2016)

	Labour income			
	Top 1	Top 5	Top 10	Average
Human healthcare activities - hospitals	23.3%	11.4%	9.0%	3.7%
Financial intermediation	8.8%	8.0%	4.5%	0.7%
General public administration	4.7%	10.9%	12.1%	8.6%
Other human health act.	2.3%	2.5%	2.6%	2.1%
Medical and dental healthcare	1.8%	1.4%	1.3%	1.0%
Non-life insurance	1.5%	1.9%	1.4%	0.4%
Other professional and scientific act.	1.5%	3.1%	3.1%	1.3%
Activities of collection agencies	1.4%	0.6%	0.4%	0.3%
Wholesale of pharmaceutical and medical goods	1.3%	1.0%	0.7%	0.3%
Manufacture of pharmaceuticals and medicinal products	1.2%	1.1%	1.0%	0.4%
	Liberal Professions			
	Top 1	Top 5	Top 10	Average
Human healthcare activities - hospitals	7.9%	5.6%	4.9%	3.7%
Non-life insurance	4.4%	3.7%	3.3%	2.9%
Construction of buildings	2.9%	2.8%	2.7%	2.8%
Medical and dental healthcare	2.9%	1.8%	1.7%	1.2%
General public administration	2.4%	2.9%	2.8%	2.2%
Other professional and scientific act.	1.8%	1.3%	1.2%	1.1%
Processing and preserving of meat	1.5%	1.7%	1.6%	1.4%
Manufacture of pharmaceuticals and medicinal products	1.4%	1.4%	1.3%	1.0%
Real estate act.	1.0%	0.8%	0.7%	0.6%
Pre-primary and primary education	1.0%	1.0%	1.0%	0.9%
	Business income			
	Top 1	Top 5	Top 10	Average
Human healthcare activities - hospitals	4.3%	3.9%	3.5%	1.8%
Activities of collection agencies	2.4%	2.5%	2.5%	1.3%
Raising of cattle	2.0%	1.8%	1.8%	1.0%
Medical and dental healthcare	1.7%	2.0%	1.9%	1.1%
Retail sale of automobile fuel	1.6%	1.4%	1.3%	0.8%
Other professional and scientific act.	1.5%	1.5%	1.3%	0.7%
Construction of buildings	1.4%	1.3%	1.2%	0.7%
Freight transport by road	1.4%	1.0%	1.0%	0.6%
Retail sale in non-specialized stores	1.4%	1.5%	1.3%	0.8%
Gambling and betting activities	1.1%	1.3%	1.3%	0.9%

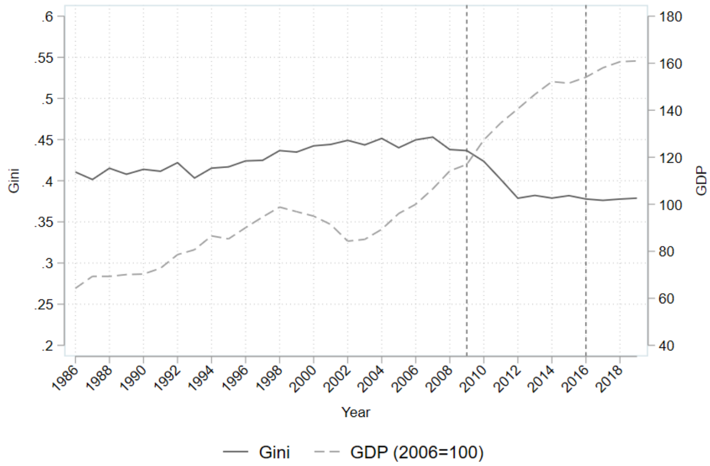
Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Table A.8: Probability of belonging to the top 1% (versus bottom 99% or centiles 90-99, 2009, 2013 and 2016. Probit estimates-marginal effects.)

	Top 1 vs bottom 99%			Top 1 vs remaining Top 10%		
	2009	2013	2016	2009	2013	2016. Probit estimates. Marginal effects
Male	0.001***	0.001***	0.001***	0.00255**	0.0175***	0.0153***
	(0.000)	(0.000)	(0.000)	(0.00125)	(0.00136)	(0.00130)
Age	0.001***	0.001***	0.000***	0.00500***	0.00679***	0.000488***
	(0.000)	(0.000)	(0.000)	(0.000261)	(0.000253)	(1.70e-05)
Age2	-0.000***	-0.000***	-0.000***	-2.81e-05***	-3.71e-05***	-1.89e-07***
	(0.000)	(0.000)	(0.000)	(2.43e-06)	(2.38e-06)	(9.39e-09)
Liberal professional	0.054***	0.104***	0.103***	0.223***	0.384***	0.332***
	(0.003)	(0.006)	(0.004)	(0.00954)	(0.0163)	(0.00994)
Capital inc. recipient	0.070***	0.031***	0.022***	0.139***	0.0536***	0.0555***
	(0.001)	(0.000)	(0.000)	(0.00141)	(0.00141)	(0.00140)
Dividends	0.026***	0.028***	0.030***	0.144***	0.184***	0.173***
	(0.001)	(0.000)	(0.000)	(0.00509)	(0.00375)	(0.00289)
Property rents	0.001***	0.009***	0.011***	0.0308***	0.0706***	0.0684***
	(0.000)	(0.000)	(0.000)	(0.00174)	(0.00177)	(0.00160)
Others	-0.001**	0.003***	0.006***	0.0258***	0.0562***	0.0557***
	(0.000)	(0.000)	(0.000)	(0.00317)	(0.00337)	(0.00301)
Labour inc.recipient	0.022***	0.008***	0.002***	0.0784***	0.0314***	-0.0133**
	(0.001)	(0.000)	(0.000)	(0.00232)	(0.00487)	(0.00554)
Pensioner	-0.009***	-0.007***	-0.000	-0.0582***	-0.0269***	0.00662***
	(0.000)	(0.000)	(0.000)	(0.00172)	(0.00181)	(0.00179)
Multi-job - Dependent	0.004***	0.004***	0.018***	0.0305***	0.0258***	0.0803***
	(0.000)	(0.000)	(0.001)	(0.00172)	(0.00187)	(0.00208)
Self-employed	0.001***	-0.001**	0.004***	0.00332	-0.00423	0.0422***
	(0.000)	(0.000)	(0.000)	(0.00306)	(0.00302)	(0.00385)
Dependent/Self-employed	0.024***	0.031***	0.031***	0.109***	0.130***	0.140***
	(0.001)	(0.001)	(0.001)	(0.00318)	(0.00341)	(0.00428)
Manufacturing	0.005***	0.006***	0.005***	0.0145***	0.0137***	0.0133***
	(0.000)	(0.000)	(0.000)	(0.00319)	(0.00426)	(0.00381)
Electricity, gas, air	0.005***	0.005***	0.003***	0.00219	0.0107*	-0.00424
	(0.001)	(0.001)	(0.001)	(0.00518)	(0.00612)	(0.00535)
Construction	-0.003***	-0.002***	-0.002***	-0.0185***	-0.0342***	-0.0245***
	(0.001)	(0.001)	(0.001)	(0.00500)	(0.00516)	(0.00464)
Wholesale and retail trade	0.006***	0.006***	0.004***	0.0250***	0.0266***	0.0172***
	(0.000)	(0.000)	(0.000)	(0.00332)	(0.00426)	(0.00380)
Transportation, Information and communication	0.002***	0.001***	0.000	-0.0215***	-0.0179***	-0.0183***
	(0.000)	(0.000)	(0.000)	(0.00341)	(0.00438)	(0.00387)
Accommodation and food services	0.002**	-0.001	-0.002**	0.0196***	-0.00574	0.00225
	(0.001)	(0.001)	(0.001)	(0.00676)	(0.00732)	(0.00717)
Real estate activities	-0.003***	-0.001*	-0.002***	-0.0315***	-0.0175**	-0.0163**
	(0.001)	(0.001)	(0.001)	(0.00695)	(0.00728)	(0.00661)
Professional and technical activities	0.003***	0.005***	0.002***	-0.00510	0.0216***	0.00381
	(0.001)	(0.001)	(0.000)	(0.00417)	(0.00481)	(0.00426)
Administrative and support service	-0.005***	-0.005***	-0.001	-0.0190***	-0.0248***	-0.00267
	(0.001)	(0.001)	(0.001)	(0.00656)	(0.00686)	(0.00540)
Public administration and defence	0.006***	0.005***	0.003***	0.0168***	0.0178***	0.0166***
	(0.000)	(0.001)	(0.001)	(0.00349)	(0.00475)	(0.00427)
Education	-0.009***	-0.003***	-0.006***	-0.0853***	-0.0461***	-0.0511***
	(0.001)	(0.001)	(0.001)	(0.00454)	(0.00508)	(0.00457)
Social work activities	0.002***	-0.008***	-0.010***	-0.00528	-0.0833***	-0.0742***
	(0.000)	(0.001)	(0.001)	(0.00358)	(0.00977)	(0.00936)
Arts, entertainment	-0.003***	-0.006***	-0.006***	-0.0269***	-0.0291***	-0.0291***
	(0.001)	(0.001)	(0.001)	(0.00522)	(0.00848)	(0.00712)
Other service activities	-0.003***	-0.001	-0.004***	-0.0407***	-0.0248***	-0.0319***
	(0.001)	(0.001)	(0.001)	(0.00618)	(0.00691)	(0.00638)
Hospital activities	0.013***	0.070***	0.071***	0.0613***	0.295***	0.274***
	(0.001)	(0.006)	(0.007)	(0.00435)	(0.0204)	(0.0201)
Medical and dentral activities	0.000	0.034***	0.026***	-0.0131**	0.168***	0.119***
	(0.001)	(0.004)	(0.004)	(0.00550)	(0.0203)	(0.0175)
Other health activities	-0.002***	0.017***	0.024***	-0.0171***	0.104***	0.125***
	(0.001)	(0.003)	(0.003)	(0.00436)	(0.0164)	(0.0168)
Financial service activities	0.013***	0.037***	0.023***	0.0414***	0.155***	0.0913***
	(0.001)	(0.002)	(0.001)	(0.00366)	(0.00887)	(0.00640)
Insurance	0.002***	0.018***	0.007***	0.0213***	0.0866***	0.0522***
	(0.001)	(0.002)	(0.001)	(0.00466)	(0.0102)	(0.00797)
Auxiliary activities to finanical ss.	0.009***	0.014***	0.012***	0.0266***	0.0630***	0.0474***
	(0.001)	(0.002)	(0.001)	(0.00639)	(0.0100)	(0.00794)
Corporations	0.009***	0.010***	0.007***	0.0168**	0.0485***	0.0175***
	(0.000)	(0.000)	(0.000)	(0.00701)	(0.00566)	(0.00599)
Public sector	0.009***	0.007***	0.006***	0.0205***	0.0205***	0.00389
	(0.000)	(0.000)	(0.000)	(0.00735)	(0.00594)	(0.00625)
Observations	1,812,067	1,904,641	1,952,876	230,281	223,133	240,044

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

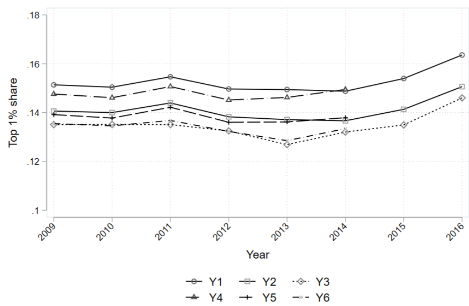
Fig. A.1. Inequality trends in Uruguay. Per capita household income. 1986-2019



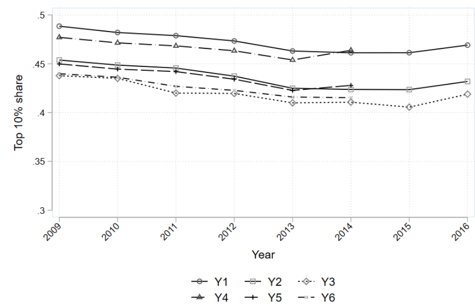
Note. Own calculations based on ECH micro-data and System of National Accounts (from Uruguay's Central Bank, *BCU*). Per-capita household income includes all cash and in-kind income sources and rental imputed income. Incomes adjusted at December 2006, based on consumer prices index. For a complete description of the household survey, see Section 4. Vertical lines indicate the period under analysis in this study.

Fig. A.2. Pre-tax top income shares, 2009-2016. Method 1. Alternative income variables.

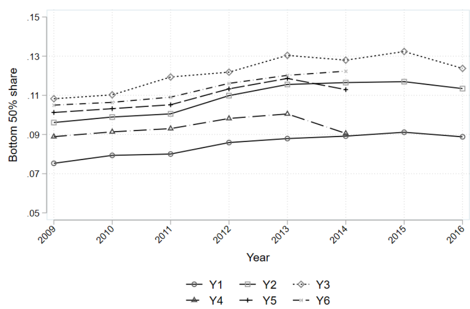
(a) Top 1%



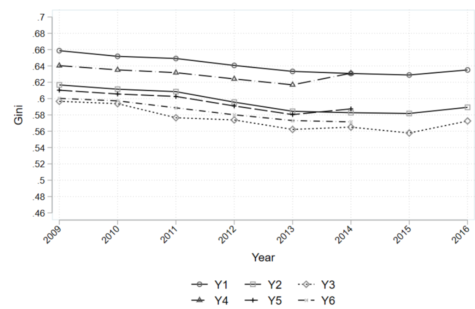
(b) Top 10%



(c) Bottom 50%

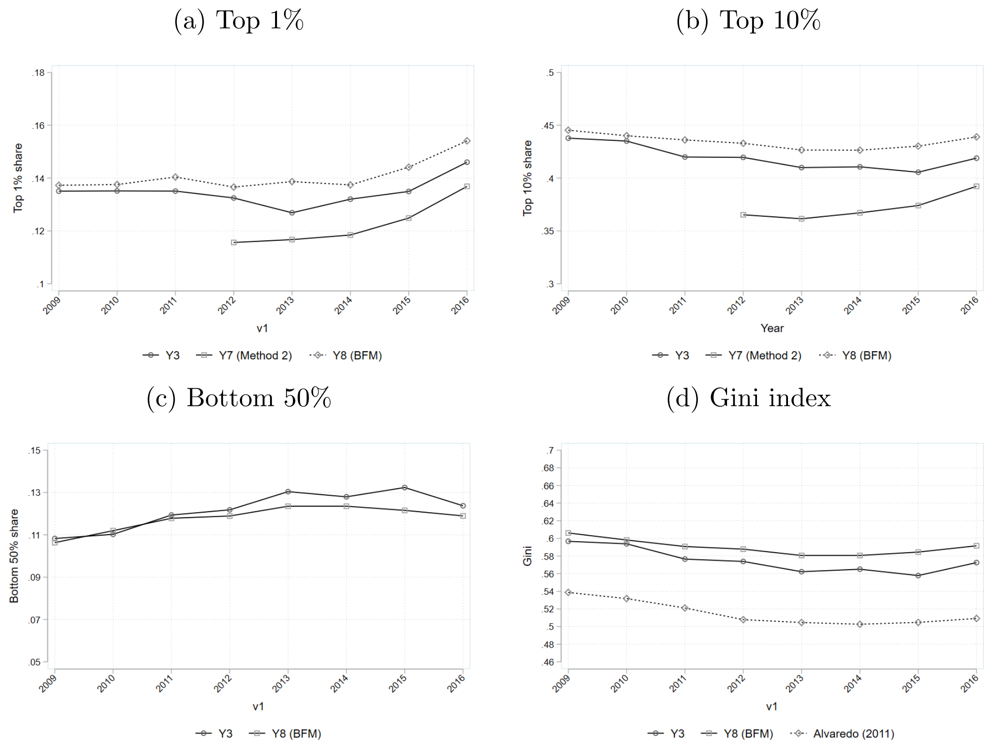


(d) Gini



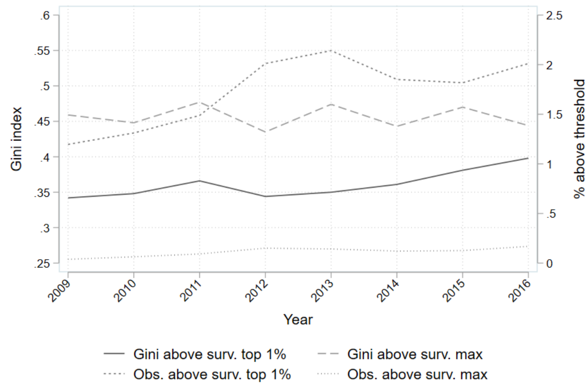
Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Fig. A.3. Pre-tax top income shares, 2009-2016. Method 2 and BFM.



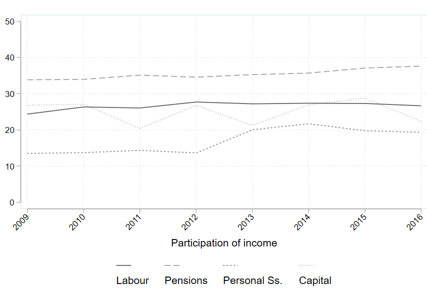
Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). BFM and Alvaredo are the Blanchet et al. (2022) and Alvaredo (2011) survey and tax corrections respectively.

Fig. A.4. Inequality trends for selected pre-tax top income groups (above survey's top 1% threshold), 2009-2016

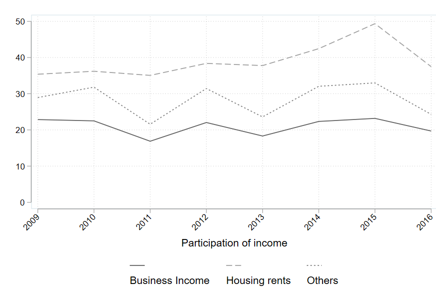


Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE) (estimates in Table A.6). Survey's highest value set at the average of the 50 higher (comparable) income, excluding the highest. All incomes at 2016 prices. The brown and blue lines illustrate the proportion of corrected tax income observations belonging to each group 1) observations with income above the 1% threshold in harmonized survey and below survey's maximum and 2) observations with income above survey's maximum. The green line represents the Gini index computed upon corrected tax income for the subset of observations beyond the survey threshold (groups 1+2)

Fig. A.5. Participation of women in the top 1% of pre-tax corrected tax income by income source, 2009-2016.



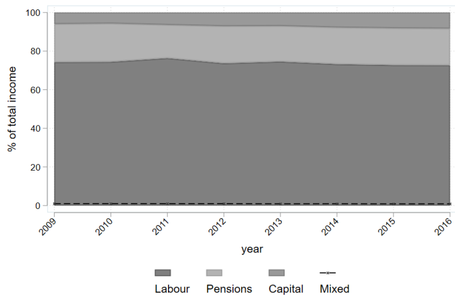
(a) Total income



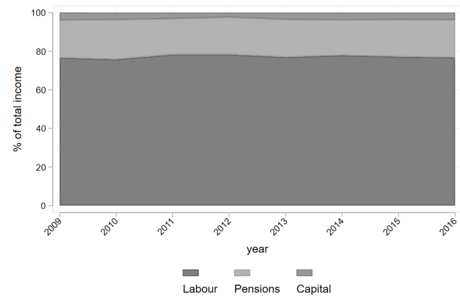
(b) Capital income

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

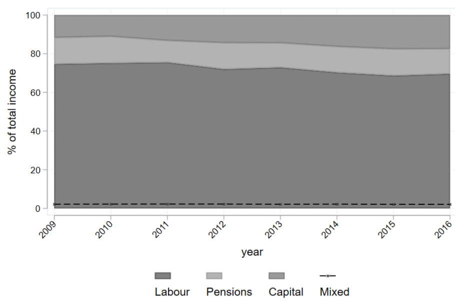
Fig. A.6. Composition of income. Pre-tax corrected tax income and survey income, 2009-2016. Average and top 10%



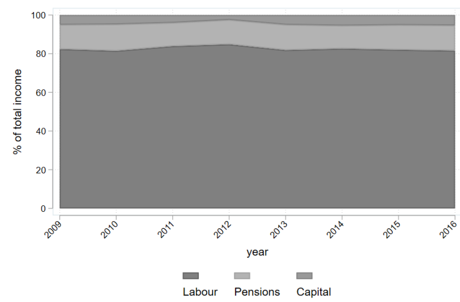
(a) Corrected tax income - Average



(b) Survey income - Average



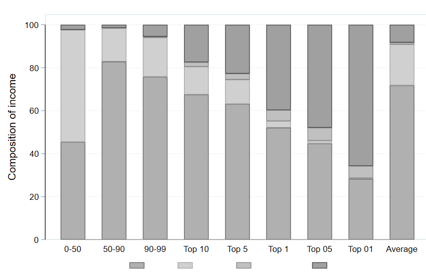
(c) Corrected tax income - Top 10%



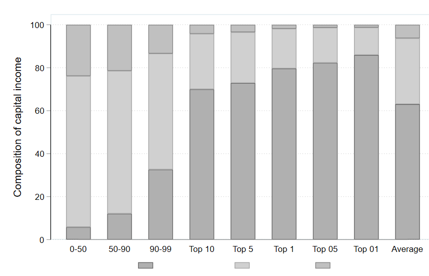
(d) Survey income - Top 10%

Note. Own calculations based on tax records (DGI) and household survey (ECH). In tax records, mixed incomes is depicted as a share of labour income for comparison purposes.

Fig. A.7. Income distribution by source and fractile



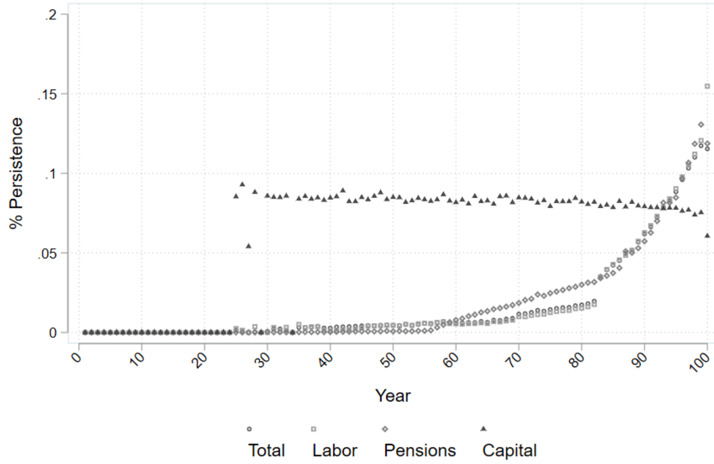
(a) Total income



(b) Capital income

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE).

Fig. A.8. Effective tax rates by income source. Pre-tax corrected tax income, 2016.



Note. Own calculations based on tax records (DGI). Effective tax rates for total income and all income sources are depicted.

B. Appendix: Additional Figures and Tables

Table B.1: Capital incomes tax rates

Capital income concept	Tax rate
Interests from bank deposits in Uruguayan currency (more than one year) and debt titles interests-3 years or more	3%
Interests from bank deposits in Uruguayan currency (less than one year)	5%
Dividends	7%
Income from Land and property	12%
Others rents (sports persons royalties, authors royalties, everlasting rents)	12%

Note. Own elaboration based on DGI (2019).

Table B.2: Labour income tax rates

Income bracket (BPC)	Tax 2009-2011	Income bracket (BPC)	Tax rate 2012-2016
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	25	900-1380	25
		¿1380	30

Note. Own elaboration based on DGI (2019).

Table B.3: Tax rates on pensions

Pension income bracket (BPC)	Tax rate
0-96	0
96-180	10
180-600	20
¿600	25

Note. Own elaboration based on DGI (2019).

Table B.4: Number of income receivers and taxpayers by income source. DGI personal income tax records

		2009	2010	2011	2012	2013	2014	2015	2016
Labour income	Total	1,187,913	1,183,629	1,237,034	1,222,505	1,272,881	1,297,408	1,313,961	1,310,285
	Taxpayers	315,300	347,001	395,207	416,318	471,838	510,567	753,705	770,127
Employed	Total	1,127,943	1,111,782	1,161,260	1,143,757	1,190,855	1,216,827	1,253,834	1,237,214
	Taxpayers	276,664	300,461	345,480	363,546	416,530	454,957	706,868	715,150
Self employed	Total	51,024	53,489	55,676	54,958	57,956	57,998	40,509	51,705
	Taxpayers	28,760	30,405	31,823	31,684	33,653	34,957	36,533	44,843
Irae	Total	3,504	3,607	3,687	3,899	4,016	4,128	3,970	4,338
	Taxpayers	3,173	3,253	3,348	3,503	3,619	3,676	3,516	3,826
Pensions	Total	639,540	661,366	627,764	684,320	690,830	698,594	709,216	715,801
	Taxpayers	102,136	112,445	111,787	137,988	148,749	158,991	170,184	173,867
Capital	Total	261,765	298,431	323,035	390,660	445,263	385,352	586,851	656,789
	Taxpayers	255,697	293,041	318,012	386,745	441,457	380,569	582,905	652,258
Dividends	Total	3,134	3,437	4,539	5,297	5,933	6,752	8,473	9,339
	Taxpayers	3,134	3,437	4,539	5,297	5,933	6,752	8,473	9,339
Real state rents	Total	55,205	55,089	57,759	58,600	61,102	66,076	70,032	73,771
	Taxpayers	50,829	50,711	54,800	57,212	59,969	65,028	69,196	72,905

Note. Own calculations based on tax records (DGI).

Table B.5: Non nominative capital income share relative to total capital incomes

	2009	2010	2011	2012	2013	2014	2015	2016
Interests bank deposits, local currency (more 1 yr.)	99,8%	100,0%	97,5%	100,0%	100,0%	100,0%	100,0%	100,0%
Interests bank deposits (no indexation)	99,9%	100,0%	98,3%	100,0%	100,0%	100,0%	100,0%	100,0%
Debt titles interests with 3 yrs. or more	41,2%	34,2%	48,1%	96,2%	74,6%	97,6%	91,1%	79,6%
Remaining financial and mobiliary capital rents	62,9%	52,2%	47,4%	59,2%	54,4%	44,3%	49,1%	48,1%
Dividends and utilities	31,3%	39,3%	42,7%	47,2%	38,7%	39,3%	36,9%	34,6%
Sportpersons royalties	10,4%	2,5%	54,0%	8,8%	13,4%	-11,8%	0,9%	-4,4%
Authors royalties	-73,0%	-73,7%	-51,8%	-70,0%	-63,0%	-62,4%	-64,3%	-64,3%

Note. Own elaboration based on tax records (DGI).

Table B.6: Dividends payments received by residents and non-residents. Absolute amount and share in total capital income

	Dividends					
Year	Non residents (1)	Residents (2)	Total Capital income (3)	(1)/(3)	(2)/(3)	(1)/(2)
2009	1.784.579.981	257.628.010	1.921.372.718	92,9%	13,4%	693%
2010	1.662.248.712	476.410.097	2.358.273.737	70,5%	20,2%	349%
2011	2.587.790.134	706.402.898	2.946.180.730	87,8%	24,0%	366%
2012	2.053.474.348	901.027.171	3.763.810.899	54,6%	23,9%	228%
2013	2.678.819.697	1.037.743.542	4.278.315.616	62,6%	24,3%	258%
2014	3.236.839.548	1.453.894.989	4.967.332.041	65,2%	29,3%	223%
2015	3.874.750.248	1.712.268.425	5.837.899.459	66,4%	29,3%	226%
2016	4.523.474.104	2.031.958.492	6.861.583.615	65,9%	29,6%	223%

Note. Based on DGI data. Taxes on dividends are 7%, while IRNR for non residents ranges from 7 to 12% depending on the type of income. Current Uruguayan pesos.

Table B.7: Income thresholds by fractile, 2009-2016. Pre-tax corrected tax income.

	2009	2010	2011	2012	2013	2014	2015	2016
Mean	188.293	210.470	250.612	285.774	339.581	371.125	417.127	465.440
P50	132.701	146.668	186.746	213.367	254.229	274.821	316.714	346.466
P90	366.068	409.350	463.200	533.994	633.251	683.249	745.044	836.371
P99	1.140.662	1.263.224	1.432.603	1.661.831	1.807.069	2.083.591	2.294.943	2.642.399
P995	1.567.457	1.741.031	1.974.811	2.284.954	2.495.197	2.919.503	3.242.469	3.839.922
P999	3.341.108	3.770.006	4.407.106	4.942.255	5.565.298	6.526.410	7.460.169	9.544.300
P9995	4.712.506	5.424.785	6.248.100	6.930.174	7.689.554	9.354.320	10.886.718	14.290.032
P9999	11.101.068	12.190.292	15.211.903	17.088.396	20.487.632	24.820.230	30.321.280	42.279.980
Mean top 000,1	32.518.506	38.539.625	54.038.430	55.624.992	73.893.026	71.746.604	92.214.760	104.682.232

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). Current annual Uruguayan pesos (1 USD approximately equivalent to 30 Uruguayan pesos in 2016).

Table B.8: Number and relative participation of individuals receiving advanced payments. 2009-2016

Number of individuals					Relative participation (%)			
Year	Sim. withdr. profits	Sim. top income earners	Additional individuals	Total	Inds. withdr./dividends	%	Inds. in tax records	%
2009	1070	3284	1552	5906	3134	188.4%	1,721,207	0.34%
2010	1611	2747	1034	5392	3437	156.9%	1,722,902	0.31%
2011	2150	3015	1350	6515	4539	143.5%	1,758,779	0.37%
2012	2280	3291	1390	6961	5297	131.4%	1,793,012	0.39%
2013	2975	3470	1435	7880	5933	132.8%	1,852,341	0.43%
2014	3430	3800	1611	8841	6752	130.9%	1,928,833	0.46%
2015	5107	4183	1865	11155	8473	131.7%	1916,230	0.58%
2016	6448	5002	2202	13652	9339	146.2%	1,923,850	0.71%

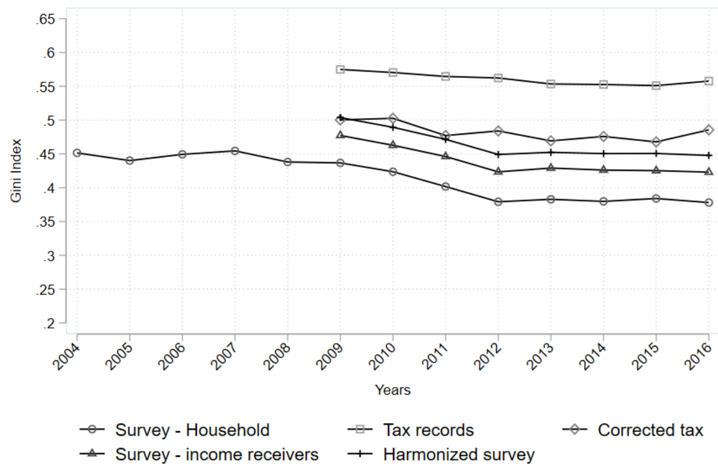
Note. The first four columns indicate the number of individuals with imputations of advanced payments, opened by their simultaneous condition in the tax records database. The fifth and seventh column respectively contain the total number of individuals in the tax records that withdrew dividends and the total number of individuals in the database. Own calculations based on tax records (DGI).

Table B.9: Inequality decomposition by income group (top 1% and bottom 99%). 2009-2016. Theil index

	2009	2010	2011	2012	2013	2014	2015	2016
	Corrected tax income (Y3)							
Theil Index	0.777	0.775	0.765	0.742	0.730	0.717	0.727	0.769
% Between	30.2%	30.3%	30.7%	30.7%	29.1%	31.6%	32.2%	34.5%
% Within	69.8%	69.7%	69.3%	69.3%	70.9%	68.4%	67.8%	65.5%
% Overlap	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Theil Bottom 99	0.554	0.545	0.511	0.505	0.488	0.486	0.469	0.487
Theil Top 1	0.463	0.510	0.655	0.574	0.721	0.525	0.643	0.600
	Harmonized survey income							
Theil Index	0.670	0.616	0.567	0.509	0.528	0.518	0.532	0.521
% Between	26.6%	23.9%	22.1%	16.4%	19.4%	19.1%	20.0%	19.2%
% Within	73.4%	76.1%	77.9%	83.6%	80.6%	80.9%	80.0%	80.8%
% Overlap	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Theil Bottom 99	0.534	0.510	0.477	0.456	0.456	0.450	0.452	0.452
Theil Top 1	0.160	0.106	0.091	0.033	0.080	0.065	0.137	0.070
	Tax records							
Theil Index	0.747	0.743	0.780	0.735	0.739	0.706	0.724	0.749
% Between	32.3%	32.4%	32.9%	32.9%	32.9%	34.3%	35.3%	37.7%
% Within	67.7%	67.6%	67.1%	67.1%	67.1%	65.7%	64.7%	62.3%
% Overlap	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Theil Bottom 99	0.509	0.497	0.496	0.476	0.454	0.452	0.437	0.440
Theil Top 1	0.487	0.536	0.688	0.605	0.755	0.540	0.662	0.613

Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). In each panel, Theil index is decomposed in *between* and *within* components, among the groups defined (bottom 99% and top 1%). Within group inequality is depicted in the last two rows of each panel.

Fig. B.1. Pre-tax income inequality (Gini index) by income definition and datasource . 2004-2016 (observations with positive income)

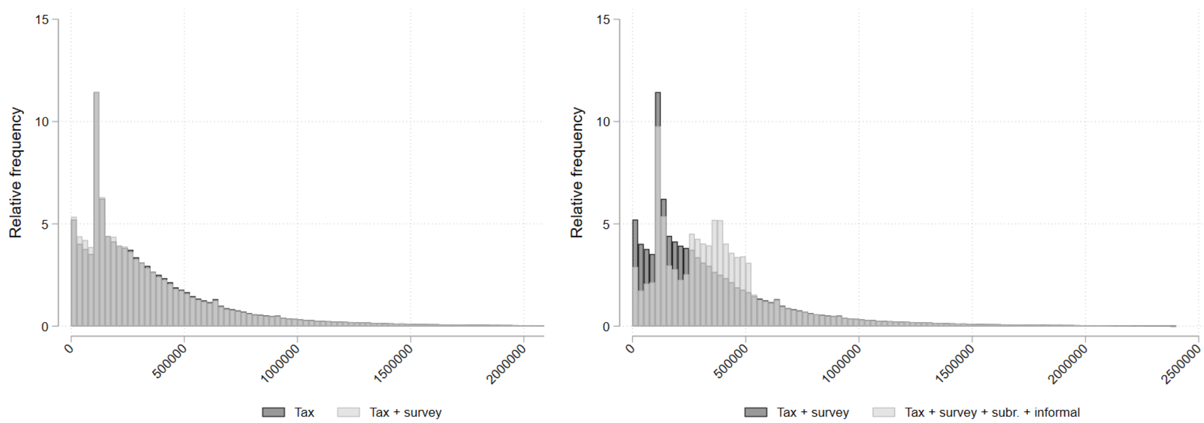


Note. Own calculations based on tax records (DGI) and household surveys (ECH). Effective tax rates for total income and all income sources are depicted.

Fig. B.2. Corrected tax income distribution in steps 2 and 4. Method 1

(a) From Y_{tax} to Step 2

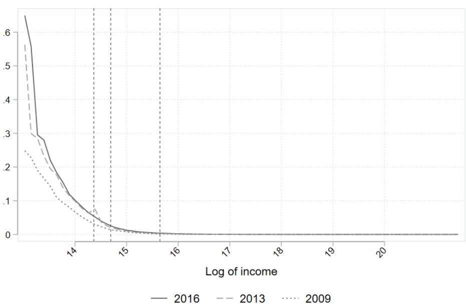
(b) From Step 2 to 4



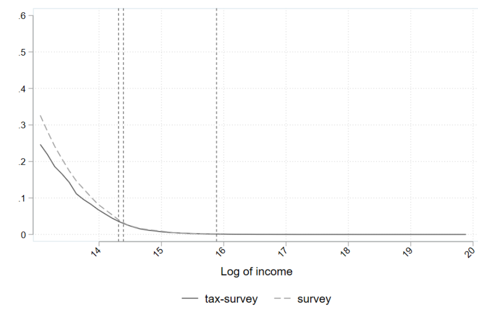
Note. Own calculations based on tax records (DGI) and household surveys (ECH) and population projections.

Fig. B.3. Distribution of the top 2% of corrected tax income. Kernel density function. 2009-2016

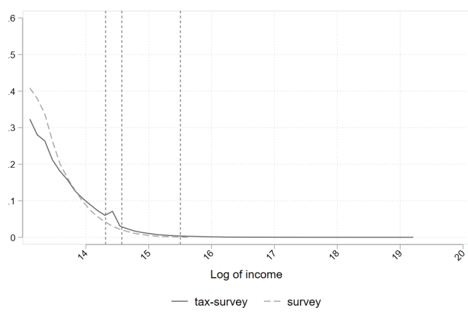
(a) Corrected tax income (2009,2013,2016)



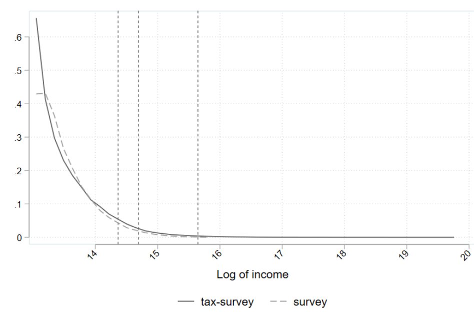
(b) Corrected tax income and survey income (2009)



(c) Corrected tax income and survey income (2013)



(d) Corrected tax income and survey income (2016)



Note. Own calculations based on population projections (CELADE-INE, 2016), tax records (DGI) and household surveys (INE). Vertical lines represent the thresholds for top 1% for survey data and tax-survey data respectively, while the last line depicts the maximum observation in the survey. All incomes at 2016 prices.

Estimating distributed profits from balance sheets information on advance payments to firm owners

We first computed the amount of undistributed profits for each firm and year. Secondly, based on the balance line indicating "share-holders/owners withdrawals in advance", we estimated the potentially undistributed profits and checked whether the firm also distributed profits during the same year or the next. If the firm had a positive value in the "potentially undistributed profits" line and, in the next year, it reported it distributed profits and the withdrawals account was equal to zero, we only considered the actual distributed profits.

Since we lacked information allowing to identify business owners or share-holders and we could only label as such those individuals withdrawing profits, we assigned "potential profits withdrawals" amounts based on three different assumptions. In the first one, we distributed these additional profits among all the individuals we could identify as firm owners based on different years withdrawals. In those cases in which we did not have this information, we created new individuals. Secondly, we distributed profit withdrawals among top labour income earners in the corresponding firm. Third, we combined the two previous criteria and created additional individuals in case the firm reported workers and profit withdrawals in the time span considered in this study. The three criteria yield to the same results, so we stick to the last one. The final number of newly created individuals was between 0,09 and 0,11% depending on the year (see Table B.8).

Chapter 4

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

Joint work with Joan Vilá.

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.

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 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, 2020), 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 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, Lustig, and Teles (2015) and Lustig, Teles, 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

[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).

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, 2008a). 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, 2020), an increasing number of DINA-based estimates for both developed (Blanchet, Chancel, and Gethin, 2019; Piketty et al., 2018; Garbinti et al., 2018) and developing countries (Piketty, Alvaredo, and Assouad, 2017; Piketty and Chancel, 2017; Novokmet, Piketty, and Zucman, 2018; Morgan, 2017b; 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 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 4 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 Figure A.1), followed by relative stagnation from 2013 to 2016 (Cornia, 2014a; 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 micro-data 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 A.2).[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 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

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

[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).

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

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 (2.2.2), 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 Figure 2.1, panel A). For the rest of the firms that report profits and for which we did not identify a shareholder or owner, we implement a probit model of the

[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.

probability of ownership. Table A.1 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 sub-group 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 Figure A.10 show, the results are not affected by the assumption made for these imputations.

Table 2.1 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.

[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 2.1: 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.

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, 2020).

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. 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

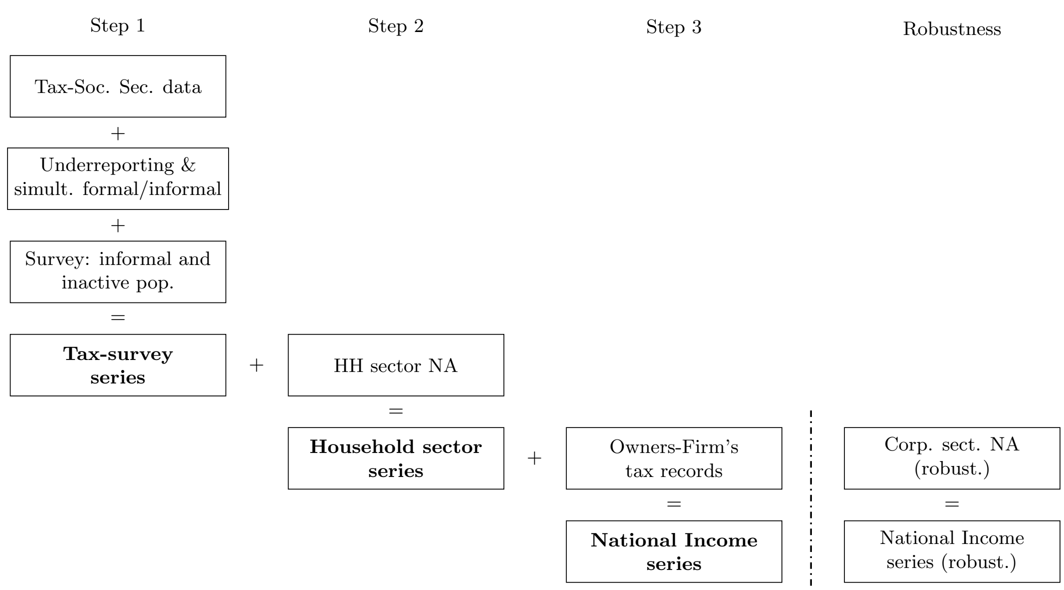
National accounts estimates are provided by the Uruguayan Central Bank (BCU) and have very recently improved from a very low baseline. Uruguay's national accounts present estimates of gross national income based on the expenditure and production approaches, but not on the income approach, except for 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 Figure 3.1.

Fig. 3.1. 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.

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 Figure A.2. 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.

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 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

[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.

[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.

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 3.1, 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 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 3.2 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.

Table 3.1: 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)	

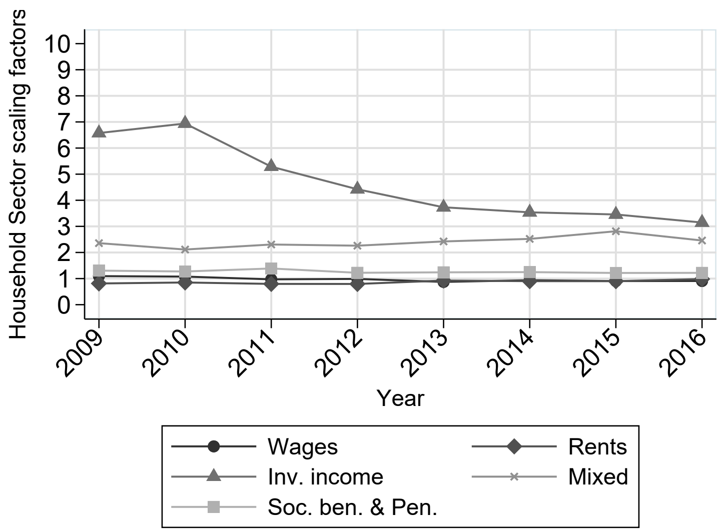
Note. Own elaboration based on similar table in Alvaredo et al. (2022). The first column depicts broad income concepts. The second and third describe their equivalents in the Tax-Survey data and their codes in SNA (along the SNA term for each). D4, D1-D61, B2, B3 and D62 are incomes received by household sector, named S14 in the SNA terminology. The fourth column lists incomes that do not exactly match. Sources: based on United Nations (2008b) and OECD (2013). All incomes are gross of capital depreciation.

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 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, Mok, and Sullivan, 2015). However, given the broad income concepts from Table 3.1 –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 Figure A.3. 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.

Fig. 3.2. Scaling factors, 2009-2016



Note. Scaling factors of tax-survey data vs household aggregates based on Table 3.1. 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 A.3.

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, 2020). 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.

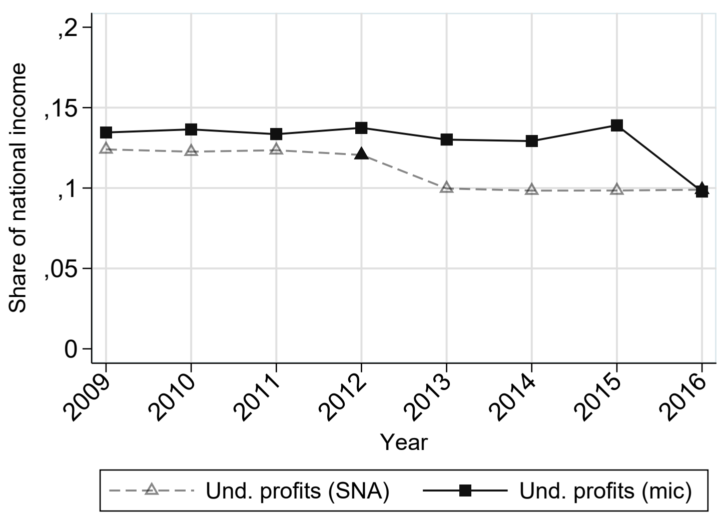
Undistributed profits can be directly calculated based on firm tax records, which are equivalent to their aggregate accounting surplus (i.e., before any tax-related adjustments), net of distributed profits and capital incomes paid to the rest of the world. Aggregate distributed profits are calculated based on individual tax records, while capital income to the rest of the world is computed based on the balance of payments (see Figure A.9, more on this below). Figure 3.3 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, Veall, Brooks, and Murphy, 2016).

[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 (2020), 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. 3.3. 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).

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 based on taxable capital (WIL, 2020). 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 Figure A.3). 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 Figure A.6. 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 Figure A.5.[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.

4. Results

4.1. The evolution of income distribution

The evolution of pre-tax income shares in the three imputation steps is depicted in Figure 4.1, 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

[13] The overall functional distribution of income is presented in Figure A.4. It depicts household incomes from Table 3.1, 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, 2020). 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.

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.

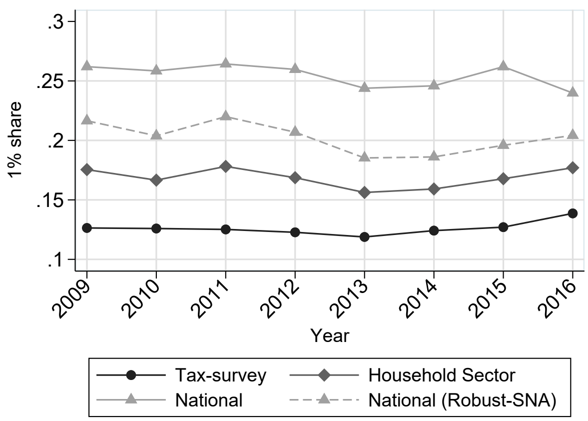
The first thing to note is that at each imputation step, period-average inequality increases, as both scaling up to household sector income and allocating undistributed profits increase the relative importance of capital income, regardless of the way it is imputed. Recalling the scaling factors from Figure 3.2, capital income is scaled up in greater proportion than other incomes and is imputed based on the distribution of dividends and interest (Figure A.3), 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 (Figure 3.3). 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 Figure A.7).

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 A.4, 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.

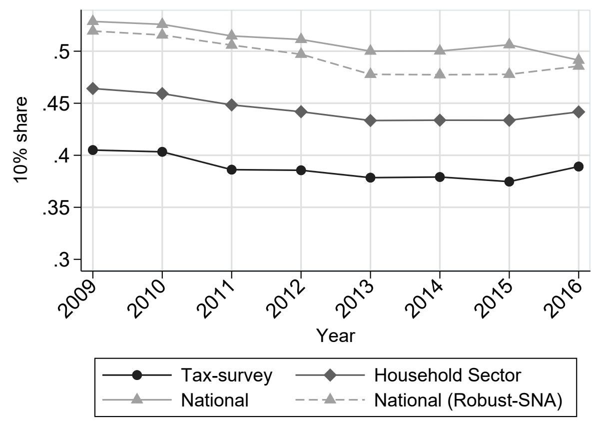
4.2. The effect of (un)distributed profits on inequality

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

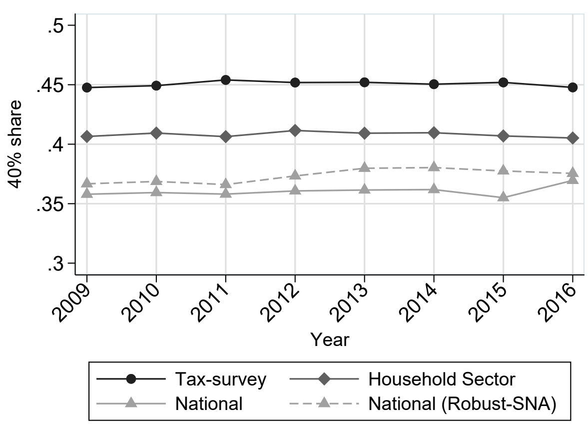
Fig. 4.1. Pre-tax income shares by imputation step, 2009-2016



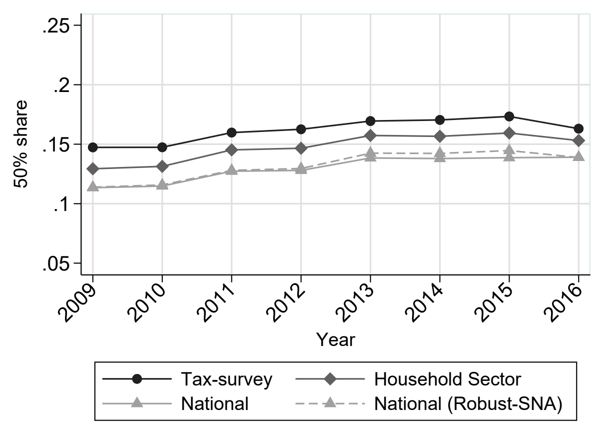
(a) Top 1%



(b) Top 10%



(c) Middle 40%



(d) Bottom 50%

Note. Own elaboration based on tax records, household surveys, and national accounts (see point estimates in Table A.4). First 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.

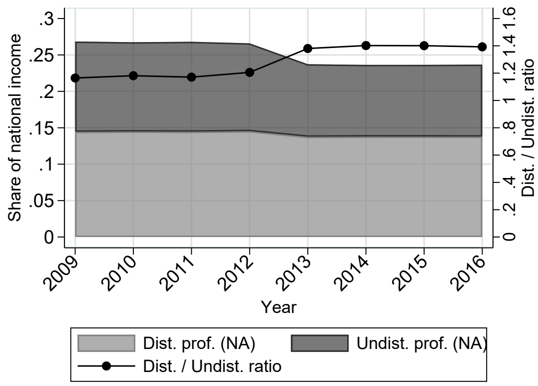
increase is mirrored by a decrease in undistributed profits. Therefore, when undistributed profits are imputed, the top 1%'s increasing share is offset and even slightly reversed. This finding highlights the importance of considering both distributed and undistributed profits in inequality analysis, 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 Figure A.9. 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 Figure 4.2, 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.

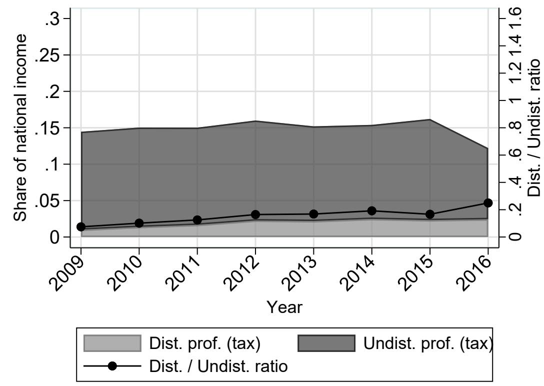
Thus, Figures A.8 and 4.2 indicate that two effects are at play: (i) firms increased their share of distributed profits; and (ii) a higher share of dividends is captured in the tax-survey data. These two combined effects result in the increase in tax-survey top income shares shown in Figure 4.1 and documented 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 4.3 shows the income composition of the top 1% in the three estimation steps, while the composition for the other income groups is included in Figures A.11, A.12 and A.13 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

Fig. 4.2. Distributed and undistributed profits by source, 2009-2016



(a) SNA



(b) Firm's and individual tax data

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 3.1), 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).

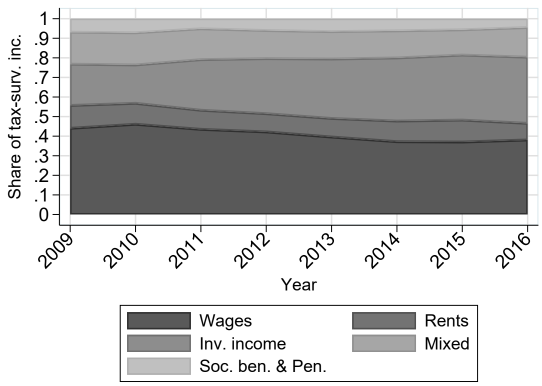
method, this source of income represents between 25 and 40% of the total income of the top 1%. At this step, capital income (investment income + rents + undistributed profits) represents at least two thirds of total income for this group. Finally, the downturn in the quantity of undistributed profits 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

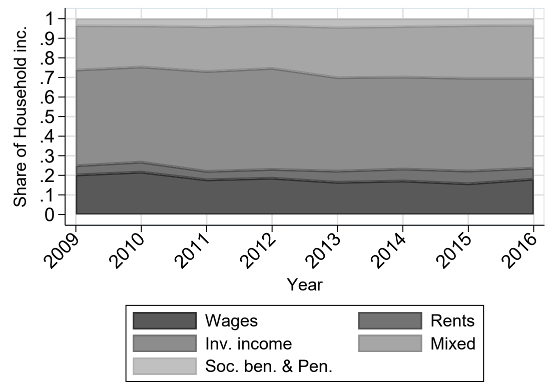
One of the most important advantages of this exercise is that in the last estimation step, the national income series provide full micro-macro consistency. This is relevant, in particular, for the analysis of growth and its distribution, since growth is typically measured in macroeconomic terms while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

Figure 4.4 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

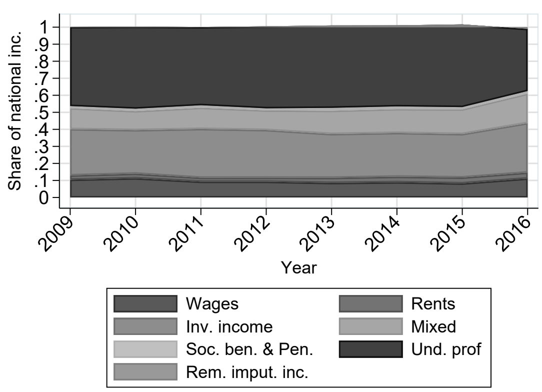
Fig. 4.3. Top 1% income composition, 2009-2016



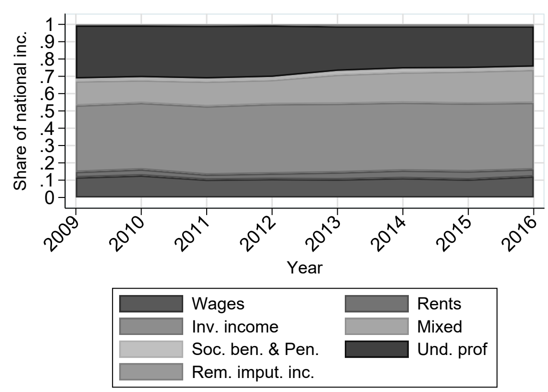
(a) Tax-survey



(b) Household sector



(c) National Income



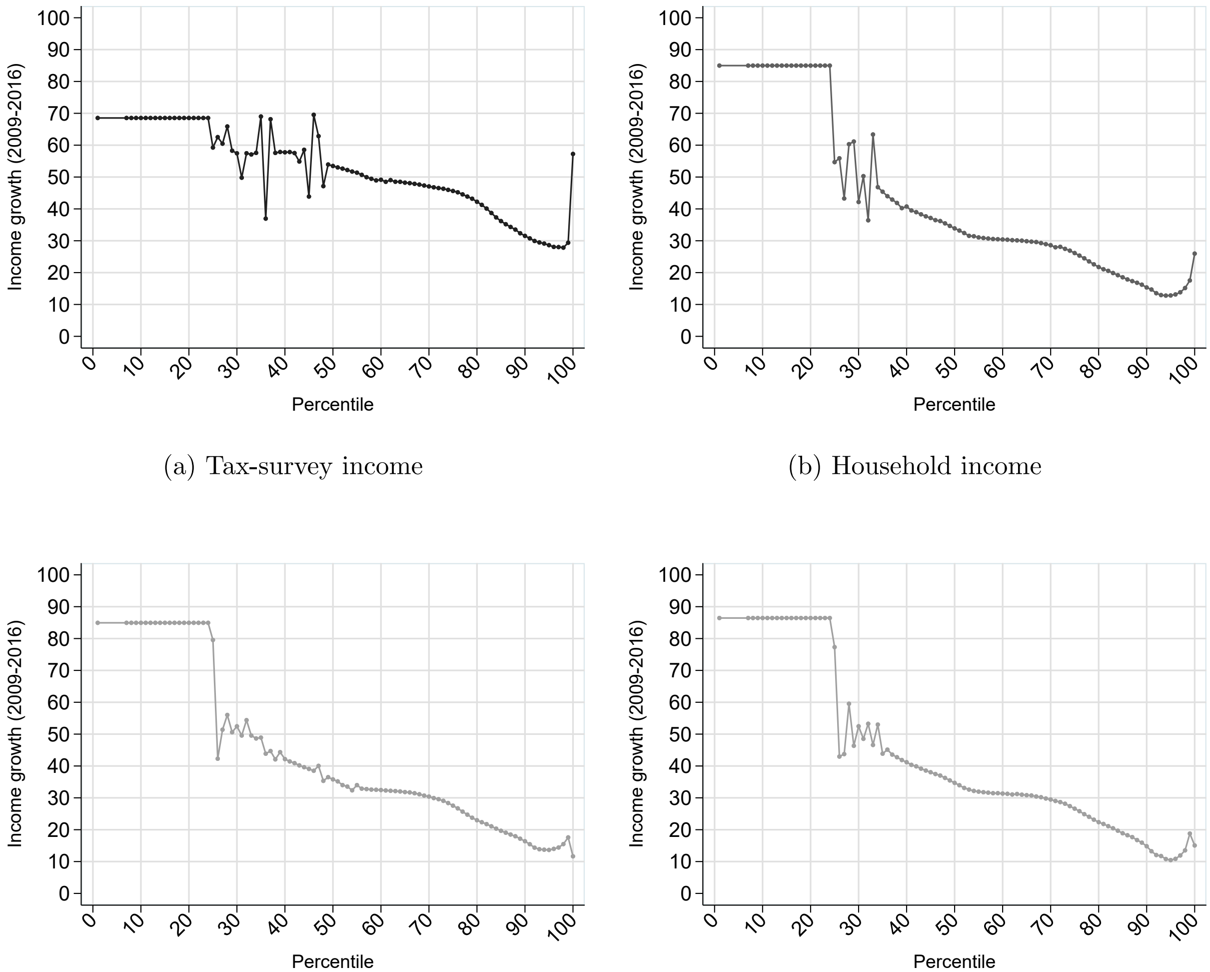
(d) Nat. Inc. (Robust - SNA)

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.

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 A.15 shows the same growth incidence curves for the national income series but for the period 2009-2015. In this case, 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.

Fig. 4.4. Growth Incidence Curves (GIC) by imputation step, 2009-2016



(a) Tax-survey income

(b) Household income

(c) National income

(d) Nat. Inc. (Robust - SNA)

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.

4.4. *Effective direct tax rates*

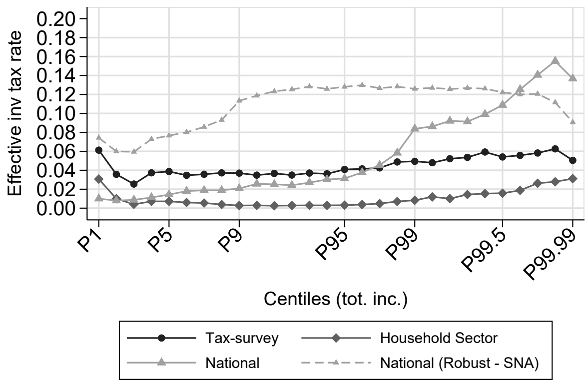
The blurry line dividing firms and their owners has consequences for income, but also for taxes paid as observed in the tax records, and therefore also for the effective tax rates estimated using these sources of information. Thus, our three-step estimation procedure allows us to calculate effective 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 4.5 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 the median income (panel d of Figure 4.5), with an increasing incidence of taxes throughout the distribution at least up to the top 1% in all estimates.

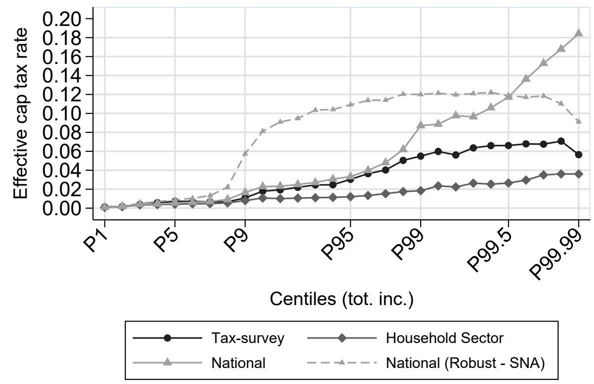
Series comparisons indicate that the scaled-up household income series, which scales incomes but not taxes since they are reported in tax records and assumed to be an accurate depiction of total revenue, results in a reduction in the average effective rate from 13 to 8% for the top 1%. The inclusion 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 Figure 4.5), 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, Steiner, and Rattenhuber, 2019) as well as other European countries (Arulampalam, Devereux, and Maffini, 2012) and between 35% (Suárez Serrato and Zidar, 2016) and 60% (Liu and Altshuler, 2013) in the United States.

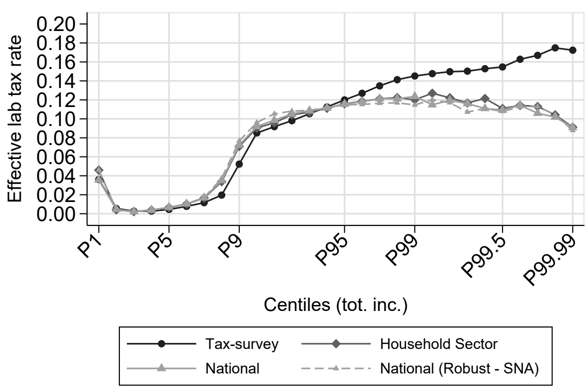
Fig. 4.5. Effective tax rates, 2016



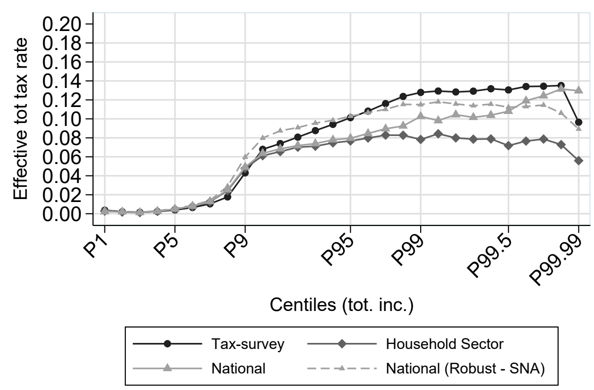
(a) Investment income



(b) Total capital incomes



(c) Labor incomes



(d) Total incomes

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.

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.

A. Appendix

Tables

Table A.1: 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.

Table A.2: 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%
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 A.3: 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 Figure 3.2.

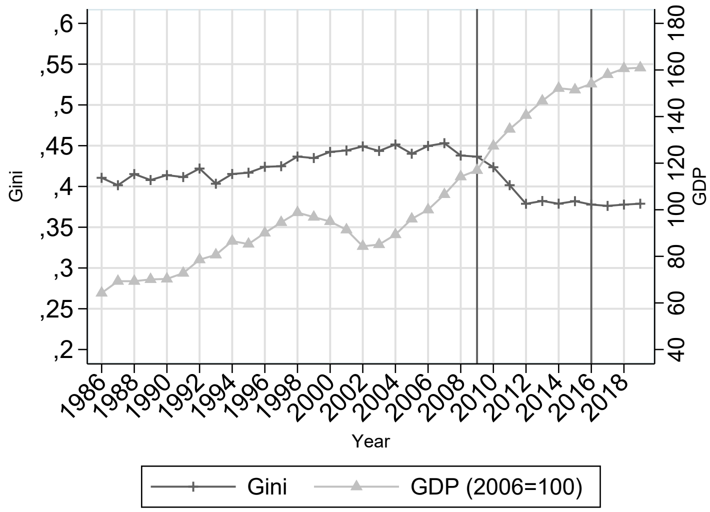
Table A.4: Income shares, 2009-2016

	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%
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.

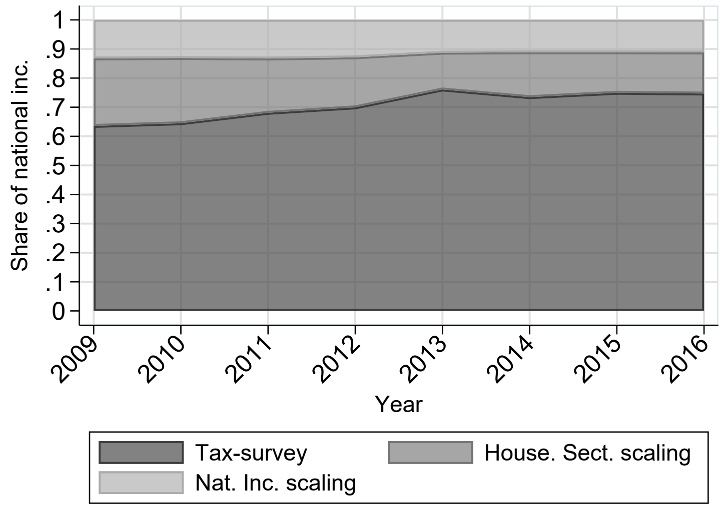
Figures

Fig. A.1. GDP and income inequality 1986-2019



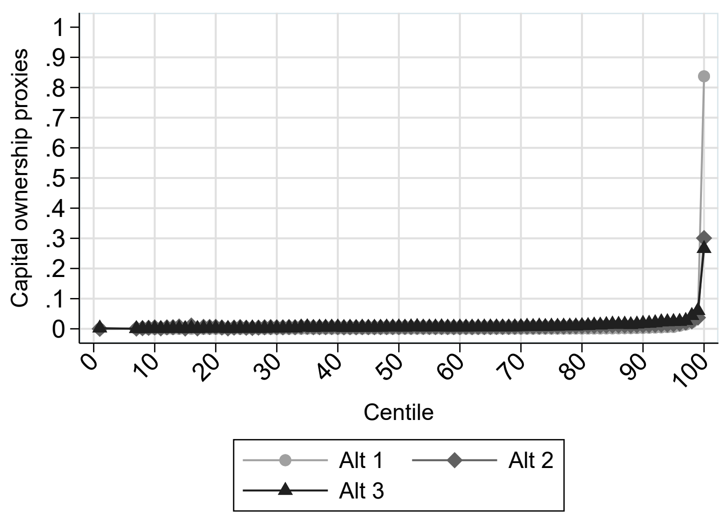
Note. In the primary axis GDP is presented with GDP 2005=100, whilst percapita household income gini index (estimated based on the household survey) is depicted on the secondary axis. During the period 2009-2016 (between vertical lines, period with tax data available), gini index dropped by about 7 points,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

Fig. A.2. Income shares by estimation step, 2009-2016



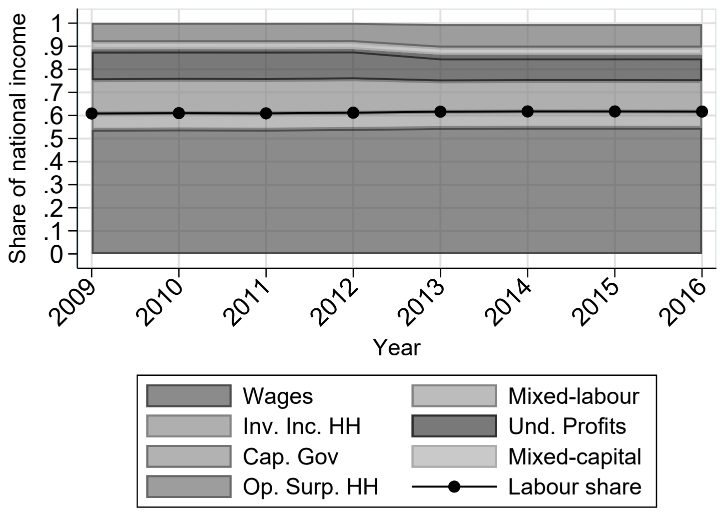
Note. Own estimates based on tax-survey data (DGI-ECH) and National Accounts 2012, 2016 (BCU). The figure depicts aggregate income by estimation step: the dark-green area is the sum of tax-survey incomes, the orange area are incomes added during scaling to household sector based on scaling factors depicted in Figure 3.2, while the blue area represents reaming imputed incomes of Figure A.6. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin American Countries.

Fig. A.3. Proxies of firm ownership, 2016



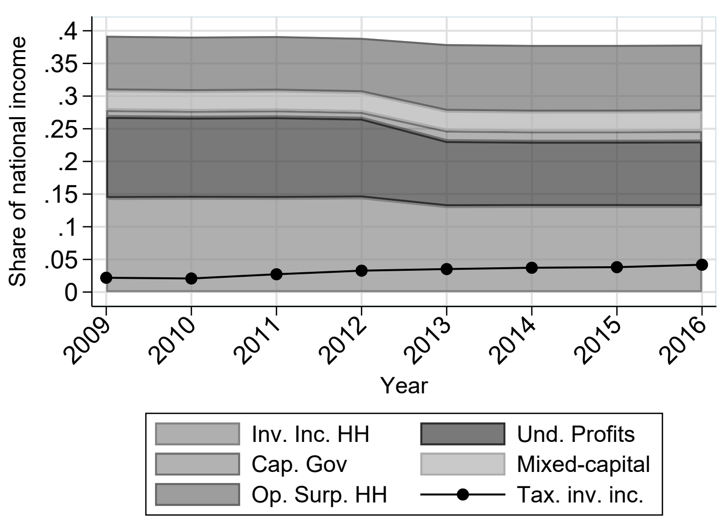
Note. Own estimates based on tax-survey data (ECH-DGI). *Alt. 1* refers to the distribution of taxable capital incomes from DGI. *Alt. 2* refers in turn to the sum of all taxable and non-taxable capital incomes, including rents and owner occupied housing rents. The preferred alternative (*Alt. 3*), is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

Fig. A.4. Functional income distribution, 2009-2016



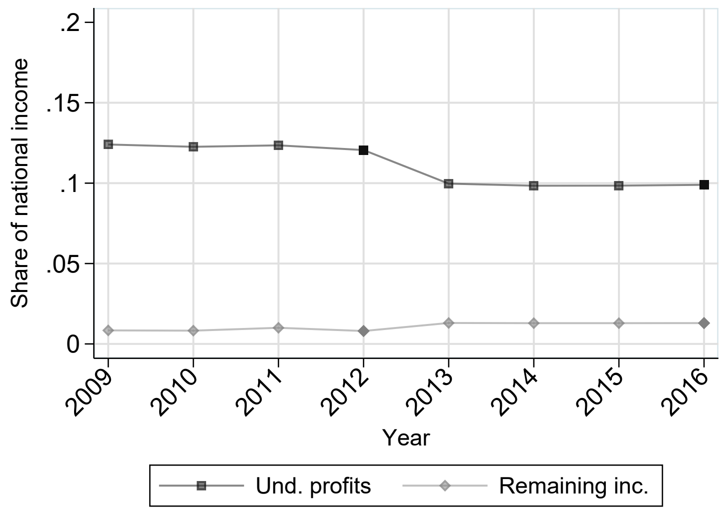
Note. Own estimates based on tax-survey data (DGI-ECH) and National Accounts 2012, 2016 (BCU). The figure presents the distribution of net national incomes in capital and labor shares and their components. All incomes from national accounts are net of depreciation, based on Wid.World data for other Latin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

Fig. A.5. Capital incomes composition



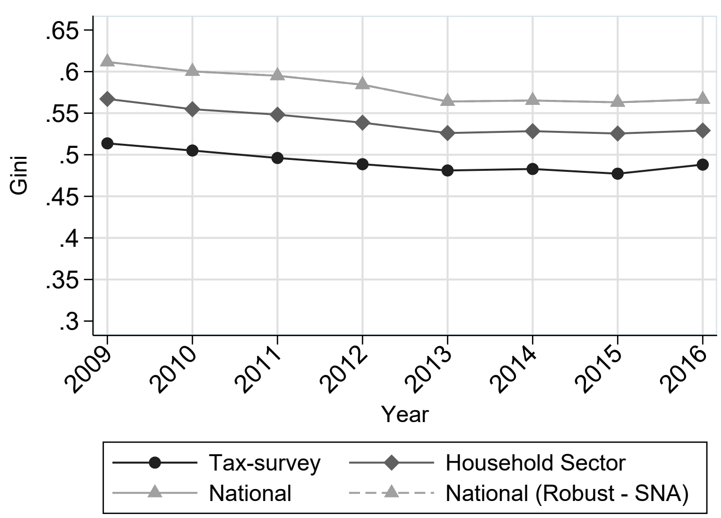
Note. Own estimates based on firm tax data (DGI), tax-survey data (ECH-DGI) and National Accounts 2012, 2016 (BCU). Solid filled areas represent national account's aggregates, while doted line depicts aggregate investment incomes (dividends, interest, etc.) from tax-survey data. This line is conceptually 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. A.6. Income aggregates of non-household sector



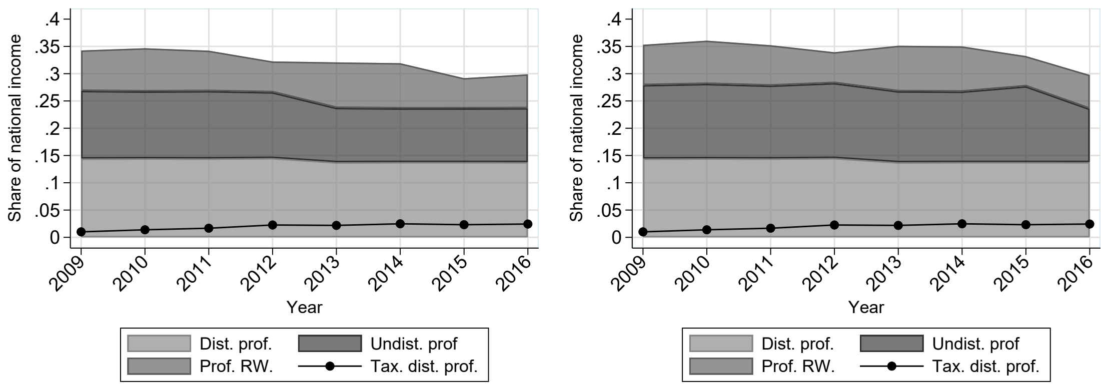
Note. Own estimates based on National Accounts 2012, 2016 (BCU). Dots in dark represent actually observed data points in national accounts. Undistributed profits are allocated based on the capital ownership proxy, while remaining components of national income are distributed proportionally to total incomes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

Fig. A.7. 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.

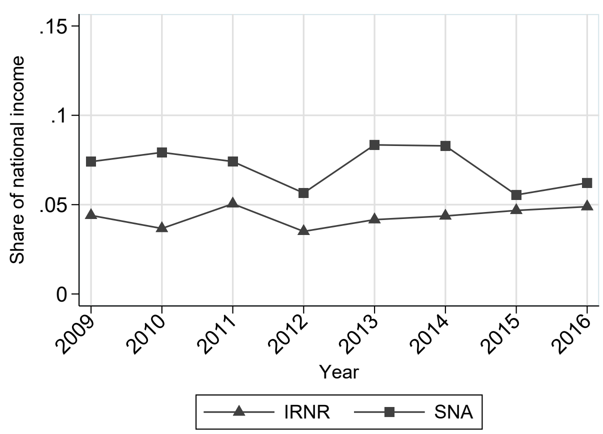
Fig. A.8. Firm's profits by alternative, 2009-2016



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

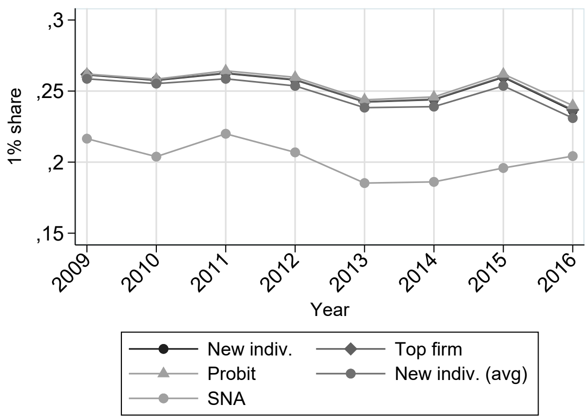
Note. Own estimates based on firm tax data (DGI), tax-survey data (ECH-DGI), National Accounts 2012, 2016 (BCU), and Balance of Payments (BCU). Both panels depict observed dividends observed in tax-survey data, investment incomes of households excluding interest, undistributed profits and capital incomes 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. A.9. Private capital incomes paid to the rest of the world

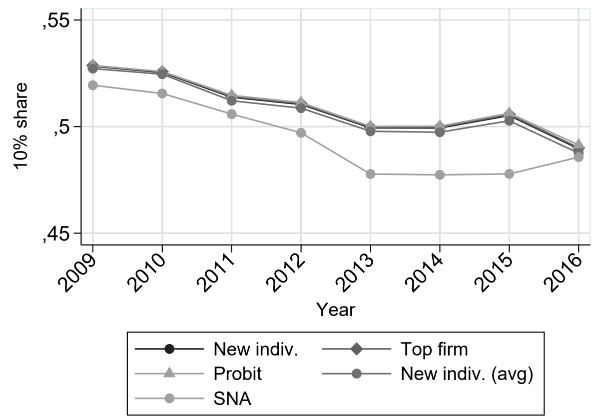


Note. Own estimates based on Balance of Payments (BCU) and *Impuesto a la Renta de los No Residentes* (IRNR) series (DGI). Balance of payments series is constructed based on Central Bank data for two periods: 2009-2012 and 2013-2016. The latter series has an updated methodology but has not been matched 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%).

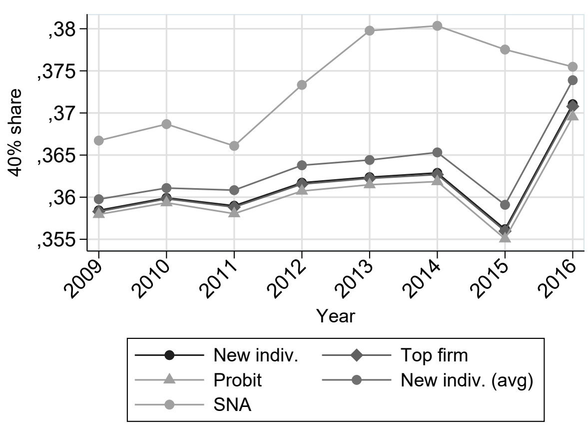
Fig. A.10. Pre-tax income shares of National Income by imputation method, 2009-2016



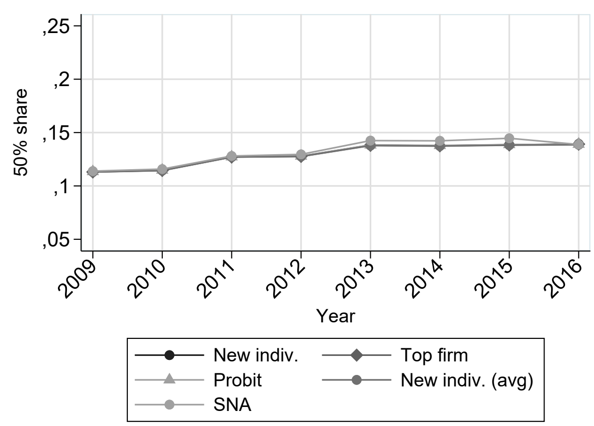
(a) Top 1%



(b) Top 10%



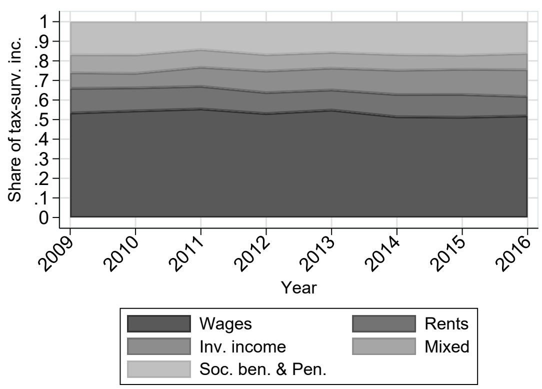
(c) Middle 40%



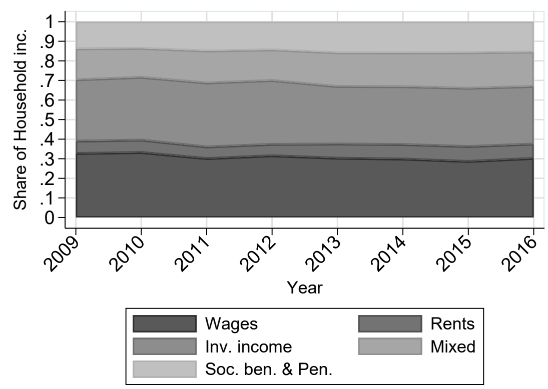
(d) Bottom 50%

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

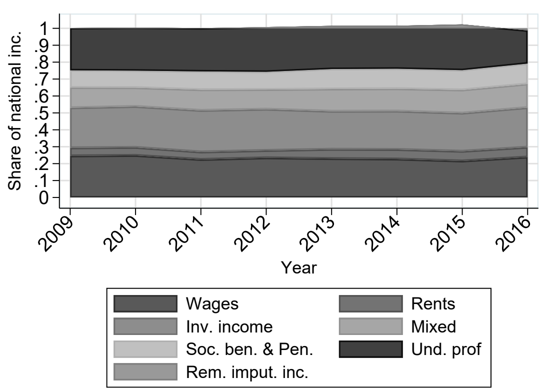
Fig. A.11. Top 10% income composition, 2009-2016



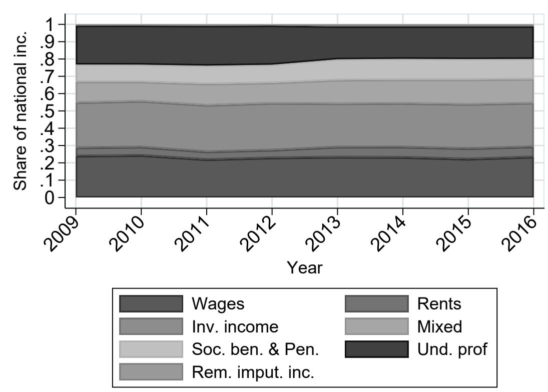
(a) Tax-survey



(b) Household sector



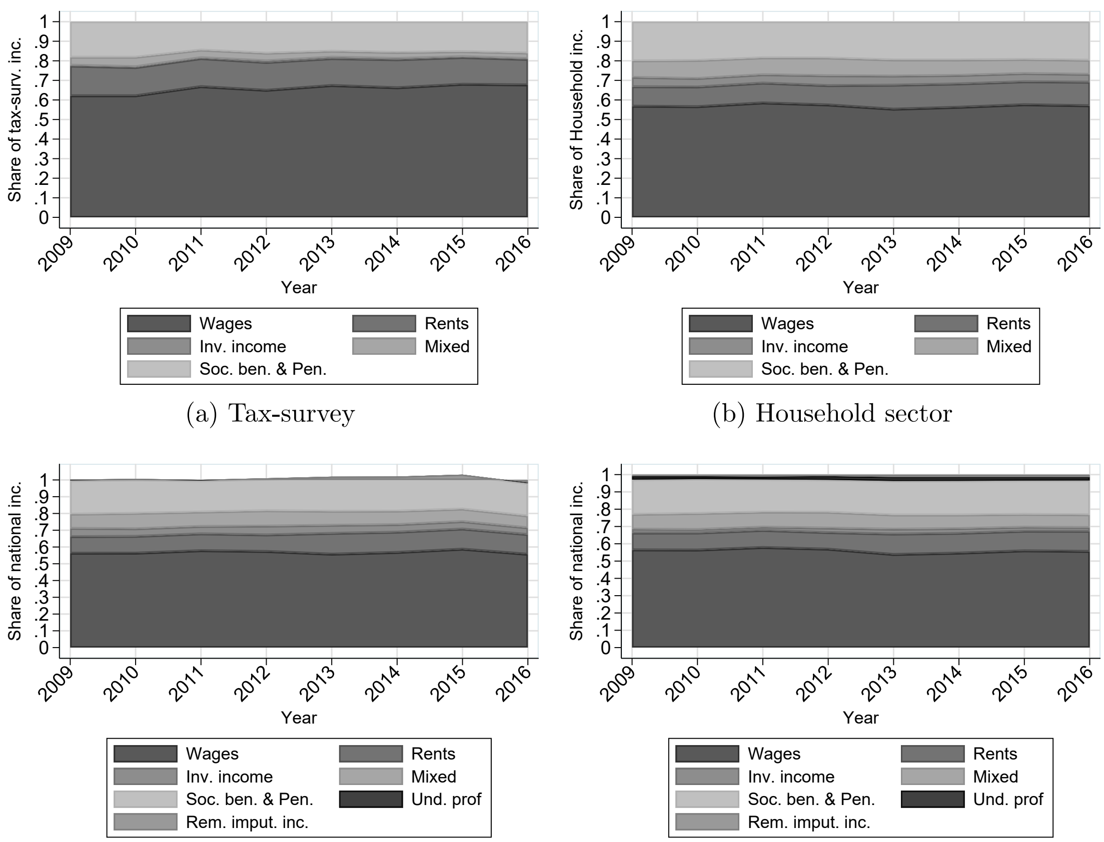
(c) National Income



(d) Nat. Inc. (Robust - SNA)

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. A.12. Middle 40% income composition, 2009-2016



(a) Tax-survey

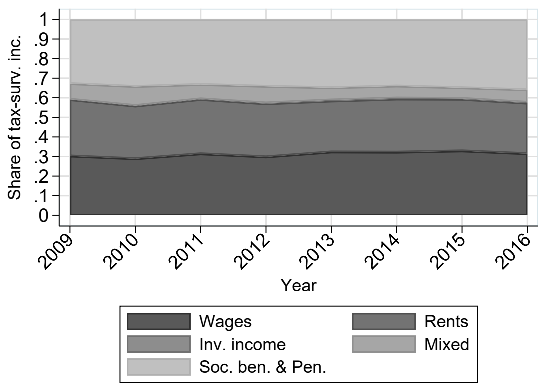
(b) Household sector

(c) National Income

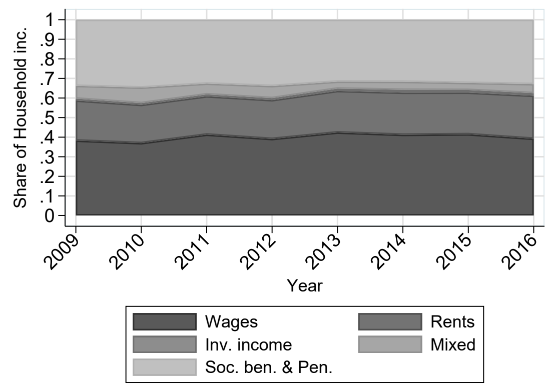
(d) Nat. Inc. (Robust - SNA)

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.

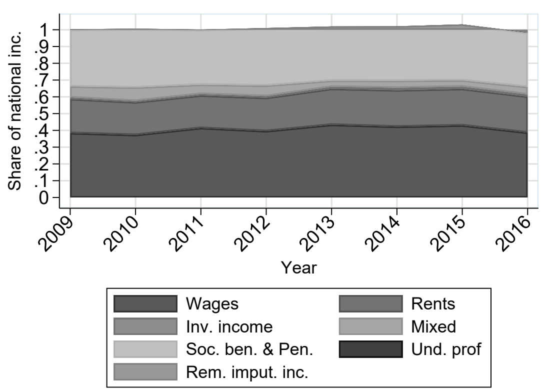
Fig. A.13. Bottom 50% income composition, 2009-2016



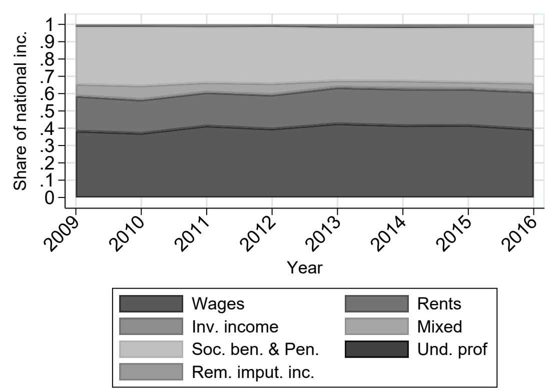
(a) Tax-survey



(b) Household sector



(c) National Income



(d) Nat. Inc. (Robust - SNA)

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. A.14. Growth Incidence Curves (GIC), 2009-2015

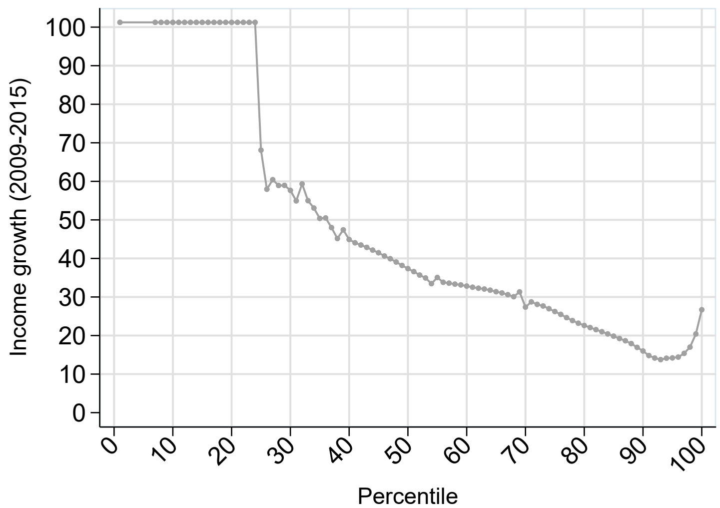


Fig. A.15. National income

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.

PART III. The distribution of wealth

Chapter 5

Wealth inequality in the south: multi-source evidence from Uruguay

Abstract

While wealth accumulation and its distribution are arguably two of the key drivers of overall economic inequality, as well as being of major importance in their own right, very little is known about them in the developing world. I contribute to filling this gap by providing a micro-macro consistent series of aggregate wealth and its distribution in Uruguay. The country's balance sheet, which is not estimated by official institutions, is constructed for the first time by combining a wide array of data sources, leading to a wealth-to-income ratio of 500%. Private wealth distribution is then estimated based on the capitalization method, taking stock of combined survey-tax-national accounts micro-data, resulting in a top 1% wealth share of 38-40%. Estimates are systematically compared with results based on the estate multiplier method, real estate wealth tax, household wealth survey, and *Forbes* billionaires list.

1. Introduction

In the economic literature on the accumulation and distribution of wealth, there is a sharp disparity between the evidence gathered in the past decade for a handful of mostly developed countries and the scarcity of estimates available for the rest of the world. The bar is set high by the literature, which provides wealth distribution estimates based on a variety of methods and sources (e.g. Saez and Zucman (2016, 2022); Alvaredo et al. (2018); Garbinti et al. (2017)). These estimates are in turn consistent with national wealth estimates (Piketty and Zucman, 2014; Blanco et al., 2021), hence providing a full account of wealth dynamics, both micro- and macro-economically. Wealth estimates for the developing world are for the most part the result of regression analyses based on countries with available data or household surveys in best case scenarios, with little reference to macro wealth or income aggregates, thus hindering our ability to provide comparable estimates and a credible narrative. I contribute to the closing of this widening gap by estimating the first fully consistent set of estimates of aggregate wealth and its distribution for a Latin American country. A wide array of surveys, detailed personal income tax micro-data, cadastre administrative data, national accounts, owner-decedent micro-data, and firm balance sheets are carefully combined to provide a consistent overview of the level, composition and personal distribution of wealth in Uruguay.

The Uruguayan case is an interesting laboratory for the Latin American setting because it is, by all accounts, the least unequal country in the region in terms of income (De Rosa et al., 2022; Lustig et al., 2011).[1] Income inequality in Latin America seems to have experienced a downturn since the early 2000s, in the context of vigorous economic growth and redistributive public policies (Cornia, 2014a), followed by halt and in some cases a reversion of that trend beginning in 2015 (Gasparini et al., 2018). The debate over whether these patterns are the result of poor performance of household surveys is still ongoing, with mixed results depending on the country, as the evidence based on income tax data accumulates (Alvaredo and Londoño Velez, 2014; Alvaredo, 2010; Morgan, 2017b; Flores et al., 2020; Burdín et al., 2022). Meanwhile, a new stream of evidence points to capital incomes and their extreme concentration as a source of divergent trends: capital incomes account for most of the gap between the sources usually used to study income inequality and macroeconomic income estimates (Alvaredo et al., 2022), and once accounted for, they entail massive consequences for both the level and trend of inequality in the region (De Rosa et al., 2022). Understanding wealth accumulation and distribution is hence particularly important in the region, yet there has been very little progress thus far.

As in most of the developing world and even in many rich countries, there are no national accounts estimates of Uruguay's balance sheet. Therefore, I estimate wealth-to-income ratios based on a wide range of secondary sources, including cadastral administrative data, prices of land and housing properties, firm tax records, and central bank financial data, among others. These estimates, although imperfect, follow the tradition of aggregate wealth estimation by independent scholars of the 17th to early 20th century. Results show that the book value wealth-to-income ratio is around 500%, comparable to what is observed in developed economies, where it is around 500-700% (Piketty and Zucman, 2014). Public net wealth is positive but decreasing, from 50% at the beginning of the century to around 25% by the end of the period under analysis. Gross domestic capital is 30 percentage points higher than net national wealth as a result of a negative net foreign asset position. Approximately half

[1] For an overview of Uruguayan background data, see appendix G.

of household wealth takes the form of financial assets (including ownership of the private corporate sector), while housing reaches 100-150% of national income.

The main distributional estimates are based on the capitalization method, which consists of estimating individual net wealth by capitalizing personal capital incomes, using capitalization factors for each type of wealth that are equivalent to the inverse of their macro rate of return, i.e., consistent with macro wealth aggregates. Capital incomes are mainly drawn from high-quality tax record micro-data—which covers 75% of adult population—merged and combined with firm tax records, household survey data, and national accounts. This capital incomes database is the result of the national income distribution series estimated in De Rosa and Vilá (2022), and it therefore provides income-wealth consistency, both at the micro and macro levels. Private personal wealth inequality is relatively stable over the period 2009-2016 but at a very high level: over 38-40% of net private wealth is owned by the wealthiest 1%, and the top 10%'s share is around 77-79%. These estimates would locate Uruguay as a relatively high wealth inequality country compared to France, closer to estimates for Spain or the US, but lower than extreme cases such as South Africa. Moreover, I characterize wealth owners in terms of age and gender, showing that men have higher wealth for all age groups (with little evidence of life-cycle accumulation patterns) and represent 70-80% of the top fractiles. I also document a high correspondence between wealth distribution and total income distribution for the top fractiles.

The inequality estimates in this article are triangulated with four other empirical approaches to provide greater certainty about the overall conclusions. First, the main results are compared with the Wealth Household Survey (*Encuesta Financiera de los Hogares Uruguayos*, EFHU) of 2013, which covers similar assets as the ones estimated with the capitalization method. The top 10% is 10 percentage points lower in the survey, with most of this difference being explained, as expected, by the top 1 and top 0.1%. Moreover, the different concentration profiles are explained by wealth composition rather than by the distribution of each asset: contrary to the results of the capitalization method, in the wealth survey, the least unequally distributed asset—housing—represents the bulk of wealth. Second, I compare data from the *Forbes* billionaires list to the very top wealth holders from the capitalization method data; these show very similar net wealth levels, which provides reassurance about the capitalization method's accuracy for individuals at the very top. Third, a novel administrative dataset of decedent real estate owners is constructed, allowing me to estimate urban and rural real estate distribution based on a simplified version of the estate multiplier method (Alvaredo et al., 2018; Berman and Morelli, 2021), which essentially entails weighting the decedent population by the average mortality rate. Results show that the top 1%'s share of real estate wealth reaches around 20-25%, whilst top 0.1% is stable around 10%. When considering urban properties only, it shows higher concentration for the top 10 and 1% than the capitalization method does (10 and 5 percentage points, respectively), but 2 points lower for the top 0.1%. The estimates are, however, remarkably close to wealth survey estimates, which suggests that capitalization method estimates represent a lower bound. Fourth, the top 0.1%'s housing share is calculated based on a wealth tax that covers a small fraction of the population (a little over 0.3%) and targets mainly real estate assets, resulting in a top 0.1% of 4-5%, almost identical to capitalization method estimates.

The contributions of this article are threefold. First, it contributes to the literature on wealth-to-income ratio estimation and wealth accumulation (Piketty and Zucman, 2014; Blanco et al., 2021; Bauluz, 2019; Del Castillo, 2017; Waldenström, 2017; Kumar, 2019), a contribution that comes not from national accounts but from my own estimates. Second, it provides estimates of wealth distribution based on combined survey-tax-national accounts data, which makes these results comparable with the growing literature on wealth distribution (Saez and Zucman, 2016; Alvaredo et al., 2018; Martínez-Toledano, 2020; Garbinti et al., 2017; Kopczuk, 2015; Novokmet et al., 2018; Chatterjee, Czajka, and Gethin, 2022) and fully consistent both in the micro-macro and income-wealth levels (WIL, 2020). Third, it contributes to the methodological debate over the different empirical approaches to the estimation of wealth distributions and the relative benefits and drawbacks of the capitalization method (Kopczuk, 2015; Bricker, Krimmel, Henriques, and Sabelhaus, 2016; Fagereng, Guiso, Malacrino, and Pistaferri, 2016; Saez and Zucman, 2016; Chatterjee et al., 2022).

The rest of the article is organised as follows. Section 3 discusses the literature on wealth-to-income ratios and wealth distribution. Section section 4 describes the data sources used throughout the article, while 2 presents the definitions and the methodological approach. National and private wealth aggregates are presented in section 5, followed by estimates of the personal distribution of private wealth based on the capitalization method, as well as a characterisation or wealth holders in terms of age, sex and location in the income distribution. Section 6, in turn, contrasts these estimates with other data sources and empirical approaches. Section 7 concludes.

2. Definitions and methodology

This section presents the main definition of wealth following the System of National Accounts, which I will use throughout the remainder of this study. It also presents the main features of the capitalization and estate multiplier methods, which are extensively used in the wealth inequality literature summarized above, and it discusses the adaptations required in the present setting.

2.1. Baseline definitions: net wealth

Following the concepts discussed in Piketty and Zucman (2015) and the Distributional National Accounts guidelines (WIL, 2020), which are in turn based on the System of National Accounts balance sheet definitions, private wealth W_t is defined as the net wealth (assets minus liabilities) owned by households.[2] These assets include "all the nonfinancial assets—land, buildings, machines, etc.—and financial assets—including life insurance and pensions funds—over which ownership rights can be enforced and that provide economic benefits to their owners" (Piketty and Zucman, 2015, p. 1309). Corporations are included in private wealth through the market value of equities and corporate bonds.

In its basic decomposition, private wealth can be decomposed into housing assets, business assets (and other non-financial assets), financial assets, and liabilities (WIL, 2020). National wealth W_{nt} results from the addition of private and public wealth, which may be divided into the same broad categories. It is also equivalent to the sum of domestic capital and net foreign assets, as depicted in equation 2.1.

$$W_{nt} = W_t + W_{gt} = K_t + NFA_t \tag{2.1}$$

W_{nt}, W_{gt}, and W_t represent net national, public, and private wealth respectively, while K_t is domestic capital and NFA_t is net foreign asset position. For most rich countries, national wealth tends to be equivalent to private wealth, since net government wealth W_{gt} is in the present close to zero (Piketty and Zucman, 2014). As for the second equivalence, it is interesting to note that, intuitively, as all national financial assets and liabilities must cancel out (including the property of corporations), national wealth W_{nt} is equivalent to the sum of all non-financial assets owned by household, corporate, and government sectors, plus the net foreign asset position.

An important issue regarding the definition of national wealth is its valuation, which may be done under two complementary perspectives, i.e. market and book value. Ideally, both

[2] To be more precise, private wealth is the sum of personal wealth, i.e. wealth owned by households, plus wealth of non-profit institutions serving households (NPISH). As this institutional sector is usually very small and—as in the Uruguayan case—often included in household sector estimates, I do not make a distinction between the two in this study.

book and market valuations should be portrayed. The market value of corporations is in practice directly observed in stock markets, or may be indirectly estimated for less frequently traded equity. The book value represents the difference between corporate assets at cost and non-equity liabilities. The "residual value of corporations" is the difference between the two, while their ratio is Tobins' Q, which is usually lower than one (WIL, 2020). When Q is bellow one, it means that firms are under-valuated in the market with respect to their book values, hence residual corporate wealth is positive and book-value aggregate wealth is higher than its market value counterpart.

With Y_t representing national income, the private and national wealth-to-income ratios (β_t and β_{nt}) are defined as:

$$\beta_t = \frac{W_t}{Y_t}, \beta_{nt} = \frac{W_{nt}}{Y_t} \tag{2.2}$$

It is important to note that personal wealth distribution estimates will refer to private wealth; i.e., the sum of all individual net wealth in the distributional estimates must add up to W_t.

2.2. The capitalization method

The main wealth distribution estimates in this article are based on the capitalization method, recently applied by Saez and Zucman (2016) to the United States[3], and the remaining methods will be used (when possible) as robustness checks. There are two main reasons to choose the capitalization method as the methodological workhorse in this setting.

First, it provides the best balance between asset and time coverage. As discussed in Section 4, the wealth survey of *Forbes* rich list has information on a larger number of assets but only for one year, whilst the estate multiplier method or the wealth tax presents a longer time span but only provides real estate distribution estimates. In any case, as will be discussed below, adding more assets does not substantially change wealth distribution estimates in the wealth survey, and the estate multiplier method's estimates only add two years to the time series. Therefore, the capitalization method is better, as it provides estimates of a complete wealth distribution, both time and asset-wise. Second, as discussed in Section 4.2.1, wealth is estimated based on the capital income distribution, which is part of national income distribution estimates (De Rosa and Vilá, 2022), therefore providing a full individual

[3] This method was originally proposed by Robert Giffen in 1913 (Fagereng et al., 2016), and applied, for instance, to the United Kingdom by Atkinson and Harrison (1978). Other recent applications can be found in Garbinti et al. (2017) for France and Chatterjee et al. (2022) for South Africa.

income-wealth database (WIL, 2020). This income-wealth database is an important product in its own right, since it allows the study of the dynamics of income growth and distribution together with wealth distribution and accumulation.

The capitalization method consists of estimating individual net wealth by capitalizing personal capital incomes, using capitalization factors for each type of wealth which are equivalent to the inverse of their rate of return. Essentially, if for certain individual i, the amount of wealth p that she owns w_{ip} yields r_p, providing her with an income flow k_{ip} (i.e. $k_{ip} = r_p * w_{ip}$); then it is possible to trace back to the wealth stock by applying a capitalization factor f_p, equivalent to the inverse of its rate of return $w_{ip} = k_{ip} * f_p$, given $f_p = 1/r_p$.

This method has some important drawbacks. The most relevant one is the fact that it assumes that for each type of wealth w_p, the capitalization factor f_p is the same for all individuals. This may not be the case, as individuals may face different rates of return r_p, thus biasing the estimations. One possible bias is associated with idiosyncratic returns, that is, that individuals who are identical in terms of observable characteristics face different rates of return for the same assets. Furthermore, it is possible that returns are positively correlated with wealth, which has been argued to be a “more serious concern” (Saez and Zucman, 2016). If return rates r_p are larger for higher income individuals, then their actual capitalization factors f_p should be lower than those used in this method. Therefore, the capitalization method is mechanically overestimating wealth concentration at the top. Estimates of r_p and f_p should be the result of estimating the rate of return of each type of wealth by comparing total wealth W_p with the sum of capital income flows, i.e. $f_p = W_p/K_p$, given $W_p = \sum w_{ip}$, $K_p = \sum k_{ip}$ and $W_t = \sum W_p$.

One of the most important advantages of this procedure is that it provides full micro-macro consistency between wealth distribution estimates and aggregate estimates. To ensure this, I follow the Distributional National Accounts guidelines (WIL, 2020) and compute return rates matching aggregate income and wealth components for four assets and liabilities: housing assets, business (and other non-financial) assets, financial assets, and liabilities.

2.3. The estate multiplier method

2.3.1. The standard application of the estate multiplier method

The estate multiplier method has been perhaps the most commonly used method for studying wealth distribution, especially from a historical perspective (Piketty, Postel-Vinay, and Rosenthal, 2006). It is based on estate tax data, which is a way to observe the wealth of individuals at the time of their death, and therefore is considered a sample of the entire population. Naturally, it is not a random sample, so it is weighted by the inverse of the individual mortality rate, thus providing a personal wealth distribution of the living population (Piketty and Zucman, 2015). Its basic inputs are estate tax records and individual mortality rates, as well as population and a wealth control totals to account for the wealth of individuals below the estate tax threshold.

Provided that there is data on the estate tax from an inheritance tax or similar, the challenge usually lies in applying an appropriate mortality rate. These mortality multipliers could in principle be relevant, but the actual extent to which they affect inequality estimations is still debated (Alvaredo et al., 2018; Saez and Zucman, 2016; Kopczuk, 2015). Sex- and age-specific mortality rates are therefore needed and usually available in estate tax data, and also some proxy for the level of wealth (e.g. income or education), since mortality can be affected by an individual's wealth. Finally, the wealth control total is taken from national accounts balance sheets and population totals from official estimates of the adult population.

2.3.2. The simplified estate multiplier method

As discussed in Section 4, almost none of these data inputs are available in the Uruguayan case. To begin with, there are no national accounts balance sheets, so the first step is to estimate a private and household real estate wealth total. Moreover, the personal wealth data used in this article comes from an administrative registry of decedents with properties, but with virtually no information about the decedents themselves. Therefore, weighting them by their specific (inverse of) probability of dying is not possible in this context.

However, as Alvaredo et al. (2018) show, at least in the case of the United Kingdom, the weighting process—the transformation from decedents to living population analysis—does not change the levels or trends of top wealth shares. In other words, estimates of decedent's wealth distribution with respect to the wealth of individuals when they passed away, and estimates of living population's wealth distribution with respect to aggregate household wealth, yield almost identical results. This is further explored by Berman and Morelli (2021), showing that it also holds for Australia, France, Italy, South Korea, and the United States. Moreover, they also show formally the conditions under which this is the case, concluding that it is possible to use average mortality rates and obtain similar results to using detailed ones. Assuming that the same holds for the Uruguayan case, I will consider that the decedent population is an adequate sample of the entire population that only differs in that they are wealthier on average than the rest. Therefore, by adjusting their wealth downward, this sample can be used to estimate the living population's real estate wealth distribution.

The procedure for estimating real estate wealth's top shares is therefore the following: (a) begin with the decedents' personal wealth micro-data; (b) weight them by (the inverse of) the average probability of passing; (c) adjust their wealth downward to account for decedent's higher wealth in relation to the living population; and (d) compute the top wealth shares based on this weighted and adjusted data, with respect to estimated aggregate household real estate wealth W^r_{ht} and the total adult population (20 or more years of age). Thus, in steps (a) and (b), we expand decedent individuals to account for the entire living population in terms of the number of people. As individuals are likely to be wealthier at the time of death due to lifetime wealth accumulation, their real estates is corrected in stage (c) by μ, which is the ratio of average wealth of decedents in relation to the living population. Once weighted and adjusted, the top real estate shares can be computed by comparing the wealthiest individuals with aggregate household real estate wealth W^r_{ht}.

3. Related literature

The last decade has witnessed a rapid expansion of studies related to wealth aggregates and its distribution. In a seminal study, Piketty and Zucman (2014) estimate the wealth-to-income ratio for a set of rich countries based on their balance sheets since 1970, as well as a longer run perspective for the United Kingdom, the United States and France (these estimates were later updated by Bauluz (2019)). They document an increase in wealth-to-income ratios since the 1970s, from 200-300% to 400-600% in the 2010s. Following this thread, Waldenström (2017) studies the Swedish case, finding that since 1970-80 the private wealth-to-income ratio has been 200% lower than other developed nations. Blanco et al. (2021) estimate the wealth-to-income ratio for Spain since 1900, resulting in a relatively stable aggregate wealth of 400-600% of national income. The ratio then spiked up to 800% in 2007 amid a housing boom, resulting in a J-shaped pattern. Baselgia and Martínez (2021) estimate the private wealth-to-income ratio in Switzerland since 1900, finding a stable 500% throughout most of the period and ending with a significant increase in the twentieth century, reaching over 700%.

Estimates for non-rich countries are scarce but increasingly available. Piketty et al. (2019) combine national accounts, survey, and tax data to account for wealth-to-income ratios in China between 1978 and 2015, finding that while the national wealth-to-income ratio grew from 350% to 700% over that period, public ownership decreased from 70 to 30%. Similarly, Novokmet et al. (2018) find that net national wealth in Russia increased from over 400% to 450% of national income since 1990, amid a large shift from public to private wealth.

Chatterjee et al. (2022) present estimates for South Africa, where the net wealth-to-income ratio declined from about 300% in 1975 to 220% by the end of the 1990s, increasing to 260% thereafter. Kumar (2019) estimate the wealth-to-income ratio in India from 1860 to 2012, finding that it had an increasing trend reaching 600% in the last decade of the period. Del Castillo (2017) represents one of the few wealth aggregate studies for Latin America. Based on the national accounts balance sheet, the study estimates that Mexico's wealth-to-income ratio was 460% in 2014, after a significant increase from a 260% in 2003.

Regarding wealth distribution, the new wave of studies also came with the use of a variety data sources and methods, which often do not provide consistent estimates, giving rise to a debate over trends in wealth inequality, especially in countries such as the United States (Kopczuk, 2016). In contrast with previous estimates (Kopczuk and Saez, 2004), Saez and Zucman (2016) show a dramatic increase in wealth concentration, particularly at the top of the distribution, with the top 0.1% share reaching 22%, rising from 7% in 1978. Kopczuk (2015) disputed this, arguing that the survey-based and estate tax methods indicate that the top 1% share did not increased significantly, while the capitalization method suggests otherwise. The debate is still active, with often not identical results depending on the methodological decisions and sources (see e.g. Sutch (2017); Smith, Zidar, and Zwick (2021); Wolff (2021)), but with increasing agreement on the general upward inequality trend (Saez and Zucman, 2022). On the other hand, estimates for France and the United Kingdom depict an increase in wealth concentration in the last few decades, but much milder than in the American case (Alvaredo et al., 2018; Piketty et al., 2006; Garbinti et al., 2017), and similar evidence has been found for northern European countries and Spain (Martínez-Toledano, 2020; Bricker et al., 2016; Fagereng et al., 2016; Roine and Waldenström, 2009). Novokmet et al. (2018), using the capitalization method, find that wealth concentration increased substantially in Russia between 1995 and 2015, with a top 1% share that fluctuates around 40%, reaching up to 45%.

Another thread of studies aims to analyse wealth inequality worldwide, relying on the combination of a variety of information sources. Davies, Sandström, Shorrocks, and Wolff (2011) use national accounts, wealth surveys[4], and secondary sources for a sample of (mostly) developed countries to fit a model that allows them to estimate the distribution of wealth in the remaining countries. The estimates were updated and improved in Credit Suisse's reports from 2010 until 2021 (see last available Shorrocks, Davies, and Lluberas (2021)),

[4] There is a growing number of studies based on wealth surveys, see Wolff (2021) for the United States, Davies and Di Matteo (2021) for Canada, and Vermeulen (2018) for Europe.

which includes survey-based estimates for Uruguay and Chile.[5] These reports include both the addition of large countries (e.g. China and India), as well as the correction of the right tail of the distribution based on rich lists such as the ones compiled by Forbes magazine and others (Davies, Lluberas, and Shorrocks, 2017), finding that 85-90% of household wealth is owned by the top 10%. This approach is also used by Vermeulen (2018) to adjust the right tail of the United States' distribution using the Survey of Consumer Finances, with similar results.

For the rest of the world the evidence is very scarce, as recent surveys show (Zucman, 2019). Chatterjee et al. (2022) estimate extremely high wealth concentration in South Africa based on the capitalization method, reaching a concentration at the top 1% of over 55%. Del Castillo (2017) estimates wealth concentration in Mexico based on a variety of household surveys, census data, national accounts, and real estate tax. He finds that the top 10%'s share in 2014 was two-thirds of total wealth, while over a third was owned by the top 1%. Gandelman et al. (2022) estimated wealth inequality based on survey data from Mexico, Colombia, Chile, and Uruguay, finding top 1% shares of 26.8, 22.1, 7.6, and 18.4% respectively. Further back in time, Torche and Spilerman (2006) use capital incomes drawn from household surveys to analyse certain asset distributions for sixteen Latin American countries. They estimate business and housing wealth distributions and find that the former is extremely concentrated (for instance, in Uruguay, 99.5% of total business assets are held by the wealthiest 10%) while housing is relatively more evenly distributed. Evidence for Uruguay in particular is scarce, yet growing.[6] Sanroman and Santos (2021) and Agustoni and Lasarga (2019) present estimates of the wealth distribution for 2013 based on the wealth household survey, with a Gini index of 0.75.[7]

4. Data

In this section, the main data sources are described. Section 4.1 provides an overview of the data sources, while Section 4.2 documents the procedures used to adjust these datasets.

[5] Uruguayan estimates are based on the same survey used in this article, see Section 4.

[6] Amarante, Brum, Fernández, Pererira, Umpiérrez, and Vigorito (2010) examine wealth inequality based on capital incomes from household surveys, arriving at results similar to the ones in Torche and Spilerman (2006)

[7] De Rosa (2018, 2019) are early attempts to estimate wealth distribution in Uruguay based on the capitalization method. The results presented in this article represent an extended and improved version of such estimates.

4.1. *Overview of the main data inputs*

For the estimation of wealth-to-income ratios and the set of personal wealth distribution estimates, a wide variety of data sources are used, listed in Table 4.1.

Table 4.1: Main data sources summary

	Years	Source	Observations
Personal income tax data	2009-2016	DGI	75% of adult population
Personal wealth tax data	2009-2014	DGI	Paid by less than 1% of indivs.
Firm tax records	2009-2016	DGI	Univ. of copr. tax-paying firms
Household survey	1986-2020	INE	Nationally representative
Wealth household survey*	2013	UR	Nat. rep., high inc. over-sampled
National Accounts	1988-2020	BCU	No balance sheet. Exp. approach except 2012, 2016
Cadastral administrative data	1999-2018	DNC	Univ. of urban and rural prop. (cad. value)
Registry of decedent's property	2007-2015	DGR	Universe of owner decedents
Real estate transactions	2009-2018	DGR-MGAP	Rural and urban real estate market prices
Demographic statistics	2007-2015	INE	Population totals decedents by age-sex
Financial sector data	2009-2016	BCU	Exchange rates and household's liabilities

Notes: Acronyms in Spanish: *Dirección General Impositiva*, DGI; *Instituto Nacional de Estadísitica*, INE; *Banco Central del Uruguay*, BCU; *Dirección Nacional de Catastro*, DNC; *Dirección General de Registros*, DGR; *Ministerio de Ganadería, Agricultura y Pesca*, MGAP; *Universidad de la República*, UR; *Ministerio de Economía y Finanzas*, MEF. (*) Wealth survey (EFHU) was a joint effort of UR, MEF and BCU.

Personal income tax records. The personal income tax record is a high quality administrative micro-database reported by the tax authority (*Dirección General Impositiva*, DGI) and covering approximately 1,800,000 individuals, that is, about 75% of Uruguay's total adult population. In addition to individual labour incomes and pensions, it also contains information about age, gender, and industry. Capital tax records in Uruguay refer to 12 capital income categories (taxed at flat rates of 7 or 12%), which can be aggregated into dividends and utilities, land and housing rent, and financial incomes. The database also includes capital gains, which are taxed when the gain is realised. For a detailed description of the database, see (Burdín et al., 2022).

Personal wealth tax records. This micro-database provides data on the personal wealth tax. It is a progressive tax, with rates that originally ranged from 0.7% to 2.75%, with and the exempted wealth threshold or about 130.000 dollars. However, the rates have a decreasing schedule which started in 2008 and ends in 2022, at which point a single tax rate of 0.1% will exist. Over the period examined in this article, rates ranged from approximately 0.7%

to 1.85%.[8] Enforcement is relatively low compared to other taxes, and very few individuals actually pay the personal wealth tax (some 8,500 individuals, just over 0.3% of adults, see Table B.3). It is intended to tax the totality of individual's assets, but it has a number of exemptions, and in practice it mainly targets real estate assets (more on this in 4.2.3). For this study, a dataset of taxable wealth and wealth tax is available for 2009-2012, while details on wealth components are available only for 2014.

Firm tax records. This micro-database is provided by the Uruguayan Tax Authority and refers to the universe of firms under the corporate taxation scheme, which excludes very small businesses. Over 100,000 firms are present in this database every year. Firms are compelled to report their total assets and liabilities, as well as the amount of their profits.

Household Wealth Survey. The Household Wealth Survey (EFHU by its Spanish acronym) is a relatively under-exploited survey. It was conducted by the central bank, the Ministry of Finance, and Uruguay's National University (BCU, MEF and UR). It surveyed 3,490 households and it is representative of the whole country. It over-samples relatively richer households, from the fourth and fifth income quantiles and households with business property (Ferre, Rivero, Sanroman, and Santos, 2016). It includes all financial and non-financial assets and liabilities of households, and also provides information about their financial behaviour. There is at the moment only one wave of the survey, covering all assets and liabilities for 2013, and two additional ones with only a specific subset of variables (which were not used for this study). The survey uses a multiple imputation procedure for missing values, but only the first set of imputations was used as not all computations could be performed using the whole set (e.g., source decomposition). The main distributional results are not substantially affected by this decision. Descriptive statistics of the main asset types are depicted in Table A.7.

Household Income Survey. The Household Survey (ECH) is a comprehensive survey of household characteristics. It has been conducted without interruption since 1981 by the National Institute of Statistics (INE). It is nationally representative since 1986, with a large sample of over 30,000 households. It accounts for a detailed desegregation of income sources for each member of the household as well as the household as a whole. In particular, it includes owner-occupied housing income, rents from real estate properties (both housing and land), profits of various types, and interest from deposits and other financial assets.

[8]For more details see art. 45, T. 14 of *Texto Ordenado 1996*.

National Accounts. National accounts are produced by the central bank, covering—in their current publications—the period from 1988 onward. They include aggregate GDP and national income, as well as data on savings and investment. From the perspective of the requirements of this study, there are two major things it does not include. First, it does not, and never has, include a balance sheet, hence the reason aggregate wealth of the country is completely unknown. Second, it has only recently presented desegregated information by institutional sectors for 2012 and 2016, but not for the remaining years (their publication was discontinued in 1999). It represents the single most challenging data restriction for the adequate study of wealth-to-income ratios and wealth distribution.

Cadastre data. Cadastre data was provided by *Dirección Nacional de Catastro* (DNC), which is part of the Finance Ministry of Uruguay. Among other tasks, they collect and update data on the universe of urban and rural properties in the country. Information on each property is documented in publicly available cadastral identity cards (*cédula catastral*), which include a wide variety of property characteristics, as can be seen in the example of Figure B.1. In particular, they present information that allows the unequivocal identification of each property (more on this in 4.2.2), the type of property (rural or various types of urban properties), its size, and the cadastral value for the present and last four years. A micro-dataset with the series of cadastral values for each property for 1999-2018 was built for this research by the DNC, for a total of 242.431 rural properties and 1.383.868 urban properties (see list of variables in Table B.1).

Registry of decedents' property. Administrative information on the properties held by each individual is registered by the *Dirección General de Registros* (DGR). Information about changes in property ownership and estates is published regularly by the State's official newspaper (*Diario Ofcial*), as can be seen in the example depicted in Figure B.8. The dataset thus has the universe of deceased individuals who owned property in the 2007-2015 period.[9] For each owner, therefore, the whole set of properties they held at death is available. The value of the properties is not included, but a set of property characteristics is, so it can be merged with the cadastre values from the National Cadastre data described above. Between 5.500 and 8.800 individuals are present in the owner decedent database, as depicted in Table B.2. The average number of properties they held at death is for most years 3-4, and the

[9] Data registers the year of death, but there may be some administrative delay in the registration, so the more recent the year the more likely to be missing some of the properties or individuals.

median in all years is 1 property. The maximum number of properties held varies from slightly over 90 to several hundred.

Real estate transactions. Market prices by square meter are constructed based on DGR administrative data for all urban real estate transactions (Lanzilotta, Souto-Pérez, and Zunino, 2020). For rural properties, the yearly publication by the Ministry of Agriculture (*Dirección de Estadísticas Agropecuarias - Ministerio de Ganadería, Agricultura y Pesca: DIEA-MGAP*) with market prices by department is used.

4.2. Adjustments to key data

4.2.1. The capital incomes database

The first step of the capitalization method is to assemble a database with all capital incomes, accounting for the full adult population. The database used in this article is estimated following Distributional National Accounts guidelines (WIL, 2020) as much as possible. The full DINA-based income distribution estimation procedure can be found in De Rosa and Vilá (2022). In a nutshell, I begin with a tax-social security dataset, which accounts for 75% of the adult population, and complete it by accounting for informal and untaxed incomes, as well as the population with zero incomes. This first step of combining this data is fully documented in Burdín et al. (2022), and includes the use of a sub-sample of tax-survey matched individuals in order to correct incomes. This is later scaled up by income components to household income aggregates, followed by an imputation of undistributed profits according to two alternative proxies of capital income ownership: (i) the sum of taxed capital incomes from tax records and untaxed capital incomes (such as dividends) from survey data and (ii) using matched firm-individual data to identify possible owners. As an additional adjustment, a stream of liabilities is simulated to match the aggregate (negative) D4-S14 and is distributed to approximately replicate the liabilities distribution from the household wealth survey.[10] The resulting capital income distribution is depicted in Figure A.3 (additional descriptive figures in appendix C).

[10]The distribution of liabilities was approximately 50 and 36% for the bottom 50 and middle 40% of total net worth; 14% for the top 10% excluding the top 1%; and 3% for the top 1%.

4.2.2. Unified Inheritance microdata

As discussed in Section 2.3, the standard requirements for this type of study involve some estate tax data, household wealth control totals from national accounts balance sheets, and a population control total. The only data input readily available in the Uruguayan case is the population estimates by age, since Uruguay has official population estimates based on census data provided by the National Statistics Institute (INE).

Meanwhile, there is no inheritance tax and no balance sheet, so a substitute needs to be found for the former, and the latter needs to be estimated. Balance sheets have never been estimated by the Central Bank of Uruguay, which only reports the government sector balance sheet and the net foreign asset position of the country in the balance of payments. Moreover, since 1974 there has been no inheritance tax. The only tax paid on estates is the *Impuesto a las Transmisiones Patrimoniales* (ITP), which is a flat tax of 3-4% on all real estate transactions, including bequests as well as sales and gifts. Unfortunately, this tax only reports data on the individuals receiving the estate, with no information on the decedent. Therefore, it is impossible to aggregate estate tax data at the individual level based on ITP.

This study is based on two main administrative data sources. The first one is cadastre data on all urban and rural properties, with a wide set of characteristics including size and cadastral value for the 1999-2018 period. The second one is a decedent-owner registry with all decedents who owned properties between 2007 and 2015. Both datasets are based on public information but were constructed especially for this study. These two datasets can be merged, thus allowing an analysis of decedents' real estate wealth. Each of them, together with supplementary data and the merging and adjusting procedures performed, are described in this section.

Cadastre property values. The cadastral value of a property may be modified for three reasons. The most common one is the annual update of cadastral values done by the DNC, which is based on a combination of the evolution of the general price index and the cost of construction index (IPC and ICC by their Spanish acronyms). The second and third reasons are related to changes to the buildings located on the properties (e.g., additional rooms built in a given house). These changes can be detected if the property is sold or if there is a general inspection and revaluation of properties in the region.

An example of a property for which the cadastral value has been reassessed is presented in

Figure B.3[11], which depicts a spike in the value and a smooth evolution thereafter. While this type of revaluation is not uncommon, it does not seem to generate any discontinuities in the aggregate evolution of cadastral value, which shows a smooth evolution for both rural and urban properties as depicted in Figure B.6. Regional revaluations could be a more serious concern if they entailed a generalized increase in the cadastral values of a given region, but they were extremely scarce and of limited reach in the study period, and no effects are visible while considering smaller geographical subdivisions.[12]

The unified real estate-decedents database Following the discussion in Section 2.3, the basic input required is a dataset with deceased individuals and the real estate wealth they held at the time of death, valued at market prices. This basic input results from the merging of the two datasets previously described, along with a number of adjustments, which are discussed hereafter.

In order to construct a dataset of the individual real estate wealth of the decedent population, it is first necessary to merge the data from the DNC (with the cadastral value of properties) and the DGR (with deceased owners of properties). The merger is performed at the property level, based on the data existing in both datasets.

In the case of urban properties, in order to single out a property, three different variables need to be considered, as depicted in Figure B.2: the department, the locality within the department, and the number of the property. The reason the property number by itself is not enough is that the numbering starts over in each locality of each department. Two more variables are needed to adequately identify a single property: unit, in the case of apartment buildings (each unit representing a flat within the building) and block, which is used in some recently urbanized localities to subdivide properties. The unit variable is available in both datasets, but the block is only available in the cadastre data, hence it was not used in the merger. In the case of rural properties, the merger is simpler, since singling out a property requires only the number and department. Figure B.10 shows that when the two datasets are merged based on these variables, 78-80% of dead individuals in DGR's raw data are merged, corresponding to 72% of rural owner-property observations and 84% of urban owner-property observations.

As depicted in Figure B.10, between 14 and 22% of each years' total decedents are accounted

[11] Figures regarding data adjustments referred in this section are depicted in the appendix B.

[12] The country is divided in 19 departments. Evolution of cadastral aggregate value for all of them is depicted in Figure B.4, which shows no discontinuities.

for. Thus, if we assume that the decedents in the DGR data are in fact the wealthiest of all decedents of that year, we have a number that allows to compute (at least) the top 10%'s share. It is worth noting, however, that the number of decedents accounted for slowly decreases. This is due to the fact that the registry may have some delay in being updated, and therefore not all of the decedents for the final years are present. This fact will be considered when analysing the results in the following sections.

Adjusting the data: individual real estate and market prices. Once we have a unified decedent-real estate database, two main adjustments need to be performed: (i) converting cadastral values into market prices and (ii) accounting for individual real estate (as opposed to household real estate).

The first and most important step is the market price adjustment, which has an effect not only on decedent real estate wealth, but also on the estimation of the national and household real estate wealth (W^r_{nt} and W^{hr}_t). For both rural and urban properties, the value of each property is adjusted by multiplying its surface area by its market value . In the case of rural properties, official data from the Agriculture Ministry (*Ministerio de Ganadería, Agricultura y Pesca*) on price by hectare by department, which is published on yearly basis, was taken.[13] In the case of urban properties, values were adjusted based on the ratio of average market prices and cadastre values by department based on Lanzilotta et al. (2020), who use real estate transactions administrative data provided by DGR (see 4.1). When computing aggregate urban and rural real estate at market prices with the aggregate cadastral values, the resulting adjustment ratios, depicted in Figure B.5, were 14-17% for rural properties and 39-44% for urban properties. These ratios, which are used to adjust each property's price, are exactly within expected adjustment values.[14] Moreover, they are very stable, which is indirect evidence that the yearly update of cadastral values is aligned with market prices' evolution in this period.

The second adjustment involves distinguishing between household and individual real estate. In the decedents registry dataset, all properties held by each deceased individual are accounted for. Nevertheless, we ignore if they are held at the same time by someone else, which is naturally problematic since a significant proportion of real estate wealth may have been

[13] See *Anuario Estadístico* by DIEA in https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/comunicacion/publicaciones/.

[14] Unofficial ratios used by the DNC in back-of-the-envelope calculations are 15% for rural properties and 40% for urban. Unfortunately, these estimates are not discussed in any official DNC document, but they were confirmed by the DNC's high-ranking officials.

accumulated jointly with a spouse. To adjust for this, two strategies were alternatively taken. In the first one, it is assumed that all real estate is accumulated jointly by two individuals, and therefore the decedent's registered real estate is in all cases actually the household's real estate. Under this assumption, real estate value is divided by two to account for the share of that wealth actually belonging to one of the spouses (defined as equal-split in the Distributional National Accounts guidelines (WIL, 2020)). Dividing by two will also reduce by half the wealth of the last surviving spouse in a given household, but this is a desirable property. As Atkinson pointed out, "The most common 'sideways' transfer is from husband to wife or vice versa. Ideally, we should like to exclude such within-generation transfers (including those from brother to sister or cousin), but this is not always possible, and to this extent the degree of inter-generational transfer is over-stated" (Atkinson, 2018, p. 143).

The preceding assumption naturally represents a lower bound estimate, since we are not considering that some individuals may have accumulated their real estate wealth alone. A less restrictive assumption would therefore be to account for individual accumulation. Based on the wealth household survey, we find that 10% of individuals live alone, while another 10% declare having wealth separated between spouses (totaling 20% of individuals). Thus, considering that a combined maximum of approximately 30% of decedents accumulated their wealth individually, a 70-30 split was performed to the data: the real estate of 70% of individuals was divided by 2, and for the remaining 30% it was left unchanged. Results reported are the average of 500 random draws from a Bernoulli distribution with these probabilities.

4.2.3. Wealth tax data

According to the Uruguayan tax code, all assets are taxed by *Impuesto al Patrimonio de las Personas Físicas* (IPPF), with the exception of rural properties used in production up to 2013 which were included after thereafter. Detailed data for 2014 has information on all asset types, including housing, land, firm's shares, deposits, assets abroad and liabilities. However, 84% of reported taxable wealth is housing.[15] Therefore, in practice, the wealth tax mainly targets real estate assets.[16] This is clear also for remaining years when comparing wealth tax payers to capital incomes tax payers (see Table B.3) Out of the individuals who pay taxes on dividends, i.e. individuals for whom we are certain that own firms' shares, only 6% pay wealth tax. Conversely, 1-4% of wealth-tax payers receive dividends. This indicates that financial assets are not well targeted by the tax. At the same time, 4-5% of individuals who

[15] This share excludes land, which in any case accounted for 11% of declared taxable wealth in 2014.

[16] This is informally accepted by Uruguayan taxation experts, but unfortunately it is not accounted for in any official document.

pay taxes rents pay the wealth tax. However, in this case almost a quarter of wealth-tax payers receive rents, which indicates that at least in part properties other than duellings are taxed. Thus, these results point at housing as the main asset type actually targeted by the wealth tax.

A number of adjustments are required to calculate housing assets at market prices. Housing is taxed based on cadaster value, and the tax schedule allows to exempt 50% of dwellings, up to the exempted wealth threshold. For rented properties, the value is set at fifteen times the yearly rent. Finally, each individual pays tax according to the share of the property they declare to own. Thus, to taxable housing wealth the allowed exemptions are added (i.e. the maximum among the exempted threshold or 50% of the declared value), 70% of the value is considered to be split with a spouse (as discussed in 4.2.2), cadaster values are adjusted to market values and fifteen times reported rents (based on income tax data) are computed as the value of rented properties. For 2014, housing value is computed with this procedure based on reported taxable housing, while for 2009-2012, these adjustments are applied to 84% of taxable wealth, i.e. the share of taxable housing wealth in total taxable wealth (excluding land).

5. Private wealth and its distribution

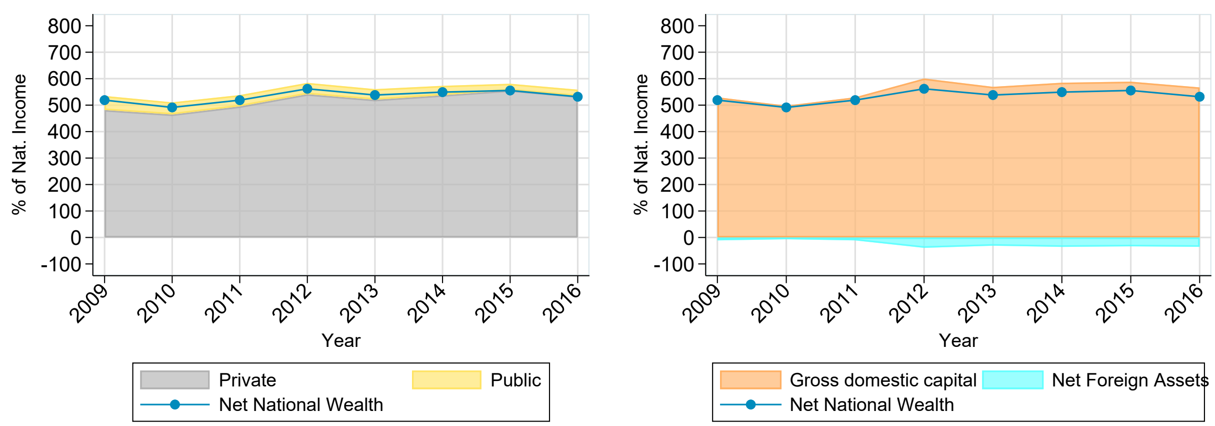
Wealth-to-income ratios for Uruguay in recent years are presented in this section, which is informative on its own and also as a key input to the capitalization method's distributional estimates. Appendix D provides a full account of the methods and sources to estimate book-value national and private wealth, as well as it components, while a full discussion of the geographical distribution of real estate wealth is presented in appendix E.

5.1. National and private wealth-to-income ratios

The resulting net national wealth-to-income ratios are depicted in Figure 5.1. In panel (a), national wealth is split into its private and public components. While aggregate net wealth is between 450-500% of national income, private wealth is some 25-50% lower, but increasing, as a result of the government's positive net wealth discussed. General government net worth at the beginning of the century was approximately 67% of national income, but by 2016 it was 21.5%. The first years of the 2000s witnessed a sharp drop and partial recovery during the crisis, a relatively stable ratio of 50% of national income up until 2009, and a steady decrease afterwards. Figure A.1 shows that it followed the same general trend as public aggregate net worth; i.e., it is not the artificial result of variations in national income not reflected in public wealth. Despite this downturn, it is worth noting that Uruguay's government net worth is significantly higher than in countries such as the US, France, Japan, or Britain, where it fluctuates around zero, and is much closer to the Chinese government's balance sheet (around 30%) (Piketty et al., 2019).

These estimates place Uruguay well within available estimates, which mostly come from rich countries (Piketty and Zucman, 2014). Although they present as their benchmark series the market-value wealth to income ratios, but book-values are remarkably close, especially in recent years. However, it is worth noting that given Uruguay's relatively low national income, average wealth is relatively low. As shown in Figure A.2, it has an increasing trend up to 2013, when it reaches US\$80.000 per adult (PPP) and declines thereafter down to US\$60.000, following a similar pattern than the level of net wealth in US dollars.[17]

Fig. 5.1. Net national wealth-to-income ratio 2009-2016



(a) Public and private

(b) Gross domestic capital and net foreign assets

Note. Panels (a) and (b) computed based on equation 2.1. Net national wealth is equivalent by definition in both cases. See net national wealth by sector and financial/non-financial assets in Figure D.6

Private wealth's decomposition is presented in Figure 5.2. Panel (b) shows that around 60% of private wealth is financial wealth[18] while roughly 25-30% is housing (which is slightly over 100% of national income, see panel a), and the rest of gross private wealth is business

[17]Section D.5 of the appendix provides some insights on the long-run evolution of the wealth-to-income ratio within a simple one-good model framework of savings and growth. It shows the key role of growth rate (in particular during economic crisis) to drive the wealth-to-income ratio.

[18]Unfortunately, I cannot distinguish between different asset classes within financial assets, which represent the bulk of private wealth. To get an approximation, it is worth noting that the private sector's net undistributed profits (B5n, S11+S12) represent 45% of total investment income, while the rest is household property rents (D4, S14) (De Rosa and Vilá, 2022). Although this does not translate directly into asset composition due to different rates of return, it does provide a general idea of what the split may look like.

assets. Private aggregate liabilities represent less than 5% of net wealth. Although these estimates present greater aggregation, their levels are very consistent with similar studies (see for instance Saez and Zucman (2016)).

Fig. 5.2. Private net wealth

(a) Level

(b) Composition

Note. Private wealth level and composition depicted in panels (a) and (b). Private net wealth refers to the aggregate depicted in panel (a) of Figure 5.1.

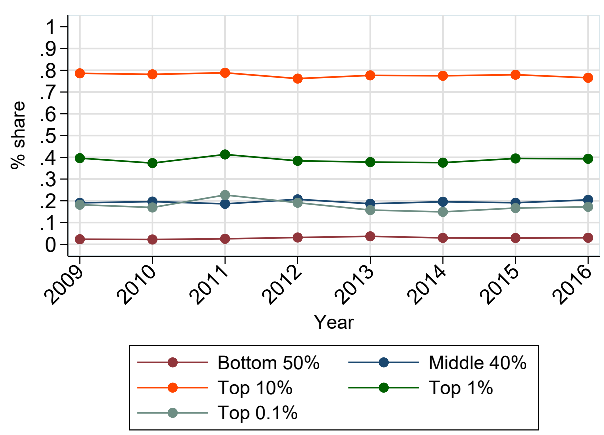
5.2. *Private wealth distribution*

In the capitalization method, the underlying wealth of each individual is estimated based on her capital incomes, which are capitalized using factors that are the equivalent to the inverse of their macro return rate. Thus, one needs a capital incomes distribution and a set of return rates, which are presented in panels (a) and (b) respectively of Figure A.3.

As discussed in Section 4, the capital incomes database is taken from the Distributional National Accounts estimates from De Rosa and Vilá (2022). Overall, capital incomes are highly concentrated (with a Gini index of about 0.7), with housing the least unequally distributed and financial incomes (which include all incomes derived from corporate sector ownership) the most concentrated. Regarding return rates, housing shows return rates of 7-8%, while financial assets have implicit return rates between 7-10%, financial liabilities 5-8%, and finally business assets have the lowest return rate at 4-5%. These returns are relatively close to the standard rates, as exemplified in WIL (2020). Pensions fund wealth is included in financial assets, but it was not capitalized; rather, it was imputed according to wealth survey estimates: approximately 31 and 41% for the bottom 50 and middle 40%, 28% for top 10% excluding the top 1%, and 8% for the top 1%. Returns are adjusted to ensure income-wealth consistency, e.g., that aggregate financial wealth divided by total financial incomes results in financial asset return rates. Given the very low share of pensions, the adjustment has no effect whatsoever.

The resulting net private wealth distribution is presented in Figure 5.3, which once aggregated is equivalent to total private wealth depicted in panel (a) of Figure 5.1. The top 10% and bottom 50% shares are remarkably similar, while the main difference is the top 1%, which is stable but 7-8 percentage points higher. The top 10% share is 77-79%, while the top 1% is stable at around 38-40%. As expected, the bottom 50%'s net wealth is under 5% throughout the period. Both the top 0.1% and middle 40% own approximately 20% of private net wealth. In the period, there is a slight downward trend for inequality, with a reduction in the top 10% and an increase in the bottom 90%. In Figure A.9, point estimates are bootstrapped to produce confidence intervals, indicating that the trend is statistically significant. The period is clearly divided in two, with a reduction in inequality up to 2012, and stability thereafter. The top 1 and 0.1%, on the other hand, do not show statistically significant upward or downward trends. This is interesting since, while the top 10% trend is similar to its income distribution counterpart, the top fractiles show stability for wealth but increasing trends for fiscal income (Burdín et al., 2022). Wealth distribution by type of asset is depicted in Figure A.4, which shows that financial wealth is by far the most concentrated, with virtually all of it held by the top 10%. At the other extreme, 40% of housing is owned by the top 10%, with over half of it owned by the middle 40%.

Fig. 5.3. Net personal wealth distribution 2009-2016



Note. Capitalized incomes, based on Distributional National Accounts estimates (De Rosa and Vilá, 2022). Capital incomes distribution and rates of return rates presented in Figures C.1 and A.3. See bootstrapped confidence intervals of wealth shares in Figure A.9. Average wealth by assets is depicted in Table A.3, thresholds in Table A.4, and percentage of asset ownership in Table A.5.

These estimates are similar to those found, based on the same methodology, for countries such as the US or Spain, but much higher than those for France. In the case of the US, Saez and Zucman (2016) estimates that the wealthiest 0.1% owned around 22% of total net wealth in 2012, while the top 1%'s share was close to 42%. In Spain, capitalization method estimates show that the top 1%'s wealth share is around 40%, while the top 10% share is 65-75% (Martínez-Toledano, 2020). Finally, in France, the top 1%'s share is 20-25%, and the top 10%'s share is 55% (Garbinti et al., 2017). Slightly lower wealth inequality is found, somewhat surprisingly, in the Mexican case, although results are not strictly comparable since the methodologies are not the same (Del Castillo, 2017).[19]

It is worth stressing at this point that these results are highly dependent on the capital incomes distribution. These, in turn, are the result of the process of Distributional National Accounts' construction, which is very sensitive to the way in which capital incomes are scaled to the household sector, and, more importantly, to the way in which undistributed profits are imputed. Indeed, when upper bound estimates from De Rosa and Vilá (2022) are considered, top 1% and 0.1% shares increase as a result of a concentration within the top 10% (which does not change its share significantly), as depicted in Figure A.8. This suggests that results should be taken with caution and compared with other data sources and methods, which is done in Section 6.

Additionally, to address the main caveat of the capitalization method, i.e., the assumption that rates of return are the same for all individuals—and in particular, are independent of wealth— two simple sensitivity analysis are performed in appendix F, showing that the assumption does not affect the main estimates substantially, except for the very top shares under some strong assumptions.

5.3. Wealth owners characterisation

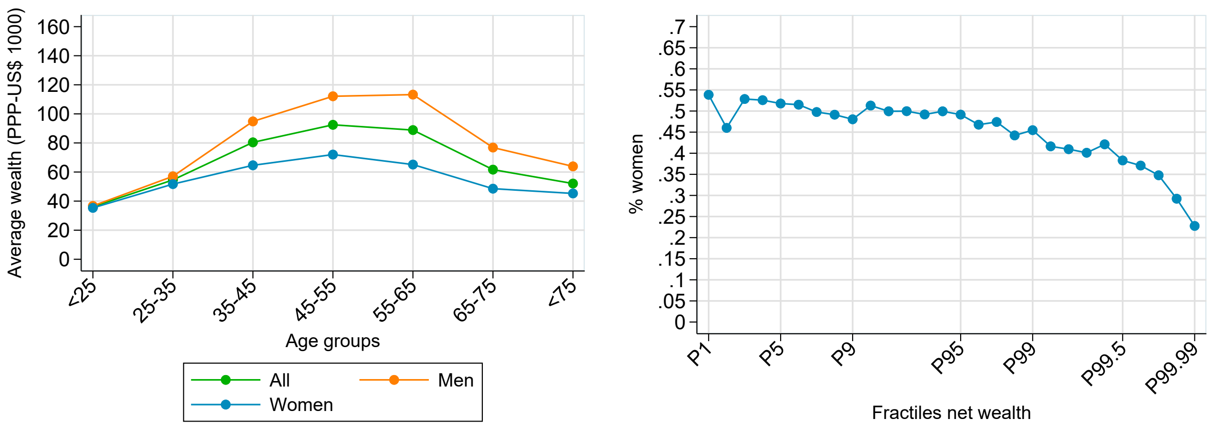
Information from the tax records on individuals is used to characterise wealth holders. In panel (a) of Figure 5.4, average wealth by sex and age group is depicted, showing that average wealth tends to increase up to retirement age and slightly decrease afterwards. This suggests there is not a clear life-cycle pattern, although it is not possible to be unequivocal since the estimates are based on cross-sectional data.

Mean wealth is higher for men in all age groups, but particularly in the 40-60 years old interval. It is interesting to observe how the gap increases as individuals grow older until

[19] For an overview of estimates available at Wid. World, see Table A.1.

approximately 60 years old, at which point both groups start to converge. This may be explained by the fact that women tend to live longer than men, and they also may inherit their partner's wealth when they die. Table A.10 shows that the proportion of women grows steadily from around 45% in the youngest age group to over 65% in the oldest. Panel (b) of Figure 5.4 depicts the percentage of women by net wealth fractile. The share of women decreases steadily from over 50% (until the seventh decile) to a little over 20% for very top fractiles. It is interesting to note that up to the 99 percentile, the share is still relatively high (about 45%), but it plunges thereafter, especially within the top 0.5%.

Fig. 5.4. Wealth by age and gender, 2016.



(a) Average wealth

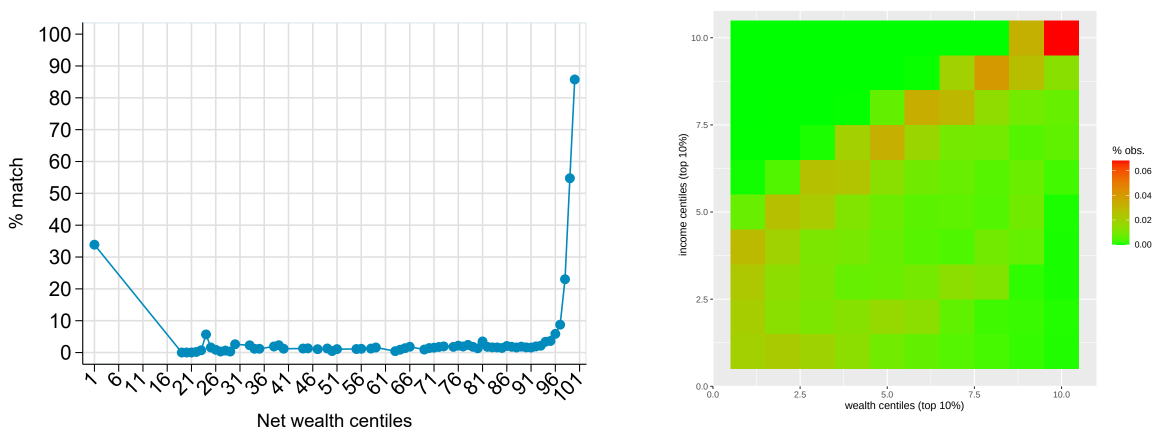
(b) Share of women

Note. Wealth averages for ten-year age groups and sex depicted in panel (a). Individuals over 20 years old. Thousands of US dollars, PPP. Similar results for each year depicted in Figure A.13. Average wealth by year in Figure A.2. Panel (b) depicts the percentage of women by net wealth fractile for 2016 (remaining years in Figure A.14).

Table A.4 depicts net private wealth thresholds in thousands of US dollars for the period, reflecting the highly skewed distribution of wealth. It is interesting to note that all thresholds move upward until 2012, and then downward from then on. This is likely the result of movements in the exchange rate, which fell by 10% between 2009 and 2012, and then increased by 49%, which is likely to impact wealth denominated in US dollars of pesos earners.

The capitalization method allows me to analyse jointly the distribution of wealth and income. Although the wealth distribution depends on the capital income distribution and is likely to be similar (as in fact it is, see Figures A.3 and A.5), they do not match for two reasons: (i) capitalized incomes are just a part of the income distribution, which includes in particular labour incomes and pensions, among others, and (ii) heterogeneity in the rates of return for different types of wealth results in individuals with the same total capital income but different composition having different estimated wealth as well. Panel (a) of Figure 5.5 depicts the percentage of individuals who belong to the same percentile in both distributions, by wealth percentile. Within the top 1%, over 70% of individuals match, i.e. the vast majority of those with top incomes are also top wealth holders, which is the result of the extreme capital income concentration at the top of the distribution (Burdín et al., 2022). The matching is relatively higher in the top 10% compared to the rest (especially in the 99 p-tile), but is not nearly as high as in the top 1%.[20] Panel (b) presents a heatmap of the top 10%. As expected given panel (a), there is a greater concentration of observations in the top percentile of both distributions, but there are also relatively more observations above the main diagonal. This means that individuals in the top percentiles of the wealth distribution tend to fall relatively high in the income distribution.

Fig. 5.5. Income and wealth distribution's matching, 2016



(a) % matching by wealth p-tile

(b) Heatmap (top 10%)

Note. Income refers to total incomes (including capital incomes, labour incomes, pensions, and other incomes). Panel (a) shows the percentage of matching percentiles of income and wealth distributions by private wealth percentile, while panel (b) depicts the heatmap of the top 10% of wealth and income distribution, i.e. the percentage of observations in each cell of income/wealth percentiles. By construction, the main diagonal of the heatmap is equivalent to the top 10% of panel (a). Estimates for remaining years yield the same results (see Figures A.15 and A.16).

[20]These results are consistent with the findings of Sanroman and Santos (2021) based on the wealth survey.

6. Triangulation of distributional evidence

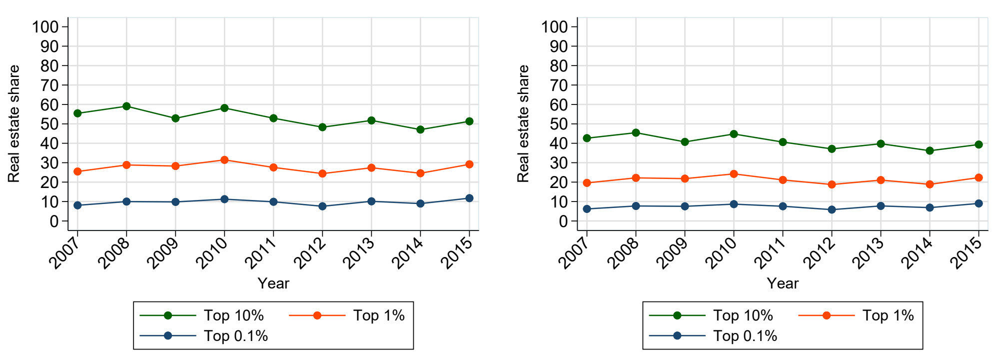
6.1. *The personal distribution of real estate wealth*

6.1.1. *The distribution of real estate based on estate-decedents data*

To estimate the personal distribution of real estate household wealth based on the estate multiplier method, I begin with the real estate-decedents database, elaborated using the procedure discussed in 4.2.2. That database is weighted by the inverse of the mortality rate in order to match the aggregate population of adults 20 years and older, and it is adjusted by μ. This ratio of the average real estate wealth of decedent and living populations is proxied based on the household wealth survey. To compute it, the average real estate wealth of individuals 78 and older (Uruguay's life expectancy) is divided by the equivalent for younger individuals. This procedure provides a $\mu = 1.4$, close to available estimates (Piketty and Zucman, 2015; Ohlsson, Roine, and Waldenström, 2020), which are between 1.4 and 1.5.

Once individual real estate wealth is computed, and considering aggregate household real estate wealth, the estimation of the top real estate wealth shares is straightforward. Panel (a) of Figure 6.1 presents the 70-30 criterion of household wealth adjustment, while panel (b) depicts the lower-bound equal-split estimates (see Section 4.2.2). Lower-bound estimates show that real estate wealth's top 10% share is around 35-45% , the top 1%'s share is 20-25%, and the top 0.1% is under 10%. Estimates under the less restrictive 70-30 split assumption show a top 10% share of 50-60%, top 1% of 25-30%, and top 0.1% of approximately 10%.

Fig. 6.1. Real estate distribution (upper and lower bounds)



(a) Personal real estate wealth distribution (70-30 criterion)

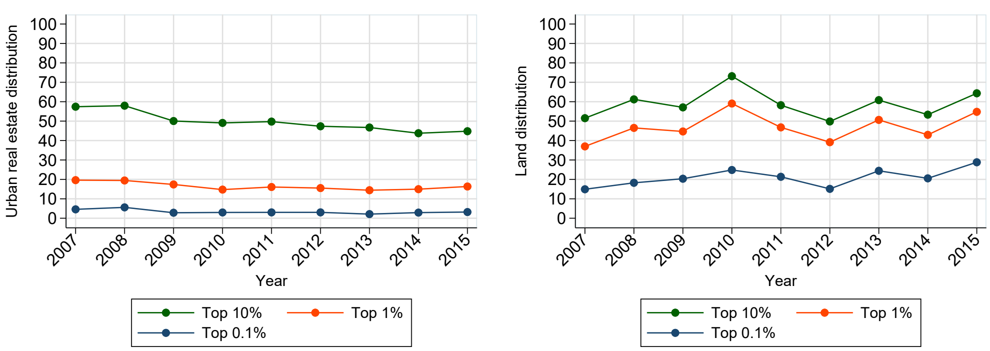
(b) Personal real estate wealth distribution (equal split)

Note. Based on DGR and DNC merged data. Decedents' wealth expanded based on average mortality rate and considering a decedent/adult population wealth ratio of 1.4, computed based on wealth survey.

As pointed out by Atkinson (2018), it would also be interesting to analyse the distribution of wealth among the decedent population, i.e. without expanding to the total population. Unfortunately, this cannot be done in this setting, because the total amount of decedents' real estate is not available. One possibility would be to use total wealth from the registry, since there is no property under the tax threshold (non-filers). However, it is unlikely that this total accounts for total real estate wealth of decedents, because only about 20% of decedents appear in the decedent owners registry as discussed above. As a reference, Atkinson (2018) documents that the United Kingdom's estate records account for 45-50% of the number of deaths, while in France it reaches 65% (Piketty, 2011).

Estimates for rural and urban real estate top shares are presented in Figure 6.2 (70-30% split, see lower bound estimates in Figure A.11). Urban real estate is more evenly distributed than rural, and it is relatively close to the total real estate top shares, given its larger share in aggregate wealth.[21] Rural distribution shows significantly higher concentration, especially in the top fractiles. The top 1%'s share is as high as 40-50% for most years (although quite unstable), the top 0.1% is around 15-25%, and the top 10% varies between 55-70%. This high level of land concentration is similar to what has been reported for other Latin American countries by Bauluz, Novokmet, and Govind (2020).

Fig. 6.2. Real estate distribution by wealth type



(a) Urban (b) Rural

Note. Based on DGR and DNC merged data. Decedents' wealth expanded based on average mortality rate and considering a decedent/adult population wealth ratio of 1.4, computed based on wealth survey. 70-30% split criterion.

Urban property distributional estimates, far more stable than those of rural properties, are compared with capitalization method housing concentration in Table A.9. It is interesting to note that the top shares of urban properties are higher than the housing share from the

[21]It is worth mentioning that the urban-rural split found in the aggregate data is also present (almost identically) in the real estate-decedents dataset, as depicted in Figure B.9 of the appendix.

capitalization method for the top 10 and 1%; it is 5 percentage points higher for the top 1% and 10% for the top 10%, but 2 points lower for the top 0.1%. Lower bound estimates, on the other hand, show almost identical shares for the top 1% and 10% (11-12% and 37-38%). While these estimates are not strictly comparable, since the capitalization method's housing may include some of the rural properties in panel (b) (houses on rural land in which people actually live), and some urban properties may not be housing (e.g., storage facilities, buildings used for closely held businesses, etc), they may still indicate an underestimation of the main estimates. Indeed, top housing shares in the capitalization method are likely to be underestimated, since they are based mainly on imputed rents declared by households (more below).

6.1.2. The distribution of real estate based on the wealth tax

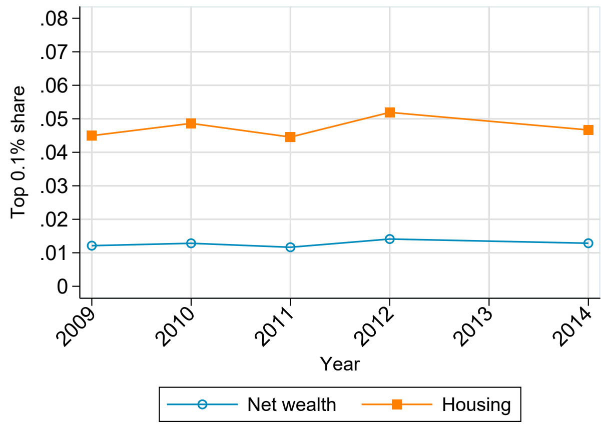
Another possible concentration robustness check is to compare the results with the wealth-tax data estimates. As discussed in Section 4, this tax is paid by approximately 0.3% of adults and mainly taxes housing, hence it is possible to use it to compare to the top 0.1% housing share. To do this, individuals from the tax records are sorted by their housing wealth, and the top 0.1% share is computed comparing the fractiles' total real estate wealth to total housing. Figure 6.3 shows the resulting share, which is 4-5% in the period under analysis. This result is very similar to the ones found in main capitalization method estimates, as shown in Table A.9. Considering the usual evasion and avoidance caveats of tax data, these results are likely to underestimates housing concentration, which is further evidence that capitalization method results also represent a lower bound. However, it is clear that the wealth tax is not a good source to study total wealth distribution. Assuming that it captures all types of assets and thus comparing it with aggregate net wealth, the top 0.1% share is a little above 1%, much lower than both the capitalization method (18% on average) and the survey (11%) as depicted in Table A.9.

6.2. The wealth household survey

The wealth household survey includes assets similar to the ones considered in the capitalization method estimates, hence providing an important insight on wealth distribution and a key data source to contrast capitalization method-based estimates.

In Table 6.1, gross and net wealth shares are presented along with their components, both for the household and per adult. The last column takes into account not only financial assets, liabilities, business, and housing, but also durable goods and jewelry, which are not accounted for in the capitalization method. The first thing to note is that the last three columns present remarkably similar estimates, which indicates that the choice of gross, net, or total wealth

Fig. 6.3. Top 0.1% share based on wealth tax, 2009-2014



Note. Based on wealth tax micro data, after adjustments discussed in Section 4.2. Net wealth refers to top 0.1%'s wealth share in aggregate net wealth, while real estate's is the same share but in aggregate real estate wealth (i.e. urban and rural properties).

does not seem to affect distributional estimates significantly. This reflects that, according to the survey, the assets excluded from the capitalization method's estimates do not have a significant impact on wealth distribution. Thus, from now on, only the comparable wealth definition of the survey is considered.

Several comments are worth making regarding the comparison with previous estimates. First, the net wealth top 10% share is much lower than that in the main estimates (61 vs 78%), and most of this difference is explained by the top 1% (Table A.9). This is not surprising, given the extensively documented issues with wealth surveys capturing the very top wealth holders (Vermeulen, 2018; Kennickell, 2019). The differences in the top tail are also clearly visible when comparing gross and net wealth thresholds from the capitalization method (Table A.4) and in the survey (Table A.7).

Second, while financial assets and businesses are more concentrated than housing in both the survey and the capitalization method, it is surprising to note that each of the asset types is more concentrated in the survey than in the capitalization method. The higher concentration of gross and net wealth in the capitalization method is hence the result of differences in composition, as the source decomposition (Lerman and Yitzhaki, 1985) of Table 6.2 shows. While housing represents 84% of gross wealth in the survey, it only accounts for 23% in the capitalization method. As for business and financial assets, they represent 22 and 56% in the main estimates respectively, while reaching just 3 and 13% in the survey. In both data sources, financial and business wealth are by far the most concentrated, while

Table 6.1: Household wealth shares, wealth survey data, 2013

	Housing	Business	Fin. Assets	Liab.	Gross wealth	Net wealth	Incl. dur.
			Household wealth				
Bottom 50%	5%	0%	0%	0%	5%	3%	5%
Middle 40%	50%	0%	9%	18%	37%	37%	37%
Top 10%	45%	100%	91%	82%	58%	60%	58%
Top 1%	13%	69%	62%	32%	26%	27%	26%
			Per adult household wealth				
Bottom 50%	5%	0%	0%	0%	5%	4%	5%
Middle 40%	46%	0%	9%	18%	36%	35%	36%
Top 10%	49%	100%	91%	82%	59%	61%	59%
Top 1%	15%	70%	62%	33%	26%	27%	26%

Note. Household wealth survey (EFHU) for 2013. Housing includes primary houses and secondary properties, for households that only own houses as real estate. "Business" refers to sole-proprietorship firms, while corporations are included in financial assets (together with deposits, bonds, etc), and may include some houses owned by households that own a pool of properties (houses, land, and business facilities, among others). Pension funds are included in financial assets. Liabilities include mortgages.

housing is more equally distributed (see more on housing below). Primarily as a result of the differences in wealth composition, financial wealth contributes to 62% of gross wealth inequality in the capitalization method, but only 11% in the survey, while the opposite happens with housing: a 10% contribution in the capitalization method vs 85% in the survey (the contribution in business wealth is comparatively similar in both sources). Moreover, each of the components' elasticity of gross wealth inequality has the same sign in both sources (only negative in the case of housing), but while the negative contribution to inequality in the capitalization method is twice as high as in the survey, its positive contribution is ten times higher. Thus, lower top shares in the survey are the result of a lack of capacity to adequately capture financial assets, mirroring what happens in income surveys (Alvaredo et al., 2022). This is further confirmed in Table A.2, which scales each wealth component to match their counterpart in aggregate household wealth, keeping their distribution unchanged. This adjustment yields larger wealth concentration, with a top 10% of 79.6%, in line with capitalization method results. This simple exercise, although not very informative on its own, does highlight the fact that the low survey concentration estimates are mainly the result of a composition miss-match.

Third, regarding housing, Table A.9 shows that survey estimates are higher than capitalization method estimates, but remarkably close to those based on the estate method (upper bound). For example, in the case of the top 1% share, the survey estimate is 15%, while it is 16%

Table 6.2: Inequality decomposition, survey and capitalization method

	Share	Gini	Cont.	Elast.
		Business		
Capit.	22%	0,96	24%	0,03
Survey	3%	0,99	3%	0,00
		Housing		
Capit.	23%	0,58	10%	-0,09
Survey	84%	0,78	85%	0,02
		Financial Assets		
Capit.	56%	0,92	62%	0,06
Survey	13%	0,86	11%	-0,02

Note. Based on EFHU. Decomposition based on Lerman and Yitzhaki (1985), for 2013.

from the estate method (12% in lower bound estimates), and 11% from the capitalization method. For the top 0.1%, the survey results in a 3% share, identical to the estate multiplier's estimates. The survey and capitalization method's shares are mirrored by the thresholds, which are lower in the survey for the bottom 50% and higher for the top 1%: 13 vs 15 and 226 vs 121 thousand dollars respectively (see Tables A.7 and A.4).[22] Note that for the household's main residence, both the survey and capitalization method rely on self-reporting, since the capitalization method capitalizes owner-occupied housing rent (from the "regular" household survey, ECH, see Section 4), but for the remaining properties, capitalization estimates use taxed capital incomes (rents), so they may be more accurate. However, if self-reporting on the value of a property is more precise than the income flow the house would provide if rented, then the wealth survey should provide a better estimate. It is interesting to note that the percentage of house ownership is 67% in the survey and 74% in the capitalization method (Tables A.7 and A.5), which explains some of the difference in inequality but implies that a larger share of people report owner-occupied housing rent (in the regular household survey) than actual house ownership (in the wealth survey). Given the much larger sample size of the household survey (ECH), it is probably more reliable. Further evidence is still needed to fully account for these differences, but in any case, it is safe to say that capitalization method housing inequality estimates are likely to represent a lower bound.

[22] In the case of the estate method's estimates, the threshold is much lower for the top 10% (23-28 thousand US$), but higher for the top 1% (247-302 thousand US$), which is consistent with relatively lower top 10% shares than top 1% shares (see Table A.6).

6.3. Rich lists

The final piece of data used to compare with the capitalization method's main estimates is the *Forbes Billionaires List* rich list. In 2022 two Uruguayans entered into the list for the first time ever, with a fortune of around 1.5 billion dollars each.[23] Although 2022 is out of the period under analysis, and two individuals' wealth alone does not enable me to compute top shares, it does allow me to compare their reported wealth with top wealth holders in the capitalized wealth data.

Table A.8 presents the top 10 net wealth holders for the last available year (2016), the position they held in the previous years, and their net wealth, based on the capitalization method. Two things are worth noting. First, as with the *Forbes* list, only two individuals enter the billionaires club, i.e. hold over one billion US dollars (at 2022 values), one holding over 1.5 billion dollars at 2 billion net wealth, and one holding less than 1.5 at 1.1 billion 1.1 , with both holding significantly more wealth than the rest of the group. Given the anonymous nature of the data and the 2016-2022 time-span, it is impossible to know if these individuals are the same as those in the rich list, but in any case, this indicates that there were in fact billionaires before 2022 and that the top fortunes in the capitalization method estimates and those of *Forbes* billionaires list are of the same orders of magnitude. If some of them were actually the same, the reason they may not have been in the *Forbes* list beforehand could be associated with the fact that the firms they own started to publicly trade on the US stock market only in 2021[24], and therefore were likely to be invisible to *Forbes* before that.

Second, exploiting the panel nature of this data, it is possible to see that out of the two billionaires, one was in the top 10 throughout the entire period, while the other entered the top 0.01% in 2013 but remained in an extremely high position (either first or fourth) thereafter. This may be compared with the probability that those in the top 0.01% of individuals in 2016 were part of that group in the previous year, depicted in Figure A.12. The figure shows that the probability decreases (as expected) as we move further back in time, starting from a nearly 30% persistence and decreasing to under 10% eight years prior. Although this simple comparison does not enable me to draw categorical conclusions, it does provide further evidence that the capitalization method's top estimates are close to what is found by completely external data sources and methods.

[23]See https://www.forbes.com/billionaires/. "Billion" dollars refers to 1.000.000.000 dollars.
[24]See shorturl.at/qvT29 (in Spanish).

7. Concluding remarks

Capital incomes are the key drivers both of the evolution of income concentration at the top of the distribution in Latin America, and of the increasing challenge of properly accounting for income inequality using official survey-based estimates. Yet, we still know very little about the accumulation and distribution of wealth. This article contributes toward filling this major knowledge gap.

The bar is set high by the rapidly expanding literature on wealth in the developed world. It is no longer enough to provide estimates of the wealth distribution among individuals or households. The challenge is to provide wealth distribution estimates based on a variety of methodological strategies and data sources and to be able to systematically compare them, while at the same time being able to make these estimates fully compatible with both national income distribution and aggregate national wealth.

The scarcity of reliable data is the single most important restriction for the analysis of wealth in almost every country. In the Uruguayan framework, as well as in most of the developing world, the complete absence of official balance sheet estimates poses an important information restriction. Wealth aggregates represent the key starting point of the capitalization method from a data viewpoint, as they are necessary for the estimation of the rates of return that ensure full micro-macro consistency. Therefore, this study has had to start from one step behind and estimate wealth-to-income ratios. Based on those results and a combination of tax micro-data, firm tax records, household surveys, and national accounts, wealth distribution is estimated. Results are compared with orthogonal data sources and methods, such as a household wealth survey, the real estate wealth tax, the estate multiplier method, and the *Forbes* billionaires list.

Results are still preliminary, and there is still room for significant improvement. They should be considered with caution at this stage. That being said, the results of this study do depict an effort to provide micro-macro consistent estimates of wealth accumulation and distribution. Results show extreme concentration of wealth in the top 1%, which is informative given Uruguay's position as one of the least unequal countries of Latin America.

A. Supplementary Tables and Figures

A.1. Supplementary Tables

Table A.1: International wealth distribution comparison, 2016

	Top 1	Top 10	Middle 40	Bottom 50
South Africa	54%	87%	16%	-3%
Uruguay	39%	77%	20%	3%
USA	36%	73%	26%	1%
Russia	46%	73%	23%	3%
India	32%	64%	30%	6%
Korea	25%	58%	36%	6%
France	25%	59%	36%	5%
United Kingdom	21%	57%	39%	4%
China	30%	67%	26%	6%

Note. Source Wid.World.

Table A.2: Survey adjustment based on household wealth aggregates, 2013

Panel (a): wealth shares		
	Survey	Adj. Surv.
Bottom 50	3,3%	1,3%
Middle 40	36,7%	19,0%
Top 10	59,9%	79,6%
Top 10	27,1%	48,0%
Top 0,1	10,8%	23,4%
Panel (b): scaling factors		
Housing	0,86	
Business	3,20	
Fin. Assets	8,82	
Liabilities	1,97	

Note. Own elaboration based on EFHU. Survey adjustment in Panel (a) based survey/aggregate scaling factors depicted in Panel (b).

Table A.3: Average wealth by fractile and asset

		Housing	Business	Fin. Ass.	Liab.	G. Wealth	Net wealth
2009	Bottom 50%	2	0	1	1	3	2
	Middle 40%	10	3	3	0	18	17
	Top 10%	44	82	174	2	265	264
	Top 1%	143	416	1071	4	1333	1329
	Total	8	8	18	1	35	34
2010	Bottom 50%	2	0	1	2	4	2
	Middle 40%	13	8	5	0	25	23
	Top 10%	54	104	265	3	366	363
	Top 1%	179	489	1509	6	1743	1737
	Total	11	11	27	2	48	47
2011	Bottom 50%	3	0	2	2	6	3
	Middle 40%	19	4	6	0	34	32
	Top 10%	70	149	411	3	522	519
	Top 1%	217	740	2531	8	2726	2718
	Total	16	15	38	2	68	66
2012	Bottom 50%	4	0	2	3	8	5
	Middle 40%	24	9	11	0	43	40
	Top 10%	81	176	397	4	588	584
	Top 1%	252	834	2449	9	2952	2944
	Total	19	18	43	3	79	77
2013	Bottom 50%	6	0	3	3	9	6
	Middle 40%	28	0	10	3	49	45
	Top 10%	90	200	476	5	697	693
	Top 1%	309	1033	2847	11	3381	3371
	Total	21	20	51	3	93	89
2014	Bottom 50%	5	0	3	3	9	5
	Middle 40%	25	2	8	0	45	42
	Top 10%	85	192	445	5	660	654
	Top 1%	283	991	2603	12	3183	3172
	Total	21	19	48	4	88	84
2015	Bottom 50%	5	0	2	3	8	4
	Middle 40%	23	0	6	0	39	36
	Top 10%	80	171	416	5	592	587
	Top 1%	301	933	2449	11	2983	2973
	Total	20	16	43	3	79	75
2016	Bottom 50%	5	0	2	3	7	4
	Middle 40%	24	1	6	0	37	35
	Top 10%	87	153	363	4	522	518
	Top 1%	310	866	2188	10	2668	2659
	Total	19	15	37	3	71	68

Note. Own elaboration based on capitalization method estimates.

Table A.4: Wealth thresholds by fractiles and asset

		Housing	Business	Fin. Ass.	Liab.	G. Wealth	Net wealth
2009	Top 1%	63	153	317	4	435	431
	Top 10%	19	7	5	2	59	57
	Top 50%	5	0	1	2	7	6
2010	Top 1%	75	193	516	6	610	604
	Top 10%	25	16	7	2	79	77
	Top 50%	7	0	2	2	10	9
2011	Top 1%	97	267	630	8	776	768
	Top 10%	35	9	9	3	110	107
	Top 50%	10	0	3	3	14	12
2012	Top 1%	115	314	672	9	895	886
	Top 10%	43	19	18	3	133	130
	Top 50%	12	0	3	3	19	16
2013	Top 1%	123	379	970	11	1293	1281
	Top 10%	46	1	15	4	159	155
	Top 50%	15	0	4	3	23	20
2014	Top 1%	119	360	895	12	1223	1211
	Top 10%	43	5	13	4	150	146
	Top 50%	14	0	4	4	21	18
2015	Top 1%	109	305	777	11	1058	1047
	Top 10%	39	0	12	4	131	127
	Top 50%	13	0	4	4	19	16
2016	Top 1%	122	253	686	10	940	930
	Top 10%	41	3	12	4	117	113
	Top 50%	14	0	4	4	20	17

Note. Own elaboration based on capitalization method estimates.

Table A.5: % ownership & distribution by asset

		Housing	Business	Fin. Ass.	Liab.	G. Wealth	Net wealth
2009	% asset ¿ 0	75%	12%	100%	100%	100%	79%
	Gini index	0,50	0,65	0,94	0,05	0,82	0,81
2010	% asset ¿ 0	74%	13%	100%	100%	100%	79%
	Gini index	0,49	0,63	0,93	0,06	0,82	0,80
2011	% asset ¿ 0	74%	11%	100%	100%	100%	79%
	Gini index	0,47	0,63	0,93	0,05	0,82	0,81
2012	% asset ¿ 0	74%	12%	100%	100%	100%	80%
	Gini index	0,45	0,62	0,92	0,06	0,80	0,78
2013	% asset ¿ 0	77%	10%	100%	100%	100%	82%
	Gini index	0,45	0,63	0,92	0,05	0,81	0,80
2014	% asset ¿ 0	77%	10%	100%	100%	100%	82%
	Gini index	0,45	0,64	0,92	0,06	0,81	0,80
2015	% asset ¿ 0	76%	10%	100%	100%	100%	80%
	Gini index	0,44	0,66	0,92	0,06	0,81	0,79
2016	% asset ¿ 0	74%	10%	100%	100%	100%	77%
	Gini index	0,43	0,68	0,91	0,06	0,80	0,78

Note. Own elaboration based on capitalization method estimates.

Table A.6: Wealth thresholds, estate method

	Lower bound			Main estimates		
	Top 10%	Top 1%	Top 0,1%	Top 10%	Top 1%	Top 0,1%
			Urban			
2007	8	73	322	10	90	406
2008	12	114	582	14	144	762
2009	10	104	631	13	128	814
2010	14	151	890	17	175	1130
2011	18	191	1236	23	241	1557
2012	21	216	1108	25	267	1419
2013	23	247	1514	28	303	2104
2014	18	222	1273	22	271	1757
2015	14	214	1347	16	269	1751
			Rural			
2007	8	74	324	10	98	402
2008	12	114	575	14	137	730
2009	11	103	630	14	127	781
2010	14	153	891	18	192	1230
2011	18	191	1199	23	246	1547
2012	21	216	1185	26	261	1540
2013	23	249	1507	30	318	1912
2014	18	223	1202	23	283	1529
2015	14	214	1438	18	266	1798

Note. Real estate wealth thresholds, based on estate multiplier method (section 6). Thousands of US dollars, PPP.

Table A.7: Descriptive statistics, wealth survey

	Housing	Business	Fin. Assets	Liab.	Gross wealth	Net wealth
	Mean values (thousands of US dollars, PPP)					
Bottom 50%	3	-	-	-	4	3
Middle 40%	30	-	2	1	35	33
Top 10%	131	80	56	15	231	227
Top 1%	408	576	384	59	1.008	1.002
Total	26	7	6	2	39	37
	Thresholds (thousands of US dollars, PPP)					
Top 1%	226	165	76	33	390	383
Top 10%	62	0	7	3	78	75
Top 50%	13	0	0	0	15	14
	% ownership & distribution					
% asset ¿ 0	67%	8%	35%	43%	79%	76%
Gini index	0,52	0,81	0,85	0,75	0,66	0,66

Note. Based on EFHU. Per adult wealth.

Table A.8: Top 10 wealthiest individuals, 2016

Net wealth ranking								Net wealth	
2016	2015	2014	2013	2012	2011	2010	2009	2016	2022
1	1	4	1	0	0	0	0	2.132	1.998
2	2	3	2	2	2	1	1	1.182	1.108
3	0	0	0	0	0	0	0	545	510
4	3	0	0	0	0	0	0	288	270
5	0	0	0	0	0	0	0	202	189
6	0	0	0	0	0	0	0	198	186
7	0	0	0	0	0	0	0	178	166
8	30	17	6	0	32	0	0	167	156
9	0	141	15	0	0	0	5	164	154
10	29	38	0	0	0	0	0	148	139

Note. The first 8 columns depict the net wealth rank of each individual in each year, only for the top 10 individuals in the 2016 database. Zero indicates that they were not part of the top 0.01% in that year. Net wealth for 2016 and 2022 is expressed in millions of USD. In the case of 2022, values in Uruguayan pesos were first converted to 2022 prices using the consumer price index.

Table A.9: Method comparison

	Capitalization method						Survey						Estate method				Tax	
	Bus.	Hous.	F. Ass.	Liab.	Gross W.	Net W.	Bus.	Hous.	F. Ass.	Liab.	Gross W.	Net W.	R. est.	R. e. LB	Urban	Urban LB	Hous.	Net W.
Top 0,1%																		
2007													8%	6%	5%	4%		
2008													10%	8%	6%	4%		
2009	19%	5%	30%	0%	30%	18%							10%	8%	3%	2%		
2010	17%	5%	26%	0%	26%	17%							11%	9%	3%	2%	4%	1%
2011	19%	5%	35%	0%	35%	23%							10%	8%	3%	2%	5%	1%
2012	18%	5%	31%	0%	31%	19%							8%	6%	3%	2%	4%	1%
2013	19%	5%	24%	0%	24%	16%	28%	3%	35%	5%	11%	11%	10%	8%	2%	2%	5%	1%
2014	19%	5%	23%	0%	23%	15%							9%	7%	3%	2%		
2015	22%	6%	25%	0%	25%	17%							12%	9%	3%	2%	5%	1%
2016	25%	5%	26%	0%	26%	17%												
Top 1%																		
2007													25%	20%	20%	15%		
2008													29%	22%	19%	15%		
2009	50%	13%	59%	59%	38%	40%							28%	22%	17%	13%		
2010	46%	13%	53%	53%	36%	37%							31%	24%	15%	11%		
2011	50%	13%	58%	58%	40%	41%							28%	21%	16%	12%		
2012	47%	12%	56%	56%	37%	38%							24%	19%	16%	12%		
2013	51%	11%	54%	54%	37%	38%	70%	15%	62%	33%	26%	27%	27%	21%	14%	11%		
2014	51%	13%	54%	54%	36%	38%							25%	19%	15%	11%		
2015	54%	14%	55%	55%	38%	39%							29%	22%	16%	13%		
2016	56%	12%	55%	55%	38%	39%												
Top 10%																		
2007													55%	43%	57%	44%		
2008													59%	45%	58%	45%		
2009	99%	40%	95%	14%	76%	79%							53%	41%	50%	39%		
2010	98%	42%	94%	14%	76%	78%							58%	45%	49%	38%		
2011	100%	41%	94%	14%	77%	79%							53%	41%	50%	38%		
2012	99%	40%	93%	14%	74%	76%							48%	37%	47%	36%		
2013	100%	38%	93%	14%	75%	78%	100%	49%	91%	82%	59%	61%	52%	40%	47%	36%		
2014	100%	40%	93%	14%	75%	77%							47%	36%	44%	34%		
2015	100%	41%	93%	14%	75%	78%							51%	39%	45%	34%		
2016	100%	34%	93%	14%	74%	77%												

Note. Each column depicts the top shares of different assets/methods, based on own asset fractiles.

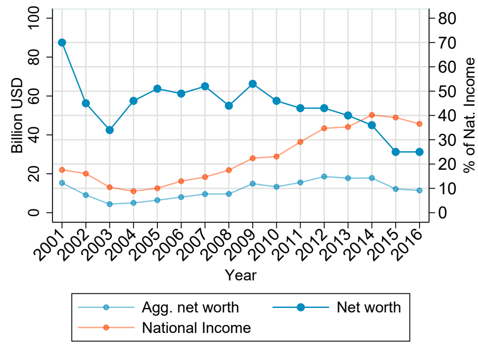
Table A.10: Proportion of women (in %) by age group.

Age group	2009	2010	2011	2012	2013	2014
20-24	54.42	56.13	55.01	44.82	41.53	44.98
25-29	41.96	44.07	43.93	46.08	44.54	44.70
30-34	43.55	45.42	45.01	47.32	45.61	46.35
35-39	44.50	46.33	45.81	48.17	46.00	47.33
40-44	45.11	46.89	46.41	48.96	46.72	48.51
45-49	46.41	47.45	46.97	49.83	47.48	49.23
50-54	47.10	48.12	47.74	50.11	48.15	49.78
55-59	48.18	48.79	48.43	50.90	48.89	50.17
60-64	49.80	49.95	49.80	52.31	50.30	51.05
65-69	51.67	51.91	51.74	53.04	52.13	52.28
70-79	53.28	53.29	53.03	56.94	53.03	53.12
80+	57.63	57.39	57.14	66.77	56.75	56.56

Note. Own elaboration based on tax income records (DGI), 2012.

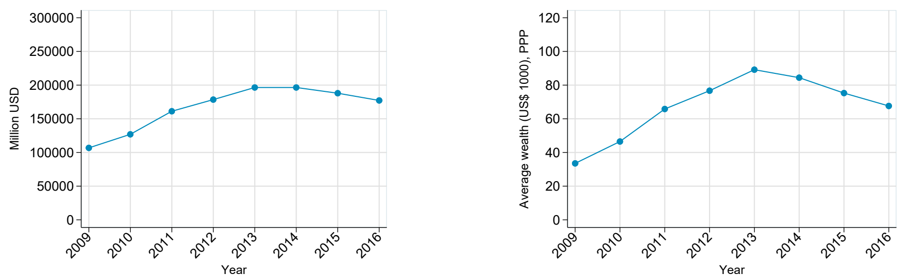
A.2. Supplementary Figures

Fig. A.1. Government sector aggregate net worth, 2001-2016



Note. Based on IMF Data Warehouse (government balance sheet as share of current GDP) and World Bank's GDP and net adjusted national income series. Primary axis depicts national income and public net worth in US billions of dollars (for aggregate government net worth and national income), while secondary axis depicts government wealth-to-income ratio.

Fig. A.2. Net wealth, 2009-2016

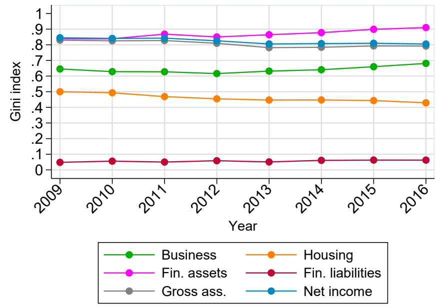


(a) Aggregate net wealth

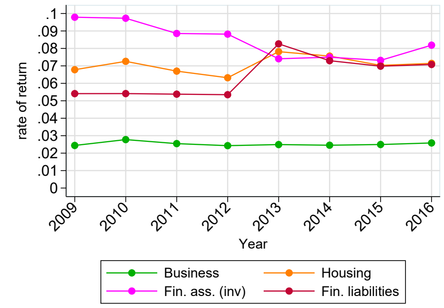
(b) Per-adult net wealth

Note. Panel (a): Net national wealth in millions of US dollars (PPP). Panel (b): Net per adult wealth, thousands of US dollars, PPP.

Fig. A.3. Net Personal capital incomes distribution and return rates 2009-2016



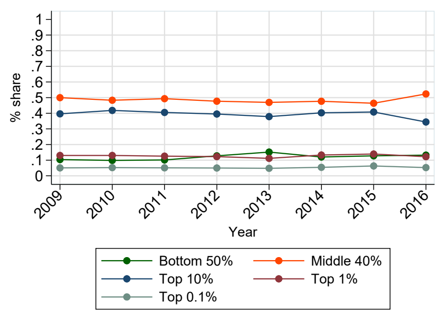
(a) Gini index



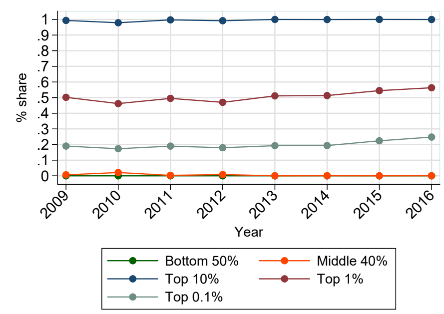
(b) Rates of return

Note. Capital income distribution based on De Rosa and Vilá (2022). Includes dividends, interest, rents, owner-occupied housing rents, and undistributed profits. Personal capital income shares depicted in Figure C.1.

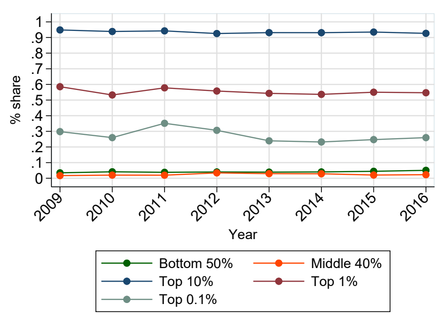
Fig. A.4. Wealth distribution by asset type, 2009-2016



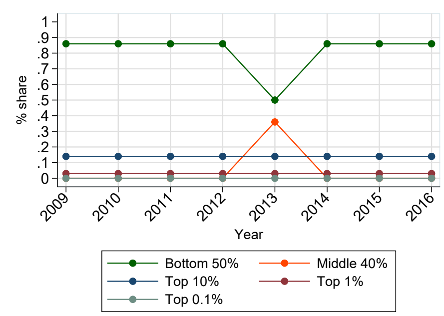
(a) Housing



(b) Business



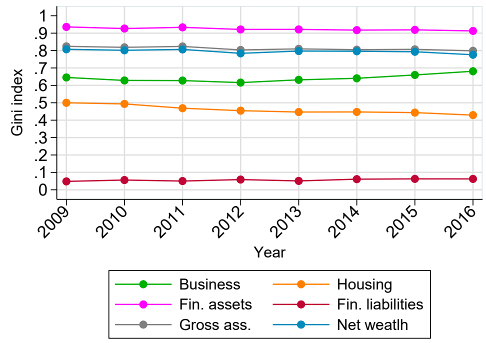
(c) Financial assets



(d) Liabilities

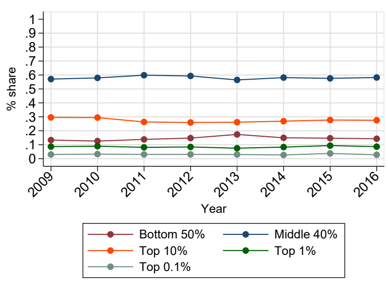
Note. Wealth shares constructed based on each asset/liability. The distribution of each asset according to total private net wealth fractiles is depicted in Figure A.6, while Gini indices by asset type are presented in Figure A.7. Figure A.10 depicts wealth portfolio by wealth group.

Fig. A.5. Net Personal wealth distribution - Gini index, 2009-2016

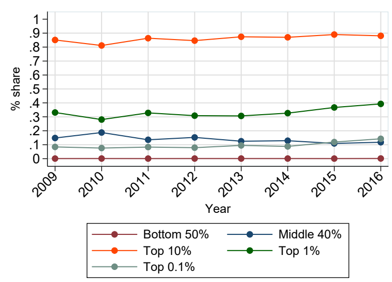


Note. Gini index of the four main wealth components. Wealth shares depicted in Figure A.6.

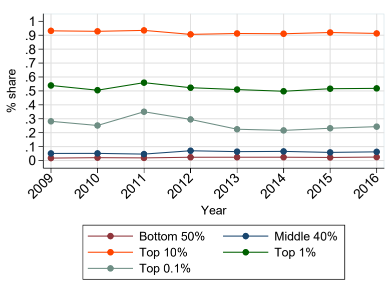
Fig. A.6. Wealth distribution by asset type in shares of total wealth, 2009-2016



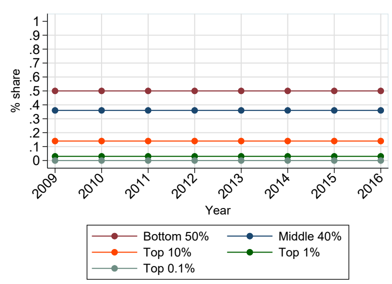
(a) Housing



(b) Business



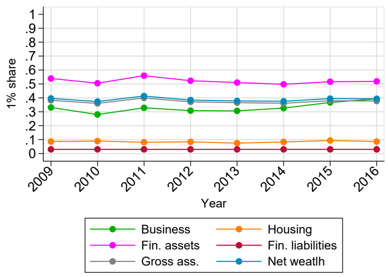
(c) Financial assets



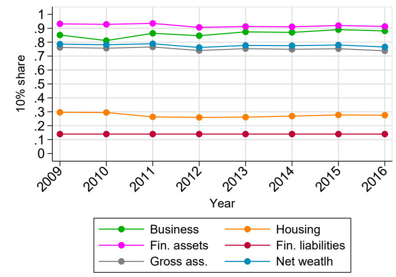
(d) Liabilities

Note. Figure A.10 depicts wealth portfolio by wealth group. Fractiles of total private net wealth.

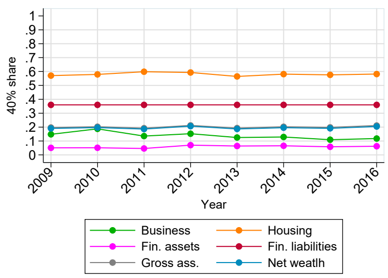
Fig. A.7. Personal wealth distribution by wealth groups, 2009-2016



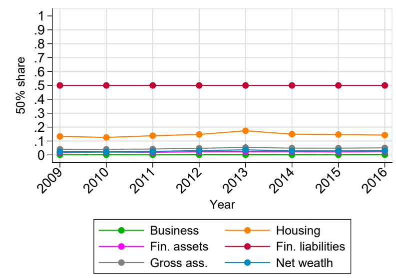
(a) Top 1%



(b) Top 10%



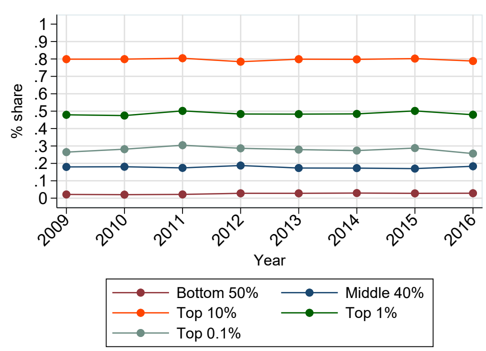
(c) Middle 40%



(d) Bottom 50%

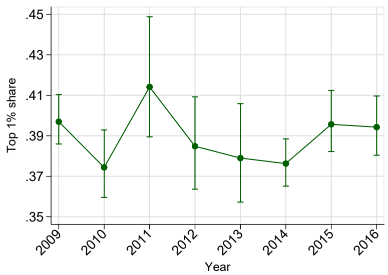
Note. Gini index by capital income types depicted in Figure C.1.

Fig. A.8. Net Personal wealth distribution (upper bound) 2009-2016

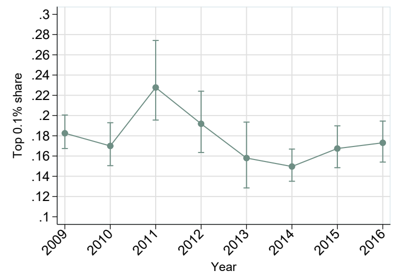


Note. Capitalized incomes, based on Distributional National Accounts estimates (De Rosa and Vilá, 2022), *alternative 2*. In this variant, undistributed profits are imputed based on matched firm-owners data, resulting in upper bound estimates.

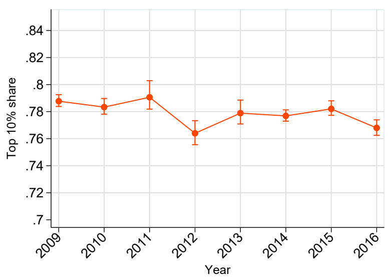
Fig. A.9. Net wealth distribution confidence intervals, 2009-2016



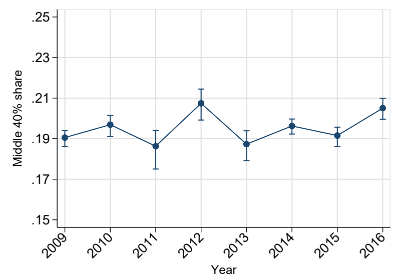
(a) Top 1%



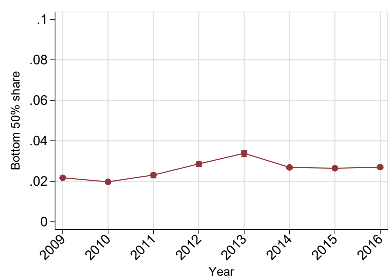
(b) Top 0.1%



(c) Top 10%



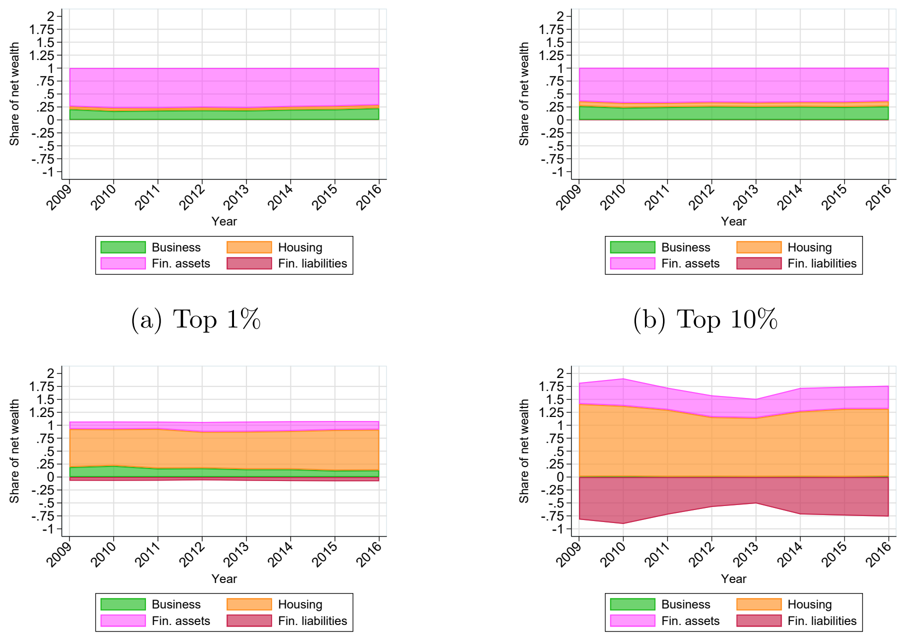
(d) Middle 40%



(e) Bottom 50%

Note. 5% confidence intervals estimated based on bootstrap (500 draws).

Fig. A.10. Personal wealth composition, 2009-2016



(a) Top 1%

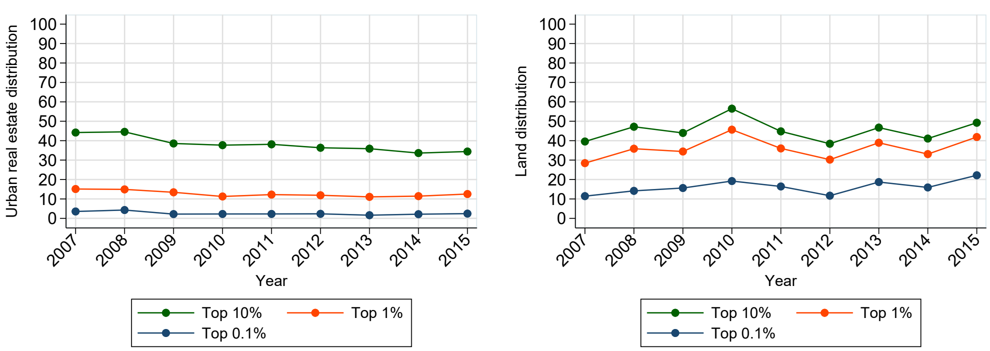
(b) Top 10%

(c) Middle 40%

(d) Bottom 50%

Note. Capital incomes composition depicted in Figure C.2.

Fig. A.11. Real estate distribution by wealth type (70-30 split)

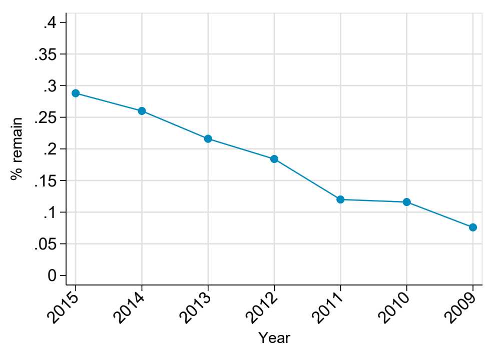


(a) Urban

(b) Rural

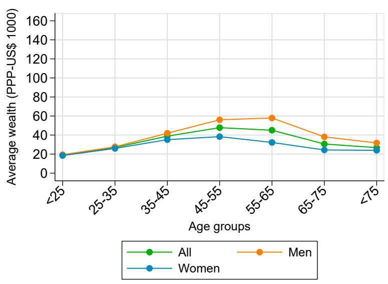
Note.Based on DGR and DNC merged data. Decedent wealth expanded based on average mortality rate and considering a decedent/adult population wealth ratio of 1.4, computed based on the wealth survey. Equal-split lower bound estimates.

Fig. A.12. Top 0.01% attrition

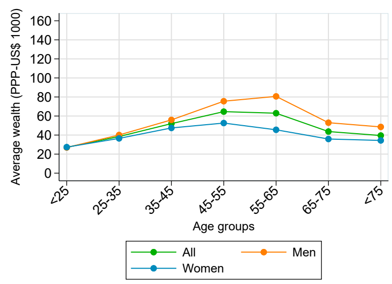


Note. Share of the 250 individuals (approximately the top 0.01%) who belonged to the same group in the previous years.

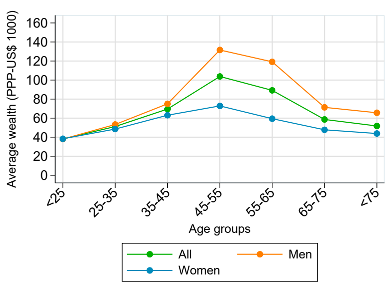
Fig. A.13. Average wealth by sex and age, 2009-2016.



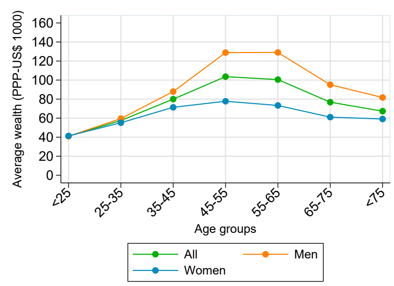
(a) 2009



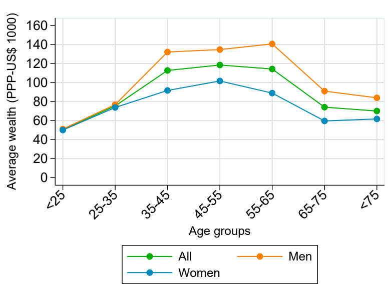
(b) 2010



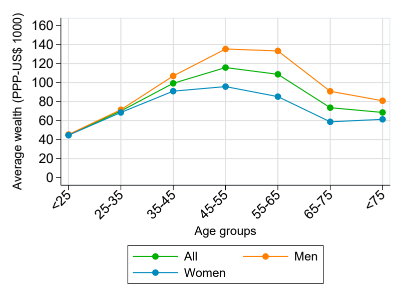
(c) 2011



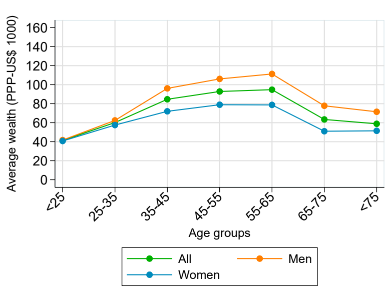
(d) 2012



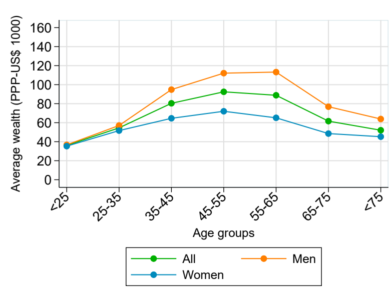
(e) 2013



(f) 2014



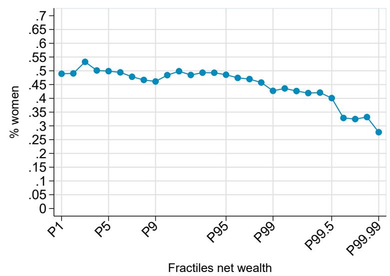
(g) 2015



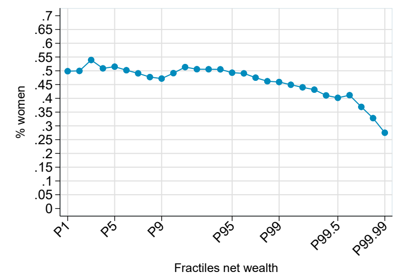
(h) 2016

Note. Wealth averages for ten-year age groups and sex. Individuals over 20 years old.

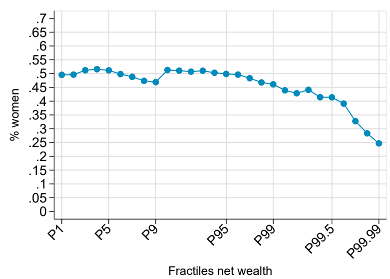
Fig. A.14. Share of women by wealth fractile, 2009-2016.



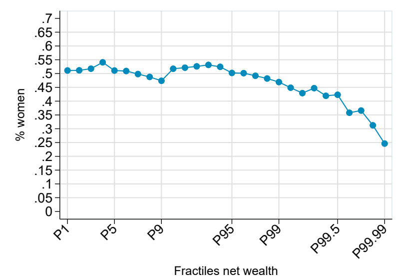
(a) 2009



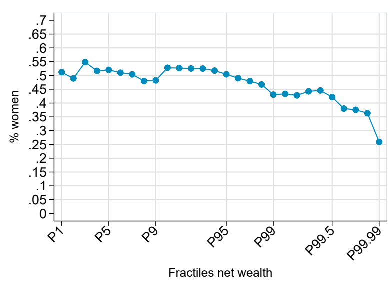
(b) 2010



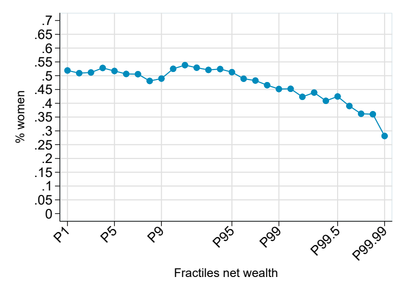
(c) 2011



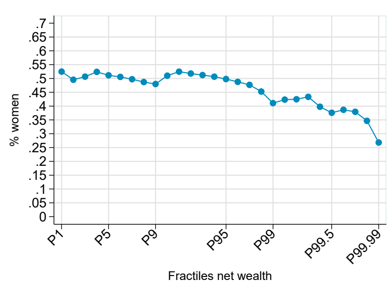
(d) 2012



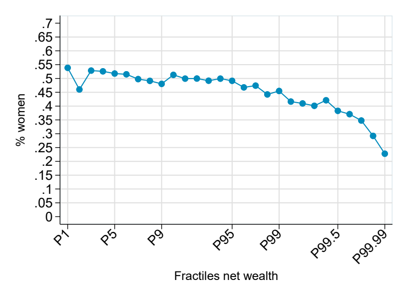
(e) 2013



(f) 2014



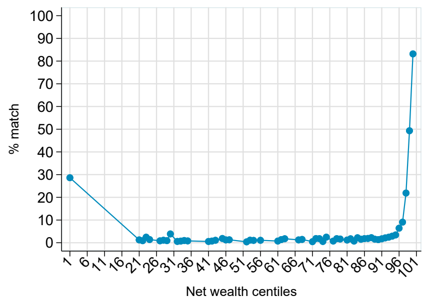
(g) 2015



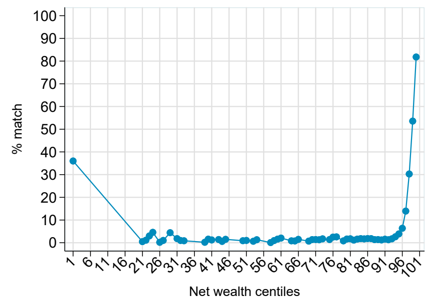
(h) 2016

Note. Percentage of women (20+ years) by net wealth fractile.

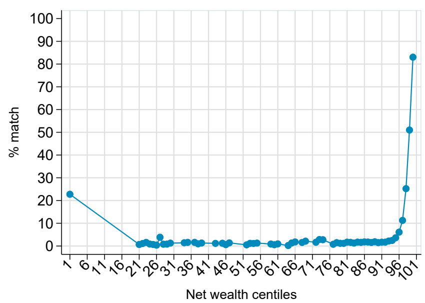
Fig. A.15. Percentage of income-wealth matching by wealth p-tile, 2009-2016



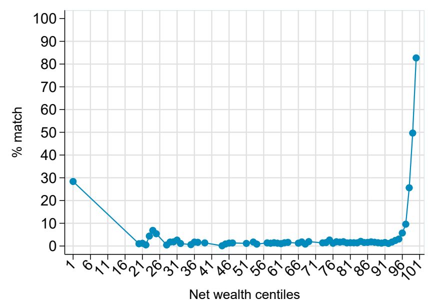
(a) 2009



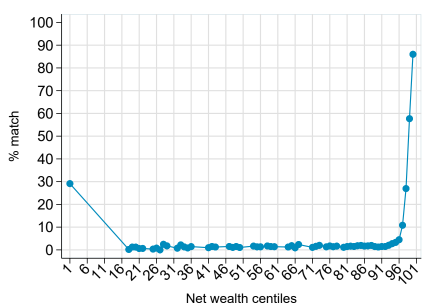
(b) 2010



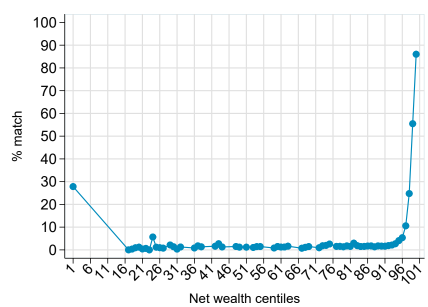
(c) 2011



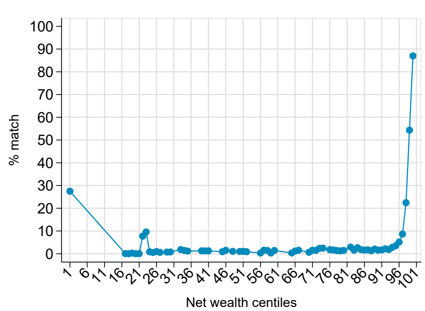
(d) 2012



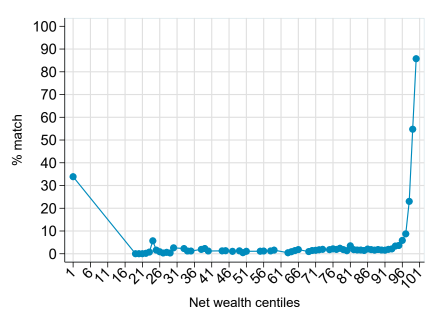
(e) 2013



(f) 2014



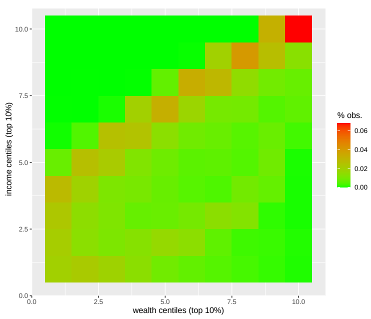
(g) 2015



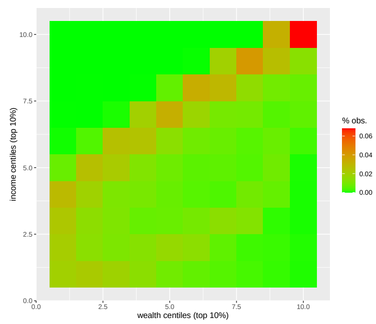
(h) 2016

Note. Income refers to total incomes (including capital incomes, labour incomes, pensions and other incomes).

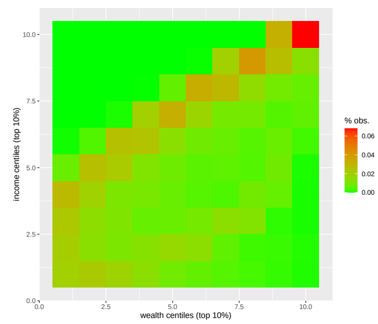
Fig. A.16. Income and wealth heatmap (top 10%), 2009-2016



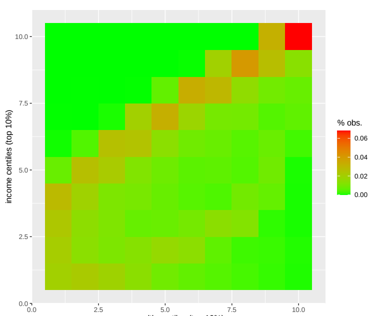
(a) 2009



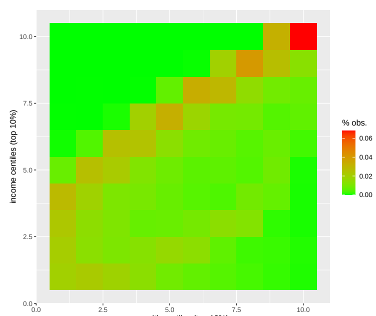
(b) 2010



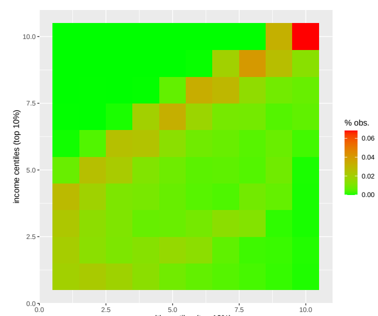
(c) 2011



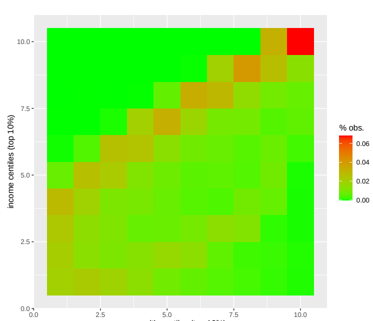
(d) 2012



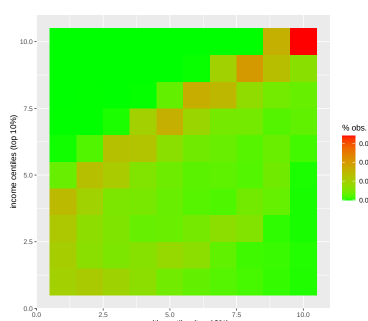
(e) 2013



(f) 2014



(g) 2015



(h) 2016

Note. Income refers to total incomes (including capital incomes, labour incomes, pensions and other incomes).

B. Data adjustments

Table B.1: Cadastre data set variables

Urban properties	Rural properties
Department (out of 19)	Department (out of 19)
Locality	Locality
Number	Number
Cadastral value (UY$)	Cadastral value (UY$)
Size (sq. mts. - building and terrain)	Size (sq.mts.)
Unit	
Block	

Note. Based on DNC data. Variables for the 1999-2018 period for 242.431 rural properties and 1.383.868 urban properties.

Table B.2: DGR decedents data-base

Year	Total decedents	Av. Num. Properties	Med. Num. Properties	Max. Num. Properties
2007	8.736	9,7	1	471
2008	8.107	4,0	1	154
2009	8.210	3,2	1	92
2010	8.342	3,8	1	94
2011	8.116	3,8	1	94
2012	7.638	4,8	1	221
2013	7.059	3,1	1	92
2014	5.959	3,6	1	99
2015	5.522	3,2	1	92

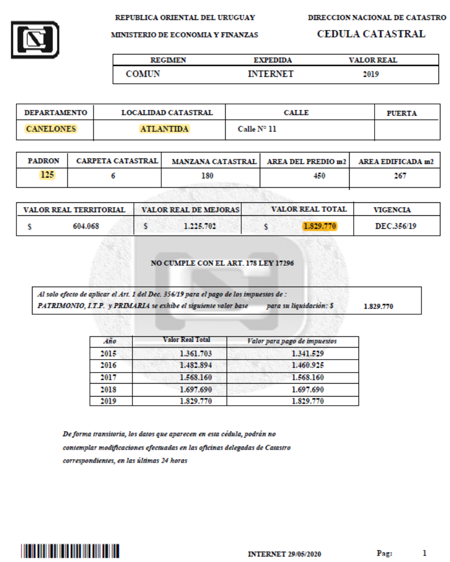
Note. Based on DGR decedents data.

Table B.3: Wealth tax - Income tax matching, 2009-2014

		Dividends - wealth tax		Rents - wealth tax	
	Num. tax-payers	100%: dividends	100%: wealth tax	100%: rents	100%: wealth tax
2009	8058	6%	1%	4%	24%
2010	8592	6%	2%	4%	24%
2011	8452	6%	2%	4%	24%
2012	8433	9%	4%	5%	36%
2014	8710	6%	4%	4%	26%

Note. Own elaboration based on wealth and income tax matched micro-data. The second and fourth column depict the percentage of dividends/rents receivers who pay wealth tax, while third and fifth columns depict the percentage of wealth tax payers who receive dividends/rents. Second column shows total number of wealth tax payers.

Fig. B.1. Raw cadastre individual data



REPUBLICA ORIENTAL DEL URUGUAY
MINISTERIO DE ECONOMIA Y FINANZAS

DIRECCION NACIONAL DE CATASTRO
CEDULA CATASTRAL

REGIMEN	EXPEDIDA	VALOR REAL
COMUN	INTERNET	2019

DEPARTAMENTO	LOCALIDAD CATASTRAL	CALLE	PUERTA
CANELONES	ATLANTIDA	Calle N° 11	

PADRON	CARPETA CATASTRAL	MANZANA CATASTRAL	AREA DEL PREDIO m2	AREA EDIFICADA m2
125	6	180	450	267

VALOR REAL TERRITORIAL	VALOR REAL DE MEJORAS	VALOR REAL TOTAL	VIGENCIA
$ 604.068	$ 1.225.702	$ 1.829.770	DEC.356/19

NO CUMPLE CON EL ART. 178 LEY 17296

Al solo efecto de aplicar el Art. 1 del Dec. 356/19 para el pago de los impuestos de : PATRIMONIO, I.T.P. y PRIMARIA se exhibe el siguiente valor base para su liquidación: $ 1.829.770

Año	Valor Real Total	*Valor para pago de impuestos*
2015	1.361.703	1.341.529
2016	1.482.894	1.460.925
2017	1.568.160	1.568.160
2018	1.697.690	1.697.690
2019	1.829.770	1.829.770

De forma transitoria, los datos que aparecen en esta cédula, podrán no contemplar modificaciones efectuadas en las oficinas delegadas de Catastro correspondientes, en las últimas 24 horas

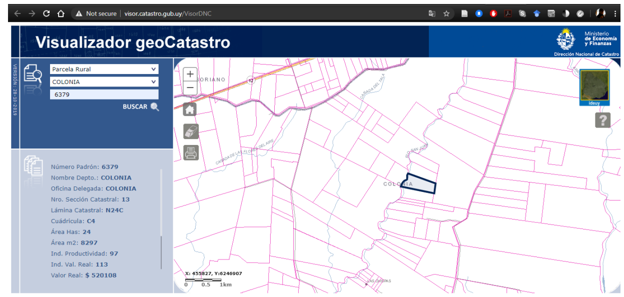
6116360

INTERNET 29/05/2020 Pag: 1

Note. Screenshot from *geoCatastro*, MEF. Raw data of *Cédula Catastral*.
http://visor.catastro.gub.uy/visordnc/. Data used for estimates (see Table B.1) is highlighted.

Fig. B.2. Cadastre property identification: example by type

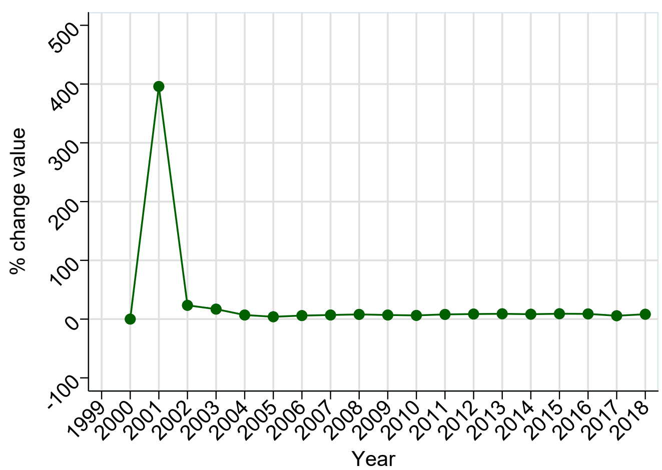
(a) Urban



(b) Rural

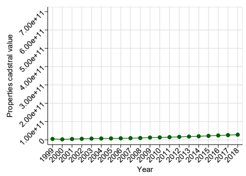
Note. Screenshots from *geoCatastro*, MEF. Urban identifier: *Departamento, Localidad, Padrón, Unidad*. *Unidad* is not an identifier of the whole property, but distinguishes individual units of real estate within the Horizontal Property Regime (including apartment buildings). Problems still remain, e.g. “block” is lost as it is not present in DGR data. Rural identifier: *Departamento, Padrón*. Close to perfect identification.

Fig. B.3. Revaluation example

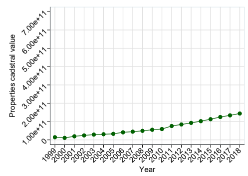


Note. Based on DNC data. The figure shows an example of a property that was revaluated in 2001, resulting in an increase of 400% of its cadastre value that year.

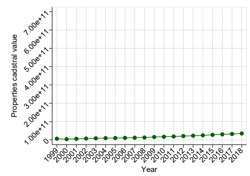
Fig. B.4. Cadastral aggregate value by department



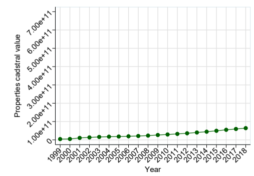
(a) Artigas



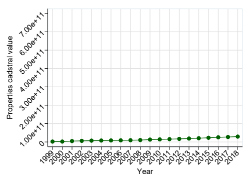
(b) Canelones



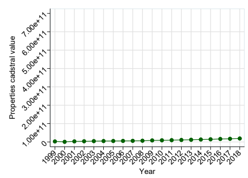
(c) Cerro Largo



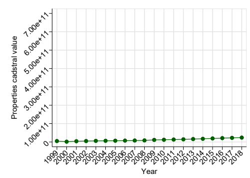
(d) Colonia



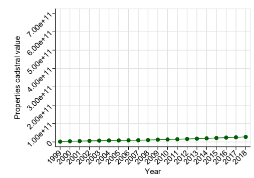
(e) Durazno



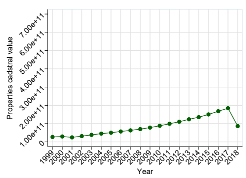
(f) Flores



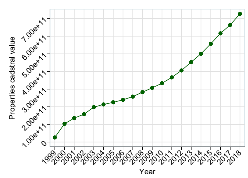
(g) Florida



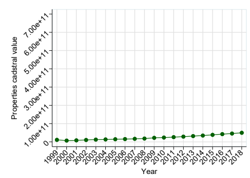
(h) Lavalleja



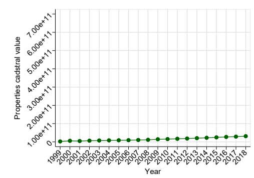
(i) Maldonado



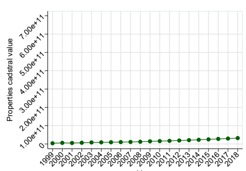
(j) Montevideo



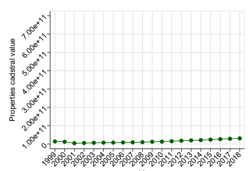
(k) Paysandú



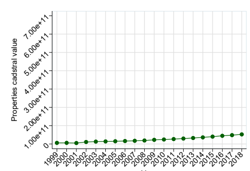
(l) Río Negro



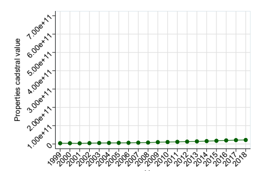
(m) Rivera



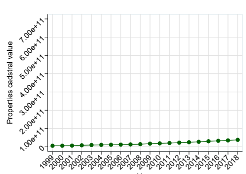
(n) Rocha



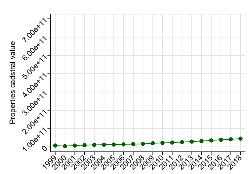
(o) Salto



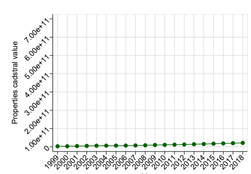
(p) San José



(q) Tacuarembó

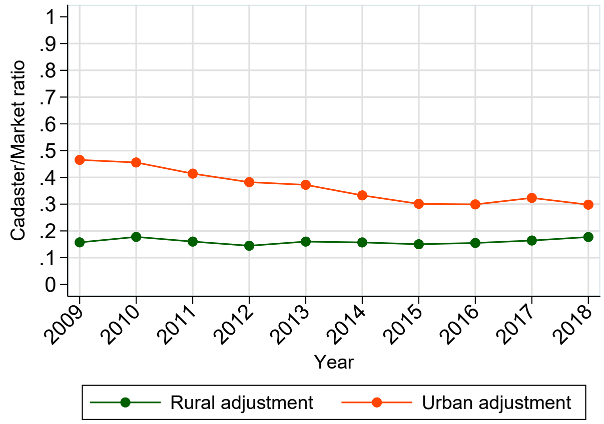


(r) Soriano



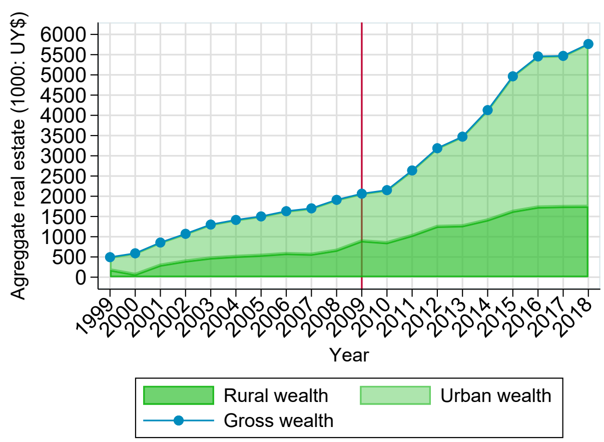
(s) Treinta y Tres

Fig. B.5. Market price adjustments, 2009-2018



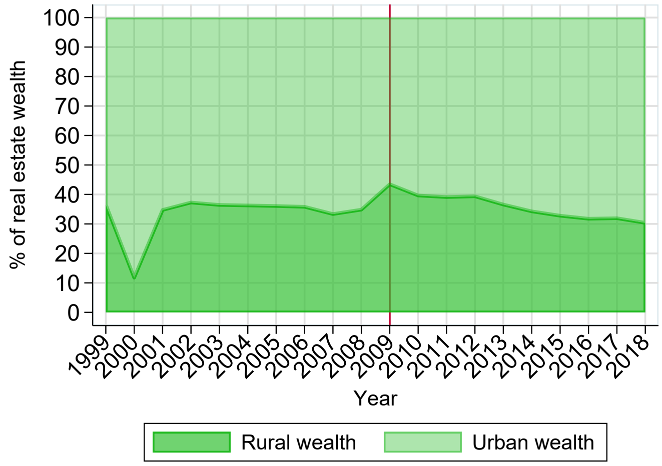
Note. Price adjustment based DNC, DGR, and DIEA data. The series depicts the ratio of urban and rural aggregate cadastre value and market prices. Rural properties adjusted by department.

Fig. B.6. Aggregate gross real estate wealth



Note. Based on cadastre data from the DNC, corrected by the market price adjustment depicted in Figure B.5. Household gross real estate share in national gross real estate is approximately 55%, based on the household wealth survey (EFHU).

Fig. B.7. Urban-rural real estate shares



Note. Based on DNC, DIEA, and DGR. The figure depicts the distribution of real estate properties owned by all institutional sectors.

Fig. B.8. Opening of inheritance process

10 *Avisos* Nº 30.434 - mayo 12 de 2020 | Diario*Oficial*

Avisos

Apertura de Sucesiones AS

PODER JUDICIAL

(Ley 16.044 Arts. 3o., 4o. y 5o.)

Los señores Jueces Letrados de Familia han dispuesto la apertura de las Sucesiones que se enuncian seguidamente y citan y emplazan a los herederos, acreedores y demás interesados en ellas, para que, dentro del término de TREINTA DÍAS, comparezcan a deducir en forma sus derechos ante la Sede correspondiente.

MONTEVIDEO

JUZGADO LETRADO DE PRIMERA INSTANCIA DE FAMILIA

PRIMER TURNO

JORGE MARIO CHAPITAL DURÁN (IUE 2-68329/2019).
Montevideo, 26 de febrero de 2020.
CAROL ANALIA ROMERO CARDOZO, Actuario Adjunto.
01) $ 6631 10/p 6883 May 11- May 25

OLGA ESTHER PAOLINI MANCIERI (IUE 2-19253/2019)
Montevideo, 29 de agosto de 2019
NANCY BEATRIZ CASTELLANO LABORDE, Actuario Adjunto.
01) $ 6631 10/p 6825 May 08- May 22

TERCER TURNO

ANIBAL MARIO SCAGNI TEJERA y EPIFANIA TEODORA ALVAREZ UMPIERREZ (IUE 2-53045/2019)
Montevideo, 16 de octubre de 2019
PATRICIA GONÇALVES FLORES, Actuario Adjunto.
01) $ 6631 10/p 6840 May 11- May 25

TANIA MABELCHOCHO BALDUINI (IUE 2-6606/2020).
Montevideo, 16 de marzo de 2020.
PATRICIA GONÇALVES FLORES, Actuario.
01) $ 6631 10/p 6789 May 08- May 22

CUARTO TURNO

ALBA GLORIA GUTIERREZ CAMACHO (IUE 2-6482/2020).
Montevideo, 10 de marzo de 2020.
NATALIA VERÓNICA MELONI MACHADO, Actuario Adjunto.
01) $ 6631 10/p 6822 May 08- May 22

MARIA ROSA GANDOLFO PEREZ (IUE 2-7098/2020).
Montevideo, 10 de marzo de 2020.
NATALIA VERÓNICA MELONI MACHADO, Actuario Adjunto.
01) $ 6631 10/p 6786 May 08- May 22

MARIA DEL CARMEN VIEITES SONEIRA (IUE 2-4402/2020).
Montevideo, 09 de marzo de 2020.
NATALIA VERÓNICA MELONI MACHADO, Actuario Adjunto.
01) $ 6631 10/p 6703 May 05- May 19

QUINTO TURNO

OCTAVO TURNO

PABLO ADEMAR MILANO DANIEL (IUE 2-1564/2020).
Montevideo, 10 de febrero de 2020.
CAROL ANALÍA ROMERO CARDOZO.
01) $ 6631 10/p 6714 May 06- May 20

NOVENO TURNO

MIGUEL ANGEL GARCILAZO (IUE 2-22338/2018).
Montevideo, 09 de marzo de 2020.
MARIA VICTORIA GIOVANELLI FERREIRA, Actuario Adjunto.
01) $ 6631 10/p 6877 May 11- May 25

BLANCA NIEVES BRITOS (IUE 2-3276/2020).
Montevideo, 13 de febrero de 2020.
MARIA VICTORIA GIOVANELLI FERREIRA.
01) $ 6631 10/p 6850 May 11- May 25

AGUSTINA DEL CARMEN IGLESIAS GONZÁLEZ (IUE 2-5001/2020).
Montevideo, 21 de febrero de 2020.
MARIA VICTORIA GIOVANELLI FERREIRA, Actuario Adjunto.
01) $ 6631 10/p 6602 Abr 30- May 14

DECIMOPRIMER TURNO

TERESITA RAQUEL SANTER LUZARDO (IUE 2-73037/2019).
Montevideo, 10 de febrero de 2020.
MARÍA de LOURDES BAPTISTA CASULLO.
Primera Publicación
01) $ 6631 10/p 6897 May 12- May 26

TERESA PERRONE PALERMO (IUE 2-73026/2019).
Montevideo, 10 de febrero de 2020.
MARÍA de LOURDES BAPTISTA CASULLO.
01) $ 6631 10/p 6817 May 08- May 22

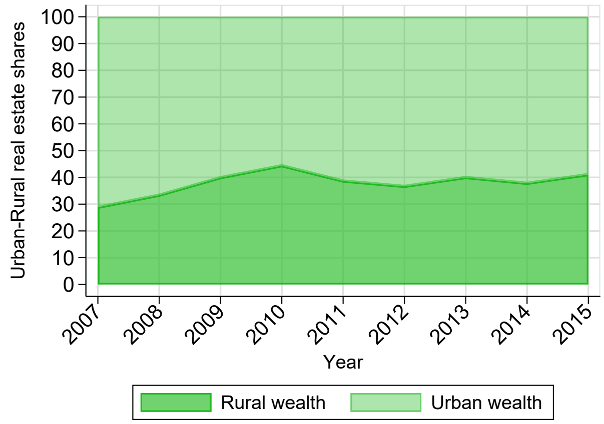
RUTH AZEREDO SILVA (IUE 2-74706/2019).
Montevideo, 10 de febrero de 2020.
MARÍA de LOURDES BAPTISTA CASULLO.
Última Publicación
01) $ 6631 10/p 6522 Abr 28- May 12

DECIMOSEGUNDO TURNO

LUIS GUARINO VARELA (IUE 2-5924/2020).
Montevideo, 02 de marzo de 2020.

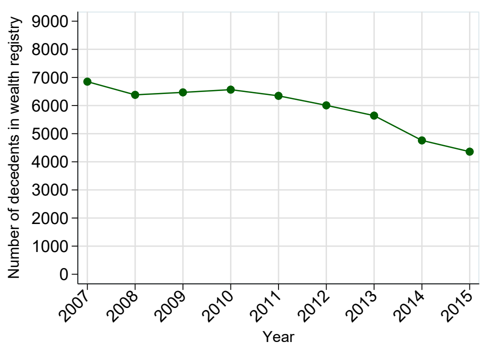
Note. Screenshot from *Diario Oficial*, IMPO. https://www.impo.com.uy/

Fig. B.9. Cadastre-decedent's urban & rural wealth

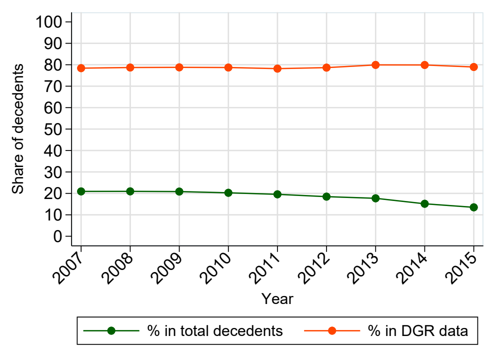


Note. Based on cadastre data from DNC and merged dataset. All values adjusted to market prices.

Fig. B.10. Decedent population in merged data, 2007-2015



(a) Number of owner decedents.

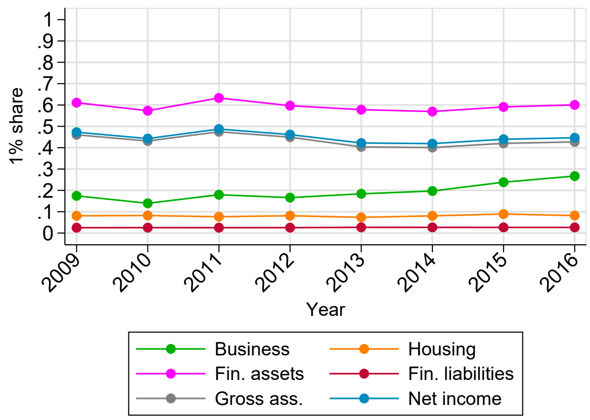


(b) % of owner decedents in total decedent population.

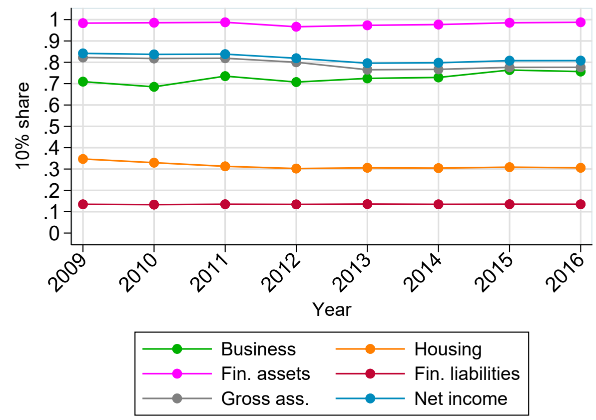
Note. Based on DGR, DNC, and INE's total decedents number. Panel (a) shows the absolute number of individuals in the decedent-real estate merged dataset. Panel (b) depicts the share of individuals in the merged decedent-real estate dataset in total decedents and in the total number of decedents in DGR raw data.

C. Capital incomes

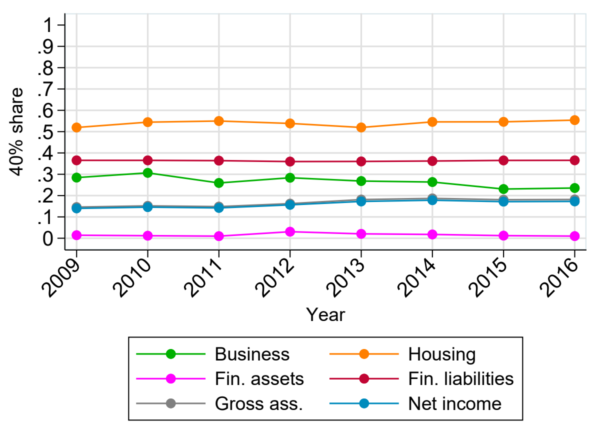
Fig. C.1. Personal capital income shares 2009-2016



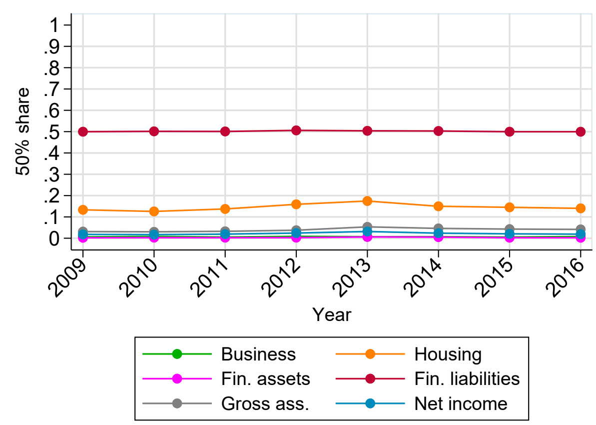
(a) Top 1%



(b) Top 10%



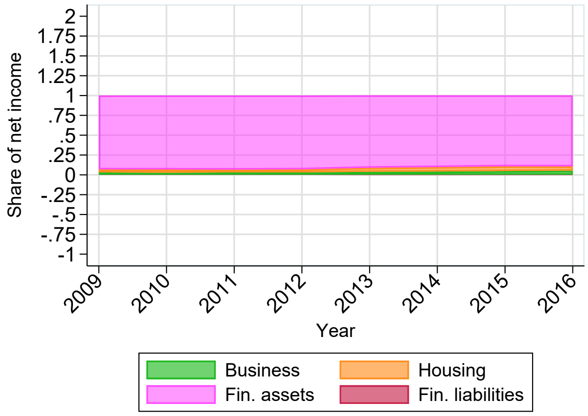
(c) Middle 40%



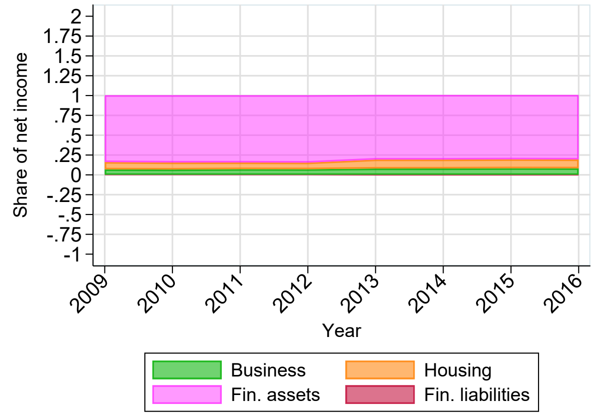
(d) Bottom 50%

Note.Capital income distribution based on De Rosa and Vilá (2022). Includes dividends, interest, rents, owner-occupied housing rents, and undistributed profits.

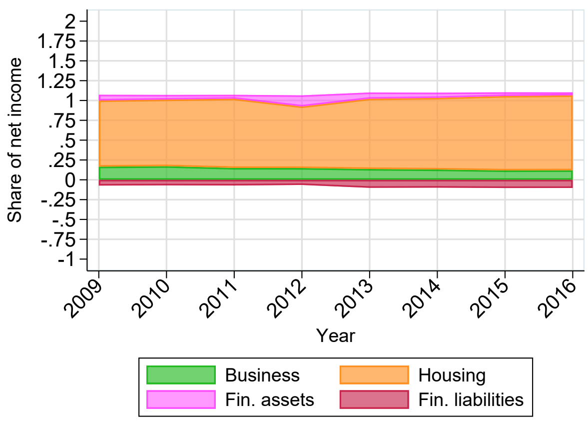
Fig. C.2. Personal capital incomes composition 2009-2016



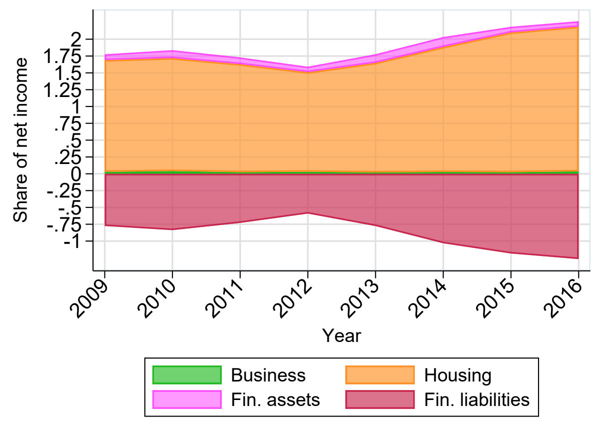
(a) Top 1%



(b) Top 10%



(c) Middle 40%



(d) Bottom 50%

Note. Capital income distribution based on De Rosa and Vilá (2022). Includes dividends, interest, rents, owner-occupied housing rents, and undistributed profits. Capital incomes composition by income group depicted in Figure A.10.

D. National wealth estimation

For the estimation of national and private wealth, I proceed in four steps. The aggregates that are unknown are in bold text, while the underlined ones represent those that can be observed or easily estimated. Steps 1 and 2 can be calculated directly from equation 2.1, while steps 3a and 3b are the different components of each sector's wealth. In the following subsections, I will describe the estimation of all the observable elements, i.e. those underlined, which will lead to the estimates presented in Section 5.1.

The four steps in the estimation procedure are:

1. **Net National Wealth** = <u>Domestic Capital</u> + <u>Net Foreign Asset Position</u>
2. Net National Wealth = **Private net wealth** + <u>Government net wealth</u>
3. Household and corporate wealth
    (a) Private net wealth = <u>Priv. non fin. assets</u> + **Priv. fin. assets** - <u>Priv. liabilites</u>
    (b) Corp. net wealth = 0 = <u>Corp. non fin. assets</u> + **Corp. fin. assets** - <u>Corp. liabilites</u>

D.1. The government sector

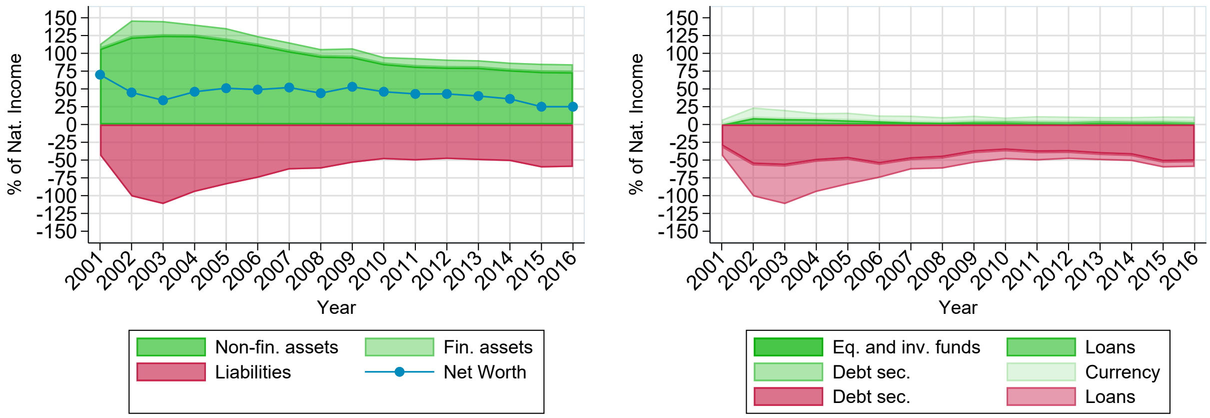
The only institutional sector with a complete balance sheet is the government sector; this balance sheet is reported annually by the Uruguayan government to the International Monetary Fund (IMF), available for 2001-2016.[25] The IMF provides data on the general government sector's net worth (S13-B90) as a percentage of GDP. Based on this series, in Figure A.1, aggregated public net worth is depicted together with the evolution of net national income (both in current USD), as well as the ratio between the two.

Following SNA's guidelines (United Nations, 2008b), the government sector's net worth consists of non-financial plus financial assets, net of liabilities. Figure D.1 (panel a) depicts these main components of government net worth as a percentage of net adjusted national income. During the early 2000s economic crisis, there was a large and rapid expansion of the government's liabilities, which had the effect of offsetting a milder increase in financial and non-financial assets. From 2003 until 2009, the ratio stabilised around 50%, slowly falling thereafter led by a decreasing share of non-financial assets and fostered by an increase in liabilities in the last few years.

Digging into non-financial assets (AN), IMF data shows that they are exclusively represented

[25] See `https://data.imf.org/`.

Fig. D.1. Government sector balance sheet, 2001-2016



(a) Balance sheet

(b) Financial balance sheet

Note. Based on IMF Data Warehouse (government balance sheet as a share of GDP) and World Bank's GDP and net adjusted national income series. Net worth is the result of financial and non-financial assets net of liabilities.

by produced non-financial assets and, in particular, by fixed assets (AN11). Non-produced non-financial assets (AN2) are missing, which is noteworthy since land is one of its components. Thus, it is not possible to present any further details about the components of non-financial assets, which is an important obstacle since they are one of the drivers of government net worth. Financial assets and liabilities (AF), on the other hand, offer greater opportunities for decomposition. Panel (b) of Figure D.1 shows that the overall dynamic of government net financial wealth is led mostly by its liabilities, which are between four and six times greater than assets. Within liabilities, debt securities (AF3) present the largest variations, and are responsible for the impressive spike of 2003-2005.

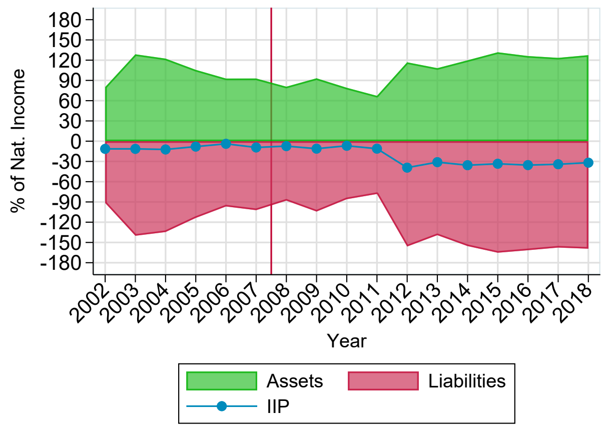
D.2. Net foreign asset position

Figure D.2 depicts the net international investment position (IIP)[26] of the country as well as its two main components, based on the balance of payments from the Central Bank of Uruguay (BCU) up until 2007, and from the IMF from 2008 onward. In 2011-2012 there is a sharp decrease in IIP, reaching around -30% of national income. This striking decline seems to be the result of a change in the way the series is constructed. The BCU series used to be presented in two sets: 2002-2011, and 2012 onward. This is what seems to be reported by the

[26]The terms international investment position and net foreign asset position will be used interchangeably throughout this study.

IMF, since the 2008-2018 IMF series matches the 'old' 2002-2011 BCU series exactly. Figure D.3 presents the variation of each of the four main components of both assets and liabilities.

Fig. D.2. International investment position (IIP), 2002-2018



Note. Based on IMF Data Warehouse (government balance sheet as share of current GDP), Uruguayan Central Bank and World Bank's GDP and net adjusted national income series. The vertical line depicts the limit between Central Bank and IMF data (left and right, respectively).

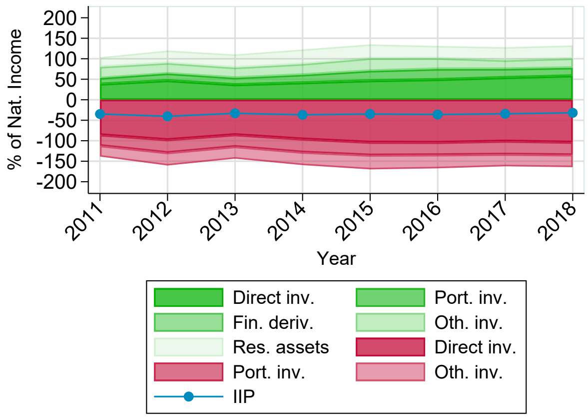
Thus, the IMF series does not seem to provide an adequate picture of the actual evolution of IIP. Unlike older version of the series, the new BCU IIP only covers the period from 2011 onward. However, it does not show a discontinuity for 2011-12, and it provides a complete decomposition, so it is a better starting point. Figure D.3 presents the IIP components both by financial instrument and by public-private split. Financial instruments include direct investment, financial derivatives, other investments, and portfolio investments, plus the country's main financial asset, which is central bank reserve assets. The main liability component is direct investment, which ranges from 85-110% of national income. When this is observed considering both the public and private sectors, it becomes clear that the negative IIP is entirely driven by private assets and liabilities, since the public sector's financial position in relation to the rest of the world is balanced or even slightly positive.

D.3. Domestic capital

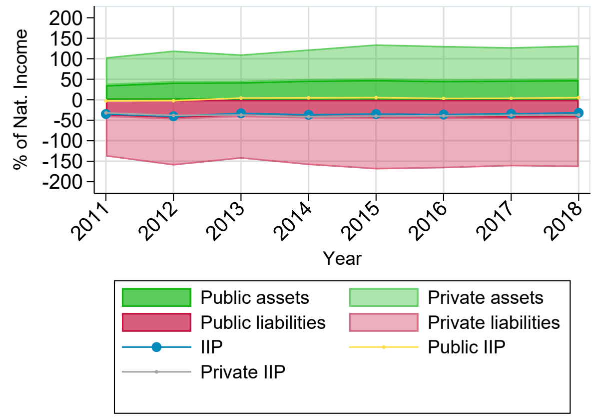
The level and composition of real estate wealth

Real estate wealth as a percentage of national income is depicted in panel (a) of Figure D.4, which results from adjusting net cadastral values to market prices as discussed in Section 4.2, including all real estate assets owned by all institutional sectors. After an initial increase in the gross real estate to income ratio in the early years of the century (which was the result of

Fig. D.3. International investment position (IIP) composition, 2011-2018.



(a) By financial instrument



(b) Public/Private

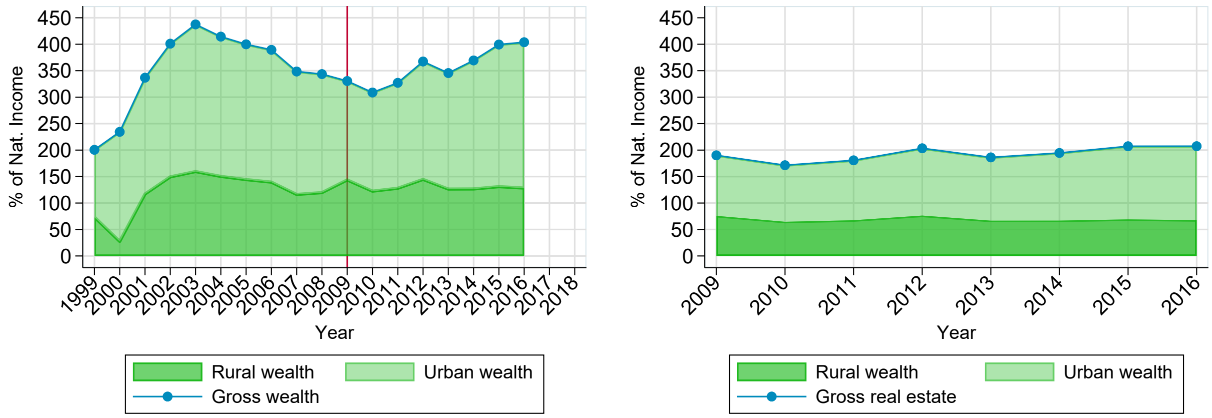
Note. Based on the central bank's (BCU) International Investment Position and World Bank's net adjusted national income series.

a massive contraction of national income of over 7% during the last major economic crisis), the ratio stabilizes around 3.5.

Panel (b) of Figure D.4 shows household gross housing and land. It is noteworthy that 30-40% of real estate wealth is rural real estate (see Figure B.7), with an urban real estate to income ratio of around 1.25 times national income, somewhat lower than what is found in rich countries. Rural real estate has an approximately equivalent magnitude to national income, which is significantly higher than estimates for the United States or Europe. Interestingly, these ratios are similar to the land to income ratios in Europe or the United States prior to World War I (Piketty and Zucman, 2014).

I compute the household sector's housing wealth based on the household wealth survey for 2013 relative to total gross urban properties, and I apply the share of real estate wealth that year to the whole series (therefore assuming the ratio is stable over time). This is further adjusted by the ratio of housing rents (owner occupied + rented dwellings) from tax-survey data to household operating surplus from national accounts (yielding a ratio very close to 1, see De Rosa and Vilá (2022)). To estimate household land ownership, I take rural real estate and adjust it by the share owned by individuals (as opposed to firms or the government) based on the 2011 agricultural census (this share is 52.8%, see DIEA (2014)). These adjustments result in agricultural land and housing owned by households equivalent to 60-70% and 110-130% of national income respectively.

Fig. D.4. Aggregate real estate wealth



(a) Real estate wealth as a % of NI. (b) National and household real estate wealth.

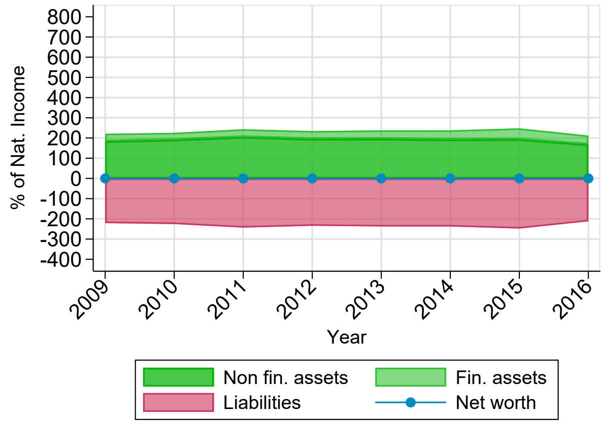
Note. Based on cadastre data from DNC, DGR, and DIEA. Market price adjustment depicted in Figure B.5 from 2009 onward (vertical red line), and 15-40% for rural and urban properties were used for years with no available estimates. Total, urban, and rural gross real estate aggregate value at market prices depicted. Urban and rural aggregate value and shares depicted in Figures B.6 and B.7.

Corporate wealth

I estimate gross corporate sector net wealth (S11-12 B90) based on firm tax records described in Section 4, which represents the book value of these firms (hence will result in book-value aggregate wealth also). Sole proprietorship firms are excluded from corporate wealth (and included as household business wealth), as are government-owned firms and direct and portfolio foreign investment, which are already included in the government's balance sheet and in the International Investment Position respectively. After these adjustments, this micro-dataset allows me to precisely calculate financial assets, non-financial assets, and liabilities for 2009-2017 period from firms' reported balance sheets. Liabilities are the sum of all liabilities recorded by firm tax records, plus a residual calculated to ensures that net corporate wealth is equal to zero, i.e. it is all owned by households.

The resulting corporate sector balance sheet is depicted in Figure D.5. Non-financial assets account for over 200% of national income. Financial liabilities, both reported by firms and imputed as the result of ownership by the household sector, account for 300% of national income, while financial assets represent (by construction) the additional 100% necessary to balance the account.

Fig. D.5. Corporate sector net wealth, 2009-2017



Note. Based on firm balance sheet administrative data (DGI). The dataset accounts for the universe of firms which pay corporate income tax. Publicly owned firms and sole proprietorships are excluded (as they are included in public and business wealth). Foreign direct and portfolio investment aggregates from Figure D.3 are also excluded.

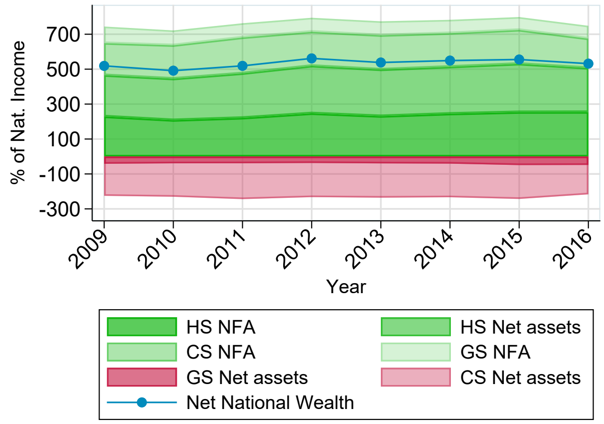
D.4. Households' wealth

After computing private net wealth following equation 2, I calculate private non-financial assets and liabilities, calculating financial assets as a residual. Within this residual, I distinguish pension funds, which are reported by the central bank, growing from 16 to 26% of net national income over the study period. Private non-financial assets are simply the result of adding the household sector's real estate wealth described in Section D.3 to the sole-proprietorship business wealth from firm tax files. Given the extent of informality in small businesses, this value is adjusted upward based on the household survey by computing the ratio of incomes of small businesses that declare not paying taxes to those that do, as a proxy of the share of informal wealth (57% of formal wealth).

The final piece of the puzzle is household liabilities. Although the central bank does not report a household balance sheet, in its yearly financial system reports, it does publish the estimated aggregate liabilities of households as a percentage of household incomes.[27] According to these reports, liabilities represent 21-27% of household incomes depending on the year. These ratios are directly applied to household incomes from national accounts described in Section 4.

[27] All reports available online (in Spanish) at `https://www.bcu.gub.uy/Servicios-Financieros-SSF/Paginas/Reporte-Anual-de-Estabilidad-Financiera.aspx`.

Fig. D.6. National wealth by sectors and financial/non-financial ssets



Note. Own elaboration. This figure is equivalent to Figure 5.1 but with greater detail.

D.5. Accounting for long run determinants of the wealth-to-income ratio

To better understand the wealth-to-income ratios estimated for 2009-2016, extended series are built by extrapolating from 2016's estimate based on a one-good accumulation model (Piketty and Zucman, 2014). The one-good framework essentially assumes that there are no price effects, and therefore wealth increases are the result of pure net savings accumulation. Naturally, this is an extremely simplifying assumption, which could be somewhat relaxed if the relative evolution of capital prices was included (i.e., assuming a two good model), but unfortunately no such series could be found. The sole purpose of this exercise is thus to understand the extent to which growth and savings might have acted on the wealth-to-income ratio in the last half a century. In this case, defining the wealth-to-income ratio as $\beta_{nt} = \frac{W_{nt}}{Y_{nt}}$, we have:

$$\beta_{nt+1} = \frac{1+g_{wst}}{1+g_t}\beta_{nt} \tag{D.1}$$

With

$$1 + g_{wst} = 1 + \frac{s_t}{\beta_{nt}} \tag{D.2}$$

and

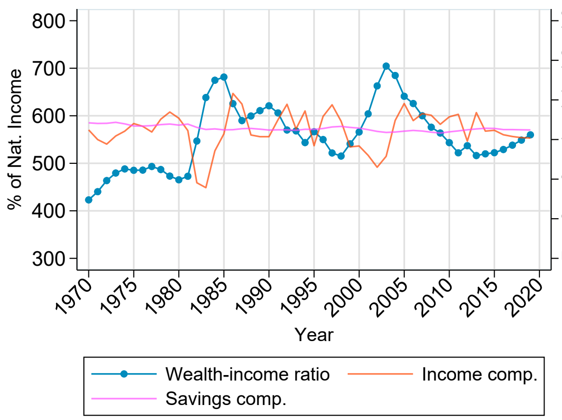
$$1 + g_t = \frac{Y_{t+1}}{Y_t} \tag{D.3}$$

The income growth component is hence given by $1 + g_t$ and equals national income growth, while the wealth growth component is given by $1 + g_{wst}$ and is equivalent to net savings.

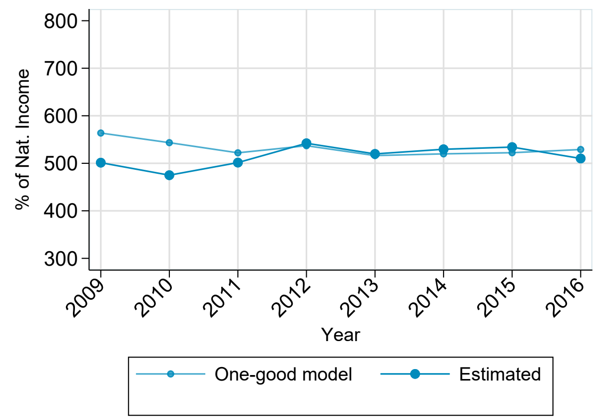
Panel (a) of Figure D.7 depicts the evolution of these two components and the wealth-to-income ratio they entail when fixing 2016's value. This analysis shows that, given investment and growth patterns, the wealth-to-income ratios observed are stable after a large decrease which followed the 2002 peak, during the last major economic crisis. Similarly, the all-time maximum was 700% after the 1982 crisis. Considering the evolution of the income and wealth growth components, it seems clear that the dynamic is mainly driven by income: in the context of a relatively stable savings rate (relative to existing wealth), what dominates the evolution of the wealth-to-income ratio is the growth of national income. In this case, the massive economic crisis the country experienced in 2001-2003 (recall Figure G.1) generated a sharp increase in the wealth-to-income ratio, which then started to slowly decrease as growth rates became positive again, and the same happened after the 1982 crisis. Thus, within this framework, wealth-to-income ratios show two peaks over the period, and are now falling from quite high levels that were the product of the collapse of national income rather than of true wealth accumulation.

Panel (b) zooms in to the recent period, depicting both the actual estimated series and the predicted (backward) evolution of the one-good model, with very similar trends, which is an indication of the absence of major price effects.

Fig. D.7. Wealth-to-income ratio in Uruguay, one-good model, 1970-2019.



(a) One-good model



(b) Model vs estimate

Note. Source: net income growth rates and consumption of fixed capital taken from `https://wid.world/`; savings rates taken from national accounts. Wealth-to-income ratio estimated by extrapolating 2016's estimate based on a one-good wealth accumulation model (Piketty and Zucman, 2014). Panel (a) considers a one-good framework, hence assuming that there are no price effects and therefore wealth increases are the result of pure net savings accumulation. Thus, if $\beta_{nt} = \frac{W_{nt}}{Y_{nt}}$, then $\beta_{nt+1} = \frac{1+g_{wst}}{1+g_t}\beta_{nt}$, with $1 + g_{wst} = 1 + \frac{s_t}{\beta_{nt}}$ and $1 + g_t = \frac{Y_{t+1}}{Y_t}$. The income growth component depicted is therefore given by $1 + g_t$ and equals national income growth, while the wealth growth component is given by $1 + g_{wst}$ and is equivalent to net savings. Panel (b) presents the one-good model estimates and the actual balance sheet estimates for the period under analysis.

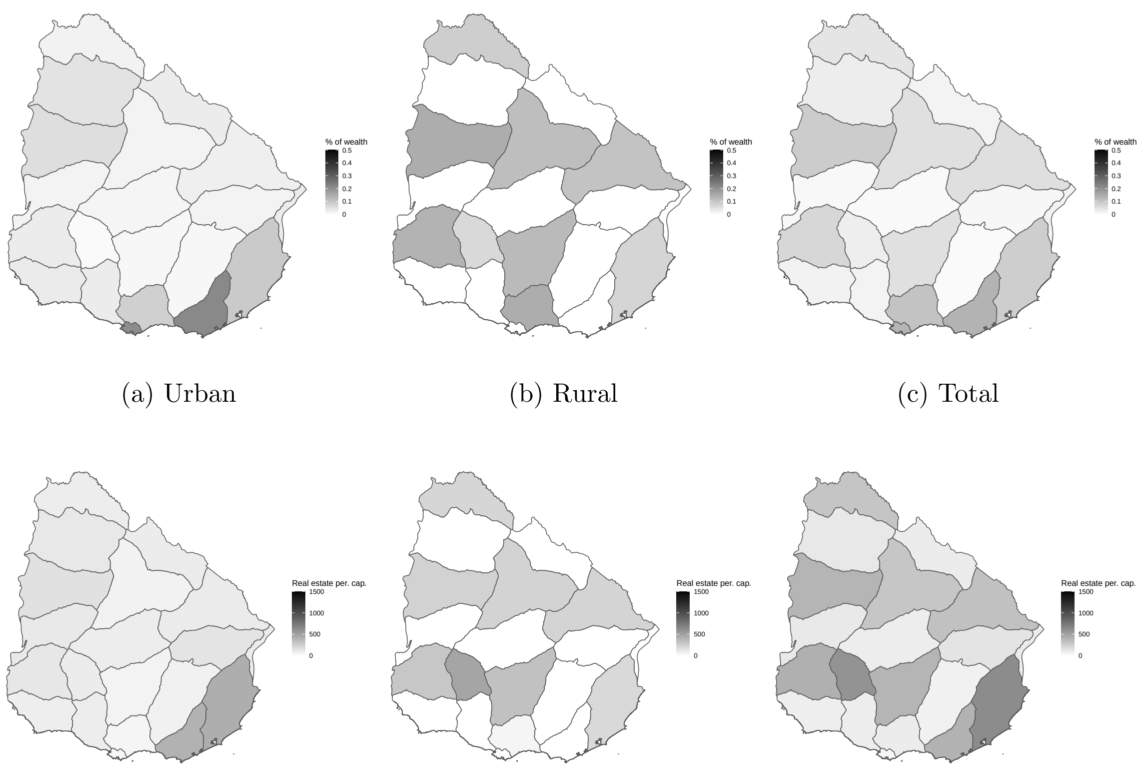
E. The geographical distribution of real estate wealth

To estimate the geographical distribution of wealth, I depart from the market price-adjusted cadastral data discussed in Section 4. This data reflects market price gross real estate wealth for the universe of rural and urban properties, regardless of the institutional sector that owns them.

Wealth distribution by department is depicted in Figure E.1 (panels a-c). Real estate wealth shares of total, rural, and urban real estate are shown, reflecting the different wealth levels (presented as a percentage of national income) depicted in Figure E.2. In the case of urban real estate wealth, it is very concentrated in the south, especially in the capital city Montevideo, the Canelones department (which hosts very large bedroom towns associated with economic activity in the capital), and Maldondado, where the high-class tourist city of Punta del Este is located. In the case of rural real estate, the distribution is more evenly spread across the territory, with radically less importance of the capital given its very small relative territory. The overall picture suggests a higher share of real estate wealth in the south of the country, and to a lesser degree in some departments from the west by the riverside.

The preceding results are heavily influenced by the size of each department (especially in the case of land) and by its population (when considering housing). For this reason, Figure E.1 (panels d-f) shows per capita real estate wealth. When doing so, the relative importance of Maldonado stands out in the case of urban real estate (given its low population relative to the value of the installed tourism capacity), and the axis of Montevideo-Canelones loses relative importance given that half of the country's population lives there. Some departments by the Uruguay river (to the west), such as Colonia (with two large cities, Colonia del Sacramento and Carmelo) also present slightly higher values. In the case of rural real estate, the relatively less populated departments of the centre emerge as the ones with higher rural wealth per capita. Overall, per capita real estate wealth seems to be higher in some departments of the south (especially Maldonado), and in the centre-west, somewhat different from the clearer L-shape pattern found in regional GDP (Rodríguez Miranda and Menendez, 2020).

Fig. E.1. Gross real estate wealth by department

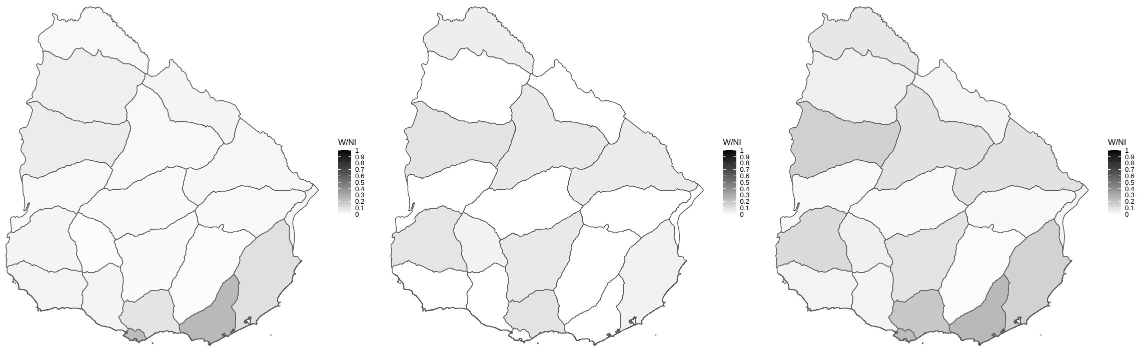


(a) Urban (b) Rural (c) Total

(d) Urban - per cap. (e) Rural - per cap. (f) Total - per cap.

Note. Based on cadastre data from DNC, DGR, and DIEA. Market price adjustment depicted in Figure B.5, 2018 values and official population projections (INE). Remaining years present very similar distributions.

Fig. E.2. Wealth-to-income ratio by department



(a) Urban (b) Rural (c) Total

Note. Based on cadastre data from DNC, 2018 values and population from 2011 census. Remaining years present very similar distributions.

F. Wealth correlated returns' sensitivity analysis

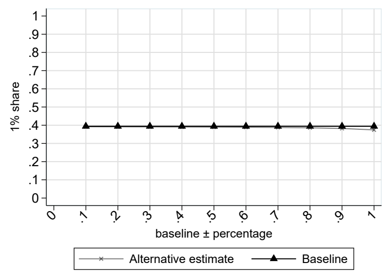
One of the most important drawbacks of the capitalization method is the assumption that rates of return for each type of wealth are identical for every individual. As explained in Section 2.2, this may not be the case since identical individuals in terms of observable characteristics may face different rates of return (idiosyncratic returns), or rates of return may be positively correlated with wealth. The first issue is probably not very important since the effects of idiosyncratic returns are likely to cancel out, but the second one may be more serious.

Saez and Zucman (2016) test this assumption based on data from *Foundations* , for which both wealth and capital income flows are observable from tax data, and conclude that the capitalization method "works well", at least in that context. However, rates of return may be larger for high income or high wealth individuals because they are better informed and advised of investment opportunities, so they are able to own safer and more profitable portfolios. Piketty (2014) argues that in fact return rates are higher for large wealth holders, based on *Forbes* global wealth rankings and publicly available data from US universities. *Forbes* rankings allow him to focus the attention on the very top fractiles of the wealth distribution, for which he observes that the growth rate of their wealth was 6.8% per year in real terms between 1987 and 2013, much higher than the average wealth growth rate (2.1%). Moreover, based on data from tax returns and actual wealth holdings from tax records for the whole Norwegian population, Fagereng et al. (2016); Fagereng, Guiso, Malacrino, and Pistaferri (2020) show that returns are indeed correlated with wealth, as in the case of the US (Smith et al., 2021).

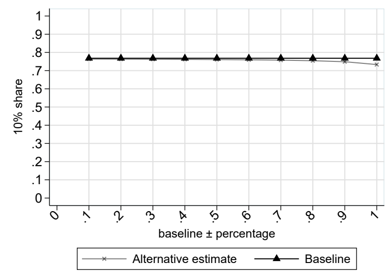
To assess the impact of the identical rates of return assumption, a simple sensitivity tests is performed. Instead of constant rates of return, increasing rates by p-tile of the capital income distribution are used, keeping constant the average rate and adjusting each wealth component to match the aggregate wealth level. Returns are assumed to be linearly increasing within each type of wealth, excluding liabilities for simplicity. Figure F.1 (panels (a)-(d)) shows the effect on wealth shares, as the gap between the rates of the bottom and top percentiles increase. In the extreme case, the rates of the top 1% are doubled, and those of the bottom 1% are brought to zero. Results show that concentration does fall as expected, but general results do not change substantially. Although this simple exercise relies on several assumptions, such as linearly increasing rates of return, it shows that wealth-correlated returns are not likely to be an important concern in this setting.

A somewhat more demanding exercise is shown in Figure F.1 (panels (e)-(h)), in which the return rates of the top 1% and the bottom 99% are increased and decreased respectively. In the extreme case, the top 1%'s rates for each asset are increased by 30%, which is mirrored by a 30% decrease in bottom 99%'s rate, resulting in the top group's rates being double the rates of the rest. This variation is close to what Smith et al. (2021) find for the US, with the top 0.01% having rates of return close to 3 times larger than the average, although within the rest of the top 1%, rates are closer. This alternative exercise pushes the top 1% share upward and the bottom 99% downward, resulting in a decrease of the top 1%'s share of close to 10 percentage points, while the change is less dramatic for larger groups such as the top 10%. This exercise shows the intuitive result that if the correlation between rates and wealth is particularly extreme at the top of the distribution (e.g., for the top 1% or smaller groups), the effect is likely to be concentrated on those groups but not in overall wealth inequality.

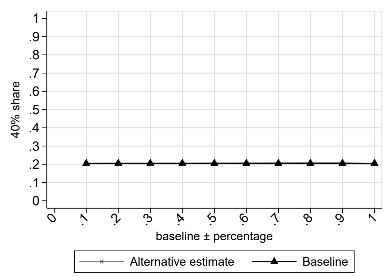
Fig. F.1. Wealth correlated returns' sensitivity analysis, 2016.



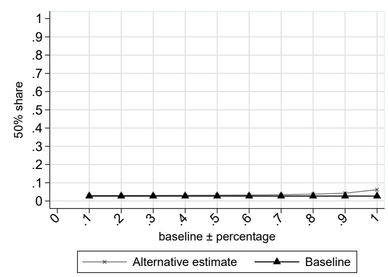
(a) Top 1%



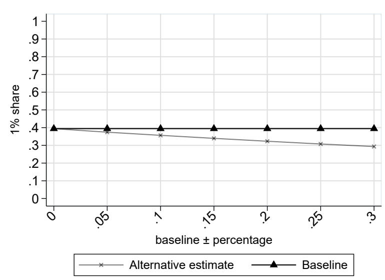
(b) Top 10%



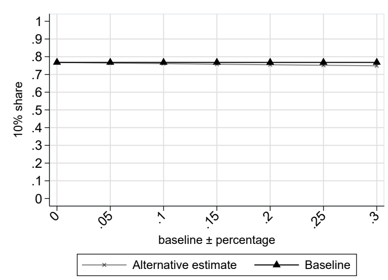
(c) Middle 40%



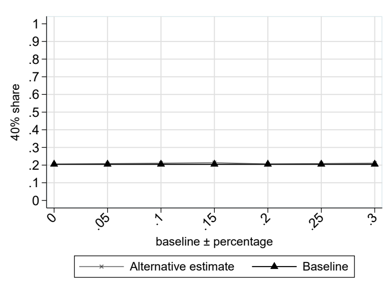
(d) Bottom 50%



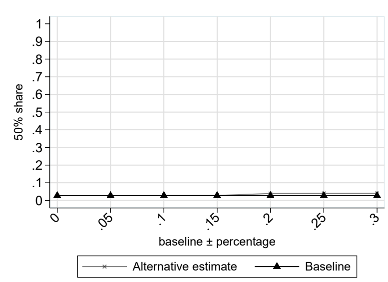
(e) Top 1%



(f) Top 10%



(g) Middle 40%



(h) Bottom 50%

Note. Panels (a)-(d): the x-axis column depicts return rates' variation ranges in relation to average return rates for extreme groups (top and bottom 1%). As an example, a value of 0.25 on the x-axis indicates that the top 1%'s rate has been increased by 25%, and reduced by 25% for the bottom 1%. Returns in the remaining percentiles vary linearly between these two points. Panels (e)-(h): the x-axis column depicts return rates' variation ranges in relation to average return rates for the bottom 99% (negative) and the top 1% (positive). As an example, a value of 0.25 on the x-axis indicates that the top 1%'s rate has been increased by 25% and reduced by 25% for the bottom 1%. Results depicted for 2016, identical conclusions are drawn from analysing remaining years (available upon request).

G. Growth and income inequality

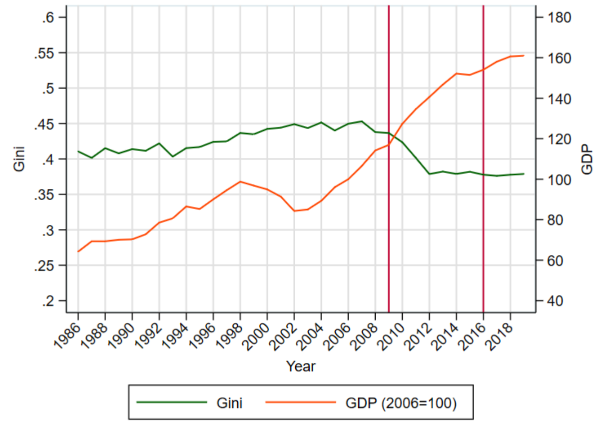
Uruguay is a small, high income country, with low income inequality in the Latin American context but still high compared to developed countries[28]. After decades of unstable economic growth and recurrent economic crisis, Uruguay has sustained an average annual growth rate of around 4% for the last fifteen years, reaching a per capita GDP of USD 21,625, about 40% above the Latin American average, but half the average of the OECD countries[29]. This economic growth, coupled with a series of relatively large reforms both in the labour market and in the tax and transfer system put in practice by a centre-left coalition in office since 2005, precipitated a significant decline in income inequality. These reforms included a major increase in the minimum wage, the restoration of centralised, collective wage bargaining, an expansion of both the coverage and amount of non-contributory cash transfers schemes, and the introduction of progressive income taxation. Based on high-quality household surveys, studies have consistently shown[30] that income inequality experienced a rapid decline between 2008 and 2012, followed by a relative stagnation from 2013 onward (Figure G.1). This income inequality decrease has been confirmed by the use of income tax records (Burdín et al., 2022) and Distributional National Accounts estimates De Rosa and Vilá (2022), so the story presented in Figure G.1 is an accurate one, yet incomplete.

[28]The population has around 3.400.000 people and has been remarkably stable over the last decades, while the survey-based Gini index has now stabilised at 0.38.

[29]Values in PPP. https://data.worldbank.org/

[30]See for instance Cornia (2014a).

Fig. G.1. Income inequality and growth in Uruguay, 1986-2019.



Notes: Gini index is based on a household survey (*Encuesta Continua de Hogares*) and refers to per capita household income. A similar trend is observed when considering inequality estimates such as the top 10% share. The survey is conducted by the National Statistics Institute (INE). GDP data is produced by Uruguay's central bank (BCU).

Chapter 6

Epilogue. On *Capital*: An essay on inequality, capital and value theory

Abstract

Capital is back at the center of the empirical distributional research agenda. New estimates of wealth accumulation, distribution and inheritance, fully consistent with national accounts' definitions and deeply rooted in standard neoclassical growth models, are now available. This provides the new inequality literature with clear-cut insights and empirical firepower. Yet while the empirical flank is increasingly well protected, the theoretical one is exposed. I revisit the debates on the underlying theory of capital and document its drawbacks, highlighting that it is particularly ill-equipped for inequality analysis and that its central problem is the theory of value. Does this mean that we should start anew? I argue on the contrary, showing that under a one-good model assumption, there is accounting correspondence with the labor theory of value, which gives room for reinterpretation of most available estimates. Moreover, it is possible to establish clear accounting links between famous drivers of the economic system such as $r > g$ and Marx's falling rate of profits. However, even under this accounting correspondence, taking distance from the scarcity theory of value has relevant implications for the inequality narrative, insofar it forces us to abandon the merit-inheritance discussion to include exploitation. The main takeaway is that empirical wealth and income inequality literature needs not get corseted in the neoclassical theoretical framework.

1. Introduction

Capital has made an astonishing comeback to the empirical distributional research agenda. Based on novel data and a variety of revised and new methodologies, the recent wealth accumulation, inheritance and inequality results are undoubtedly better estimated than ever before. Most of the new estimates of wealth accumulation and distribution are anchored to the System of National Accounts (SNA) framework, hence fully consistent with internationally accepted definitions. Thus, there is an increasing amount of evidence on wealth distribution (see for instance Saez and Zucman (2016); Alvaredo et al. (2018); Garbinti et al. (2017)), which is at the same time consistent with national wealth estimates (e.g. Piketty and Zucman (2014); Blanco et al. (2021); Bauluz (2019)). Moreover, SNA-based wealth estimates are fully consistent by construction with national income and its distribution, hence providing a complete depiction of income and wealth dynamics, both micro and macro-economically. This allows researchers to simultaneously account for variables such as growth, the capital share, the wealth to income ratio, inheritance flows, rate of return and wealth distribution under SNA's framework, endowing the empirical inequality literature with renewed firepower.

Consistently estimating all these variables for Latin American countries was the main objective of this thesis. In its second part and particularly in 5, I estimated all key wealth variables for Uruguay under the Distributional National Accounts framework (WIL, 2020), in an attempt to catch up with the new literature and to provide comparable estimates for a developing country, which are extremely rare. In all these five articles, we showed that both as an estimation challenge as well as an inequality driving force, *capital* was the gravity-center of our empirical endeavors.

These variables and capital itself are not only estimated consistently with National Accounts, but can also be easily linked with standard neoclassical growth models, hence providing the theoretical foundation for the wealth accumulation dynamics (Piketty and Zucman, 2015). While there is no 'unified theory of inequality' (Atkinson and Bourguignon, 2000), for macro-distributional purposes standard neoclassical growth models are the main reference point. These models are admittedly limited, in particular the "one-good, perfect competition model is not a very satisfactory model, to say the least " (Piketty, 2015, p. 81). Nevertheless, they do provide important insights and intuitions on these macro variables and their likely evolution in the future (Piketty, 2014), and have been extensively debated (see e.g. Acemoglu and Robinson (2015); Jones (2015); Piketty (2015)). Even if only used as a general reference point, it is still *the* theoretical reference that the vast majority of the empirical literature uses, for better or worse. But is it adequately equipped for distributional analysis? In case it is not, can we do better? To explore these questions is the aim of this epilogue, in which I ponder on the economic theory that underlies the empirical effort my co-authors and I undertook.

Debates over these broad set of models are not new, as the standard neoclassical growth theory is rooted in extremely questioned assumptions and definitions. To begin with, the very definition of *capital* is problematic and has been subject to a significant amount of controversy (Hodgson, 2014). The last of the great controversies that raged during the 1950s until mid 1980s, i.e. the famous Cambridge Capital controversies, was primarily focused on how to measure capital. Yet, that was only the corollary: the main issue at stake was the very essence of what capital was (Harcourt, 2014). Kick-started by the call to arms of Joan Robinson's assault on the existence of a production function (Robinson, 1954), it was shown that neoclassical growth theory was unable to provide convincing explanations for the main driving variables of the capitalist system, especially for the rate of return r (Cohen and Harcourt, 2003). In particular, it was shown that once one leaves the one-commodity-model assumption, it is not longer possible to determine the rate of return to capital, hence turning it impossible to provide an explanation for the macro-distribution of income.

Although the exact significance of the overall conclusion was not settled (and nor will I try to do so in this essay), it was indeed admitted even by the neoclassical side that standard growth models were unable to produce an adequate theory of factor prices, i.e. a theory of income distribution. In the final paragraph of an overview of capital theory, Robert Solow claimed:

> "Very little has been said in this survey about income distribution (in other words, about the determination of factor prices). That is because there is no special connection between the neoclassical model of growth and the determination of factor prices. The usual practice is to appeal to the same view of factor pricing that characterizes static neoclassical equilibrium theory. If the working assumption that all markets clear were to be lifted, an alternative theory of factor prices would certainly be needed. Much else would change besides" (Solow, 2000, p. 378).

During the Cambridge Capital controversies, the debate focused on very specific details of the production function such as *reswitching* and *capital reversing*, which allegedly entailed the impossibility of the determination of r. Yet under the surface, the true underlying difficulty of the neoclassical theory lies in the theory of value upon which it is founded. The scarcity theory of value can be traced back to the irruption of the Marginalist revolution, which entailed a very different departure point than the previous economic thinkers. Following Cohen and Harcourt (2005), in the classical vision of political economy, the fundamental economic problem is the allocation of surplus output, to which social class is the fundamental unit of analysis, and consumption's purpose is to satisfy production. The rate of profits arises from social relationships in production and is the result of the expansion of capital. In the canonical neoclassical models, the rate of return is in turn just the result of decreasing marginal productivity of capital and households maximising inter-temporal utility, which is the driver of economic activity.

Under a one-good model world, these contradictory 'visions' of the fundamental functioning of the economic system may still provide similar conclusions, while true divergence emerges only once one leaves such an assumption (Cohen and Harcourt, 2005). In fact, most of the neoclassical growth model distributional results are accounting identities, true by definition, either in all settings or in the steady state (Piketty and Zucman, 2015). Nevertheless, even if by assuming simple underlying one-good models the inequality literature manages to dodge the main criticisms to the capital theory, it still remains rooted in a scarcity approach to value and to a narrow view of the economic process. The question that I try to address is whether it is possible to use the classical political economy approach to account for the distributional and growth estimates, given the same one-good model assumption, and if so what can we learn.

There are a number of different alternatives to do this. Cohen (1989) shows that both the classical and neoclassical approaches are associated with "robust results that hold without exception" within one-good models such as Samuelson's surrogate production function (see 3.2) and the neo-Ricardian corn model. Tobón and Ríos (2020) show in turn that long run distributional results from one good-models such as the ones used by Piketty (2014) could be routed funded in post-keynesian models (i.e. with no production function), using as the key link the "Cambridge Equation" (Pasinetti, 1979). I will in turn focus on Marx, since it allows me to better discuss the theory of value implications in a more tractable way.

Marxian insights have been highlighted by many authors who were not themselves Marxists as Veblen, Schumpeter or post-Keynesians such as Sraffa (Bellofiore, 2008) or Robinson (Alves, 2022). Indeed, in her *essay on Marxian economics*, Robinson said that if "the orthodox notion of a definite supply price of capital thus disintegrates upon examination, we are left with nothing but Marx's notion that capital is accumulated and maintained because capitalists are forced to accumulate in order to survive" (Robinson, 1942, p. 61). Following this thread, I explore the accounting links between the neoclassical macro distributional theory and Marx's labour theory variables (Marx, 1867), in an attempt to better understand their differences and the implications for empirical research. The aim is simply to understand the conditions under which there is *accounting correspondence* between these two approaches.

The starting point is the recognition of the conceptual gaps between similarly named variables. In particular, I contrast the definitions of the rate of return r and Marx's rate of profits P and, more importantly, I discuss the fact that –unlike neoclassical theory– Marx's capital is closer to a flow than to a stock. These two points alone help clarify much of the confusion which often result from these terms. Moreover, I show that under a one-good model assumption, accounting correspondence does exist between available estimates and Marx's labour theory of value, which can in turn be linked with main variables of standard neoclassical growth model and hence to the empirical inequality literature.

After documenting the theoretical differences between the definition of capital and the rate of return (profits), I show that, under a one-good model and closed economy's assumption, it is possible to interpret both labour theory of value and the new empirical literature in a simple unified accounting framework. Moreover, it is possible to establish the accounting links between famous drivers of the capitalist system such as $r > g$ and Marx's falling rate of profits. This is not surprising, since under a one-commodity model, many different theories find their place. Thus, "while heuristically valuable, the insights one-commodity models provide do not allow us to distinguish between competing theories that view interest or profits as payment for the marginal productivity of capital, or as exploitation of workers" (Cohen and Harcourt, 2005). What I show is that under such a model, $r > g$ and the falling rate of profits yield perfectly consistent results. Specifically, I show that a stable rate of return r larger than the growth rate g, results in both an increase in the wealth to income ratio and the capital share, as well as in falling rate of profits P, which is offset at the beginning by the increase in the capital share but falling in the long run regardless.

Schumpeter (1942) famously discussed different 'Marxs' –the prophet, the sociologist, the economist and the teacher–, I argue that by also considering 'Marx the accountant'[1], there is room to re-interpret important empirical results in light of a theory that better allows to discuss distributional issues. Indeed, Marx's theory of value has at least the merit of being quite intuitive form a strictly accounting perspective (Sweezy, 1942; Shaikh et al., 1997; Bryer, 1999). Labor theory of value is by no means exempt from criticisms, but the case can be made that "adverse judgment or even exact disproof, by its very failure to injure fatally,

[1]Marx's accounting has been highlighted by the literature, see Bryer (1999) for a discussion, where he argues that Marx provides a general theory of accounting.

only serves to bring out the power of the structure" (Schumpeter, 1942, p. 3). Moreover, it is nonetheless true that most of the criticism to the labour theory of value emerges in a multi-sector model, just as in the case of the neoclassical approach. In the best case scenario, this simple exercise provides a bridge with existing empirical literature and classical political economy, including Marx. The main takeaway is that empirical wealth and income inequality literature needs not to get corseted in a restrictive neoclassical framework. However, I argue that while at the one-commodity-model level there is overall accounting consistency between the different approaches, the shift from a scarcity theory of value to a the classical approach entails relevant implications. One immediate consequence is that to the accumulation of wealth mechanisms one should add exploitation to work and inheritance, and that completely changes the narrative.

This essay is organized as follows. Section 2 presents the System of National Accounts' wealth and capital definition and provides an overview of the main neoclassical growth model distributional results, focusing on the steady state determination of r and its gravitating effect in both wealth accumulation, the capital share and the personal wealth distribution. In Section 3, the literature on the definition of capital and the main caveats of the neoclassical growth model are briefly discussed, highlighting its distributional implications and tracing them back to the scarcity theory of value. As an alternative, classic political economy and Marx's labour theory of value are presented, showing how it too faces theoretical difficulties once the one-good model assumption is abandoned. The accounting equivalence between the models of sections 2 and 3 is discussed in Section 4. Section 5 concludes the essay and points at possible ways forward.

2. Wealth and capital theory

In this section, the national accounts' definition of wealth is presented, showing its link to the standard production function (2.1), followed by a discussion of the determinants of wealth accumulation and distribution, focusing on the rate of return r and the growth rate g (2.2). In section 2.3, the steady state determination of the key variable r in the standard neoclassical model is briefly summarized.

2.1. Wealth and capital from a National Account's perspective

The departure point of this analysis are the standard definitions of wealth and capital. The 'handbook definition' presents a powerful and coherent framework for its analysis, rooted in the general definitions of the System of National Accounts (SNA). Based on SNA balance sheet's definitions, private wealth W_t is defined as the net wealth (assets minus liabilities) owned by households (United Nations, 2008b; WIL, 2020). These assets include all the nonfinancial and financial assets financial assets over which ownership rights can be enforced and that provide economic benefits to their owners (Piketty and Zucman, 2015). Private wealth can be decomposed in housing assets, business assets (and other non-financial assets), financial assets and liabilities, while National wealth W_{nt} results from the addition of private and public wealth. It is also equivalent to to the sum of domestic capital and net foreign assets, as depicted in equation 2.1.

$$W_{nt} = W_t + W_{gt} = K_t + NFA_t \tag{2.1}$$

W_{nt}, W_{gt} and W_t represent net national, public and private wealth respectively, K_t domestic capital and NFA_t net foreign asset position. As for the second equivalence it is interesting to note that, intuitively, as all national financial assets and liabilities must cancel out (including the property of corporations), national wealth W_{nt} is equivalent to the sum of all non-financial assets owned by household, corporate and government sectors, plus the net foreign asset position.

With Y_t being the national income, the private and national wealth to income ratios (β_t and β_{nt}) are hence defined as:

$$\beta_t = \frac{W_t}{Y_t}, \beta_{nt} = \frac{W_{nt}}{Y_t} \tag{2.2}$$

National wealth W_{nt} and domestic capital K_t have different magnitudes only if the net foreign asset position NFA_t is not zero; otherwise they are equivalent. In fact, in a closed economy or at the world level, where $NFA_t = 0$ by definition, aggregate wealth and capital coincide. In such a setting, capital and wealth are interchangeable, and not only $\beta_{nt} = \beta_t$, but also they are equivalent to the capital - output ratio ${}^{K_t}/_{Y_t}$. This is the case because $W_{nt} = K_t$, but also because output Y_{dt} is equivalent to net national income Y_t, given that $Y_t = Y_{dt} + r_t \cdot NFA_t$.

$$Y_t = Y_{dt} = F(K_t, L_t) \tag{2.3}$$

The capital - output ratio is important, since it can be directly be traced back to the production function, as in equation 2.3 (which assumes a closed system), where L_t is the aggregate labour input (Piketty and Zucman, 2015). This capital-output ratio has been one of the main concerns of growth theory, since the early attempts to model growth (Harrod, 1939; Domar, 1946), and it is further discussed in the following section.

2.2. *Growth, capital and inequality theory*

When modelling the economic activity with a production function such as in equation 2.3, i.e. assuming a one-good model, some important steady-state conclusions may be easily drawn. First, in this one-good model setting, wealth growth is $W_{t+1} = W_t + S_t$, given S_t being the aggregate net savings rate. Considering equation 2.2 and noting that $s_t = {}^{S_t}/_{Y_t}$, we have:

$$\beta_{t+1} = \frac{\beta_t + s_t}{1 + g_t} \tag{2.4}$$

From equation 2.4, it is straightforward to show the steady state formula for β, with $g_t \to g$ and $s_t \to s$:

$$\beta_t \to \beta = \frac{s}{g} \tag{2.5}$$

This is a Harrod-Domar-Solow-Swan result (see section 2.3), which is an accounting definition that holds in the steady state of any micro-founded, one-good model of capital accumulation, independently of the exact nature of saving motives (Piketty and Zucman, 2015).[2] Thus, independently from the long-run fundamentals of s and g, it is true by definition that β will increase the higher s and the lower g. In particular, in a low growth context, the amount of wealth (or capital) relative to income will be higher, which is the case for rich countries from 1970 onward (Piketty and Zucman, 2014).

The level of capital relative to output (or wealth to income) has a direct effect on the functional distribution of income. Again by construction, we have that the capital share of the economy α_t if given by:

$$\alpha_t = r_t \cdot \beta_t \tag{2.6}$$

where r_t is the average rate of return of the economy. In the standard neoclassical framework,

[2] Authors show that these general conclusions hold in a variety of micro-funded general equilibrium models, where s is endogenous and depends on g Piketty and Zucman (2015, p. 1344-1347).

r_t is equivalent by its marginal product. Thus, considering a general and relatively flexible Constant Elasticity of Substitution (CES) production function, where $Y = F(K, L) = \left(a \cdot K^{\frac{\sigma-1}{\sigma}} + (1-a) \cdot L^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma-1}{\sigma}}$, then r_t is given by $r = F_K = a\beta^{-1/\sigma}$, being σ the capital-labour elasticity of substitution. As r_t inversely depends on β, which results from an inverse relation in the quantity of capital K_t and its price r_t the direction of changes in α_t will depend on the magnitude of the variations, which is given by σ.

$$\alpha = a \cdot \beta^{\frac{\sigma-1}{\sigma}} \tag{2.7}$$

Thus, if $\sigma > 1$ then an increase in β_t will result in a less-than-proportional decrease in r_t, hence pushing α_t upward. Naturally, the inverse is true when $\sigma < 1$, while $\sigma = 1$ results in the special case of the Cobb-Douglass production function, where changes in K_t (β_t) are exactly neutralized by opposite changes in the marginal product of capital (which equals r_t), hence leaving the capital share α_t unaffected.

Finally, it can be shown that these macro-results have relevant implications on micro (i.e. personal) distribution of wealth. In a wide variety of models, the concentration of wealth increases as the gap between r and g widens. Piketty and Zucman (2015) show that all models with multiplicative random shocks in the wealth accumulation process give rise to distributions with Pareto upper tails.[3] Being z_{ti} the normalized individual wealth[4] and ω_{ti} a i.i.d multiplicative shock with mean $\omega = E\left(\omega_{ti}\right) < 1$ and ε_{ti} an additive shock, we have:

$$z_{t+1i} = \omega_{ti} \cdot z_{ti} + \varepsilon_{ti} \tag{2.8}$$

Aggregating individuals modeled as in equation 2.8 results in a distribution with a Pareto upper tail with a coefficient a, which must solve $E\left(\omega_{ti}^{a}\right) = 1$. Thus, for a given average $\omega < 1$, as the variance of the shock increase (goes to infinity), wealth concentration also does $(a \to 1)$.

Regardless of the nature of the individual-level shocks, the shape of the upper tail (i.e. a) depends on $r - g$. Assuming a generation length H (e.g. 30 years) and each period of the

[3] They also present an example with closed form formulas, in which individual utilities are modeled as a Cobb-Douglas function, while the production function is a CES. The results represent a particular case of this general one.

[4] The model assumes a stationary population $N_t = [0, 1]$, made of continuous agents, which results in equivalent aggregate and average wealth and income variables, i.e. $W_t = w_t$ and $Y_t = y_t$. From there, it z_{ti} is defined as $z_{ti} = {}^{w_{ti}}/_{w_t}$.

model as lasting this long, ω can be written as (with r and g instantaneous rates)

$$\omega = s \cdot \frac{1+R}{1+G} = s \cdot \mathrm{e}^{(r-g)H} \tag{2.9}$$

with $1+R = \mathrm{e}^{rH}$ and $1+G = e^{gH}$ being the generational return and growth rates. Under binomial shocks, the inverted Pareto coefficient is $b = (\log(1/p))/(\log(1/\omega))$. This coefficient is thus extremely sensitive to variations in ω and therefore in $r-g$, which entail substantial concentration increases relative to minor $r-g$ upturns. Thus, as famously pointed out by Piketty (2014) and further discussed in Piketty (2015), if the inequality $r > g$ holds for sufficiently long periods, wealth concentration is likely to increase.

2.3. *The steady-state and the role of r*

The discussion of the steady-state is important since it allows us to distinguish the main forces at play under the neoclassical growth model. To be sure, it is a useful abstraction but does not necessarily reflect real world economies, especially in the short run (Piketty and Zucman, 2014, 2015). However, it may provide insights as to what may happen in the 21st century, provided there are no massive shocks as the ones of the previous one (Piketty, 2014).

The models briefly summarised in section 2.2 show that the gap $r-g$ is central to understanding the main macro and micro-distribution of capital, especially after a long-run cumulative process. In particular, it was shown that a low growth rate g result in high wealth to income (or capital-output in closed systems) ratio β. In that setting, given a capital-labour elasticity of substitution of $\sigma > 1$, the rate of return r decreases less than proportionally to the increase in β, resulting in a higher capital share α. Moreover, the gap $r-g$ produces thick-tailed wealth distributions under a wide variety of models.[5]

The growth rate g is probably one of the most studied macro economic variables of all times, and has been modeled in a variety of ways.[6] In the basic Harrod-Domar model (Harrod, 1939; Domar, 1946), output is equal to the sum of savings S_t and consumption C_t from an accounting perspective, so $Y_t = C_t + S_t$. It then follows that output may be expressed as the sum of consumption plus investment I_t, so $S_t = I_t$. As by definition total capital is the result of the interaction between investment and depreciation, we get $K_{t+1} = (1-\delta)K_t + I_t$, and

[5] This general distributional conclusions have been questioned by a number of scholars. For an interesting discussion see Jones (2015), and for a not-so-interesting one see Acemoglu and Robinson (2015).

[6] See section 1.4 of Barro and Sala-i Martin (2004) for a brief summary.

thus is the capital-output the one depicted in equation 2.5.[7] In this model, the savings rate is of paramount importance as it determines investment and growth, but is assumed to be exogenous, and if it is sufficiently high, it may lead to an extremely high level of β.[8]

The Solow-Swan version of the model (Solow, 1956; Swan, 1956) makes β endogenous, by introducing profit-maximizing firms facing a production function with constant returns to scale but diminishing returns to each factor (especially to capital), which "chokes-off per-capita growth in the steady-state" (Ray, 1998). As a result, in such a model long-run growth rate is the result of population growth n and a (exogenous) productivity growth component h so that $g = n + h$.[9] Moreover, in this simple model the rate of return, which is equivalent to the marginal product of capital, is equal to the growth rate of the economy. This is known as the *golden rule* of capital accumulation (Phelps, 1961). It essentially says that if equal weight is given to present and future generations, capital will be accumulated (and hence output increased) until $r = g$, so that each generation's consumption is the same and the highest possible for all t.

In the Harrod-Domar-Solow-Swan models, the savings rate s_t is exogenous, which is a major drawback. Following the seminal contributions by Ramsey (1928), in the Cass-Koopmans version of the neoclassical model (Koopmans, 1965; Cass, 1965), an infinite-horizon household who maximizes dynastic consumption is also introduced into the model, providing an endogenous savings rate. Firms will use capital until its marginal productivity equals r, which is in turn set by the interaction of this demand with the supply price of capital, resulting form a combination of the reward of waiting and the risk of lending capital. This is the essence of the *modified golden rule*, which results from the steady state equilibrium of profit-maximizing firms and inter-temporal consumption-maximizing households (Barro and Sala-i Martin, 2004)[10]:

$$r = \theta + \gamma g \tag{2.10}$$

where γ is the curvature of the utility function, and θ is the rate of time preference. The model requires that the transversality condition that (net of depreciation) return rate r is higher

[7]In the standard model, it is expressed as ratio is $\beta = {}^{s}/_{g+\delta}$, therefore it is important to bare in mind that equation 2.5 is net of depreciation

[8]It is important to note, at this point and considering what follows, that the original Harrod-Domar's macro-dynamic model actually confronted with the dominating Marginalist mainstream (Pasinetti, 1983).

[9]If $L_t = N_t \cdot h_t$, with $N_t = N_0 \cdot (1+n)^t$ and $h_t = h_0 \cdot (1+h)^t$, then $1 + g = (1+n) \cdot (1+h)$, i.e., $g \approx n + h$ (Piketty and Zucman, 2015).

[10]The notation is not the same: it was adapted for consistency reasons.

than the growth rate g, so that household can follow an optimal inter-temporal consumption path given their relative impatience θ. This is argued to be one possible explanation of the relative stability of the rate of return in the long run, which might be interpreted as an interval of psychologically plausible time preference parameters (Piketty, 2015). The return rate r in turn determines the level of capital, since it is accumulated until its marginal product equals the rate of return. The key to determine capital per-capita k is hence the diminishing return's assumption.

Thus, under this general neoclassical growth model, the inequality $r > g$ is a condition for inter-temporal efficiency: g and r are determined by population and productivity increases on the one hand, and the interaction of profit-maximizing firms with diminishing returns to capital and infinite-horizon consumption maximizing households on the other. However, the determination of the rate of return is possible only within a number of highly restrictive assumptions. Moreover, while these assumptions may work in a one-good-model setting, it has been extensively argued that they do not hold under slightly more realistic assumptions. Indeed, once we leave the one-good model, it becomes essentially impossible to theoretically determine r, hence substantially compromising the theoretical foundations of the neoclassical inequality predictions. This is discussed in the following section.

3. Capital, distribution and value theory

In this section, controversies over the definition of capital are briefly summarized (3.1), in order to introduce the Cambridge capital controversies and the impossibility of determining r outside the one-commodity model (3.2). Section 3.3 discusses to what extent this is the result of the underlying scarcity theory of value, as well as the classical political economy alternative; while 3.4 focuses on Marx's theory of value, presenting the basic outlay of his value theory, necessary for the discussion of the accounting equivalence of section 4.

3.1. A controversial definition

In section 2, I presented what appears to be the growing consensus on how to measure wealth (and capital) accumulation and distribution, which anchors the main definitions in the System of National Accounts, providing a number of important advantages discussed above. However, these definitions have been the object of an recurring debate. Are capital and wealth really interchangeable? Is capital a sum of money or a bunch of stuff? Is it an accounting magnitude, a historically-specific process or the reflect of investment's time-structure? The discussion goes far beyond the definition, as it is both a manifestation and a driver of our understanding of what is valuable, how that value is distributed, what is the underlying justification for such a distribution and the very nature of the economic system itself. The sheer number of heavy-weight scholars who dedicated time and effort in engaging in this two-and-a-half century long debate is telling of its scope and relevance. The sometimes confusing nature of the debate, especially in the twentieth century reveals that, as Solow bluntly said, "there is also an intrinsic reason for the controversial character of capital theory: it is very complicated and very difficult" (Solow, 1963, p. 11).

Wealth and Capital

As discussed in section 2.1, although capital and wealth are not equivalent in the standard definition, under relatively general assumptions (i.e. that the Net Foreign Asset Position is zero[11] or at the world level), they may well be considered interchangeable in the national balance sheet. However, there is room for differentiating capital and wealth from a SNAs perspective. Following WIL (2020), such distinction can be traced back to the sequence of non-financial accounts, since operating surplus and mixed income represent the flow of income accruing to capital stock-owners, while property income is received by owners of financial or non-produced assets. This separates assets used in production (capital) from other assets such as equity, bonds or land (rest of wealth). Under this perspective, wealth includes capital but also other non-produced assets.

Distinguishing capital and wealth based on its use in production is possible but potentially tricky, since many assets can be used for more than one purpose. For instance, some buildings may have residential uses, while others used as offices or storage facilities, and there are many similar examples, for which the all-encompassing accounting wealth definition may be less ambiguous and more pragmatical. Considering wealth and capital simply as the sum of all the (net) assets that may be bought or sold, presents the enormous advantage of dodging the need to discuss the nature of each of them (Piketty, 2014).

Nevertheless, it may be argued that it is important to keep capital and wealth as distinct concepts. While it is clear that wealth generates an income flow accrued by wealth holders, allowing consumption smoothing when income declines (Davies and Shorrocks, 2000), and more importantly contributing to shape income distribution, the notion of the power dimension of wealth is very relevant. As Atkinson pointed out, "wealth is important because it gives not only income (interests, dividends and rent) but also security, freedom of maneuver, and

[11] Which is actually quite close to reality, at least for most developed countries (Piketty and Zucman, 2014).

economic and political power" (Atkinson, 1973, p.239). The power dimension of wealth has a long-standing tradition in economics, which can be traced back to Adam Smith, who famously claimed that on top of being a source of a stream of income, wealth was a synonym of power, which was in turn closely linked with his theory of value.

> "Wealth, as Mr. Hobbes says, is power. (...) The power which that possession immediately and directly conveys to him, is the power of purchasing; a certain command over all the labour, or over all the produce of labour (...) the exchangeable value of everything must always be precisely equal to the extent of this power which it conveys to its owner" (Smith, 1776, p.51).

It is precisely the power dimension which may be behind the conceptual difference between wealth and capital, since again following Atkinson "wealth is now quite widely distributed, but much of the wealth that people own conveys little or no control over the productive activities of the economy beyond their own front door" (Atkinson, 2015, p. 95). In marxian terms, it is not only that capital is power, but it is a very specific type of power also linked to value theory, which in the case of labour theory of value is related with surplus labour (see 3.4). On Smith's notion, Marx says that "capital, therefore, is not only the command over labour, as Adam Smith thought. It is essentially the command over unpaid labour. All surplus-value, whatever particular form (profit, interest or rent) it may subsequently crystallize into, is in substance the materialization of unpaid labour-time. The secret of the self-valorization of capital resolves itself into the fact that it has at its disposal a definite quantity of the unpaid labour of other people" (Marx, 1867, p. 672).

Money and Stuff

Marshal, building on Adam Smith, presents a definition of aggregate capital that is closer to the one discussed in section 2. When considering capital from an aggregate perspective, he says that "this brings us to consider the use of the term capital from the point of view of inquiries into the material well-being of society as a whole. Adam Smith said that a person's capital is that part of his stock from which he expects to derive an income. And almost every use of the term capital, which is known to history, has corresponded more or less closely to a parallel use of the term Income: in almost every use, capital has been that part of a man's stock from which he expects to derive an income" (Marshall, 1890, p. 53). He linked capital to income producing assets, excluding land, which is related to the fact that he, like the classics, is thinking of the tree main classes associated with capital, labour and land. It is interesting to note that Marshall clearly had in mind the link between capital and national income, which is behind National Accounts' definition (see section 3.1) and is the backbone of wealth distribution methodologies such as the capitalization method (Saez and Zucman, 2016).[12]

If Marshall represents a bridge between classical and neoclassical theories, it is with Fischer that capital is really placed within neoclassical walls and acquires the features that we now attribute to the canonical capital theory. In probably one of the most widely read economics handbooks of the twentieth century, Samuelson and Nordhouse (2009) argue that the modern idea of the capital stock (and interest rate) being determined by the the interplay between psychological preferences and diminishing marginal returns presented in section 2.3 is attributable to Fischer (1906, 1907). This locked capital theory well within the roam of scarcity theory of value (see section 3.3), hence away from any association between capital, power and class.

Following Cohen and Harcourt (2005), most of capital controversies originate in the dual nature of capital. i.e. whether it is a set of items used in production or an sum of money. The classics, beginning with The Wealth of Nations (Smith, 1776, ch. 9), talked about the 'stock' when referring to capital. Hodgson (2014) complains that it was with Smith that the term capital was corrupted forever, abandoning the use given for centuries by businessmen, for whom capital meant a collateral, and replaced it for 'stuff'. However, in the classics the idea of capital is also associated with a 'fund' that the capitalists advance to begin production and which also includes wages (Dobb, 1937).

For Marx, the discussion of what capital is is inseparable from the theory of value. It is related to the process of creation and appropriation of surplus value, and as we will discuss below and has been extensively argued, it is a process, not a thing. As such, it can be either money or commodities, depending on the purpose of its use –as opposed to its nature– and on which stage of the Money-Commodity-Money cycle one is located (see 3.4).

> "If we pin down the specific forms of appearance assumed in turn by self valorizing value in the course of its life, we reach the following elucidation: capital is money, capital is commodities. In truth, however, value is here the subject of a process in which, while constantly assuming the form in turn of money and commodities, it changes its own magnitude, throws off surplus-value from itself considered as

[12] He makes it explicit by stating that "labour together with capital and land thus defined are the sources of all that income of which account is commonly taken in reckoning up the National Income" (Marshall, 1890, p. 54-55).

> original value, and thus valorizes itself independently. (...) By virtue of being value, it has acquired the occult ability to add value to itself. It brings forth living offspring, or at least lays golden eggs" (Marx, 1867, p. 255).

It is interesting to note that, while capital may well be a sum of money in Marx's view, or even the net worth of a firm's balance sheet, a pure accounting definition –i.e. the sum of owned net assets– is not enough to define capital.[13] However, as pointed out by Paul Sweezy, capital is for Marx essential one-dimensional, i.e. "'capital' is not simply another name for means of production; it is means of production reduced to a qualitatively homogeneous and quantitatively measurable fund of value. The concern of the capitalist is not with the means of production as such, but with capital, and this necessarily means capital regarded as a quantity, for capital has only one dimension, the dimension of magnitude" (Sweezy, 1942, p. 338).

In his *Essentials of Economic Theory*, Clark (1907) distinguished 'pure capital' and 'capital goods', the first referring to the value of all capital and the second to the actual equipment. Veblen (1908), debating Clark but also Fischer, questioned them for their use of 'capital value' and 'capital'[14], which are "conceptually distinct, tho substantially identical". He stresses their contradictions (e.g. when discussing the transfer of capital, which makes sense as value but not necessarily as goods), but also the fact that they do not include 'intangible assets' or 'immaterial wealth', which are a part of pecuniary definitions of capital. He has the classic's and Marx's notion of power in mind, when he claims that owners are able to "corner the wisdom of the ancients and the accumulated experience of the race" (Veblen, 1908, p. 154), and that this explains why any "natural law" of the share of capital in product is misleading.

> The returns actually accruing to him under competitive conditions would be a measure of the differential advantage held by him by virtue of his having become legally seize of the material contrivances by which the technological achievements of the community are put into effect. (Veblen, 1908, p. 167)

A few years later, the debate re-emerged, this time with Knight defending the neo-classical position and Hayek attacking it from the Austrian perspective who, unlike his predecessor Böhm-Bawerk, did not considered capital as heterogeneous items but as a fund of value that is

[13]"If I state, like for example Say, that capital is a sum of values, then I state nothing more than that capital = exchange value. Every sum of values is an exchange value, and every exchange value is a sum of values. I cannot get from exchange value to capital by means of mere addition. In the pure accumulation of money, as we have seen, the relation of capitalizing [Kapitalisieren] is not yet posited"(Marx, 1857).

[14]These are the terms used by Fischer, in Clark's writings they are capital and capital-goods.

'malleable and perpetual' (Hodgson, 2014).[15] But it was in the 1950 where these long-standing controversy took-off and acquired epic proportions. Joan Robinson wrote an incendiary article which questioned the definition of capital behind which neoclassical economists had been barricaded. She famously went after the neoclassical production function, such as the one presented in equation 2.3 of section 3.1, arguing at the very start of her piece that:

> "The production function has been a powerful instrument of miseducation. The student of economic theory is taught to write $O = f(L, C)$ (...). He is instructed to assume all workers alike, and to measure L in man-hours of labour; he is told something about the index-number problem involved in choosing a unit of output; and then he is hurried on to the next question, in the hope that he will forget to ask in what units C is measured. Before ever he does ask, he has become a professor, and so sloppy habits of thought are handed on from one generation to the next" Robinson (1954, p. 81).

This triggered a sequence of offensives and counter-offensives with no evident victor (Cohen and Harcourt, 2003), but with potentially important consequences anyway, which will be discussed in section 3.2.

Time and History

To discuss the theoretical implications of the above-mentioned heterogeneity of capital goods, it is necessary to understand the time-dimension of capital. Capital is intrinsically interlinked with time since (i) capital is invested in a production process that takes time; (ii) capital is consumed within one or several periods and (iii) the income stream resulting from capital investments are spread through several periods. This time-dimension is at the base of the *Wicksell effects*, which when interacted wit heterogeneous capital goods creates serious problems for the theoretical determination of the rate of return r, as I will discuss in section 3.2. The dimension of time, and the related but conceptually different dimension of change throughout time and more generally of history, where an important part of the capital controversies.

In the classical tradition, the only important distinction in capital is precisely the one between circulating and fixed capital, i.e. the part of capital that lasts more than one period and the one that is consumed within the production process. In Mill's words, the former is the one "which fulfills the whole of its office in the production in which it is engaged, by a single use",

[15] For a survey of the Hayek-Knight debate, see Cohen (2003).

and the latter "exists in a durable shape and the return to which is spread over a period of corresponding duration" (Mill, 1848). In the second book of *Capital*, Marx argues against this distinction, since for him conceals the true origin of value, but also discusses capital's rotation period, i.e. the time it takes for an investment produce surplus value and how it affects the rates of profit (more on this in 3.4).

It was the Austrians who used the concept of time in the most radical way, merging it with capital in such a way that capital was time, reaching its most developed form in Hayek (1941). The idea was that profits were a product of time itself and the different investment's time-structure was central for the Austrians, although heavily criticized. Solow for one claimed that "I think the Austrian school overdid the identification of capital itself (and capital theory) with time –it was an inspired simplification that didn't work– but the need for a theory of capital does arise only when we try to take account of production processes which involve time in some essential way (Solow, 1963, p. 11)."[16]

But the dynamic effect was also big part of the assault to the neoclassical definition of capital, especially during the more advanced stages of the Cambridge Capital Controversy. Post-Keynesians such as Joan Robinson criticized capital and growth theory in general from the methodological standpoint, arguing against the very idea of an equilibrium and, more importantly, that a series of equilibrium can adequately capture the path between them. She set a distinction between logical time and historical time, which was rooted in her readings of Marx (Alves, 2022). In Robinson's words, "the problem of the 'measurement of capital' is a minor element in the criticism of the neo-classical doctrines. The major point is that what they pretend to offer as an alternative or rival of the post-Keynesian theory of accumulation is nothing but an error in methodology—a confusion between comparisons of imagined equilibrium positions and a process of accumulation going through history" (Robinson, 1974, p. 213).

Finally, the historical specificity of capital is another dividing line between competing theories. Marx is one of the main advocates of this view, arguing that capital is not only "the all-dominating economic power of bourgeois society" (Marx, 1857), but also that is specific to it. Thus, there is no capital outside capitalism –neither theory of value (Sweezy, 1942). This notion of capital inextricably linked with the theory of value and hence specific to the capitalist mode of production is quite different from the current consensus, which focus on things that can be owned and sold rather than in the social relations of production. Thus,

[16] On Hayek's Pure Theory of Capital, Mirowsky said that it was "a deeply flawed attempt to explicate the Austrian theory of capital using simple deterministic geometric models" (Mirowski, 2002).

both financial assets and slaves can be considered as part of the evolution of aggregate capital and compared through time (Piketty and Zucman, 2014). However, this does not seem to represent a rupture with the classical political economy prior to Marx, which also considered capital as things that were useful for production regardless the historical phase. Capital hence existed in all places and times, since Smith's "rude state of society" (Smith, 1776, ch. 6). As Ricardo made explicit, "even in that early state to which Adam Smith refers, some capital, though possibly made and accumulated by the hunter himself, would be necessary to enable him to kill his game" (Ricardo, 1817, p. 17). The recourse to the rude state of society is typical of classical political economy but survived it, very clearly so for instance since the early chapters of Clark (1907), but has been repeatedly criticized. For instance, when analyzing Clark's work, Veblen says, emphasizing by the way the role of immaterial assets:

> "The best excuse that can be offered for these excursions into 'primitive life' is that they have substantially nothing to do with the main argument of the book, being of the nature of harmless and graceful misinformation. (...) The 'capital' possessed by such a community-as, e.g., a band of California 'Digger' Indians-was a negligible quantity, more valuable to a collector of curios than to any one else, and the loss of which to the 'Digger' squaws would mean very little. What was of 'vital concern' to them, indeed, what the life of the group depended on absolutely, was the accumulated wisdom of the squaws, the technology of their economic situation (...) growth contemplated by Mr. Clark. The 'natural' system of free competition, or, as it was once called, 'the simple and obvious system of natural liberty', is accordingly a phase of the development of the institution of capital; and its claim to immutable dominion is evidently as good as the like claim of any other phase of cultural growth" (Veblen, 1908, p. 152-154).

3.2. The Cambridge capital controversy and the determination of r

In the previous section, a number of issues concerning the definition of capital were discussed. Two of them, i.e. the fact that capital is actually composed of very heterogeneous items, and the fact that it operates and that 'produces' income throughout time, creates theoretical problems for its valuation, and (perhaps more importantly) for the determination of r. This was only one of the vast and sometimes confusing number of issues debated during the Cambridge Capital Controversies, but since it is the key for the present discussion, I will focus on it. The participants included scholars such as Joan Robinson, Luigi Pasinetti, Pierangelo Garegnani and Piero Sraffa on one side, versus Paul Samuelson, Robert Solow, Christopher Bliss and Frank Hahn on the other, in a debate that, for all practical purposes, seems to have been the last time the classical political economy tradition offered and organized defense.

The problem is best explained recurring to 'Samuelson's parables'. While still arguing that capital theory could actually be developed assuming capital is composed of heterogeneous items, i.e. without recurring to the "Clark-like concept of aggregate 'capital' " (see section 3.1), Samuelson (1962) offered a "simple neoclassical capital models in a rigorous and specifiable sense can be regarded as the stylized version of a certain quasi-realistic MIT model of diverse heterogeneous capital goods' processes" (Samuelson, 1962, p. 201-202). He used surrogate production functions or *as if* productions functions, which predict how complex heterogeneous capital models can work as if they had come from a simple generating production function.[17] Based on this approach, he produced the famous three parables of the neoclassical production function[18], which are: (i) the real return on capital is determined by diminishing marginal productivity of capital; (ii) a greater quantity of capital leads to a lower marginal product of additional capital and thus to a lower rate of interest, and the same inverse, monotonic relation with the rate of interest also holds for the capital/output ratio; (iii) the distribution of income between laborers and capitalists is explained by relative factor scarcities and marginal products.

These parables are true under the surrogate production function or, what is by all practical purposes the same, under a one-commodity model. However, they do not hold once the one-good assumption is lifted. The basic problem is that once one considers heterogeneous capital goods, they need to be valuated in order to be aggregated and incorporated in the model. Valuation, in turn, is interlinked with time as discussed in section 3.1, since whether value is measured as cost of production or as present value of future output, time is involved. Thus, one needs the interest rate to perform the valuation, but the interest rate is itself endogenous, i.e. is the result of the quantity of 'capital'. There is an inherent circularity, by which one needs the interest rate to determine the interest rate, or the rate of return r for that matter. This problem is behind *reswitching* and *capital reversing*, which were two of the main technical objections to the neoclassical production function and capital theory in general.

Reswitching occurs when the same physical capital/labour ratio is preferred at two or more interest rates, while a different one is preferred at intermediate ones, therefore switching to a different technique (capital/labour) and then reswitching back. Capital-reversing implies that a lower capital/labour ratio is associated with a lower interest rate, which is equivalent

[17] As Samuelson says, Nicholas Kaldor –making fun of his own terminology– wrote to him that he was trying to pretend Clark could be defended as a 'stylized Samuelson'.

[18] It is interesting that Samuelson claims that, at this point, they should be called 'neo-neo-classical'.

to having lower interest rates when capital is more scarce, and thus an upward slope demand curve for capital. Reswitching thus violates parables 1 and 2, and capital-reversing violates parables 2 and 3. The main takeaway is that changing quantities are no longer associated with unambiguously signed price effects.

Samuelson (1966) provides a toy example, which in his words "tells more simply the full story of the twenty-fifth and eighth degree polynomials of the Sraffa-Pasinetti example of reswitching", in which Champagne is produced using only labour and time. When different amounts of labour are organized in different time arrangements (in the example, 7 labour units in $t-2$ versus 2 labour units in $t-3$ and 6 in $t-1$) to produce the same output, the cost-minimizing technique in not unambiguously determined. At either very high or very low interest rates, the first technique is chosen, while for intermediate ones the second one is preferred. This problem crates the theoretical possibility for multiple equilibria, which make the one-way scarcity principle useless to explain r and thus the income distribution. Moreover, reswitching "leads to a pattern of capital/output ratio that fails to move in one direction". This is very important in the context of the our discussion, since these patterns are precisely the foundation of the model that determines r, which is in turn critical in explaining macro and micro income distributions, as discussed in section 2.3.

Naturally, there were a number of replies and counterattacks, which were inconclusive (Cohen and Harcourt, 2003) and are beyond the scope of this essay. The key point is that a relevant and substantial critique was delivered to the neoclassical model, which especially affects its ability to unambiguously determine key distributional variables such as r, α and β presented in 2.2. More importantly, these points were taken by the neoclassical side, with unequivocal clarity in most cases (see e.g. Solow (2000)). As Samuelson himself pointed out in his 'summing up', "if all this causes headaches for those nostalgic for the old time parables of neoclassical writing, we must remind ourselves that scholars are not born to live an easy existence. We must respect, and appraise, the facts of life" (Samuelson, 1966, p. 583).

While important at the moment and for my purpose in this essay, these technical issues were only at the surface of the debate, which in reality was about deeper concerns by the critiques of the neoclassical school. In the end, for people such as Joan Robinson, it mattered more what capital was than its corollary, i.e. how to measure it (Harcourt, 2014). To dive into that level, one needs to address the underlying problem of the theory of value, which is behind the theory of capital and of our whole understanding of the functioning of the economy.

3.3. Capital and the theory of value

To understand the effect of the theory of value in neoclassical capital theory, it is necessary to start by explicitly establishing the aim and requisites for a theory of value. This is not easy, since the focus of mainstream economics is not on value, but on prices. Price theory establishes how production, consumption, distribution and exchange interact to determine prices under a general equilibrium framework (Cohen and Harcourt, 2005). They may depend on preferences, endowments and technology as in the neoclassical case, or in costs and production as in the classical. Value theory, on the other hand, aims at accounting for the underlying forces at stake, and therefore of the ultimate determinant of prices, as opposed to its mutual interdependence.

Following Dobb (1937), a theory of value needs to solve the problem of value of commodities and the distribution of income at the same time, based on some external variable that is itself not a value in a general long-run equilibrium setting. It should be sufficiently obvious that the value of something cannot be explained, i.e. determined in a causal way, by other values, which are at the same time determined by the former. In partial short-run equilibrium models, this may not be difficult, since a number of elements can be considered as unalterable, such as wages or the amount of capital. But in the long run, all these elements of the system may change. What can be considered as such an external variable? The classical political economy found this independent variable in an objective element of the production process, i.e. labour, while the neoclassical in a subjective factor underlying consumption and demand; the former resulting in the labour theory of value, and the latter in the scarcity theory of value.[19]

The father of the labour theory of value was Ricardo, after an insightful but unclear early discussion in The Wealth of Nations (Smith, 1776). For his discussion of value, Ricardo focused his attention on those commodities that could be produced by human effort, explicitly leaving aside the relatively unimportant –both numerically and conceptually– scarce goods: "in speaking then of commodities, of their exchangeable value, and of the laws which regulate their relative prices, we mean always such commodities only as can be increased in quantity by the exertion of human industry, and on the production of which competition operates without restraint" (Ricardo, 1817, p. 9). Neoclassic economists, in turn, focused their attention precisely on those goods which were taken out of the analysis by the classics. i.e. scarce goods.

[19]The scarcity theory of value is also called *subjective* theory, but in this point I will follow the names given by Cohen and Harcourt (2005), since they provide greater clarity to the present discussion.

Pasinetti (1983) argues that this distinction can be traced back to the trade and industry phases: one static, focused on one time-gains and subject to maximization; the other dynamic, focused on learning and production. This opinion, regardless of its validity, does show what are the main dividing lines between the two, i.e. one focused on maximization of utility given an endowments restriction, and the other on production and the conditions for expanding it. These two approaches were in fact so opposed to each other that manifested into different definitions of economics: one that defines the object of economics as the study of efficient allocation of scarce goods, popularized by Robbins (1932), which was at odds with the classical view of political economy as the study of the economic process (see e.g. Mill (1836)).

Following Pasinetti (1983), it is interesting to note that the classics would probably not have objected the neoclassical approach, and would not have denied that there exists a rational problem of allocation of scarce goods. But they would not have made it more than a minor problem, and an easy one to solve . However, neoclassic economists went too far and began to "advance a disproportionate claim", i.e. to apply this to all branches of the economic analysis. In the 1890s, the theory of marginal productivity emerged as an extension of the field of optimum allocation of scarce goods. In particular, they end up considering capital as equivalent to land (or labour) as a scarce good: they were all factors of production now.

Under this approach, capital is considered as a factor of production and its price is determined by the relative scarcity of capital. Moreover, "the scarcity theory of value entails, *ceteris paribus*, a unique inverse relationship between a commodity's quantity and its price" (Cohen and Harcourt, 2005). Thus, by extending the scarcity theory to the production, factorial income distribution becomes a subset of the general theory of price determination. As I discussed in section 3.2, the Cambridge controversies showed that there is no unequivocal one-direction relation between scarcity of capital and its price, i.e. the interest rate. It ends up in a circular argument under relatively more general models than the one-good model. Thus, the scarcity theory of value cease to be effective as an explanation of value *and* of price. Prices can be determined by solving a number of simultaneous equations, but this is not enough to *explain* them.

Thus, the extension of the scarcity theory of value to all areas of the economic analysis and, in particular, to the capital theory, entails major difficulties for our understanding of capital and its relation to income distribution. The irruption of such a theory, and its blitzkrieg-like success is in itself an study matter (Mirowski, 2002), but it is hard not to associate it with the almost simultaneous challenge proposed by Marx's *Capital* (Pasinetti, 1983; Dobb, 1937).

Whether Marx was a classical economist or not is debated[20], yet it seems evident that it had an impact on economic theory, both as a result of his studies but probably more so because of the reaction it generated, since:

> "Marx's overall arguments were not easy to challenge. The obvious procedure to follow would have been to question the premises. But this is precisely what was so difficult. Marx's premises were exactly the same as those of Smith and Ricardo, i.e. of all established economists" (Pasinetti, 1983, p. 13).

Certainly, a theory of value which could make production virtually invisible could do the trick. By introducing profit-maximization behavior to firms, the neoclassical model defines a correspondence between demand of commodities and demand of (scarce) production factors. Everything in between is downplayed, and "for all essential purposes, the model had, so to speak, eliminated the process of production from the analysis" (Pasinetti, 1983, p. 16). In the next section, I explore Marx's labour theory of value and some of its main caveats, in order to establish the background for the accounting correspondence discussion of section 4.

3.4. Marx the accountant

In this section I go over some of the main definition of Marx's theory of value, which is explicitly intertwined with his understanding of capital. Only a brief account of main variables is presented, not an in depth explanation (which would be impossible), nor a discussion of the validity of the assumptions or predictions. Out of the vast marxian contributions, only the definitions and issues that are directly related to the main point of this essay will be discussed. Given the multitude of often contradictory interpretations of Marx's work, I rely mostly on Marx's Capital, making the strong assumption suggested by Sweezy (1942) that he actually meant what he said.

In the foreword of her 1941 edition of the *essay on Marxian economics*, Joan Robinson described the spirit of this section, when she wrote:

> "the purpose of this essay is to compare the economic analysis of Marx's *Capital* with current academic teaching. The comparison is, in one sense, a violent anachronism, for the development of Marx's thought was influenced by controversy with his own contemporaries, not with mine. But if we are interested, not in the

[20] In one view, "objectively, Karl Marx was a Classical economist in the full sense of the word", yet "subjectively [...] he used classical theory for purposes which were diametrically opposite to those of the classical economist." (Pasinetti, 1983, p. 12).

historical evolution of economic theory, bit in its possible future progress, this the relevant comparison to make" (Robinson, 1942).

On top of the 'violent anachronism', I focus on Marx's accountancy of value and capital, stripping it out of most of his most interesting insights. The starting point to understand Marx's accounting is the fact that capital necessarily describes a cycle of Money-Commodity-Money, or M-C-M for short, in which the capitalist has money, buys inputs and hires labour, to sell the product in exchange for money again. As discussed in section 3.1, capital is not money nor commodities, but this movement: "money which describes the latter course in its movement is transformed into capital, becomes capital, and, from the point of view of its function, already is capital" (Marx, 1867, p. 248). This circulation is not for consumption, as would be the case in the cycle C-M-C, where the entry point is a use-value[21] and exit with a different one. In the capital movement, the beginning and ending of the cycle are qualitatively the same (money), so they can only differ in terms of their exchange value. Only a quantitative difference justifies such an operation, hence capital describes the M-C-M', where $M' = M + \Delta M$, being ΔM the *surplus value*.

> "In simple circulation, the value of commodities attained at the most a form independent of their use-values, i.e. the form of money. But now, in the circulation M-C-M, value suddenly presents itself as a self-moving substance which passes through a process of its own, and for which commodities and money are both mere forms. But there is more to come: instead of simply representing the relations of commodities, it now enters into a private relationship with itself, as it were. It differentiates itself as original value from itself as surplus-value, just as God the Father differentiates himself from himself as God the Son, although both are of the same age and form, in fact one single person; for only by the surplus-value of £10 does the £100 originally advanced become capital, and as soon as this has happened, as soon as the son has been created and, through the son, the father, their difference vanishes again, and both become one, £ 110" (Marx, 1867, p .256).

In this sense, Marx establishes a clear difference between the concepts of capital and what we now call wealth. A sum of money, not invested to result in more money, is wealth but not capital in the marxian sense. This is important, because it is not even related to the *power* over the production process as in Atkinson (2015). It is related to the movement and its aim,

[21]In Marx, commodities have use-values insomuch they satisfy any human need or want, while exchange value –or plain value– is an exchange relation with other commodities.

since “if the £110 is now spent as money, it ceases to play its part. It is no longer capital. Withdrawn from circulation, it is petrified into a board, and it could remain in that position until the Last Judgment without a single farthing accruing to it (Marx, 1867, p .252)”.

Capital is intertwined with value, insofar as to explain M-C-M' one needs to explain the origin of ΔM. As discussed in section 3.3, in order to explain exchange values, a theory of value needs an external element that is not itself a value. For Marx, as a ‘pupil of Ricardo’ (Schumpeter, 1942), this element is labour. The intuition is that value cannot emerge from circulation alone (i.e. from buying and selling more dearly), since anyone’s gains cancel-out with someone else’s loss. Value hence emerges while combining commodities in the production process, as a part of M-C-M'. “Capital cannot therefore arise from circulation, and it is equally impossible for it to arise apart from circulation. It must have its origin both in circulation and not in circulation (Marx, 1867, p. 268)”. Classical political economy, especially Ricardo, found a circularity when trying to explain value as labour, which lead them to slide to the more general concept of ‘cost of production’. In an introduction to Marx’s ‘Wage Labour and Capital’, Engels explains that “as soon as the economists applied this determination of value by labour to the commodity ‘labour’, they fell from one contradiction into another. How is the value of ‘labour’ determined? By the necessary labour embodied in it. But how much labour is embodied in the labour of a labourer of a day a week, a month, a year. If labour is the measure of all values, we can express the “value of labour” only in labour” (Engels, 1891). Marx’s solution to this problem, was to notice that the capitalist does not hire labour, but labour-power, which, like every other commodity has a price.

> “The rock upon which the best economists were stranded, as long as they started out from the value of labour, vanishes as soon as we make our starting-point the value of labour-power. Labour-power is, in our present-day capitalist society, a commodity like every other commodity, but yet a very peculiar commodity. It has, namely, the peculiarity of being a value-creating force, the source of value, and, moreover, when properly treated, the source of more value than it possesses itself” (Engels, 1891).

Thus, if the exchange value of every commodity is its embedded labour, then labour-power’s value also is: the labour needed to reproduce the labour-power, i.e. the worker herself. Since the value of the commodities needed to maintain and reproduce workers is lower than what they actually produce during a working day, a surplus value emerges. This surplus value is a product of surplus labour, which results from the part of the day in which the worker is producing commodities that surpass the value of her labour-power. Labour-power, as every other commodity, is paid at its value[22], but is the only commodity that has the ability to produce more than it costs.

Marx thus defines two components of capital: constant capital C and variable capital V. Constant capital includes all the equipment and inputs used in production, while variable capital is the hired labour-power. Constant capital does not generate value, only transfers it in its interaction with labour. "Here, the dead labor that is present alongside living labor in the production process appears as an alien and hostile power — as capital. (...) Capital is now more than a claim on surplus; it has become a tangible force that drains the worker of all energy and cripples all his talents" (Elster, 1986, p. 55). Variable capital, in turn, "does undergo an alteration of value in the process of production. It both reproduces the equivalent of its own value and produces an excess, a surplus-value" (Marx, 1867, p. 317). Labour, by producing more value than it costs to the capitalist, generates a surplus value S. Total value is therefore the sum of C, V and S.

From this components of value, three key relations are drawn. The first is the rate of surplus value $S' = {}^{S}/_{V}$, which is equivalent to the rate of exploitation, in the sense that it provides an account of how much the capitalist keeps of the value produced by workers.[23] If the working day is eight hours long and the necessary work is 4 hours, then $S' = {}^{S}/_{V} = {}^{4}/_{4} = 100\%$. A second important relation is the organic composition of capital $Q = {}^{C}/_{V}$[24], which is a measure of the extent of inputs and equipment used in relation to labour-power. Finally, Marx's rate of profit $P = {}^{S}/_{C+V}$, which is the surplus value the capitalist keeps in relation to the totality of his investment, in both labour-power and inputs/equipment.

Noting that the rate of profit can also be written as $P = {}^{S'}/_{1+Q}$, is straightforward to visualize Marx's law of the falling tendency of the rate of profit. If one assumes a constant rate of surplus S', then P varies inversely with the organic composition Q. Assuming that capitalist development entails an upward tendency of Q, given by growing need for equipment and inputs in relation to labour, then it results that P falls. Marx enumerates a number of 'counteracting forces' to this law, from which one of the most important ones for my purposes is that the rate of surplus S' may increase, i.e. that workers may be more exploited. The assumption of constant rate of surplus is very demanding, since there are a number of reasons why it may increase, even as part of capitalist development. In fact, Marx stresses over and

[22]This is the key assumption of the first volume of Capital.

[23]It is interesting to note that Joan Robinson fully incorporated the idea of exploitation in her essay on marxian economics (Robinson, 1942), without ever accepting the labour theory of value as such (Alves, 2022)

[24]Some authors express this relation as $Q = {}^{C}/_{V}$ (Sweezy, 1942), which does not change this results, nor the ones presented in 4.

over again that the very "same causes that bring about a fall in the general rate of profit provoke counter-effects that inhibit this fall, delay it and even paralyze it (Marx, 1894, p. 346). In this sense, Sweezy (1942) is right in saying that he should have included these counteracting factors as part of the law itself, given their importance. In any case, it is worth pointing out that Marx is evidently thinking of a tendential law, in the sense of the classical methodological approach (Mill, 1836) and also the ones that followed, typically Marshall (1890).

One final important remark is that, just as neoclassical economists, Marx's theory of value is challenged when abandoning the one-commodity model. In fact, the problem emerges as soon as one adds a multi-sector production (or departments, as Marx would call them). Again, given S' is the same across sectors, since workers would move from one industries were they are less exploited in the sense described above, then if Q varies across sectors, P should be different too. But this is a contradiction insofar capitalist would move to those sectors with the higher rate of profit. The theory would require, as in Ricardo (1817), the assumption that the organic composition is the same across sectors, which is clearly untrue.

In the second part of the third volume, Marx provides an explanation based on 'production prices', which are the results of the *average* rate of profit of the system to the capital invested $(C + V)$. In this way, the assumption that goods are sold at their values is lifted and now commodities are exchanged at their prices. Thus, capitalist with higher organic compositions Q, would be compensated by appropriating not only the surplus value produced in their firms, but also that of others, who will be the losing end of the thread. This procedure entails that in the end each capitalist will obtain a share of the overall surplus value produced by the system according to his own total capital investment. This 'solution' has been criticized and several alternatives were proposed (Sweezy, 1942), but I will not go beyond this succinct presentation.

It is worth mentioning, following Dobb (1937), that it is not that Ricardo or Marx were not aware of the potential problems of different capital compositions (organic compositions) generated, it is more likely (and more relevant to us here) that they considered them relatively unimportant given their objectives. They were not interested in determining the prices of specific commodities, but rather on "macroscopic problems" (today we would say macroeconomic), i.e. determining rents, profits and wages and values of commodities in relation to those. It is at this high aggregation level that I discuss the accounting equivalence in section 4.

4. The one-commodity-model accounting equivalence

4.1. The one commodity model

I now focus on establishing the accounting equivalence between the macro distributional model presented in section 2 and the Marx's value categories of section 3.4. It is worth stressing that these are by no means conceptual equivalences, in the sense that they do refer to different (and often opposite) ways of understanding the functioning of the economic system, the underlying theories of value and the very definition of what *economics* is. The aim is more modest, and is simply to understand the conditions under which there is accounting equivalence between these two approaches.

The setting under which some degree of equivalence holds is the one-commodity model. This is not a novelty, insofar "by simplifying so extensively, one-commodity models allow many competing theories to demonstrate their results" (Cohen and Harcourt, 2005, p. 44). However, clearly comparing the two may help understand their conceptual differences while at the same time establishing if some of their predicted outcomes are incompatible or not.

More importantly, the one-commodity model is very important in particular for the marxian analysis, since it is not just a simplifying assumption as may be in the neoclassical model. The one-commodity model is a synonym in this case of considering the system as whole, and it is at the system level where the rate of profit P is determined, independently of what individual rate of profit each capitalists would face if commodities were to be sold at their values. At this level, profits and surplus value coincide, and so does aggregate production measured with prices or values. The one-commodity model is the key link between value and price, and therefore not a working assumption to be lifted and abandoned at later stages, but a central feature for understanding of the economic process. Thus, the one-commodity-model has a paramount conceptual importance in Marx: it is were the theory of value manifests directly. Indeed, the point could be made that regardless of the way in which each class or even each individual ends up getting their share of aggregate net income and of how large is such a share, at the system level the theory of value holds true and best captures some basic characteristics of the functioning of a modern economy, as noted by Sweezy.

> "The entire social output is the product of human labor. Under capitalist conditions, a part of this social output is appropriated by that group in the community which owns the means of production. This is not an ethical judgment, but a method of describing the really basic economic relation between social groups. It finds its most clearcut theoretical formulation in the theory of surplus value.

> As long as we retain value calculation, there can be no obscuring of the origin and nature of profits as a deduction from the product of total social labor. The translation of pecuniary categories into social categories is greatly facilitated. In short, value calculation makes it possible to look beneath the surface phenomena of money and commodities to the underlying relations between people and classes (Sweezy, 1942, p. 129)".

At this point, it is worth stressing two important conceptual differences between some of Marx's categories of section 3.4 and the variables discussed in Section 2, which will be key in understanding the accounting equivalence of the reminder of this section. First, besides the obvious fact that capital in Marx includes wages (as in most classical political economy) while the neoclassical theory does not, it is important to note that Marx's capital is more a flow than a stock. This is not only because for Marx capital describes a certain 'movement' M-C-M', but also because being the sum of $C + V$ is literally closer to output than to net worth in SNA's terminology. Marx's capital in its money form accounts for wages paid to workers, inputs of production and the wearing out of equipment in the period of production. These are all flow-like variables. The stock of equipment is behind the depreciation of the equipment, but does not enter in calculation as a stock. In Marx's words:

> "The value component c, which represents the constant capital *consumed* in the course of production, is not the same thing as the value of the constant capital *applied* in production. The materials of production are certainly completely consumed, and their value is therefore entirely transferred to the product. But only a part of the *fixed* capital is entirely consumed, its value thereby being transferred to the product. Another part of the fixed capital in machines, buildings, etc. continues to exist and to function just as before, even if its value is diminished by the annual wear and tear. This part of the fixed capital that continues to function does not exist for us when we consider the value of the product. It forms a part of the capital value that is independent of this newly produced commodity value and is present alongside it." (Marx, 1885, p. 472, emphasis in original).

The preceding comment results in the second point, which is that the neoclassic rate of return r is a return on the *stock of capital* which does not include wages, while for Marx is a return on the invested flow. Theses two variables are, although similarly named, quite different from a conceptual view point. Moreover, form a quantitative standpoint it is not evident at first sight which should be higher. Since r is expressed in terms of a stock, one could think that it should be lower than P. However, although P is the surplus value in relation to a flow, this denominator is very large insofar it includes all the economy's wage bill, plus depreciation plus all inputs. In the reminder of this section, I compare them more systematically in order to understand their role in their respective theories.

4.2. Equivalent aggregate variables

Before establishing the equivalence, I lay out the main assumptions and recall the variables to be compared from each approach.

Assumptions. We assume that there is no increase in population or hours worked and a one good model. Moreover, I assume that there are no inputs, i.e. there is only fixed capital in constant capital. The later assumption can be lifted with no relevant consequences, as shown in Appendix A. Moreover, it makes sense to exclude it since inputs (or circulating capital) are not part of the neoclassical production function, which considers only new output or added value.

The neoclassical variables. We begin by considering a generic production function.

$$Y = F(K, L)$$

F can be any production function. From an income perspective, considering Y_g gross income (eq. 4.1) and Y_n net income (eq. 4.2), we get:

$$Y_g = wL + rK + \delta K \tag{4.1}$$

where L and K are labour and capital, w and r are the wage and interest rate, while δ is the depreciation rate in terms of aggregate capital.

$$Y_n = wL + rK \tag{4.2}$$

Considering a depreciation rate d, which is very similar to the SNA´s concept of consumption of fixed capital[25], expressed in terms of Y_g, i.e $d = CFK/Y_g$, we get:

[25] There are some differences which are unimportant in this context, see WIL (2020).

$$Y_g(1-d) = Y_n$$

Normalizing equation 4.1 in Y_g, we get:

$$\frac{Y_g}{Y_g} = 1 = \frac{wL}{Y_g} + \frac{rK}{Y_g} + \frac{\delta K}{Y_g}$$

Hence

$$\frac{wL}{Y_g} + \frac{rK}{Y_g} + \frac{\delta K}{Y_g} = \frac{wL}{\frac{Y_n}{(1-d)}} + \frac{rK}{\frac{Y_n}{(1-d)}} + \frac{\delta K}{\frac{Y_n}{(1-d)}} = 1$$

Re-arranging:

$$(1-\alpha)(1-d) + \alpha(1-d) + \delta\beta(1-d) = 1 \tag{4.3}$$

With α being the capital share of net income, i.e. $\alpha = \frac{rK}{Y_n}$, and β being the capital to income ratio $\beta = \frac{K}{Y_n}$.

The labour theory of value variables.

Aggregate value W_{agg} is:

$$W_{agg} = S_{agg} + V_{agg} + C_{agg}$$

These values may be expressed in monetary terms to facilitate exposition, but actually they refers to embedded labour, so their magnitudes are unchanged at any level of production. This means that there is no Y_g or Y_n in Marx (as in Ricardo (1817, ch. 20)), only value. Normalizing,

$$\frac{W_{agg}}{W_{agg}} = 1 = \frac{S_{agg} + V_{agg} + C_{agg}}{W_{agg}} = S + V + C$$

we get a macro-distributional definition, in which S, V, and C are shares of aggregate value.

Expressing the different components or value as shares and not embedded labour does not affect computations, since we are always expressing these quantities in relation to others. Moreover, by presenting them as shares the distributional implications of the analysis are best captured.

The accounting correspondence

It has been long acknowledged that there is room for mapping marxian categories to standard economic variables, at least at high level of abstraction. Following Sweezy, it is clear that "Marx's theory of value has the great merit, unlike some other value theories, of close correspondence to the actual accounting categories of capitalistic business enterprise" (Sweezy, 1942, p. 63), as $V + S$ corresponds to net national income, which "includes payments to individuals plus business savings".

This holds no only overall, but also in V and S separately. "If the labour-power has performed its function, then the capital no longer consists of labour-power on the one hand and means of production on the other. The capital value that was laid out on labour-power is now value which has been added to the product (together with surplus-value)" (Marx, 1885, p. 300). For Marx, $V + S$ represent the value equivalent of net national income, since C is the equivalent of consumption of fixed capital and inputs. This is discussed extensively in Shaikh et al. (1997), which represents the most thorough analysis of correspondence between National Accounts and Marxian variables. Although his aim goes far beyond the scope of this analysis[26], he establishes a direct link between key value components in the two approaches under the assumptions of "production sectors alone", which is close to the one-good model assumed here. In this setting, it is hence straightforward that:

$$\frac{wL}{Y_g} = V = \frac{wL}{\frac{Y_n}{(1-d)}} = \frac{wL(1-d)}{Y_n} \rightarrow \boxed{(1-\alpha)(1-d) = V} \tag{4.4}$$

$$\frac{rK}{Y_g} = S = \frac{rK}{\frac{Y_n}{(1-d)}} = \frac{rK(1-d)}{Y_n} \rightarrow \boxed{\alpha(1-d) = S} \tag{4.5}$$

$$\frac{\delta K}{Y_g} = C = \frac{\delta K}{\frac{Y_g}{(1-d)}} = \frac{\delta K(1-d)}{Y_n} \rightarrow \boxed{\delta\beta(1-d) = C} \tag{4.6}$$

[26] His objective is to "provide an alternate foundation for the measurement of the production of nations." This is done by mapping Marxian and national account's variables, and –more importantly– by systematically including the classical distinction of production and consumption activities, which excludes sectors such as the military or the police from production sector.

4.3. The rate of return r and the rate of profit P

From the preceding accounting correspondences, it is possible to establish a comparison of the main driving forces of the system, i.e. $r > g$ in the neoclassical macro-distributional models and the falling tendency of the rate of profit P in the Marxian framework.

We start by deriving the organic composition of capital Q and the rate of surplus S'. First, we have that:

$$Q = \frac{C}{V} = \frac{\delta\beta(1-d)}{(1-\alpha)(1-d)} = \frac{\delta\beta}{(1-\alpha)} \rightarrow Q = Q(\overset{+}{\beta}, \overset{+}{\alpha}) \tag{4.7}$$

On the other hand:

$$S' = \frac{S}{V} = \frac{\alpha(1-d)}{(1-\alpha)(1-d)} = \frac{\alpha}{(1-\alpha)} \rightarrow S' = S'(\overset{+}{\alpha}) \tag{4.8}$$

Combining equations 4.7 and 4.8, we get that the system's profit rate P is:

$$P = \frac{S}{C+V} = \frac{S'}{1+Q} = \frac{\alpha}{(1-\alpha)+\delta\beta} \rightarrow P = P(\overset{-}{\beta}, \overset{+}{\alpha}) \tag{4.9}$$

The result from equation 4.7 is straightforward. The organic composition of capital C/V increases both when β or α grow, since in the first case it entails a larger C[27] through the effect of depreciation, while a larger α entails a higher capital share, i.e. lower V. The same idea applies to equation 4.8, since higher α is mechanically associated with increasing rate of exploitation in this setting. In the case of the system's rate of profit P, equation 4.9 shows that it is consistent with Marx´s insights of the tendency of the rate of profit to fall. When β increases, it pushes Q up and hence P down, but the rise in α increases Q too (and also S'), and thereby pressing P upwards.[28] Note that this is one of the main counteracting forces of the tendency of the rate of profits discussed in Marx (1894): the increase in exploitation S' offsets the tendency of the profit rate to fall. The net effect is hence unknown.

Three things are worth noting. First, under the Cobb-Douglas assumption, the increase in

[27] If depreciation is stable or increasing. In the case of decreasing depreciation, it will depend on the rate of increase of β and δ.

[28] This happens because the increase in α has more effect in S' than in Q, as in both the effect on the denominator is the same, but in the case of S' it also increases the numerator.

β has no effect in α insofar it pushes r downwards in the exact proportion so as to keep α unaffected (sec. 2.2), thus Marx's downward trend for P prediction holds. In the end, the Cobb-Douglas production function is the equivalent of a constant exploitation rate S', which resonates in Marx's assumption throughout *Capital*. However, in the more flexible CES production function assumption, the increase in α may push P upwards, given the elasticity of substitution of capital σ is larger than 1. Naturally, α has an upper limit (probably even before its absolute maximum of 1), hence Marx's point re-emerges, as eq. 4.10 shows:

$$\lim_{\alpha\to 1} P = \lim_{\alpha\to 1} \frac{\alpha}{(1-\alpha)+\delta\beta} = \frac{1}{\delta\beta} \tag{4.10}$$

Second, if an increase in β ends up forcing P downwards, it is interesting to note that this can happen with a constant r, provided it is larger than g. In Marx's terms, this is through its effect on $C = \delta\beta(1-d)$, i.e. the part of the stock of physical capital and inputs that goes into the flow of value C. This point is entirely missed if one focuses on neoclassical rate of return r instead of Marx's profit rate P. What is missed in Marx, in turn, is the fact that β (hence C, given d and constant or increasing δ) tends to grow mechanically when $r > g$, as discussed in section 2.3. This is absent in his analysis, because despite acknowledging that there is a stock of capital, it is not in his accounting framework (only the part that passes on to value C) and –even worse– he is not interested in accounting for changes in aggregate production (which he assumes as a feature of the capitalist system) and therefore of g.

This means that if $r > g$, and as a result both β and α increase (given the right elasticity of substitution), then in the long run the falling rate of profits P persists, even after accounting for the effect of the counteracting force of increasing α. This is clear when comparing r and P. Recalling that $\alpha = r \cdot \beta$ and equation 4.9, we have that:

$$\frac{r}{P} = \frac{\frac{\alpha}{\beta}}{\frac{\alpha}{(1-\alpha)+\delta\beta}} = \frac{\beta}{(1-\alpha)+\delta\beta} = \frac{\beta}{(1-\alpha)} + \frac{1}{\delta}$$

Except for extremely high values of α and implausibly low values of β (lower than one), $r > P$. Moreover, the gap between the two widens with increasing β and/or α. Thus, the world envisioned by Piketty (2014), with low growth and a constant r but $r > g$, entails both his result of growing β and α, but also the falling rate of profit P.

Third, if β varies inversely with the growth rate g (section 2.3), then low growth societies will mechanically amplify the downward pressure on the system's rate of profit P. Moreover, this is also true if we think of a higher obsolescence rate, which will increase the depreciating of capital δ, likely to happen in the next decades, and also if production becomes more inputs-intensive, which is the trivial yet important result of the generalization of the formula in Appendix A.

Once the differences between the rate of return and the rate of profits are presented, it is worth commenting on some implications for empirical distributional analysis. In particular, many studies use the (inverse of) the rates of return on different assets to compute the distribution of those assets based on observed capital incomes, in what is called the 'capitalization method' (see e.g. Saez and Zucman (2016); Garbinti et al. (2017); Martínez-Toledano (2017), as well as the wealth distribution chapter of this thesis). From an accounting perspective, these rates of return do in fact represent the stream of incomes accruing the owners of the assets in the Smith-Marshall sense (see 3.1), and therefore the estimates do provide important insights on the distribution of wealth and capital. What changes once one assumes that the labour theory of value is operating behind the scenes, is the way in which we understand it, i.e. not as the remuneration of the factor capital, but the appropriation of surplus value. This has important consequences from the view point of the justification for this remuneration. In Robinson's words, "with the notion of the supply price of capital, the moral justification of profit as a necessary cost of production disappears, and the whole structure of the orthodox apology falls to the ground." Robinson (1942, p. 62).

This idea is discussed in length by Marx at the beginning of the third volume of capital, in what he describes as describes the 'mystification of capital', which refers to the fact that even the capitalist himself misunderstands the origin of his profit as the result of his overall investment. He does not see that this profit results from surplus value, because he is not seeing variable capital as such, only a mass of investment which results in a flow of income. For Marx, this was also the case for classical economists such as Smith or Ricardo, who by considering fixed and circulating capital instead of constant a variable capital (see section 3.1), were unable to see the true origin of value. More generally, this results from the twin ideas of alienation and fetishism, which applied to capital result in the notion that capital produces value by itself. As Elster summarized:

> "Alienation-as-subjection, though closely linked to exploitation, is not equivalent to it. Alienation adds to exploitation a belief on the part of the workers that the capitalist has a legitimate claim on the surplus, by virtue of his legitimate ownership of the means of production. (...) Capital fetishism is the belief that capital's power to produce is a faculty inherent in it, not one it owes to the labor

process" (Elster, 1986, p. 56-57)

The phenomenon capital fetishism captures the main point of this essay. Under a one-commodity assumption there is consistency in some macro-distributive Marxian and neoclassical results as the one best described by Piketty and Zucman (2015). Yet while by using Marx's variables one is able to grasp the notion of where value is actually created and by whom, in the neoclassical setting this is obscured. The debate over the validity of Marx's theory of value is way beyond the scope of this short essay, but I do believe that allowing ourselves to look at empirical results on β, α, r and g through these lenses has the potential to open a fruitful conversation.

5. Concluding remarks

In this essay, I have tried to show that while our empirical distributional estimates are far better than they used to be, they are still anchored to a very questioned theory. It is a theory which, however useful for many purposes, is especially poorly equipped to deal with long-run macro distributional issues. This hinders our ability to understand past trends and look into the future.

It is true that it is not mandatory to buy the full neoclassical story to use it. Some general models, such as the one-commodity long run growth model, are useful as tools for thinking about some of the main economic variables and contribute to a certain data-based historical narrative. We can use these models as subsidiary to our empirical efforts without necessarily trusting them blindly. It is a very productive way forward, in which a lot of progress is still to be done. This thesis was in fact a part of such an effort.

But theories are important. We need them to interpret, understand and predict whenever possible. And however problematic they may be, they do need to grasp the main features of the economic system. The scarcity approach to value, which is behind theoretical models and especially the conceptualization of capital, is useful for many purposes. Yet it hardly accounts for some very basic facts about which are the drivers of the economic system. The functioning of modern economies cannot be modeled in the traditional marginalist fashion. The central variable to the system, which is the return to capital, cannot be theorized based on ahistorical atomistic individuals with perfect information maximizing their infinite-horizon utility –for whatever that is. But, as Pasinetti would say, "marginialism had the advantage of synthesis. For they have always clearly presented their arguments around a unifying problem (optimum allocation of scarce resources) and a unifying principle (the rational process of maximisation under constraints)" (Pasinetti, 1983, p.19).

Classical political economy, although outdated and flawed in many ways, still offers a much simpler starting point: that the economic system is driven by production decisions, heavily determined by profits and by the never-ending need to grow or disappear of capitalists, molded by the conflicting interaction of more or less stable sets of individuals who share certain roles in such a production. The return to capital in this context is the appropriation of socially produced value by the individuals who are in the position to call the shots. This was obvious for Shcumpeter, Robinson, Veblen, for most pre-marginalist economist of the 19th century and most importantly for Marx. Even though this essay was not about discussing Marx's theory, its potential or its limitations, I do want to stand by the idea that –as Robinson would say when comparing him with neoclassical school– Marx's "intellectual tools are far cruder, but his sense of reality is far stronger, and his argument towers above their intricate constructions in rough and gloomy grandeur" (Robinson, 1942, p. 2). In the end, it just seams more reasonable to assume that is human activity which is behind all human production and exchange –and therefore it determines both value of things and social relations–, instead of isolated individuals harmonically interacting with stuff in a conceptual vacuum.

The only point I wish to make is that there is room for thinking in political economy terms with the estimates we already have. That, for instance, if we observe an $r > g$ which results in an increasing β, that probably means that the rate of profit P is falling, which in turn can be offset by an increase in α but only for so long. And this opens a whole new conversation. More importantly, if we believe that the labor theory of value better explains these estimates than the scarcity theory, then it is no longer true that wealth can be accumulated through work and inheritance alone, insomuch the main (though not the only) way of accumulating wealth would be exploitation, and this indeed changes our narrative. In Solow's quote in the introduction, he warned us that if we abandoned the ad-hoc assumption that factors were priced at their marginal productivity "much else would change besides", and he was right.

A. Appendix: accounting correspondence with intermediate consumption

As above, aggregate value W_{agg} is:

$$W_{agg} = S_{agg} + V_{agg} + C_{agg}$$

Normalizing,

$$\frac{W_{agg}}{W_{agg}} = 1 = \frac{S_{agg} + V_{agg} + C_{agg}}{W_{agg}} = S + V + C$$

We assume that there is both fixed and circulating capital in C, and therefore we should also include intermediate consumption in the accounting framework. Thus, being I the intermediate consumption of the economy, i.e. the inputs of production, we consider output Y_o:

$$Y_o = Y_g + I = Y_n + \delta K + I$$

We proceed again by establishing that:

$$Y_o = Y_n(1 - d' - i)$$

being i the share of intermediate consumption in total output ($i = \frac{I}{Y_o}$), and d' as the depreciation rate in terms of output[29], i.e. $d' = \frac{\delta K}{Y_o}$.

[29] Note that it´s different from d defined above, since this is defined in terms of output and not of gross income Y_g.

In this setting, we get:

$$\frac{wL}{Y_o} = V = \frac{wL}{\frac{Y_n}{(1-d'-i)}} = \frac{wL(1-d'-i)}{Y_n} \rightarrow \boxed{(1-\alpha)(1-d'-i) = V} \tag{A.1}$$

$$\frac{rK}{Y_o} = S = \frac{rK}{\frac{Y_n}{(1-d'-i)}} = \frac{rK(1-d'-i)}{Y_n} \rightarrow \boxed{\alpha(1-d'-i) = S} \tag{A.2}$$

$$\frac{\delta K + I}{Y_o} = C = \frac{\delta K + I}{\frac{Y_n}{(1-d'-i)}} = \frac{(\delta K + I)(1-d'-i)}{Y_n} \rightarrow \boxed{\delta\beta(1-d'-i) + \frac{I(1-d'-i)}{Y_n} = C} \tag{A.3}$$

Under these conditions, note that S' does not change, but Q does.

$$S' = \frac{S}{V} = \frac{\alpha(1-d'-i)}{(1-\alpha)(1-d'-i)} = \frac{\alpha}{(1-\alpha)} \tag{A.4}$$

$$Q = \frac{C}{V} = \frac{\delta\beta(1-d'-i) + \frac{I(1-d'-i)}{Y_n}}{1-\alpha)(1-d'-i)} = \frac{\delta\beta + i'}{1-\alpha} \tag{A.5}$$

Being $i' = \frac{I}{Y_n}$, which is very similar to i, but the denominator is Y_n instead of Y_o. Thus, considering equations A.4 and A.5, we get:

$$P = \frac{S}{C+V} = \frac{S'}{1+Q} = \frac{\frac{\alpha}{(1-\alpha)}}{\frac{\delta\beta + i'}{(1-\alpha)}} = \frac{\alpha}{\delta\beta + i'} \tag{A.6}$$

Note that equation A.6 is very similar to equation 4.9, with the addition of i' in the denominator. This entails the trivial result that a larger share of inputs in the production process will mechanically reduce the rate of profit, and leaves the main conclusions of the evolution of G with respect to β and α.

Bibliography

Aaberge, R. and M. Mogstad (2015). Inequality in current and lifetime income. *Social Choice and Welfare 44*(2), 217–230.

Abowd, J. M. and M. H. Stinson (2013). Estimating measurement error in annual job earnings: A comparison of survey and administrative data. *Review of Economics and Statistics 95*(5), 1451–1467.

Acemoglu, D. and J. Robinson (2015). The rise and decline of general laws of capitalism. *Journal of economic perspectives 29*(1), 3–28.

Adriaans, J., P. Valet, and S. Liebig (2020). Comparing administrative and survey data: Is information on education from administrative records of the german institute for employment research consistent with survey self-reports? *Quality & Quantity 54*(1), 3–25.

Agustoni, B. and E. Lasarga (2019). Incidencia de la herencia en la distribución de la riqueza real bruta en Uruguay. *DIE 05/19, Serie Documentos de Investigación Estudiantil, Instituto de Economía.*.

Alstadsæter, A., M. Jacob, W. Kopczuk, and K. Telle (2017). 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.

Altimir, O. (1975). Estimaciones de la distribución del ingreso en América Latina por medio de encuestas de hogares y censos de población: una evaluación de confiabilidad. CEPAL.

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

Alvaredo, F. (2010). The rich in Argentina over the Twentieth Century, 1932-2004. In A. B. Atkinson and T. Piketty (Eds.), *Top incomes: A global perspective*, pp. 253–298. Oxford University Press.

Alvaredo, F. (2011). A note on the relationship between top income shares and the Gini coefficient. *Economics Letters 110*(3), 274–277.

Alvaredo, F., A. B. Atkinson, and S. Morelli (2018). Top wealth shares in the UK over more than a century. *Journal of Public Economics 162*, 26–47.

Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez (2013). The top 1 percent in international and historical perspective. *Journal of Economic Perspectives 27*(3), 3–20.

Alvaredo, F., M. De Rosa, I. Flores, and M. Morgan (2022). The inequality (or the growth) we measure: data gaps and distributions of incomes. https://doi.org/10.31235/osf.io/fs5jn.

Alvaredo, F., S. Garriga, and F. Pinto (2017). Household Surveys, Administrative Records, and National Accounts in Mexico 2009–2014. Is a Reconciliation Possible? Paris School of Economics.

Alvaredo, F. and L. Gasparini (2015). Recent trends in inequality and poverty in developing countries. In *Atkinson A. and Bourguignon F. (eds.)Handbook of income distribution*, Volume 2, pp. 697–805. Elsevier.

Alvaredo, F. and J. Londoño Velez (2014). High income and income tax in Colombia, 1993-2010. *Revista de Economía Institucional 16*(31), 157–194.

Alves, C. (2022, May). Joan Robinson on Karl Marx: “His Sense of Reality Is Far Stronger”. *Journal of Economic Perspectives 36*(2), 247–64.

Amarante, V. (2016). Income inequality in Latin America: A factor component analysis. *Review of income and wealth 62*, S4–S21.

Amarante, V., R. Arim, and G. Salas (2007). Impacto distributivo de la reforma impositiva. background paper for poverty and social impact analysis (psia) of Uruguay.

Amarante, V., R. Arim, and M. Yapor (2016). Decomposing inequality changes in uruguay: the role of formalization in the labor market. *IZA Journal of Labor & Development 5*(1), 1–20.

Amarante, V., M. Brum, A. Fernández, G. Pererira, A. Umpiérrez, and A. Vigorito (2010). La distribución de la riqueza en Uruguay: elementos para el debate/. *Art. 2*.

Amarante, V. and A. Carella (1997). La captación del ingreso en las encuestas de hogares. *Informe proyecto CSIC de iniciación a la investigación*.

Amarante, V., M. Colafranceschi, and A. Vigorito (2014). *Uruguay's Income Inequality and Political Regimes over the Period 1981–2010.* Cornia A. (ed.) Falling Inequality in Latin America. Policy Changes and Lessons, WIDER Studies in Development Economics. Oxford University Press.

Amarante, V. and J. P. Jiménez (2015). Desigualdad, concentración y rentas altas en América Latina. In J. P. Jiménez (Ed.), *Desigualdad, concentración del ingreso y tributación sobre las altas rentas en América Latina*, pp. 13–47. Santiago: CEPAL.

Anand, S. and P. Segal (2017). Who are the global top 1%? *World Development 95*, 111–126.

Arulampalam, W., M. P. Devereux, and G. Maffini (2012). The direct incidence of corporate income tax on wages. *European Economic Review 56*(6), 1038–1054.

Atkinson, A. B. (1973). *Wealth, income and inequality.* Penguin Education.

Atkinson, A. B. (2007). Measuring Top Incomes: Methodological Issues. In *Top Incomes over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*, pp. 18–42. Oxford University Press.

Atkinson, A. B. (2015). *Inequality: What Can Be Done?* Harvard University Press.

Atkinson, A. B. (2017). Pareto and the upper tail of the income distribution in the uk: 1799 to the present. *Economica 84*(334), 129–156.

Atkinson, A. B. (2018). Wealth and inheritance in Britain from 1896 to the present. *Journal of Economic Inequality 16*(2), 137–169.

Atkinson, A. B. and F. Bourguignon (2000). Introduction: Income distribution and economics. *Handbook of income distribution 1*, 1–58.

Atkinson, A. B., A. Casarico, and S. Voitchovsky (2018). Top incomes and the gender divide. *The Journal of Economic Inequality 16*(2), 225–256.

Atkinson, A. B. and A. Harrison (1978). *Distribution of personal wealth in Britain.* Cambridge University Press.

Atkinson, A. B. and E. Marlier (2010). *Income and living conditions in Europe.* Luxembourg: Publications Office of the European Union.

Atkinson, A. B. and T. Piketty (2007). *Top incomes over the twentieth century: a contrast between continental european and english-speaking countries.* oup Oxford.

Atkinson, A. B., T. Piketty, and E. Saez (2011). Top incomes in the long run of history. *Journal of economic literature 49*(1), 3–71.

Barro, R. J. and X. Sala-i Martin (2004). Economic growth.

Barros, R. P. d. O., M. N. O. Foguel, and G. O. Ulyssea (2006). Desigualdade de renda no Brasil: uma análise da queda recente.

Baselgia, E. and I. Martínez (2021). A safe harbor: Wealth-income ratios in Switzerland over the 20th century and the role of housing prices.

Bauluz, L., F. Novokmet, and Y. Govind (2020). Global Land Inequality. *Wid.World Working Paper series 2020/10*.

Bauluz, L. E. (2019). Revised national income and wealth series: Australia, Canada, France, Germany, Italy, Japan, UK and USA. *WID. world Working Paper* (2017/23).

Bellofiore, R. (2008). Sraffa after Marx: an open issue. In *Sraffa or an alternative economics*, pp. 68–92. Springer.

Bergolo, M., G. Burdin, M. De Rosa, M. Giaccobasso, M. Leites, and H. Rueda (2019). Earnings Responses of Top Labor Incomes to Tax Changes: Evidence from a Tax Reform in Uruguay.

Berman, Y. and S. Morelli (2021). On the distribution of estates and the distribution of wealth: Evidence from the dead. *National Bureau of Economic Research*.

Bértola, L. and J. Williamson (2017). *Has Latin American inequality changed direction?: Looking over the long run*. Springer.

Bertrand, M. and S. Mullainathan (2001). Are CEOs rewarded for luck? The ones without principals are. *The Quarterly Journal of Economics 116*(3), 901–932.

Bittencourt, G., G. Carracelas, A. Doneschi, and N. Reig Lorenzi (2009). Tendencias recientes de la inversión extranjera directa en Uruguay. *Documento de Trabajo/FCS-DE; 27/09*.

Bivens, J. and L. Mishel (2013). The pay of corporate executives and financial professionals as evidence of rents in top 1 percent incomes. *Journal of Economic Perspectives 27*(3), 57–78.

Blanchet, T., L. Chancel, and A. Gethin (2019). How Unequal Is Europe? Evidence from Distributional National Accounts, 1980-2017. *WID. world Working Paper 6*.

Blanchet, T., I. Flores, and M. Morgan (2022). The weight of the rich: Improving surveys using tax data. *Journal of Income Inequality 20*(1), 1–32.

Blanco, M. A., L. Bauluz, and C. Martínez-Toledano (2021). Wealth in Spain 1900–2017 A Country of Two Lands. *The Economic Journal 131*(633), 129–155.

Boschini, A., K. Gunnarsson, and J. Roine (2020). Women in top incomes–Evidence from Sweden 1971–2017. *Journal of Public Economics 181*, 104115.

Bourguignon, F. (2015, 12). Appraising income inequality databases in Latin America. *The Journal of Economic Inequality 13*(4), 557–578.

Bourguignon, F. (2018). Simple adjustments of observed distributions for missing income and missing people. *The Journal of Economic Inequality 16*(2), 171–188.

Bricker, J., J. Krimmel, A. Henriques, and J. Sabelhaus (2016). Measuring income and wealth at the top using administrative and survey data. *Brookings Papers on Economic Activity 2016*(SPRING), 261–331.

Bryer, R. A. (1999). Marx and accounting. *Critical Perspectives on Accounting 10*(5), 683–709.

Bucheli, M., N. Lustig, M. Rossi, and F. Amábile (2013). *Social spending, taxes and income redistribution in Uruguay.* The World Bank.

Burdín, G., F. Esponda, and A. Vigorito (2014). Inequality and top incomes in Uruguay: a comparison between household surveys and income tax micro-data. *World Top Incomes Database Working Paper 1.*

Burdín, G., M. De Rosa, A. Vigorito, and J. Vilá (2022). Falling inequality and the growing capital income share: Reconciling divergent trends in survey and tax data. *World Development 152*, 105783.

Burkhauser, R. V., S. Feng, S. P. Jenkins, and J. Larrimore (2012). Recent trends in top income shares in the united states: reconciling estimates from march cps and irs tax return data. *Review of Economics and Statistics 94*(2), 371–388.

Burkhauser, R. V., N. Hérault, S. Jenkins, and R. Wilkins (2018). Survey under-coverage of top incomes and estimation of inequality: What is the role of the UK's SPI adjustment? *Fiscal Studies 39*, 213–240.

Canberra Group (2001). *Expert Group on Household Income Statistics, The Canberra Group : final report and recommendations*. Canberra Group.

Cano, L. (2015). Income mobility in Ecuador new evidence from individual income tax returns. UNU-WIDER Working Paper No. 2015/040.

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *The Review of economic studies 32*(3), 233–240.

Castro, F. A. and M. Bugarin (2017). A progressividade do imposto de renda da pessoa física no Brasil. *Estudos Econômicos 47*(2), 259–293.

CEPAL (2020). Brechas de género en los ingresos laborales en el Uruguay.

CEQ (2021). CEQ Data Center on Fiscal Redistribution. Commitment to Equity (CEQ) Institute, Tulane University.

Charpentier, A. and E. Flachaire (2019). Pareto models for top incomes.

Chatterjee, A., L. Czajka, and A. Gethin (2022). Wealth Inequality in South Africa, 1993–2017. *The World Bank Economic Review 36*(1), 19–36.

Chudnovsky, D. and A. López (2007). Inversión extranjera directa y desarrollo: la experiencia del Mercosur. *Revista de la CEPAL*.

Clark, J. B. ([2007]1907). *Essentials of economic theory*. Ludwig von Mises Institute.

Clifton, J., D. Díaz-Fuentes, and J. Revuelta (2020). Falling Inequality in Latin America: The Role of Fiscal Policy. *Journal of Latin American Studies 52*(2), 317–341.

Cohen, A. J. (1989). Prices, capital, and the one-commodity model in neoclassical and classical theories. *History of Political Economy 21*(2), 231–251.

Cohen, A. J. (2003). The Hayek/Knight capital controversy: the irrelevance of roundaboutness, or purging processes in time? *History of political economy 35*(3), 469–490.

Cohen, A. J. and G. C. Harcourt (2003). Retrospectives: Whatever Happened to the Cambridge Capital Theory Controversies? *Journal of Economic Perspectives 17*(1), 199–214.

Cohen, A. J. and G. C. Harcourt (2005). Introduction. capital theory controversy: scarcity, production, equilibrium and time. *Capital theory 1*, xxvii–lx.

CONADE (1965). *Distribución del Ingreso y Cuentas Nacionales en la Argentina, Vols. I-V.* Buenos Aires: CONADE-CEPAL.

Cornia, G. A. (2014a). *Falling inequality in Latin America: Policy changes and lessons.* OUP Oxford.

Cornia, G. A. (2014b). Inequality trends and their determinants: Latin America over 1990 - 2010. In G. A. Cornia (Ed.), *Falling Inequality in Latin America: Policy Changes and Lessons.* Oxford University Press.

Cowell, F. A. (2011). *Measuring Inequality* (Third ed.). Oxford: Oxford University Press.

Cowell, F. A. and E. Flachaire (2015). Statistical methods for distributional analysis. In *Handbook of income distribution*, Volume 2, pp. 359–465. Elsevier.

Davies, J. B. and L. Di Matteo (2021). Long run Canadian wealth inequality in international context. *Review of Income and Wealth 67*(1), 134–164.

Davies, J. B., R. Lluberas, and A. F. Shorrocks (2017). Estimating the level and distribution of global wealth, 2000–2014. *Review of Income and Wealth 63*(4), 731–759.

Davies, J. B., S. Sandström, A. Shorrocks, and E. N. Wolff (2011). The level and distribution of global household wealth*. *The Economic Journal 121*(551), 223–254.

Davies, J. B. and A. F. Shorrocks (2000). The distribution of wealth. In *Handbook of income distribution*, Volume 1, pp. 605–675.

De la Torre, A., E. L. Yeyati, G. Beylis, T. Didier, C. R. Castelan, and S. Schmukler (2014, 10). Inequality in a Lower Growth Latin America. Technical report, World Bank.

De Rosa, M. (2018). Wealth distribution in uruguay: capitalizing incomes in the dark. *Serie Documentos de Trabajo; 7/18*.

De Rosa, M. (2019). Wealth accumulation and its distribution in uruguay: first estimates of the untold half of the story. *PPD Master Thesis, Paris School of Economics*.

De Rosa, M., I. Flores, and M. Morgan (2020). Inequality in Latin America Revisited: Insights from Distributional National Accounts. WIL Technical Notes.

De Rosa, M., I. Flores, and M. Morgan (2022). More unequal or nor as rich: revisiting the Latin American exception. *SocArXiv*.

De Rosa, M., S. Sinisclachi, J. Vilá, A. Vigorito, and H. Willebald (2018). La evolución de las remuneraciones laborales y la distribución del ingreso en Uruguay; futuro en foco. *Cuadernos Sobre Desarrollo Humano: Montevideo, Uruguay*.

De Rosa, M. and J. Vilá (2022). Beyond tax-survey combination: inequality and the blurry household-firm border. *Serie Documentos de Trabajo; 10/22*.

Deaton, A. (2005, 2). Measuring Poverty in a Growing World (or Measuring Growth in a Poor World). *Review of Economics and Statistics 87*(1), 1–19.

Del Castillo, M. (2017). La distribución y desigualdad de los activos financieros y no financieros en México. *CEPAL*.

DIEA (2014). Censo general agropecuario 2011. Technical report, Dirección de Estadísticas Agropecuarias.

Dobb, M. ([1961]1937). *Economía Política y Capitalismo*. Fondo de Cultura Económica.

Domar, E. D. (1946). Capital expansion, rate of growth, and employment. *Econometrica, Journal of the Econometric Society*, 137–147.

Domínguez-Amorós, M., K. Batthyány, and S. Scavino (2021). Gender Gaps in Care Work: Evidences from Argentina, Chile, Spain and Uruguay. *Social Indicators Research*, 1–30.

Dwenger, N., V. Steiner, and P. Rattenhuber (2019). Sharing the burden? empirical evidence on corporate tax incidence. *German Economic Review 20*(4), e107–e140.

Elster, J. (1986). *An Introduction to Karl Marx*. Cambridge University Press.

Engels, F. (1891). Introduction to Karl Marx's wage labour and capital.

Espino, A., F. Isabella, M. Leites, and A. Machado (2017). Do women have different labor supply behaviors? Evidence based on educational groups in Uruguay. *Feminist Economics 23*(4), 143–169.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016). Heterogeneity in returns to wealth and the measurement of wealth inequality. *American Economic Review 106*(5), 651–55.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2020). Heterogeneity and persistence in returns to wealth. *Econometrica 88*(1), 115–170.

Fairfield, T. and M. Jorratt De Luis (2016, 8). Top Income Shares, Business Profits, and Effective Tax Rates in Contemporary Chile. *Review of Income and Wealth 62*(S1), S120–S144.

Feenberg, D. R. and J. M. Poterba (1993). Income inequality and the incomes of very high-income taxpayers: evidence from tax returns. *Tax policy and the economy 7*, 145–177.

Fernandes, R. C., B. P. C. Diniz, and F. G. Silveira (2018). The distributive impact of income taxes in Brazil. Technical report, International Policy Centre for Inclusive Growth (IPC-IG), Working Paper, No. 171.

Ferre, Z., J. I. Rivero, G. Sanroman, and G. Santos (2016). Encuesta financiera de los hogares uruguayos (EFHU-2): descripción y resultados. *Documento de Trabajo/FCS-DE; 16/06*.

Ferreira, F. H., N. Lustig, and D. Teles (2015). Appraising cross-national income inequality databases: An introduction. *The Journal of economic inequality 13*(4), 497–526.

Fesseau, M. and M. L. Mattonetti (2013). Distributional measures across household groups in a national accounts framework. OECD Working Papers.

Fischer, I. (1906). *The Nature of Capital and Income*. Macmillan.

Fischer, I. (1907). *The Rate of Interest: Its Nature, Determination and Relation to Economic Phenomena*. Macmillan.

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

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

Flores, I. (2018). Income under the carpet: What gets lost between the measure of capital shares and inequality. `http://precog.iiitd.edu.in/people/anupama`.

Flores, I. (2021). The capital share and income inequality: Increasing gaps between micro and macro-data. *The Journal of Economic Inequality 19*(4), 685–706.

Flores, I., C. Sanhueza, J. Atria, and R. Mayer (2020). Top Incomes in Chile: A Historical Perspective on Income Inequality, 1964–2017. *Review of Income and Wealth 66*(4), 850–874.

Gandelman, N., R. Lluberas, et al. (2022). Wealth in Latin America.

Garbinti, B., J. Goupille-Lebret, and T. Piketty (2017). Accounting for wealth inequality dynamics: Methods, estimates and simulations for France (1800-2014).

Garbinti, B., J. Goupille-Lebret, and T. Piketty (2018, 6). Income inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA). *Journal of Public Economics 162*, 63–77.

Gasparini, L., J. Bracco, L. Galeano, and M. Pistorio (2018). Desigualdad en países en desarrollo: ¿ajustando las expectativas? *Documentos de Trabajo del CEDLAS*.

Gobetti, S. and R. Orair (2017). Taxation and distribution of income in Brazil: new evidence from personal income tax data. *Revista de Economia Política 37*(2), 267–286.

Goñi, E., J. H. López, and L. Servén (2011). Fiscal redistribution and income inequality in Latin America. *World development 39*(9), 1558–1569.

Goolsbee, A. (2000). What happens when you tax the rich? evidence from executive compensation. *Journal of Political Economy 108*(2), 352–378.

Groskoff, R. (1992). Análisis y ajuste de los ingresos investigados por las encuestas de hogares. *Documento de Trabajo*.

Hanni, M., R. Martner Fanta, and A. Podestá (2015). El potencial redistributivo de la fiscalidad en América Latina. *Revista cepal*.

Hansen, N.-J. H., K. Harmenberg, E. Öberg, and H. H. Sievertsen (2021). Gender disparities in top earnings: measurement and facts for denmark 1980-2013. *The Journal of Economic Inequality*, 1–16.

Harcourt, G. C. (2014). Comment: Lance Taylor on Thomas Piketty's World as Seen Through the Eyes of Maynard Keynes and Luigi Pasinetti. *International Journal of Political Economy 43*(3), 18–25.

Harrod, R. ([1972]1939). An essay in dynamic theory. In *Economic Essays*, pp. 254–277. Springer.

Hayek, F. A. ([2009]1941). *The pure theory of capital.* The Ludwig von Mises Institute.

Higgins, S., N. Lustig, A. Vigorito, et al. (2018). The rich underreport their income: Assessing bias in inequality estimates and correction methods using linked survey and tax data. *Ecineq WP 475*.

Hodgson, G. (2014). What is capital? Economists and sociologists have changed its meaning: should it be changed back? *Cambridge Journal of Economics 38*(5), 1063–1086.

IBGE (2016). *Sistema de Contas Nacionais Brasil: Ano De Referência 2010.* Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística.

ILO (2019). The Global Labour Income Share and Distribution. ILO Department of Statistics, Methodological description.

Instituto Nacional de Estadística (2013). Encuesta de nutrición, desarrollo infantil y salud, ENDIS.

Instituto Nacional de Estadística (2021). Encuestas continuas de hogares.

Jantti, M., V.-M. Törmälehto, and E. Marlier (2013). *The use of registers in the context of EU? SILC: challenges and opportunities.* Luxembourg: Publications Office of the European Union.

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

Jenkins, S. P. (2017). Pareto Models, Top Incomes and Recent Trends in UK Income Inequality. *Economica 84*(334), 261–289.

Jiménez, J. P. and D. Rossignolo (2019). Concentración del ingreso y desigualdad según registros tributarios en América Latina: el caso de la Argentina. Technical report, Centro de Estudios para el Cambio Estructural (CECE), Buenos Aires.

Jones, C. I. (2015). Pareto and Piketty: The macroeconomics of top income and wealth inequality. *Journal of Economic Perspectives 29*(1), 29–46.

Kennickell, A. B. (2019). The tail that wags: differences in effective right tail coverage and estimates of wealth inequality. *The Journal of Economic Inequality 17*(4), 443–459.

Kingston, J. and L. Kingston (1972). A distribuição da renda no Brasil, 1960-1970. *Revista Brasileira de Economia 26*(4), 241–256.

Koopmans, T. (1965). On the concept of optimal growth, the econometric approach to development planning. *Econometric approach to development planning, 1st edn. North Holland, Amsterdam*, 225–287.

Kopczuk, W. (2015). What do we know about the evolution of top wealth shares in the United States? *Journal of Economic Perspectives 29*(1), 47–66.

Kopczuk, W. (2016). Comment on Bricker et al, 2016. *Brook. Pap. Econ. Act.(Spring)*, 321–327.

Kopczuk, W. and E. Saez (2004). Top wealth shares in the United States, 1916–2000: Evidence from estate tax returns. *National Tax Journal 57*(2), 445–487.

Kopczuk, W. and E. Zwick (2020). Business incomes at the top. *Journal of Economic Perspectives 34*(4), 27–51.

Kumar, R. (2019). The evolution of wealth-income ratios in India 1860-2012. *Available at SSRN 3111846*.

Kuznets, S. (1953). *Shares of Upper Income Groups in Income and Savings.* NBER.

Kuznets, S., L. Epstein, and E. Jenks (1941). *National Income and Its Composition, 1919–1938, Volume I.* National Bureau of Economic Research.

Lakner, C. and B. Milanovic (2016). Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession. *The World Bank Economic Review 30*(2), 203–232.

Langoni, C. (1973). *Distribuição da renda e desenvolvimento economico do Brasil.* Rio de Janeiro: Editora Expressao e Cultura.

Lanzilotta, B., A. Souto-Pérez, and G. Zunino (2020). Propuesta metodológica para la determinación del monto imponible de contribución inmobiliaria urbana. *Oficina de Planeamiento y Presupuesto, Uruguay*.

Leigh, A. (2006). Using Panel Data on Top Income Shares to Analyse the Causes and Effects of Inequality. Discussion paper, Australian National University.

Leigh, A. (2007). How closely do top income shares track other measures of inequality? *The Economic Journal 117*(524), F619–F633.

Lemieux, T. and W. C. Riddell (2015). Who are canada's top 1 percent? *Income Inequality: The Canadian Story*, 1–53.

Lerman, R. I. and S. Yitzhaki (1985). Income inequality effects by income source: A new approach and applications to the United States. *The review of economics and statistics*, 151–156.

Liu, L. and R. Altshuler (2013). Measuring the burden of the corporate income tax under imperfect competition. *National Tax Journal 66*(1), 215–237.

Londoño-Vélez, J. (2012). Income and Wealth at the Top in Colombia: An Exploration of Tax Records 1993-2010. Master's dissertation, Paris School of Economics, Paris.

López, R., E. Figueroa, P. Gutiérrez, et al. (2013). La 'parte del león': Nuevas estimaciones de la participación de los súper ricos en el ingreso de chile. *Serie Documentos de Trabajo 379*, 1–32.

López-Calva, L. F. and N. C. Lustig (2010). *Declining inequality in Latin America: A decade of progress?* Brookings Institution Press.

Lustig, N. (Ed.) (2018). *Commitment to Equity Handbook: Estimating the Impact of Fiscal Policy on Inequality and Poverty*. CEQ Institute at Tulane University and Brookings Institution Press.

Lustig, N. et al. (2019). The missing rich in household surveys: Causes and correction approaches. Technical report, Tulane University, Department of Economics.

Lustig, N., L. F. López-Calva, and E. Ortiz-Juarez (2011). The decline in inequality in Latin America: How much, since when and why. *Since When and Why (April 24, 2011)*.

Lustig, N., C. Pessino, and J. Scott (2014). The impact of taxes and social spending on inequality and poverty in Argentina, Bolivia, Brazil, Mexico, Peru, and Uruguay: Introduction to the special issue. *Public Finance Review 42*(3), 287–303.

Lustig, N., D. Teles, et al. (2016). Inequality convergence: How sensitive are results to the choice of data? Technical report.

López, R., E. Figueroa, and P. Gutiérrez (2013). La Parte del León: Nuevas Estimaciones de la Participación de los Súper Ricos en el Ingreso de Chile. Serie Documentos de Trabajo de la Universidad de Chile Working Paper nº SDT379.

López-Calva, L. F. and N. C. Lustig (2010). *Declining Inequality in Latin America: A Decade of Progress?* Brookings Institution Press and United Nations Development Programme.

Marshall, A. ([2009]1890). *Principles of economics: unabridged eighth edition*. Cosimo, Inc.

Martínez-Toledano, C. (2017). Housing bubbles, offshore assets and wealth inequality in spain. *World Wealth and Income Database Working Paper 19*.

Martínez-Toledano, C. (2020). House price cycles, wealth inequality and portfolio reshuffling. *WID. World Working Paper 2*.

Marx, K. ([1976]1867). *Capital. A Critique of Political Economy. Volume I.* Penguin.

Marx, K. ([1978]1885). *Capital. A Critique of Political Economy. Volume II.* Penguin.

Marx, K. ([1981]1894). *Capital. A Critique of Political Economy. Volume III.* Penguin.

Marx, K. ([1997]1857). *Grundrisse: Foundations of the critique of political economy.* Penguin Books in association with New Left Review.

Medeiros, M., J. de Castro Galvão, and L. de Azevedo Nazareno (2018). Correcting the Underestimation of Top Incomes: Combining Data from Income Tax Reports and the Brazilian 2010 Census. *Social Indicators Research 135*(1), 233–244.

Medeiros, M., P. Souza, and F. A. d. Castro (2015). O Topo da Distribuição de Renda no Brasil: Primeiras Estimativas com Dados Tributarios e Comparação com Pesquisas Domiciliares (2006-2012). *Dados 58*(1), 7–36.

Mendive, C. and A. Fuentes (1996). Diferencias en la captación del ingreso por fuente" en ine. *Taller de expertos sobre medición de pobreza*.

Messina, J. and J. Silva (2017). *Wage Inequality in Latin America: Understanding the Past to Prepare for the Future.* The World Bank.

Meyer, B. D., W. K. Mok, and J. X. Sullivan (2015). The under-reporting of transfers in household surveys: Its nature and consequences.

Mill, J. S. (1836). On the definition and method of political economy. *The philosophy of economics*, 41–58.

Mill, J. S. ([1900]1848). Principles of political economy. vol. 1.

Mirowski, P. (2002). Machine dreams. *Cambridge Books*.

Morgan, M. (2015). Income concentration in a context of late development: an investigation of top incomes in Brazil using tax records, 1933-2013. Master's dissertation, Paris School of Economics, Paris.

Morgan, M. (2017a). Desigualdade de Renda, Crescimento e Tributação da Elite no Brasil: Novas Evidências Reunindo Dados de Pesquisas Domiciliares e Dados Fiscais. In J. R. Afonso, R. O. Orair, and F. G. Silveira (Eds.), *Tributação e Desigualdade*. Rio de Janeiro: Letramento & FGV Direito Rio.

Morgan, M. (2017b). 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.

Morgan, M. (2018). *Essays on Income Distribution: Historical, Methodological and Institutional Perspectives.* Phd dissertation in economics, Paris School of Economics, Paris.

Morgan, M. and P. Souza (2019). Distributing Growth During Late-Development: New Evidence on Long-Run Inequality in Brazil. CLADHE IV, Universidad de Santiago, Chile. July 23-25 2019.

Mortara, G. (1949). Representação analítica das distribuiçoes dos contribuintes e das respectivas rendas líquidas determinadas para a aplicação do imposto de renda, em função do valor da renda líquida. *Revista Brasileira de Economia 3*(2), 7–34.

Nolan, B., M. Roser, and S. Thewissen (2019, 9). GDP Per Capita Versus Median Household Income: What Gives Rise to the Divergence Over Time and how does this Vary Across OECD Countries? *Review of Income and Wealth 65*(3), 465–494.

Novokmet, F., T. Piketty, and G. Zucman (2018). From soviets to oligarchs: inequality and property in Russia 1905-2016. *The Journal of Economic Inequality 16*(2), 189–223.

Ocampo, J. A. (2017). Commodity-led development in latin america. In *Alternative pathways to sustainable development: Lessons from Latin America*, pp. 51–76. Brill Nijhoff.

OECD (2013). *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth.* OECD Publishing.

OECD (2020). Distributional Information on Household Income, Consumption and Saving In Line With National Accounts – Guidelines. Prepared by the Statistics and Data Directorate of the OECD.

OECD/ECLAC/CIAT/IDB (2022). *Revenue Statistics in Latin America and the Caribbean 2022.* Paris: OECD Publishing.

Ohlsson, H., J. Roine, and D. Waldenström (2020). Inherited Wealth over the Path of Development: Sweden, 1810–2016. *Journal of the European Economic Association 18*(3), 1123–1157.

Pasinetti, L. (1979). Growth and income distribution: essays in economic theory. *Cambridge Books*.

Pasinetti, L. (1983). Structural change and economic growth. *Cambridge Books*.

Peluffo, A. (2015). Foreign direct investment, productivity, demand for skilled labour and wage inequality: An analysis of Uruguay. *The World Economy 38*(6), 962–983.

Phelps, E. (1961). The golden rule of accumulation: a fable for growthmen. *The American Economic Review 51*(4), 638–643.

Piketty, T. (2003). Income inequality in france, 1901–1998. *Journal of political economy 111*(5), 1004–1042.

Piketty, T. (2011). On the long-run evolution of inheritance: France 1820-2050. *Quarterly Journal of Economics 126*(3), 1071–1131.

Piketty, T. (2014). *Capital in the Twenty First Century*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Piketty, T. (2015). Putting Distribution Back at the Center of Economics: Reflections on" Capital in the Twenty-First Century". *The Journal of Economic Perspectives*, 67–88.

Piketty, T., F. Alvaredo, and L. Assouad (2017). Measuring inequality in the middle east 1990-2016: The world's most unequal region?

Piketty, T., A. Bozio, B. Garbinti, J. Goupille-Lebret, and M. Guillot (2020). Predistribution vs. Redistribution: Evidence from France and the US.

Piketty, T. and L. Chancel (2017). Indian income inequality, 1922-2014: From british raj to billionaire raj?

Piketty, T., G. Postel-Vinay, and J.-L. Rosenthal (2006). Wealth concentration in a developing economy: Paris and France, 1807-1994. *American economic review 96*(1), 236–256.

Piketty, T., E. Saez, and G. Zucman (2018, 5). Distributional National Accounts: Methods and Estimates for the United States. *The Quarterly Journal of Economics 133*(2), 553–609.

Piketty, T., L. Yang, and G. Zucman (2019). Capital accumulation, private property, and rising inequality in china, 1978–2015. *American Economic Review 109*(7), 2469–96.

Piketty, T. and G. Zucman (2014). Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly Journal of Economics 129*(3), 1255–1310.

Piketty, T. and G. Zucman (2015). Wealth and inheritance in the long run. In *Handbook of income distribution*, Volume 2, pp. 1303–1368. Elsevier.

Ramsey, F. P. (1928). A mathematical theory of saving. *The economic journal 38*(152), 543–559.

Ravallion, M. (2003, 8). Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree? *Review of Economics and Statistics 85*(3), 645–652.

Ray, D. (1998). *Development economics*. Princeton University Press.

Ricardo, D. ([2001]1817). *On the principles of political economy and taxation*. Batoche Books.

Robbins, L. ([2007]1932). *An essay on the nature and significance of economic science*. Ludwig von Mises Institute.

Robeyns, I. (2019). What, if anything, is wrong with extreme wealth? *Journal of Human Development and Capabilities 20*(3), 251–266.

Robinson, J. (1954). The Production of Function Capital and the Theory. *Review of Economic Studies 21*(2), 81–106.

Robinson, J. ([1967]1942). *Essay on Marxian Economics*. Springer.

Robinson, J. (1974). History versus equilibrium. *Indian Economic Journal 21*(3), 202.

Rodríguez-Castelán, C., L. F. López-Calva, N. Lustig, and D. Valderrama (2016). Understanding the dynamics of labor income inequality in latin america. Technical report, Policy Research Working Paper 7795, The World Bank.

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

Rodríguez Miranda, A. and M. Menendez (2020). Desigualdades regionales, crecimiento económico y cambio estructural en uruguay: 1983-2017. Technical report, Instituto de Economia-IECON.

Roine, J. and D. Waldenström (2009). Wealth concentration over the path of development: Sweden, 1873-2006. *Scandinavian Journal of Economics 111*(1), 151–187.

Rossignolo, D., N. Oliva, and N. Villacreses (2016). Cálculo de la concentración de los altos ingresos sobre la base de los datos impositivos: un análisis para el Ecuador. Serie Macroeconomía del Desarrollo No. 184.

Saez, E. and G. Zucman (2016, May). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data *. *The Quarterly Journal of Economics 131*(2), 519–578.

Saez, E. and G. Zucman (2020). The rise of income and wealth inequality in America: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives 34*(4), 3–26.

Saez, E. and G. Zucman (2022, August). Top wealth in america: A reexamination. Working Paper 30396, National Bureau of Economic Research.

Samuelson, P. A. (1962). Parable and realism in capital theory: the surrogate production function. *The Review of Economic Studies 29*(3), 193–206.

Samuelson, P. A. (1966). A summing up. *The quarterly journal of economics 80*(4), 568–583.

Samuelson, P. A. and V. Nordhouse (2009). Economics 19th ed.

Sánchez-Ancochea, D. (2021). The surprising reduction of inequality during a commodity boom: what do we learn from Latin America? *Journal of Economic Policy Reform 24*(2), 95–118.

Sanhueza, C. and R. Mayer (2011). Top Incomes in Chile using 50 years of household surveys: 1957-2007. *Estudios de economía 38*(1), 169–193.

Sanroman, G. and G. Santos (2021). The joint distribution of income and wealth in uruguay. *Cuadernos de Economía 40*(83), 609–642.

Schumpeter, J. A. ([1950]1942). *Capitalism, Socialism and Democracy.* Harper Perenial Modern Thought.

Shaikh, A. M. et al. (1997). Measuring the wealth of nations. *Cambridge Books*.

Shorrocks, A., J. Davies, and R. Lluberas (2021). Credit suisse research institute global wealth report 2021. *Zurich: Credit Suisse AG*.

Shorrocks, A. F. (1981). Income stability in the United States. In *The statics and dynamics of income*. Citeseer.

Shorrocks, A. F. (1999). Decomposition procedures for distributional analysis: a unified framework based on the shapley value. Technical report, mimeo, University of Essex.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Kelley.

Smith, M., D. Yagan, O. Zidar, and E. Zwick (2019). Capitalists in the twenty-first century. *The Quarterly Journal of Economics 134*(4), 1675–1745.

Smith, M., O. M. Zidar, and E. Zwick (2021). Top wealth in america: New estimates and implications for taxing the rich. Technical report, National Bureau of Economic Research.

Solow, R. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics 70*(1), 65–94.

Solow, R. (1963). *Capital theory and the rate of return.* North-Holland Publishing Company Amsterdam.

Solow, R. (2000). The neoclassical theory of growth and distribution. *BNL Quarterly Review* (215).

Souza, P. (2016). *A desigualdade vista do topo: a concentração de renda entre os ricos no Brasil, 1926–2013.* Phd dissertation in sociology, Universidade de Brasília, Brasília.

Souza, P. (2018). A history of inequality: top incomes in Brazil, 1926–2015. *Research in Social Stratification and Mobility 57*, 35–45.

Souza, P. and M. Medeiros (2015). Top income shares and inequality in Brazil, 1928–2012. *Sociologies in Dialogue 1*(1), 119–132.

Stiglitz, J., A. K. Sen, and J.-P. Fitoussi (2009). The measurement of economic performance and social progress revisited: reflections and overview. OFCE n2009-33.

Suárez Serrato, J. C. and O. Zidar (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review 106*(9), 2582–2624.

Sutch, R. (2017). The One Percent across Two Centuries: A Replication of Thomas Piketty's Data on the Concentration of Wealth in the United States. *Social Science History 41*(4), 587–613.

Swan, T. W. (1956). Economic growth and capital accumulation. *Economic record 32*(2), 334–361.

Sweezy, P. M. ([1970]1942). *The theory of capitalist development.* Modern Reader Paperbacks.

Székely, M. and M. Hilgert (1999). What's behind the inequality we measure? an investigation using latin american data. *An Investigation Using Latin American Data (December 1999). IDB-OCE Working Paper* (409).

Tobón, A. and Y. S. Ríos (2020). Los fundamentos clásicos de la acumulación del capital en Piketty. *Sociedad y Economía* (41), 109–120.

Torche, F. and S. Spilerman (2006). *Household Wealth in Latin America.* Number 2006/114. Research Paper, UNU-WIDER, United Nations University (UNU).

Törmälehto, V.-M. (2011). LIS and national accounts comparison. LIS Technical Working Paper Series No. 2.

Torregrosa-Hetland, S. (2020). Inequality in tax evasion: the case of the spanish income tax. *Applied Economic Analysis 28*(83), 89–109.

United Nations (2008a). System of national accounts. https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf.

United Nations (2008b). *The System of National Accounts 2008.* European Commission, IMF, OECD, UN, WB.

Veblen, T. (1908). Professor Clark's Economics. *The Quarterly Journal of Economics 22*(2), 147–195.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth 64*(2), 357–387.

Waldenström, D. (2017). Wealth-income ratios in a small, developing economy: Sweden, 1810–2014. *The Journal of Economic History 77*(1), 285–313.

WIL (2020). Distributional National Accounts Guidelines: Methods and Concepts Used in the World Inequality Database. WID.world Working Paper series.

Winkelried, D. and B. Escobar (2020). Declining inequality in Latin America? Robustness checks for Peru. MPRA Paper No. 106566.

Wolff, E. N. (2021). Household wealth trends in the united states, 1962 to 2019: Median wealth rebounds... but not enough. Technical report, National Bureau of Economic Research.

Wolfson, M. C., M. R. Veall, W. N. Brooks, and B. B. Murphy (2016). Piercing the veil: Private corporations and the income of the affluent.

Zucman, G. (2013). The missing wealth of nations: Are europe and the us net debtors or net creditors? *The Quarterly journal of economics 128*(3), 1321–1364.

Zucman, G. (2014). Taxing across borders: Tracking personal wealth and corporate profits. *Journal of economic perspectives 28*(4), 121–48.

Zucman, G. (2019). Global Wealth Inequality. *Annual Review of Economics 11*(1), 109–138.

Zuniga-Cordero, A. (2018). Income Inequality in Costa Rica: An Inquiry Towards Distributional National Accounts (DINA). Master's dissertation, Paris School of Economics, Paris.

Zuniga-Cordero, A. (2022). Distributional National Accounts Costa Rica: inequality measures from the national to the local level, 2000-2020. Mimeo.

Zwijnenburg, J. (2019, 5). Unequal Distributions: EG DNA versus DINA Approach. *AEA Papers and Proceedings 109*, 296–301.

Zwijnenburg, J. (2022). The use of distributional national accounts in better capturing the top tail of the distribution. *The Journal of Economic Inequality*, 1–10.

Zwijnenburg, J., S. Bournot, F. Giovanelly, and F. Giovannelli (2017). Expert Group on Disparities in a National Accounts Framework: Results from the 2015 Exercise. OECD Statistics Working Papers.