From Housing Gains to Pension Losses: New Methods to Reveal Wealth Inequality Dynamics in Chile*

Bastián Castro Nofal† Ignacio Flores‡ Pablo Gutiérrez Cubillos†

April 14, 2025

Abstract

This paper examines wealth inequality dynamics in Chile from 2007 to 2021, focusing on two key macroeconomic events: the sharp rise in housing prices after the introduction of a real estate value-added tax in 2016 and the substantial liquidation of pension assets through early withdrawals during the pandemic. We introduce a methodological innovation that aims to improve the measurement of wealth inequality by integrating administrative pension fund records into household wealth surveys using machine learning techniques. Our results reveal extreme levels of wealth concentration, with the top 10% holding approximately two-thirds of national private wealth. However, inequality slightly declined over the period, particularly after 2016, as the outcome of two opposing forces: housing appreciation, which benefited middle-class households, and pension fund withdrawals, which disproportionately reduced wealth at the lower end of the distribution.

Keywords: Fiscal policy and household wealth distribution, Pension funds
JEL Classification codes: H31, H55, D31, G51, E21

*Corresponding author: Ignacio Flores, email: ignacio.flores@psemail.eu.
†Universidad de Chile, Diagonal Paraguay 257, Santiago, Chile.
‡Paris School of Economics & Roma Tre, 48 Boulevard Jourdan, 75014, Paris, France.

1 Introduction

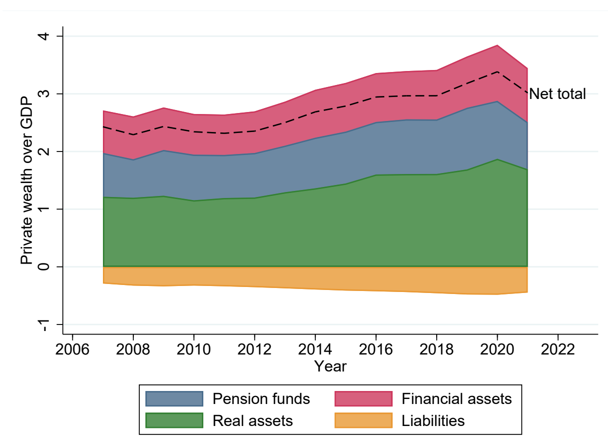
The new wealth literature has transformed our understanding of wealth concentration and its historical dynamics by introducing methods such as the income capitalization approach and estate tax multipliers. These tools allow for more comprehensive measurements, particularly for the top end of the wealth distribution (Piketty, 2014; Saez and Zucman, 2016; Zucman, 2019). However, their reliance on detailed tax data and comprehensive national accounts limits their applicability, leaving a majority of countries understudied due to insufficient statistical infrastructure. To address these limitations, alternative approaches have emerged, including combining household financial surveys with billionaires lists, which are more widely available and offer an effective substitute (Vermeulen, 2016, 2018). Despite these advances, some challenges persist. Surveys adjusted with billionaire data are still prone to misreporting, often influenced by behavioral biases and financial illiteracy. This is crucial for certain assets such as private pensions funds, which are notoriously misreported and affect the entire wealth distribution. This problem is especially relevant in countries with mandatory private pension systems, where private pension funds constitute a substantial part of national wealth.[1]

This paper addresses these challenges by developing a method to impute pension fund balances from administrative records into household surveys. This approach is independent yet complementary to top-tail adjustments such as those proposed by Vermeulen (2018). Using Chile as a case study, we present the first set of macro-consistent estimates of wealth inequality for the country. These estimates provide new insights into a period of accelerated wealth growth marked by two major macroeconomic events resulting from rare policy interventions. The first event was the surge in housing prices following the introduction of a value-added tax on real estate transactions in 2016. The second was the liquidation of a substantial portion of pension assets through early withdrawals permitted during the COVID-19 pandemic. We evaluate the distributional outcomes of both events.

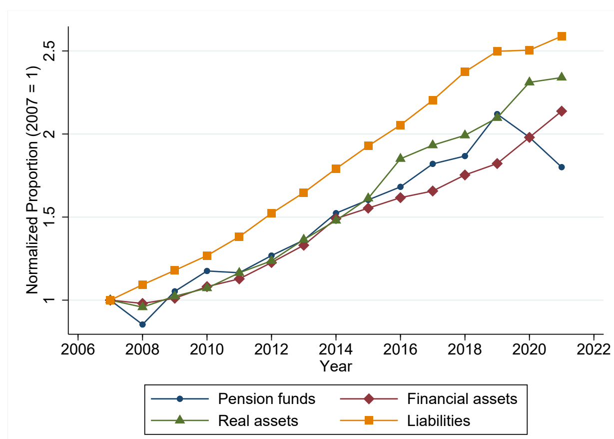
[1] In the new wealth literature, the inclusion of pension funds—particularly in public pay-as-you-go systems—remains contested. While privately funded pensions have clear individual valuations and may sometimes be inherited, public systems offer only future income. In the system of national accounts, they are valued as expected pension revenue to avoid understating wealth in these cases. This distinction has implications: anticipated public pensions reduce saving, supporting an "augmented" wealth concept (Roine and Waldenström, 2015). However, the standard definition of wealth in this literature excludes non-tradable assets like unfunded pensions from their definitions (Alvaredo et al., 2016).

To enhance the measurement of pension fund assets, we use a 5% random panel issued from the universe of accounts held by Chilean pension fund administrators, which precisely tracks the evolution of individual assets. We impute balances from the panel into the billionaire-adjusted financial survey based on a range of socio-economic characteristics, with the help of a machine-learning model inspired by Bloise et al. (2021). To measure housing assets, we construct a set of macroeconomic estimates of housing wealth, which are not officially reported in the local national accounting. Such estimates are based on Flores et al. (2018), who provide a solid estimates –yet for a shorter period– of the stock of housing assets, using comprehensive data on transactions, along with the national cadastre, thanks to hedonic pricing. We extend their estimates using administrative data on construction permits, including information on the average size of dwellings and the evolution of market prices by square meter. To validate our top wealth adjustment, we use individual administrative data from the Chilean tax authority, which provide a detailed snapshot of the top end of the wealth distribution for a single year, including both direct and indirect firm ownership.[2]

Figure 1: The evolution of Chilean private net wealth (2007-2021)



(a) Wealth over GDP



(b) Evolution by Asset

Notes. Elaborated by the authors using data from financial accounts (Central Bank), the housing price index (Central Bank and the National Tax Agency), and Flores et al. (2018). Panel A: The market value of total wealth, net of liabilities, is approximately three times the GDP in 2021. The "Real Assets" category includes formal housing and agricultural land, while the "Financial Assets" category comprises, stocks, cash and other financial assets. Panel B: All categories are expressed in 2021 billions of nominal USD and are normalized such that the value for 2007 equals 1.

[2] Access to this data was granted as part of a tax reform study conducted by the Chilean Finance Ministry under the supervision of Damian Vergara. The estimates were used to construct the denominator of top wealth shares, which were then complemented by aggregate wealth estimates from macroeconomic data, presented further in this paper.

As a preview of the macroeconomic setting for our distributional study, figure 1 presents the evolution of national wealth, net of liabilities, over a fourteen-year span. Panel (a) depicts a remarkable increase in volume from 2.4 times the GDP to over 3 times the GDP between 2007 to 2021. Real assets are the largest category, taking close to half the volume of total private assets, while the other half is roughly split into equal parts between pension funds and financial assets throughout the period. Panel (b) showcases the growth of each category, as a normalized index. Overall, all assets followed a similar growth path during the first half of the period. The faster growth of liabilities seems to have compensated such growth, leaving aggregate net wealth mostly unchanged relative to GDP. It is only after 2016, the implementation year of the 15% rate housing VAT, that net wealth growth accelerated, mainly driven by real assets. Pension funds experienced positive growth until 2019, after which they fell abruptly. Although the total growth rate between 2007-2019 was almost identical to that of real assets, a decline of 8% is observed after the coronavirus pandemic. Such decline was primarily driven by pension fund withdrawals permitted as part of financial aid measures to mitigate the economic impact of the pandemic. We expect pension funds and real assets to be among the main driving forces of distributional outcomes, since they stand out both in terms of total volume and in the size of their variations.

The distributional results presented in this paper reveal extreme levels of wealth concentration by international standards. They also reveal a slight decrease in wealth inequality, likely starting in 2016, which accelerated further during the pandemic. Two opposing forces seem to be behind these dynamics, the one increased the value of assets, while the other liquidated them. First, increasing prices in the housing market –although widespread– were faster towards the bottom of the distribution, provoking an equalizing effect. Second, the distributional outcome of pension fund withdrawals, which is less straightforward. After the withdrawals, we observed diminished pension assets across the whole distribution, but more markedly towards the bottom of the distribution. This fact alone could count for an increase in inequality, given its regressive distribution; however, it seems that the bottom 50% predominantly used these withdrawals to repay debt, while the middle 40% retained them as cash or invested in financial instruments. For the top 10%, decreases in pension and financial assets likely resulted from higher consumption or negative returns on risky investments. Although summary measures like the Gini coefficient remained stable over that period, it seems that top shares, especially the top 1% share dropped as an outcome.

We contribute to two strands of the literature. First, we extend and complement the longstanding research on wealth inequality measurement (Saez and Zucman, 2016, 2020; Kopczuk and Saez, 2004; Kopczuk, 2015; Smith et al., 2023; Jakobsen et al., 2020; Acciari et al., 2024; Garbinti et al., 2021; Alvaredo et al., 2018; Martínez-Toledano, 2020; Martínez-Toledano et al., 2023; De Rosa, 2025) by providing new evidence on wealth distribution dynamics in developing economies. Specifically, we document wealth inequality trends in Chile over a 14-year period, offering insights into the distributional effects of the housing VAT policy and early pension fund withdrawals.

Second, we contribute to the literature on imputation methods in wealth inequality research (Vermeulen, 2016; Cowell and Kerm, 2016; Alvaredo et al., 2018; Saez and Zucman, 2016) by implementing a two-stage approach, similar to Kenedi and Sirugue (2023), that leverages administrative pension data to predict household private pension balances. This approach employs machine learning techniques following Bloise et al. (2021). Finally, our imputation procedure is enhanced by incorporating the national accounts-consistent imputation method developed by Vermeulen (2018). This method enables the estimation of reliable wealth inequality trends in contexts with data limitations, particularly when private pension funds constitute a major component of household wealth.

The paper is organized as follows: Section 2 presents a review of the wealth inequality measurement literature and imputations techniques for missing data. Section 3 describes the methodology used to derive our results. Secrtion 4 provides details about the national context and our main data sources. Results are presented in Section 5, which includes a discussion. Finally, Section 6 synthesizes the results and provides a brief conclusion.

2 Literature review

Our paper is related to two strands of the literature, each of which we address in this section. We begin by reviewing the literature on wealth inequality measurement. Considering the limited financial literacy concerning Chilean pension funds, we also include a review of the literature on imputation techniques, particularly focusing on machine learning methods used in the intergenerational income mobility literature.

2.1 Wealth inequality mesurement

Significant progress has been made in measuring wealth inequality in recent years across high-income countries, especially in ensuring the correct measurement of top wealth, thanks to the use of tax data. For instance, Saez and Zucman (2016) and Smith et al. (2023) analyze the wealth distribution in the United States using capitalized income data, while Garbinti et al. (2021) applies a combination of capitalized data and survey data for France. Similarly, Alvaredo et al. (2018) examines wealth distribution in the UK using estate data and mortality multipliers, with Acciari et al. (2024) and Davies and Di Matteo (2021) conducting comparable analyses for Italy and Canada, respectively. For Spain, Alvaredo and Saez (2009) relies on wealth tax return statistics and administrative data, as do Roine and Waldenström (2009) for Sweden and Fagereng et al. (2020) for Norway. While some controversy remains over specific methodological aspects, a consensus has emerged that wealth concentration among top percentiles declined during the last century until around 1980, after which it has risen, albeit at varying rates across regions. For example, Blanchet and Martínez-Toledano (2023) provide evidence that wealth inequality is increasing more rapidly in the United States than in Europe.

Where tax data is absent, other methods have emerged as alternatives to adjust for the missing rich at the top of the wealth distribution, such as Vermeulen (2016) and Vermeulen (2018). These correction methods incorporate a few extreme observations from billionaires lists to data obtained from household financial surveys, assuming a Pareto distribution in between. Several studies have used these methods in different contexts. For instance, Bach et al. (2019) focused on France, Germany, and Spain, while Wodrich (2020) investigated Canada. Recent papers include Waltl and Chakraborty (2022), who used both corrective methodologies and analytical simulation approaches to reconstruct the micro-structure in the final dataset, leading to more accurate wealth distribution estimations for Austria and Germany, and Cantarella et al. (2023), who incorporated calibration techniques into the Pareto distribution methods. Across these studies, a common finding emerges: household surveys tend to underestimate the top 1% wealth share, leading to erroneous conclusions about inequality dynamics.

Very few studies on wealth inequality have focused on developing countries, mainly due to data limitations. For example, Carranza et al. (2025) highlight the scarcity, incompleteness, and contradictions in reliable wealth data in Latin America. However, where it can be measured, the concentration-levels are extreme, where the top 1% share of national private wealth ranges between 37% and 40.6%, placing it among the highest globally. Similarly, Gandelman and Lluberas (2024) analyze the relationship between wealth indicators and sociodemographic characteristics of household heads in Latin American countries such as Chile, Colombia, Mexico, and Uruguay. De Rosa (2024) estimate both aggregate wealth and its distribution in Uruguay, while Chatterjee et al. (2020) provide a similar analysis for South Africa, and Piketty et al. (2019) conducts a related study for China. Our work draws on the research by Flores and Gutiérrez (2021), who estimate an aggregate wealth series for the country using the aforementioned methods. Their findings reveal that household surveys tend to underestimate the true level of national wealth. In this context, we contribute to this literature by estimating trends in both aggregate wealth stock and wealth inequality for Chile, leveraging a combination of micro and macro data sources.

2.2 Imputation techniques for missing data

Our paper is closely connected to the literature on imputation techniques for missing data, which plays a significant role in both inequality measurement and intergenerational mobility studies.

In privately funded pension systems, ownership is clearly defined. However, individuals are not always aware of the exact amounts in their funds or how these accumulate over time. This lack of knowledge may stem from financial illiteracy, a challenge that is more pronounced in developing countries. Although it is typically measured by the understanding of concepts such as inflation and interest rates, financial illiteracy often correlates with knowledge about one's own financial assets. As documented by Hastings et al. (2013), developed countries tend to exhibit higher levels of financial literacy compared to developing countries, and as Karakurum-Ozdemir et al. (2019) suggests, financial literacy tends to increase with higher levels of education.

Evidence from Chile shows that respondents with higher financial literacy—measured using the OECD INFE index and other indicators of financial behavior, knowledge, and use of digital tools—report loan-related data more accurately compared to administrative records (Madeira and Margaretic, 2022). Additionally, respondents with greater financial literacy exhibit smaller discrepancies in debt ownership and reported amounts (Madeira et al., 2022). This pattern may also extend to pension funds. Consequently, in cases where households suffer from financial illiteracy, imputing pension fund values becomes a necessary approach.

The use of imputation methods in inequality literature is increasing, particularly regarding the utilization of machine learning techniques. Recent papers have employed these methods in the measurement of inequality of opportunity (IOP). For instance, Brunori et al. (2023) developed a framework of conditional inference regression trees and forests to estimate inequality of opportunity. Similarly, Brunori and Neidhöfer (2021) utilized conditional inference regression trees to identify relevant interactions of circumstances ("Romerian types") to estimate IOP in Germany. Additionally, Salas-Rojo and Rodríguez (2022) employed conditional inference algorithms to analyze how inheritances received contribute to measuring opportunities to accumulate wealth, assessing IOP in the United States, Canada, Italy, and Spain.

Moreover, recent studies in the field have used machine learning techniques for the imputation of wealth and income. For example, Suss et al. (2023) employed machine learning techniques to develop accurate predictive models of household wealth. Additionally, Brunori et al. (2023) explored various methods of correcting missing observations, revealing that sample reweighting based on probabilities of non-response is an effective approach for addressing these issues. More recently, Lucchetti et al. (2024) introduces a Least Absolute Shrinkage and Selection Operator with multiple imputation by Predictive Mean Matching (LASSO-PMM) method to estimate intra-generational income dynamics.

Our work also builds on Bloise et al. (2021), whose study examines different machine learning techniques used to handle low-quality data on intergenerational income data, providing insights to the classical Two-Sample Two-Stage Least Squares (TSTSLS) literature, which originated from Björklund and Jäntti (1997). Our contribution here lies in the utilization of those machine learning methods employed in the literature to predict missing data in household surveys. By integrating these techniques into the Vermeulen's methodology, we enhance the accuracy of wealth distribution measurements. This novel approach allows us to overcome data limitations and address issues related to item non-response bias, particularly about pension funds and financial assets data, resulting in a more comprehensive and reliable estimation of wealth distribution.

3 Methodology

In this section, we detail our methodological contributions. We begin by explaining our imputation of pension funds from administrative records to the household survey, addressing both the selection of individuals across data sets and the enforcement of global coherence with benchmark aggregates. We then discuss our version of the right-tail adjustment, which draws on common methods in the literature. We finish this section discussing the construction of wealth aggregates.

3.1 The imputation of pension funds

The imputation process is divided in two steps. First, individual observations are selected for imputation, based on the measured probability of having pension funds. Second, we rely on the Two-Sample Two-Stage Least Squares (TSTSLS) model, a method frequently used to predict incomes based on observable characteristics across two data sets.

3.1.1 Step 1: the selection of households

Let us assume we have knowledge of the true proportion of households with pension funds. In this case, the challenge lies in deciding which households to impute data to. One potential solution to this problem involves ranking individuals according to their likelihood of having pension funds. This can be achieved by leveraging an auxiliary sample that establishes the relationship between pension fund balances and individual characteristics. Subsequently, we can estimate the probability of having pension funds as a function of these individual characteristics.

Given this approach, we formulate the problem as estimating the probability of having pension funds, denoted as $Pr[\{P_{it} > 0\} = 1]$, which is a function of individual characteristics x_{it}. We represent this relationship as:

$$Pr[\{P_{it} > 0\} = 1] = g(x_{it}'\gamma). \tag{1}$$

We can estimate this model using any binary regression such as a Linear Probability Model (LPM) or a Probit regression, enabling us to predict the probability of having pension funds based on individual characteristics. Once we have estimated the probabilities, our solution involves ordering individuals from the most probable to the least probable of having pension funds. Subsequently, we impute data for these individuals until we achieve a proportion of individuals with pension funds that matches the known proportion in the population.

3.1.2 Step 2: the two sample two stage method

To estimate pension funds, we use a two-sample estimation procedure to address biases in self-reported data, which primarily result from non-response and misreporting.[3]

First, assume we have an auxiliary sample with data on pension fund balances, denoted by P_{it}. It is reasonable to argue that these balances are a function of individual characteristics. Specifically, as Equation 2 shows, if we had a panel of labor income and worked months by individual, the calculations might be direct: let $P_{i,t+j}$ represent the pension funds accumulated in the period $t+j$, and y^L represent labor income. Assuming a constant interest rate r, and a proportion of saved income κ we can approximate:

$$P_{i,t+j} = \sum_{k=0}^{j}(1+r)^k \cdot \kappa \cdot y^L_{i,t+j-k}. \tag{2}$$

Unfortunately, data at that level of specificity is practically non-existent. Given these relationships, one may suppose that $P_{it} = \sum_{k=0}^{j}(1+r)^k \cdot \kappa \cdot y^L_{i,t+j-k}(x_{it}) = f(y^L_i(x_{it}))$, where y^L_i represents permanent income and x_{it} is a vector of characteristics such as income, age, gender, and other relevant factors. These factors are significant because labor income is subject to different gaps associated with such characteristics, such as the well-documented gender wage gap (Goldin, 2014).[4] Utilizing this relationship, we can suppose that $P_{it} \equiv \phi x_{it}$ and then estimate

$$P^{ps}_{it} = \phi x^{ps}_{it} + \varepsilon_{it}, \tag{3}$$

$$\hat{P}^{p}_{it} = \hat{\phi} x^{p}_{it}. \tag{4}$$

In our proposed approach, we note that we only need the imputed pension funds, $\hat{\phi} x^p_i$, to proceed with our estimations of wealth inequality. Therefore, we can replace the data of individuals who

[3]The two sample procedures are standard in the inter-generational mobility literature, often used to predict the inter-generational elasticity of income in cases where parental income data is unavailable. See for instance (Björklund and Jäntti, 1997; Bloise et al., 2021).

[4]Additionally, there can be periods of time with gaps where workers do not save, such as during unemployment or maternal leaves in informal contexts.

have $P_i^p = 0$ with the imputed values $\hat{P}_i^p$.[5]

To estimate $\hat{\phi}$, we adopt the approach presented by Bloise et al. (2021), Brunori et al. (2023), and Meinshausen (2007) that use a relaxed-LASSO methodology. The objective is to predict the pension funds balances of individuals who report 0 with the smallest possible squared error, which can be formulated as follows:

$$\min\Big\{\mathrm{E}\Big[\big(P_{it}^p - \hat{P}_{it}^p\big)^2\Big]\Big\} = \min\Big\{\mathrm{E}\Big[\big(P_{it}^p - \hat{f}(x_{it}^{ps})\big)^2\Big]\Big\} \tag{5}$$

In this equation, $\hat{P}_{it}^p$ represents the imputed pension funds balances, which we aim to predict accurately. The squared error is then decomposed into variance and bias. Bloise et al. (2021) argue that a tradeoff exists between complex models, which have low bias but high variance, and simple models, which have high bias but low variance. Their findings suggest that the relaxed LASSO approach outperforms other methods, such as OLS, Elastic net, and Random forests, in predicting the missing pension funds balances with improved accuracy.[6]

3.2 Selection of the imputed households

Up to this point, and for each year, a Relaxed LASSO equation is estimated using an auxiliary sample that includes personal characteristics of surveyed individuals. The estimated coefficients are then used to impute pension fund balances. However, not all individuals in the household survey require imputation.

Data from the Chilean Superintendency estimate that approximately 48% to 58% of the population has pension funds (see Table D.1). Consider the year 2007: in that year, 33% of surveyed individuals reported having pension funds greater than zero, whereas administrative records indicate that 48.75% of the population held pension funds. A natural approach in this case would be to impute balances for the difference –15.75% of the surveyed individuals– so that the proportion of individuals with pension funds aligns with the actual population share. This, however, raises the question of whom to impute balances to.

[5]This argument can also be extended to the imputation of financial assets, although the underlying rationale may not be as straightforward. Nevertheless, one could contend that households maintain savings in financial assets relative to their income, and other individual characteristics may similarly influence this relationship, much like how they affect pension fund balances.

[6]Refer to Appendix A for more details about the relaxed Lasso procedure.

A direct approach is to use the same set of characteristics specified in equations 3 and 7 to estimate a Linear Probability Model (LPM) using the auxiliary sample. In this model, the dependent variable takes a value of one if an individual has pension funds and zero otherwise. The model then predicts the probability of an individual holding pension funds based on fundamental characteristics such as gender, age, personal labor income, and interaction terms. This yields an estimated probability of each individual having a pension fund balance greater than zero in the main sample.

Within the main sample, individuals are then ranked as follows: first, those who report positive pension fund balances (as shown in Table D.1, this corresponds to 33% of the sample for 2007); second, those who report zero pension funds, sorted in descending order based on their predicted probability of having pension funds. Finally, we replace reported zero pension balances with the imputed values obtained through the Relaxed LASSO model until the total proportion of individuals with pension funds matches the estimated population proportion.

3.3 Upper tail adjustment with billionaire lists

We use Vermeulen's (2016, 2018) methodology to handle missing wealth data at the upper end of the distribution. This approach entails leveraging two distinct samples: one from household surveys, which effectively captures the lower spectrum of the wealth distribution, and another from auxiliary sources, typically containing data on individuals representing the extreme upper levels of the distribution, such as billionaire lists. The goal is to bridge the gap between these two samples and produce a comprehensive picture of the entire wealth distribution.

We adopt the six-step procedure proposed by Vermeulen (2016) to adjust the national wealth measured in the household survey, which is typically notably lower than the aggregated national wealth measured by other methods, such as national accounts.

In the first and second steps, we construct adjustment factors for real and financial assets as well as liabilities. These adjustment factors are used to scale each household's assets and liabilities, respectively. This process aligns with the approaches taken in several studies, including Vermeulen (2018), Bach et al. (2019), Wodrich (2020), and Waltl and Chakraborty (2022), where adjustments were made for these three general categories.

However, we include an additional fourth category in our specification, which separates financial assets from pension funds. This distinction is particularly important in the context of countries with private pension systems because the dynamics of mandatory pension funds may differ from those of financial assets: the latter might be closely related to the capacity for savings and accessibility to capital income, while the former is closely linked to labor income, as discussed in Section 3.1.2. Indeed, poorer households may not own any financial assets, but as long as they are in the formal labor force, they may be accumulating pension funds.

Up to this point, we have an "underreporting adjusted survey". In step three, we estimate the Pareto tail using ordinary least squares (OLS).[7] However, other authors, such as Waltl and Chakraborty (2022), suggest alternative methods like quantile regression, which may be an alternative approach with favorable properties. We then create synthetic households with wealth distributed according to the Pareto distribution, placing them between the richest household in the household survey and the poorest household in the billionaires list. Steps four and five involve calculating the ratio of wealth components for the entire tail, assuming that the synthetic and billionaires households have ratios similar to the top of the household survey. Next, we compute the total real national wealth for that year, breaking it down by each component: real assets, financial assets, and liabilities. The last step includes comparing the aggregate value for real assets, financial assets (and pension funds if used), and liabilities with the national data. If the values do not match, we adjust the preliminary adjustment factors and repeat steps 2 to 6 iteratively until the results converge and align with the actual national data. In the following section we apply this methodology to estimate wealth inequality in Chile.

4 Data and context

In this section, we provide details on the Chilean case, standing as a country with a predominantly privately-capitalised pension system. We then comment on the main data inputs of our study, some of which we built ourselves, to then construct our macro-consistent wealth distribution estimates.

[7]See Appendix B for a detailed explanation of the Pareto tail estimation.

4.1 The Chilean context

4.1.1 The 2014 reform on housing transactions

The 2014 tax reform in Chile introduced the application of Value Added Tax (VAT) on real estate transactions (Laws 20.780 and 20.899, and Decrees 910 and 824). These legal modifications broadened the VAT's scope beyond direct sales by construction companies to include transactions by habitual sellers—those who sell properties within one year of acquisition or construction. Prior to the reform, VAT applied only to direct purchases from construction firms, with a partial tax credit available under the Special Tax Credit for Construction Companies (CEEC). The reform gradually reduced CEEC benefits, limiting them to properties valued at up to 74,000 USD starting in 2017 while maintaining a maximum credit of 8,350 USD per unit. Additionally, rental agreements with purchase options became subject to VAT. Certain exemptions remained, including properties with building permits issued before January 1, 2016, social housing supported by government subsidies, and transactions involving inheritance or donations. As Lozano and Idrovo (2024) documents, the VAT reform increased the price of new dwellings by 9.6% to 12.6%. This price increase led to changes in wealth distribution, which play a key role in explaining our distributional results.

4.1.2 The local pension system

Introduced during the Chilean dictatorship in 1981, the Chilean pension system underwent a significant transformation from a pay-as-you-go system to a fully funded one. This transition primarily impacted civilian workers, while military personnel kept their system unchanged. Over time, the system evolved into a mixed one, especially with the introduction of a public means-tested pillar in 2008. This component was further reformed in 2022 to provide monthly transfers to pensioners[8]. The design of the individually funded system includes pension fund administrators, which are defined as private for-profit entities. The oversight and regulation of the market is conducted through a public entity called *Superintendencia de Pensiones* or Superintendency of Pension Funds. Under this system, employees have the autonomy to chose their fund's administrator (called AFP, *Administradoras de Fondos de Pensiones*) and allocate 10% of their pre-tax monthly wages into a designated pension account.

[8] See Troncoso (2022), for a detailed description

Continuing with the notation employed in Section 3.1.2, we approximate the pension funds ($P_{i,t+j}$) as follows.

$$P_{i,t+j} = \sum_{k=0}^{j} (1+r)^k \cdot 0.1 \cdot y^L_{i,t+j-k}. \tag{6}$$

Two notable simplifications are essential to consider in this context. Firstly, the interest rate applied to pension funds varies depending on age and economic expectations, given that workers have the flexibility to choose their AFP and allocate their funds across five risk categories. These categories are typically correlated with age, as it is advised to invest in less risky funds as one approaches retirement. Secondly, mandatory contributions are capped at approximately 3000 USD per month.

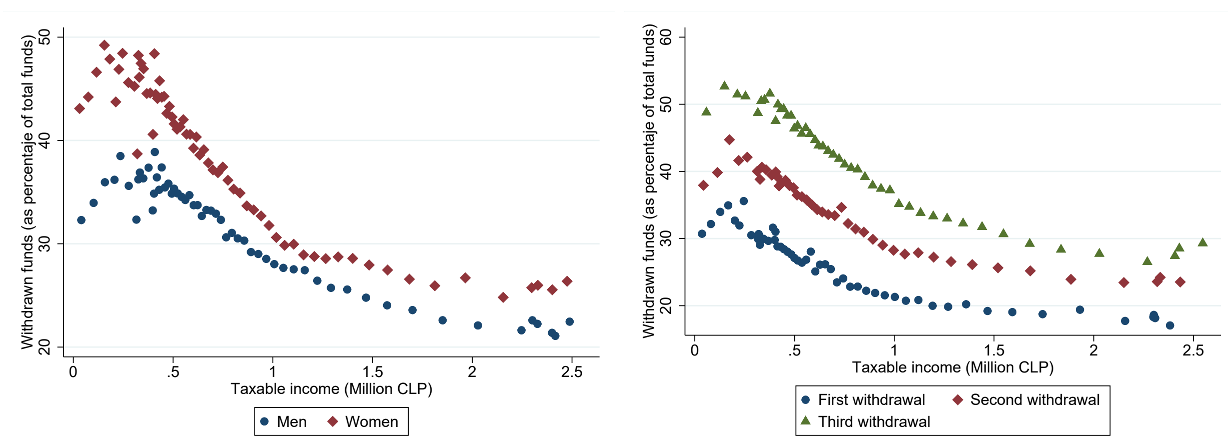
Regarding eligibility for old-age pensions, men and women become eligible at the ages of 65 and 60 years old, respectively. Pensioners are offered the choice between a programmed withdrawal or an annuity, with the former being inheritable.

Upon the death of an individual, their pension funds can be inherited and distributed either through survivor pensions or as an inheritance. Survivor pensions are allocated to specific beneficiaries as defined by law. These beneficiaries include the spouse or civil partner of the deceased, children under the age of 18 (or up to 24 if they are enrolled in full-time education), children with disabilities regardless of age, and parents, provided they were financially dependent on the deceased. If no eligible survivors exist, or if a balance remains after the survivor pensions have been paid, the remaining funds may be withdrawn by the legal heirs as a lump sum. This ensures that pension savings are either allocated to dependent family members or, in their absence, transferred through inheritance.

Given the notion of ownership embedded in the Chilean pension system, a constitutional reform in July 2020 facilitated the withdrawal of up to 10% of pension funds to support families adversely affected by the pandemic. Subsequent withdrawals took place in December 2020 and May 2021, collectively amounting to 19% of Chile's GDP. These withdrawals necessitated significant liquidity policy interventions by the Chilean central bank, leading to tighter labor market conditions and an average reduction of 21% in contributory pensions for future retirees (Briones et al., 2023; Inzunza and Madeira, 2023).

Figure 2 illustrates withdrawn amounts, as a percentage of accumulated funds, at different income levels. We follow Cattaneo et al. (2024) in the construction of our binned scatter plot, which allows showing the conditional means of withdrawn funds across taxable income bins, while controlling for variables such as the number of contribution periods, and analysing heterogeneity across other variables. As shown in the figure, these withdrawals were more pronounced in the lower percentiles of taxable income. They also had a greater impact on women than on men, and subsequent withdrawals seem to have been increasing in their regressivity. Given the expected strong correlation between declared income and wealth levels, we expect to observe a regressive impact of such withdrawals on the wealth distribution too. In some cases, withdrawals accrued to more than 30% of total funds. This pattern is not only a result of fluctuations in withdrawal amounts over time but also reflects the legal provisions governing withdrawals. Specifically, the law stipulated that individuals with less than 35 UF (approximately 1,000 USD) could withdraw the entirety of their funds, while those with over 1,500 UF (approximately 43,000 USD) were limited to a maximum withdrawal of 150 UF (approximately 4,300 USD). As a result, in certain cases, individuals in specific percentiles withdrew over 50% of their total funds.

Figure 2: Withdrawn Funds by Taxable Income



(a) By Gender

(b) By Withdrawal Round

Notes. Own elaboration based on the Administrative Database Sample of Requests for the First, Second, and Third Withdrawal of Pension Funds provided by the Superintendency of Pensions.

In summary, the Chilean pension system is characterized by two key features. First, pension funds are considered the property of individual affiliates, making them inheritable. Second, although pension funds were not initially regarded as a liquid asset, specific political developments—along with ongoing debates regarding access for non-pensioners—are increasingly shaping wealth distribution within the country.

4.2 Wealth aggregates

We define wealth in terms of household disposable net wealth, which is the market value of all assets minus liabilities. To quantify aggregate wealth across various categories, we extend the time coverage of estimates first introduced in Flores and Gutiérrez (2021), from period 2008-2019 to 2007-2021. Such estimates rely on sound data from the Central Bank to value financial assets, as well as composite estimates for non-financial assets. Housing assets, in particular, are anchored to estimates from a study applying hedonic-pricing based on a comprehensive data set of national housing transactions, along with an exhaustive national cadastre. We use anchor estimates to extrapolate values using administrative data on construction permits, including information on the average size of dwellings and the evolution of market prices by square meter (see Appendix C for details).

4.3 Distributional wealth data

4.3.1 Wealth surveys

For distributional estimates, one of the building blocks of our estimates is the Household Finance Survey (*Encuesta Financiera de Hogares* in spanish) conducted by the Central Bank of Chile. This survey provides comprehensive information on various wealth components, including real and financial assets and liabilities, for the years 2007, 2011, 2014, 2017, and 2021. It is representative of all urban households in the country, making it a valuable data source for our analysis of wealth distribution dynamics in Chile across these specific years. Regarding pensions, the questionnaire distinguishes between voluntary and mandatory pension fund contributions[9].

As a preliminary analysis, we measured top shares and other distributional estimates directly from the un-adjusted survey (see appendix D). Our primary concern regarding these results stems from their inconsistency with international evidence, since they indicate a higher concentration of income compared to wealth, which is unseen and unreasonable given how concentrated wealth

[9]These pension fund balances, however, are reported by the individual responding to the survey, typically the household head, and represent individual balances rather than the aggregate funds of the households. We will elaborate on this point later.

tends to be. The literature on top income shares suggests a top 1% income share well above 22% between 2007-2021, while the survey-based top wealth share is closer to 15% for most of the period (Flores et al., 2020; Fairfield and Jorratt De Luis, 2016; Lopez et al., 2016).[10]

The observed problems in household wealth surveys stem from several factors. While these surveys offer a direct means of observing household wealth, they encounter common challenges in accurately capturing wealth information for the top wealth holders. Some of the key issues include sampling errors, where the sample may not be representative of the entire population, leading to inaccurate survey results. Response errors also play a role, occurring when families inaccurately report the value of their assets or liabilities. Additionally, a significant concern is the phenomenon of differential unit non-response, where high-net-worth families may be less likely to participate in surveys, resulting in an underrepresentation of their wealth.

Considering all the challenges previously discussed, it is reasonable to expect that the household survey (EFH) may not accurately measure the wealth composition and distribution. This expectation aligns with the observations depicted in Figures D.2 and D.3. In Figure D.2, we observe that the EFH underestimates the aggregate national wealth. Furthermore, one might anticipate that wealthy households hold a substantial proportion of their assets in financial assets. However, this pattern is not fully represented in Figure D.3, potentially due to limitations in the survey's ability to accurately capture the asset composition of the wealthiest individuals. As a result, the EFH severely underestimates aggregate financial wealth. These problems challenge the reliability and comprehensiveness of wealth household surveys, particularly in capturing the wealth of the top percentiles of the population. Therefore, it is crucial to employ corrective methodologies to mitigate these issues and provide more accurate wealth distribution measurements.

This observation of the EFH dis-aligning with actual aggregate wealth also motivates a thorough analysis of the survey data itself, focusing specifically on understanding how households respond to the survey questions. In particular, we concentrate on item non-response related to financial assets and pension funds, as we have previously observed that the survey fails to accurately measure these components compared to real assets.

[10] As a reference, Saez and Zucman (2016) for the US, the top 1% holds a share of 16% in terms of income and over 40% in terms of wealth. In Colombia, Londoño-Vélez and Ávila-Mahecha (2021) provides evidence indicating top 1% shares of 20% for income and 45% for wealth in Colombia. In South Africa, Chatterjee et al. (2022) presents findings of 55% for wealth and 19% for income.

Tables D.1 and D.2 presents data on the number of households (in millions) with non-real assets greater than zero. For instance, in 2007, out of a total of 3.8 million households surveyed, 14.93% reported having financial assets greater than zero. We conduct separate analyses with and without cash, as this item wasn't included in the surveys of 2007 and 2011.

Two key findings emerge from Table D.2. Firstly, the percentage of individuals who declare having financial assets is notably lower in the first two years compared to the following three years. This suggests the presence of item non-response in the survey data. Specifically, the proportion of respondents reporting financial assets in 2011 is particularly low. If those who do respond are concentrated in the top deciles of wealth, it could lead to a potential overestimation of wealth inequality, which appears plausible. A similar argument can be made for 2007 and 2014, although to a lesser extent.

Secondly, we have clear evidence of item non-response, particularly concerning pension funds. For instance, taking the example of 2014, we observe that the proportion of individuals who remember their pension fund balances is low compared to the total surveyed. In December 2014, a total of 9,746,467 individuals were affiliated with the AFP system,[11] which corresponds to approximately 55% of the total population. However, only 15.7% of surveyed individuals reported having pension funds. This discrepancy highlights a significant underrepresentation of pension fund holders in the survey, with the reported proportion far below the actual 55% affiliation rate.

4.3.2 Forbes data

As we discussed, Vermeulen's methodology involves combining survey data with a few extreme observations to mitigate the downward bias when estimating a power law distribution represented by a Pareto distribution—which is accurate for the top of the income and wealth (Blanchet et al., 2022). These extreme observations are typically obtained from national or international billionaire lists, with a well-known example being the annual Forbes World Billionaires list, which includes individuals with estimated wealth above US$1 billion.

Several critiques have been raised regarding the use of the Forbes list, including concerns about inconsistent data collection from various sources and the questionable completeness of these lists (Bach et al., 2019). Additionally, ambiguity arises due to fluctuations in exchange rates, which can

[11] See this source and Table D.1.

impact the accuracy of wealth estimations (Waltl and Chakraborty, 2022). National lists, on the other hand, offer a more comprehensive approach as they do not solely focus on billionaires and are not susceptible to exchange rate issues. However, for the Chilean case, these national lists only include data between 2011 and 2017, limiting their utility for a more extended time frame analysis. Considering these factors, we have chosen to utilize the Forbes list for our analysis, despite its limitations.

The EFH survey is typically conducted at the conclusion of each year, and as a result, we rely on the Forbes database for the subsequent year. For instance, the survey for 2007 was conducted between November 2007 and January 2008, while the Forbes list for 2008 was published in March 2008. Therefore, we use the data from the latter publication instead of the former. The Forbes list generally includes information on market closings and exchange rates, which we utilize to convert billionaire data from US dollars to Chilean *pesos*. For example, we refer to the data corresponding to the exchange rate on February 11th for the year 2008.

4.3.3 Administrative data on pension funds

We use an administrative database on pension funds provided by the Superintendency of Pension Funds as the auxiliary sample for imputation. Although the administrative database lacks detailed personal data compared to the household survey, it allows us to provide imputations for all the years of the survey.

Since April 2007, the Superintendency has been collecting a monthly random sample of 5% of active affiliates to the pension system. This results in between 200,000 and 500,000 individuals each year, providing panel data for every month, accounting for between 2,000,000 and 5,000,000 observations annually. This data includes information about the accumulated funds in each person's account, along with characteristics such as gender, age, number of months worked formally, nationality, and formal labor income, from which the 10% contribution is extracted.

This allows us to impute funds based on personal characteristics, with the imputation performed separately for each year. We use the accumulated funds at the end of each year as the measure of individual pension funds and consider the average of monthly personal income to avoid the effects of unemployment on the estimation.

5 Results

5.1 Baseline distributional results

We begin our inequality analysis by examining top wealth shares. That is, the share of total wealth held by the top 1% and the top 10% wealthiest households. Our estimated series are showcased in figures 3a and 3b, as well as in table 1.

Table 1: Top wealth shares (%), 2007-2021

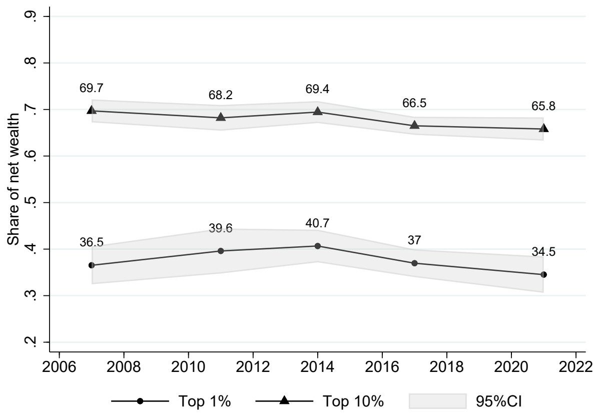
Year	Top 1%		Top 10%	
	With Pension Funds	Without Pension Funds	With Pension Funds	Without Pension Funds
2007	36.5	47.5	69.7	80.4
	(2.1)	(2.5)	(1.3)	(1.2)
2011	39.6	49.5	68.2	77.8
	(2.5)	(3.1)	(1.4)	(1.5)
2014	40.7	51.9	69.4	79.7
	(1.8)	(2.2)	(1.2)	(1.2)
2017	37	46.3	66.5	75.5
	(1.5)	(1.9)	(1)	(1)
2021	34.5	43.4	65.8	72.8
	(2)	(2.3)	(1.3)	(1.3)

Notes. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*), Forbes list of the year, Vermeulen (2018) methodology, and the relaxed LASSO imputation proposed in this article. Both specifications include imputation of financial assets. Standard errors in parenthesis. **Lecture.** In 2014, the top 1% of the population held 40.7% of the total wealth when pension funds are included and 51.9% when they are excluded.

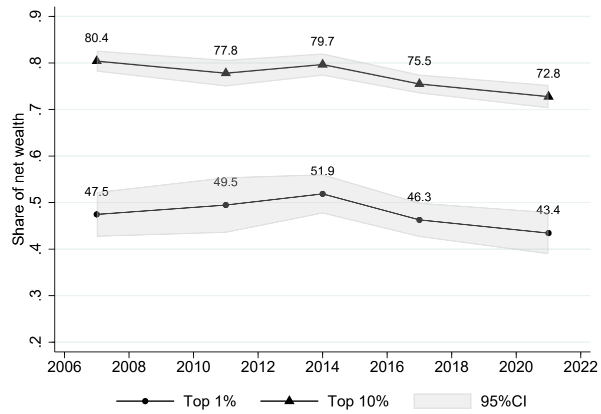
The first notable result is how massive these shares are. Our benchmark estimates, including private pension funds, indicate that the top decile of the distribution controls more than 2/3 of national private wealth, with almost half of that amount concentrated within the top percentile. For comparison, we also report results from the wealth distribution excluding pension funds, which is conceptually comparable to the situation in countries with pay-as-you-go pension systems, such as many European countries.[12] Excluding pension funds form the definition of total wealth, the wealthiest top 1% controls close to half of total wealth, while the top 10% controls close to 3/4 of it for most of the period. This suggests that pension funds play a crucial role in shaping the overall wealth distribution in Chile, and their inclusion or exclusion in the analysis can have a

[12]In such cases, retired individuals receive pension revenue without having to accumulate private funds. This conceptual mismatch usually makes mandatory privatized systems seem more equal in comparison, since they make individual pension wealth visible. However, even in privatized systems, pension funds remain a controversial category, mostly because control over the assets and freedom to liquidate them is often extremely limited.

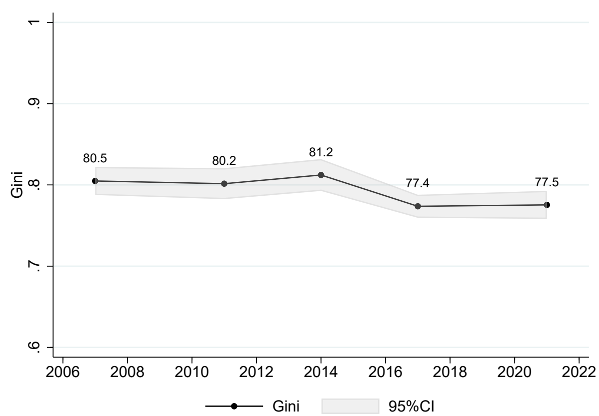
Figure 3: Adjusted top 1% top 10% wealth shares, 2007-2021



(a) Including pension funds



(b) Excluding pension funds



(c) The Wealth Gini coefficient (with pensions)

Notes. Adjusted series include the right-tail adjustment using Vermeulen (2018) methodology. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*), Forbes list of the year, and relaxed LASSO imputation proposed in this article.

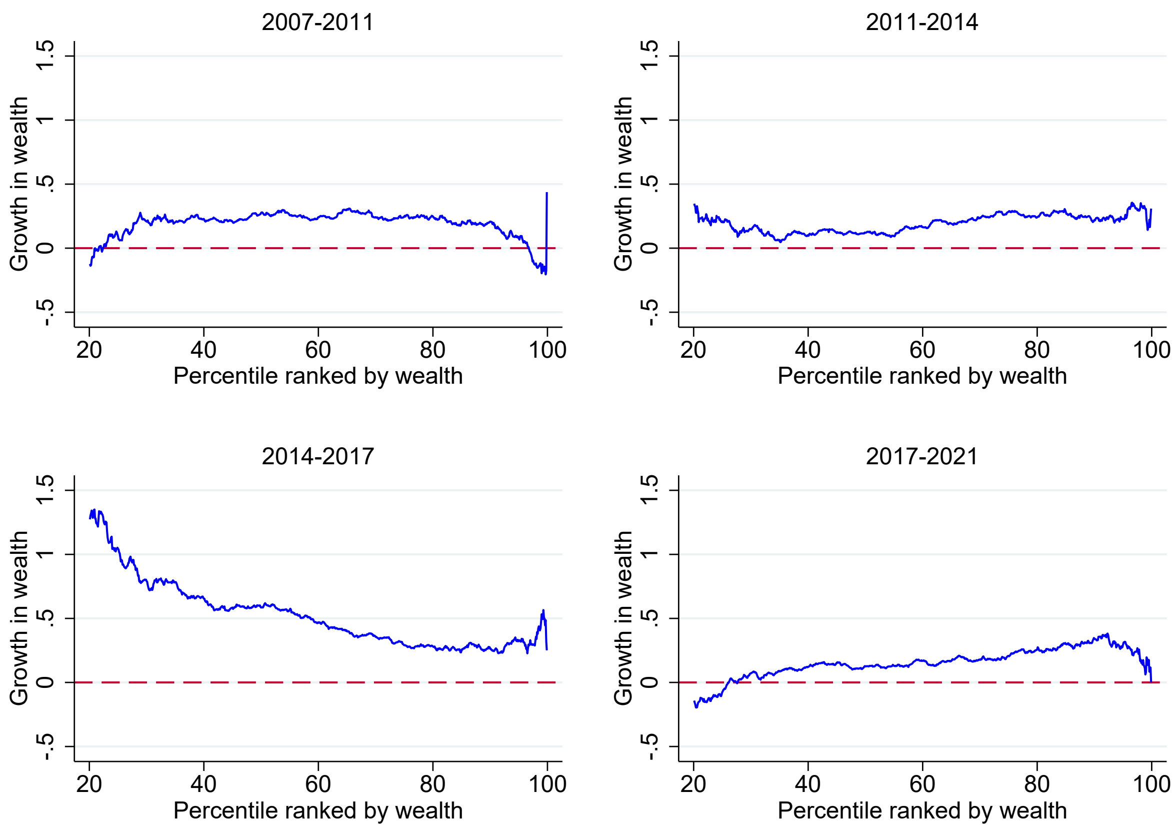
substantial impact on the measured inequality. In our benchmark estimates, wealth inequality in Chile is significantly higher than in Western Europe, where estimates mostly range between 20-25% (Chancel et al., 2022). However, our estimates fall below the extreme levels observed in South Africa, where values approach 50% (Chatterjee et al., 2022). The closest comparison to our estimates would be the United States or Colombia (see figure E.2).

The trend described by our estimates indicates a slight decrease in wealth inequality in Chile between 2007 and 2021 in all scenarios, while staying at extreme levels. When measured as the share of the top 10%, the decrease starts early in the series, while the top 1% increases between 2007-2014, only falling below its initial level at the very end of the series, in 2021. This suggests that, the top

1% share experienced different dynamics than the following 9% (P90-P99). A similar pattern is also reflected in Figure 3c, illustrating that the Gini coefficient –a more comprehensive measure of inequality– remains stable or even slightly increases until 2014 and only decreases thereafter. From a variety of angles, inequality appears to have slightly decreased during the period, which could be attributed to various factors.

We now turn to a more detailed analysis of the growth distribution across wealth percentiles, regardless of composition. Figure 4 presents Growth Incidence Curves as defined by Ravallion and Chen (2003), showing wealth growth as a function of percentiles.

Figure 4: Wealth growth incidence curves



Notes. This figure depicts the growth incidence curves for each period in our benchmark series, including private pension funds (see figure 3). Percentiles below P15 are ignored to avoid including households with near-zero or negative net, where changes in appear much larger in magnitude, but are less pertinent for our analysis (see figure G.1).

From 2014 to 2017 there is a more progressive trend in wealth growth, where wealth growth decreases convexly across percentiles, except for the top 5% where it increases again. The period from 2017 to 2021 exhibits a heterogeneous pattern: wealth growth is regressive between the 20th and 90th percentiles, and then progressive thereafter. While these findings are consistent with Chancel et al. (2022), indicating that billionaires increase their wealth (as seen in the highest wealth household each year), growth rates for the top 10% and top 1% are lower than those for preceding percentiles. This indicates that the decline in wealth inequality is explained by an increase in the wealth of the upper-middle of the distribution, in parallel with a slower than average growth at the top end.

5.2 From housing gains to pension losses

To examine the distribution of gains and losses, we classify households into four groups: the bottom 50% of the distribution, the middle 40%, the top 10%, and the top 1%. It is important to note that our data do not track individuals over time, and wealth rankings are cross-sectional. Therefore, this analysis provides an approximation of how gains and losses are distributed. Given the size of these groups and the expected limited mobility between them in this period, we consider them to be broadly representative. The first four panels in figure 5 represent the share of total assets and liabilities held by every group over time, while the last panel shows the aggregate size of each item relative to total net wealth.

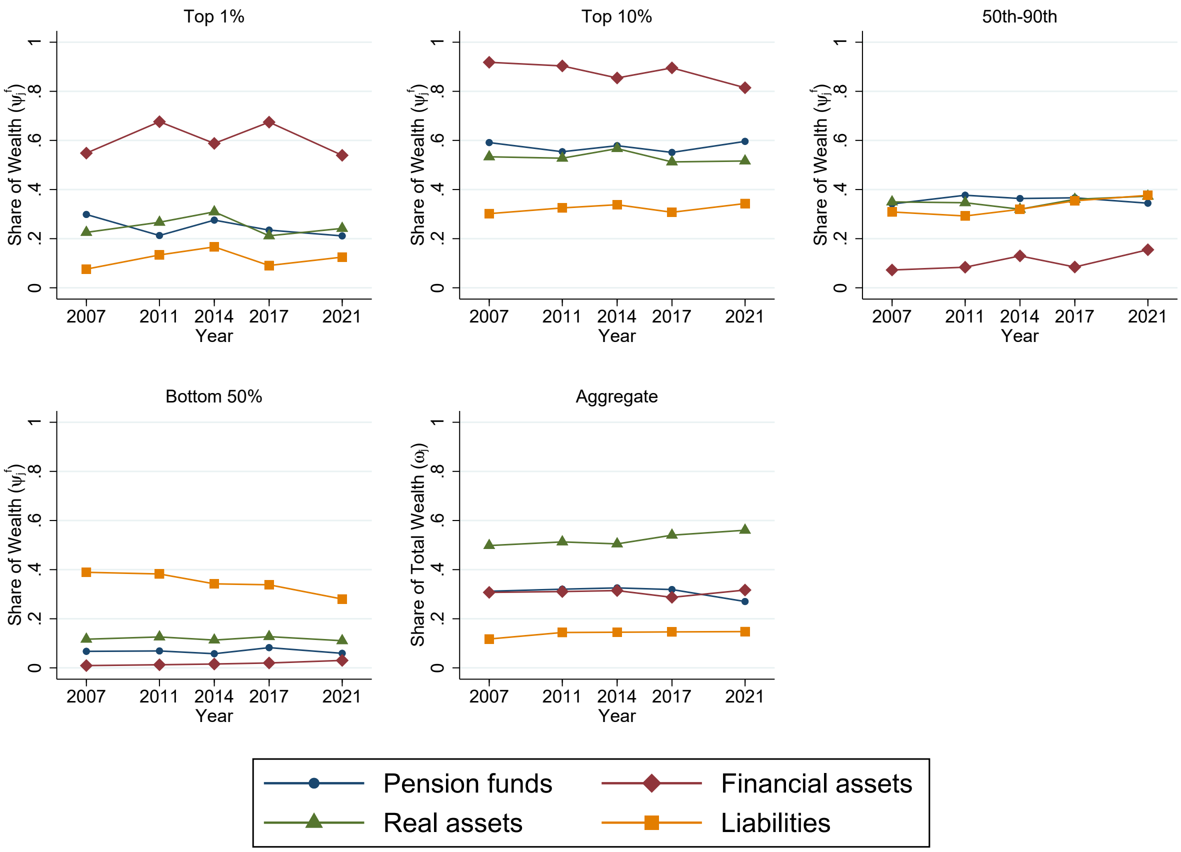
The distribution of real assets remains relatively stable over time across all groups, with approximately 50% held by the top 10%, 35% by the middle 40%, and 15% by the bottom 50%. This suggests that the rise in housing prices driving real asset growth was distributed somewhat proportionally to initial endowments. Although the top 10% holds the largest share of these assets and benefits the most from their appreciation, the overall effect was a slight decrease in inequality. This occurs because real assets constitute a larger share of total wealth for lower groups, whereas the top 10% also holds substantial shares of other assets, which grow at a slower pace (see figures 5 and 6).

The share of financial assets held by the bottom 90% appears to increase over the period, while the share held by the top 10%, particularly the top 1%, declines. This trend also moves in the direction of lower inequality, yet our confidence in this finding is lower due to greater measurement error in household reports of financial assets. Table D.2 shows that the share of households reporting positive financial assets, excluding cash, increased from 15% to 30% over the period. While this shift may partly reflect greater access to financial instruments among higher-middle-income households, it could also result from improved measurement of these assets.

Pension fund shares remain stable for most of the period, but notable changes occur toward the end. In aggregate terms, the pension fund withdrawals allowed between 2017 and 2021 account for a significant decline in pension wealth (Figure 1). During this period, the top 10% was the only group to see an increase in its share of pension assets, while all other groups experienced a decline. Even the top 1% saw a decline, indicating that the increase only affected high-wealth households (p90–p99) but not the top 1%. This suggests that the bulk of pension losses were concentrated in the bottom 90% of the wealth distribution.

Regarding liabilities, we observe a progressive reduction in the debt held by the bottom 50% throughout the period, with a steeper decline toward the end. This may indicate that a portion of the pension withdrawals was used to reduce debt rather than for consumption. However, the share of total debt increased for the middle 40% and the top 10% of the household distribution.

Figure 5: Share of assets and liabilities by wealth groups, 2007-2021



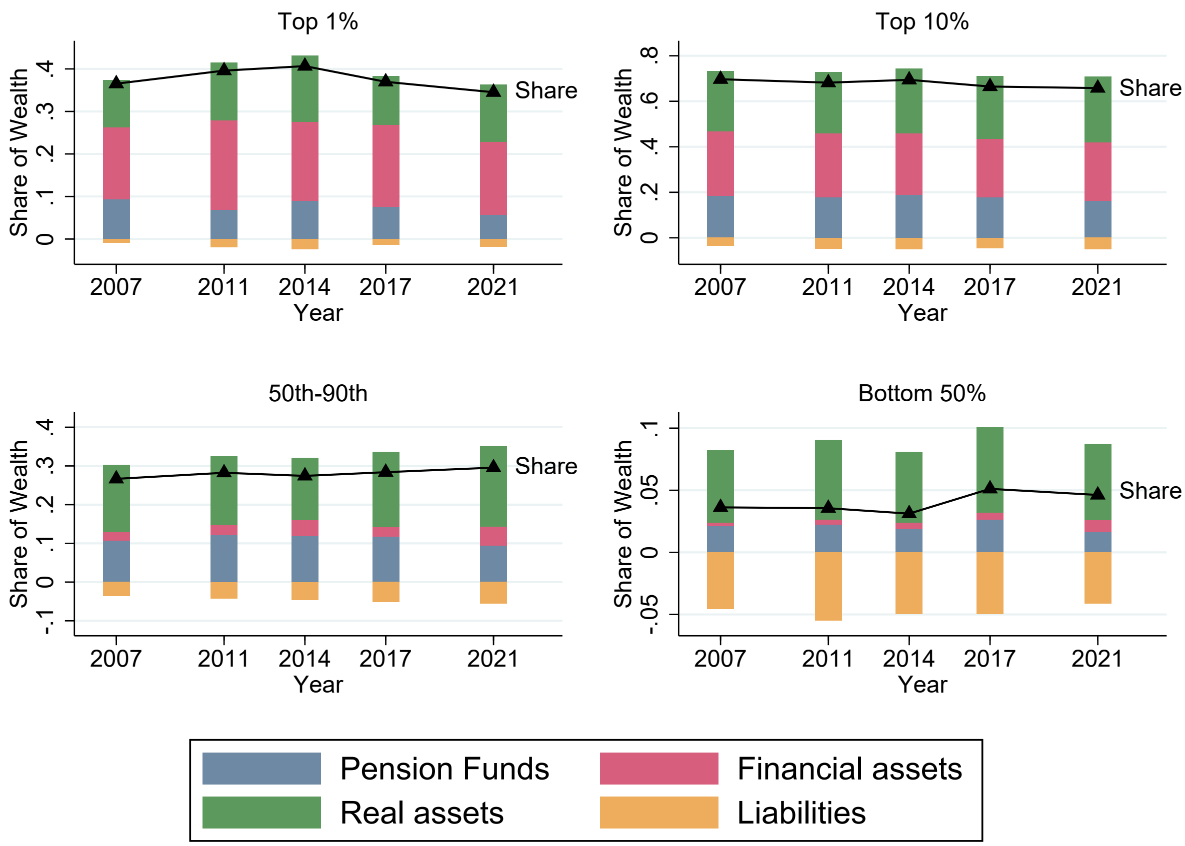
Notes. In the first four panels, shares represent the part of each wealth category held by a certain group. In the last panel, shares are expressed as parts of total wealth.

Thus, in wealth inequality dynamics between 2007 and 2021 can be attributed to four main factors: (i) housing gains were distributed proportionally to initial endowments, (ii) pension losses between 2017 and 2021 were mainly concentrated in the bottom 90%, (iii) financial wealth may have become more significant for the middle 40% during the same period, and (iv) debt declined for the bottom 50%.[13] These trends align with two underlying dynamics at the macroeconomic level: the rising value of housing and agricultural assets, which resulted from the introduction of the housing VAT, and the effects of pension fund withdrawals, which were either consumed, used for debt repayment, or retained as savings or cash. The distributional patterns suggest that the bottom 50% primarily allocated these funds to debt repayment, reducing their liabilities, while

[13] Table G.1 presents detailed share decompositions between 2017 and 2021 for our preferred fractiles.

the middle 40% held them as cash or invested in financial instruments such as savings or mutual funds. Meanwhile, the decline in both pension funds and financial assets among the top 10% may reflect increased consumption or negative returns on high-risk assets, such as stocks or pension fund portfolios.

Figure 6: Portfolios of Wealth groups over time



Notes. The sum of the height of all assets minus the height of liabilities, corresponds to each group's wealth share.

5.3 Robustness checks

To address potential concerns about the specific imputation method used in our analysis, the relaxed LASSO, we conducted further analyses using alternative machine learning methods. We utilized a specification where both pension funds and financial assets are imputed, with the former using administrative data and the latter using survey data. Tables F.1 and F.2 presents the results of varying the imputation method for estimating wealth inequality in Chile. We considered alternative machine learning methods, including LASSO regression, Random Forest regression, Elastic Net regression, and Ridge regression, in addition to the Relaxed LASSO used in our main analysis.

For the top 1%, the results from the tables show that the choice of imputation method has a minor impact on the estimated wealth inequality. For example, in 2007, the estimated share of wealth for the top 1% ranges from 32.5% (LASSO) to 33.9% (Random Forests), all of which are slightly lower than the 36.5% shown in the baseline. By 2011, the estimates range between 39.1% and 39.6%, closely aligned with the 39.6% shown in the baseline. These differences indicate that while the imputation method introduces some variability, the overall trends remain consistent.

For the top 10%, the tables reveal slightly higher estimates for the earlier years compared to the baseline. In 2007, the wealth share for the top 10% ranges from 63.2% (LASSO) to 64.2% (Random Forests), which is significantly lower than the 69.7% shown in the baseline. By 2011, the estimates range between 65.5% and 66.7%, again lower than the 68.2% reported in the baseline. However, between 2014 and 2021, the estimates from the tables are closer to those in the baseline. For example, in 2021, the top 10% wealth share is estimated between 64.0% and 64.3%, while the baseline shows 65.8%. These differences suggest that imputation methods can slightly underestimate wealth shares for the top 10%, particularly in earlier years, though the general dynamics remain similar.

We also present the results of two alternative specifications: one that includes only the imputation of pension funds, and another that imputes pension funds, financial assets, and cash as separate categories (in contrast to our original specification, where cash was included as part of financial assets). In this context, the original specification, which imputes pension funds and financial assets together, can be viewed as an intermediate specification in terms of the level of detail in wealth category separation.

As shown in Tables F.3 and F.4, the results are consistent with our original findings. However, the specification including only pension funds tends to increase the estimated shares slightly in some cases, while the specification separating pension funds, financial assets, and cash slightly reduces those shares. This pattern is observed for both the top 1% and the top 10%. Despite these minor differences, the overall results, the gaps, and the trends towards the end remain unchanged.

6 Conclusion

This study examines how housing appreciation and pension fund withdrawals shaped Chile's wealth distribution between 2007 and 2021. By incorporating administrative data into household wealth surveys, we provide a more detailed assessment of wealth concentration and its evolution over time. While Chile exhibits high levels of wealth inequality, our findings suggest a slight decline, primarily driven by two contrasting forces: rising housing prices, which benefited middle-income households, and pension fund withdrawals, which reduced wealth across most of the distribution. Lower-wealth households seem to have primarily used these withdrawals to repay debt, while wealthier individuals retained them as savings or invested in financial instruments. These results illustrate how policy decisions and macroeconomic shifts can influence wealth dynamics.

Our methodological approach contributes to ongoing efforts to refine wealth distribution estimates, particularly in economies with privatized pension systems. We combine a micro-adjustment where we impute for non-reported pension funds using machine learning methods commonly used in intergenerational income mobility literature, and a macro-adjustment where we correct the upper tail of the distribution assuming a Pareto distribution. The proposed method yields two interesting results: first, there is a substantial disparity in inequality if we consider or not the pension funds as a category, with specifications that include them reducing inequality (measured as the share of the top 1% or 10%) by approximately 10 percentage points. Second, economic crises can be seen as moments of significant changes in the wealth distribution, as evidenced by the changes observed after 2007 and during the coronavirus pandemic.

Acknowledgements

We thank Isaías Cataneo, Raimundo Smith, David Valencia, Javier Cortés Orihuela, and Damian Vergara for their contributions to early methodological discussions. We also acknowledge the participants of SECHI 2024 and the Seminar of the Chilean Superintendency of Pensions for their valuable feedback.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Acciari, P., F. Alvaredo, and S. Morelli (2024). The concentration of personal wealth in italy 1995–2016. *Journal of the European Economic Association*, jvae002.

Alvaredo, F., A. Atkinson, L. Chancel, T. Piketty, E. Saez, and G. Zucman (2016). Distributional national accounts (dina) guidelines: Concepts and methods used in wid. world.

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. and E. Saez (2009). Income and wealth concentration in spain from a historical and fiscal perspective. *Journal of the European Economic Association 7*(5), 1140–1167.

Bach, S., A. Thiemann, and A. Zucco (2019). Looking for the missing rich: Tracing the top tail of the wealth distribution. *International Tax and Public Finance 26*(6), 1234–1258.

BCCh (2014). Indice de Precios de Vivienda en Chile: Metodología y Resultados.

Björklund, A. and M. Jäntti (1997). Intergenerational income mobility in sweden compared to the united states. *The American Economic Review 87*(5), 1009–1018.

Blanchet, T., J. Fournier, and T. Piketty (2022). Generalized pareto curves: theory and applications. *Review of Income and Wealth 68*(1), 263–288.

Blanchet, T. and C. Martínez-Toledano (2023). Wealth inequality dynamics in europe and the united states: Understanding the determinants. *Journal of Monetary Economics 133*, 25–43.

Bloise, F., P. Brunori, and P. Piraino (2021). Estimating intergenerational income mobility on sub-optimal data: a machine learning approach. *The Journal of Economic Inequality 19*(4), 643–665.

Briones, L., G. Carlomagno, and P. García (2023). Inflation and labour markets in the wake of the pandemic: the case of chile. Technical report, Central Bank of Chile.

Brunori, P., P. Hufe, and D. Mahler (2023). The roots of inequality: Estimating inequality of opportunity from regression trees and forests. *The Scandinavian Journal of Economics 125*(4), 900–932.

Brunori, P. and G. Neidhöfer (2021). The evolution of inequality of opportunity in germany: A machine learning approach. *Review of Income and Wealth 67*(4), 900–927.

Brunori, P., P. Salas-Rojo, and P. Verne (2023). Estimating inequality with missing incomes. Unpublished.

Cantarella, M., A. Neri, and M. G. Ranalli (2023). Estimating the distribution of household wealth: Methods for adjusting survey data estimates using national accounts and rich list data. *Review of Income and Wealth*.

Carranza, R., M. De Rosa, and I. Flores (2025, 03). Wealth inequality in latin america (2000–2020): data, facts and conjectures. *Oxford Open Economics 4*(Supplement), i36–i54.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng (2024). On binscatter. *American Economic Review 114*(5), 1488–1514.

Chancel, L., T. Piketty, E. Saez, and G. Zucman (2022). *World inequality report 2022*. Harvard University Press.

Chatterjee, A., L. Czajka, and A. Gethin (2020). *Estimating the distribution of household wealth in South Africa*. Number 2020/45. WIDER Working Paper.

Chatterjee, A., L. Czajka, and A. Gethin (2022). Wealth inequality in south africa, 1993–2017. *The World Bank Economic Review 36*(1), 19–36.

Cowell, F. A. and P. Kerm (2016). Wealth inequality: A survey. *A Collection of Reviews on Savings and Wealth Accumulation*, 93–137.

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.

De Rosa, M. (2024). Wealth inequality in the south: Multi-source evidence from uruguay. *Review of Income and Wealth*.

De Rosa, M. (2025). Wealth inequality in the south: Multi-source evidence from uruguay. *Review of Income and Wealth 71*(1), e12683.

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). Top income shares, business profits, and effective tax rates in contemporary c hile. *Review of Income and Wealth 62*, S120–S144.

Flores, I. and P. Gutiérrez (2021). Riqueza en chile: Medición y tributación. In *Impuestos justos para el Chile que viene: Diagnósticos y desafíos tributarios para un nuevo pacto fiscal*. Fondo de Cultura Económica.

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.

Flores, R., J. Pérez, and F. Uribe (2018). Valoración de la tierra de uso residencial y su contribución al valor de mercado de la vivienda en Chile.

Gandelman, N. and R. Lluberas (2024). Wealth in latin america: Evidence from chile, colombia, mexico and uruguay. *Review of Income and Wealth 70*(1), 4–32.

Garbinti, B., J. Goupille-Lebret, and T. Piketty (2021). Accounting for wealth-inequality dynamics: Methods, estimates, and simulations for france. *Journal of the European Economic Association 19*(1), 620–663.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American economic review 104*(4), 1091–1119.

Hastings, J. S., B. C. Madrian, and W. L. Skimmyhorn (2013). Financial literacy, financial education, and economic outcomes. *Annu. Rev. Econ. 5*(1), 347–373.

Inzunza, A. and C. Madeira (2023). The impact of the covid pension fund withdrawals in chile on the future retirement income of the social security affiliates and their households. Technical report, Central Bank of Chile.

Jakobsen, K., K. Jakobsen, H. Kleven, and G. Zucman (2020). Wealth taxation and wealth accumulation: Theory and evidence from denmark. *The quarterly journal of economics 135*(1), 329–388.

Karakurum-Ozdemir, K., M. Kokkizil, and G. Uysal (2019). Financial literacy in developing countries. *Social Indicators Research 143*, 325–353.

Kenedi, G. and L. Sirugue (2023). Intergenerational income mobility in france: A comparative and geographic analysis. *Journal of Public Economics 226*, 104974.

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. 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.

Londoño-Vélez, J. and J. Ávila-Mahecha (2021). Enforcing wealth taxes in the developing world: Quasi-experimental evidence from colombia. *American Economic Review: Insights 3*(2), 131–148.

Lopez, R. E., E. Figueroa, and P. Gutiérrez (2016). Fundamental accrued capital gains and the measurement of top incomes: An application to chile. *The Journal of Economic Inequality 14*, 379–394.

Lozano, F. J. and B. J. Idrovo (2024). Shocks regulatorios al mercado inmobiliario de chile:¿ cuánto del iva a la vivienda se transfiere al precio de venta? *Economic Analysis Review 39*(2), 35–59.

Lucchetti, L., P. Corral, A. Ham, and S. Garriga (2024). An application of lasso and multiple imputation techniques to income dynamics with cross-sectional data. *Review of Income and Wealth*.

Madeira, C. and P. Margaretic (2022). The impact of financial literacy on the quality of self-reported financial information. *Journal of Behavioral and Experimental Finance 34*, 100660.

Madeira, C., P. Margaretic, F. Martínez, and P. Roje (2022). Assessing the quality of self-reported financial information. *Journal of Survey Statistics and Methodology 10*(5), 1183–1210.

Martínez-Toledano, C. (2020). House price cycles, wealth inequality and portfolio reshuffling. *WID. World Working Paper 2*.

Martínez-Toledano, C., A. Sodano, and S. Toussaint (2023). Wealth inequality in the netherlands, 1894–2019: Updated series with and without pension wealth. *WID. World Technical Note* (2023/13).

Meinshausen, N. (2007). Relaxed lasso. *Computational Statistics & Data Analysis 52*(1), 374–393.

Piketty, T. (2014). *Capital in the twenty-first century.* Harvard University Press.

Piketty, T., L. Yang, and G. Zucman (2019). Capital accumulation, private property, and rising inequality in china, 1978–2015. *American Economic Review 109*(7), 2469–2496.

Ravallion, M. and S. Chen (2003). Measuring pro-poor growth. *Economics letters 78*(1), 93–99.

Roine, J. and D. Waldenström (2009). Wealth concentration over the path of development: Sweden, 1873–2006. *Scandinavian Journal of Economics 111*(1), 151–187.

Roine, J. and D. Waldenström (2015). Long-run trends in the distribution of income and wealth. *Handbook of income distribution 2*, 469–592.

Saez, E. and G. Zucman (2016). 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.

Salas-Rojo, P. and J. G. Rodríguez (2022). Inheritances and wealth inequality: a machine learning approach. *The Journal of Economic Inequality 20*(1), 27–51.

Smith, M., O. Zidar, and E. Zwick (2023). Top wealth in america: New estimates under heterogeneous returns. *The Quarterly Journal of Economics 138*(1), 515–573.

Suss, J., T. Kemeny, and D. Connor (2023). Measuring sub-national wealth inequality in the united states using ensemble learning. Unpublished.

Troncoso, P. (2022). Employment effect of means-tested program: Evidence from a pension reform in chile. *Available at SSRN 4247302*.

Vermeulen, P. (2016). Estimating the top tail of the wealth distribution. *American Economic Review 106*(5), 646–650.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth 64*(2), 357–387.

Waltl, S. R. and R. Chakraborty (2022). Missing the wealthy in the hfcs: micro problems with macro implications. *The Journal of Economic Inequality 20*(1), 169–203.

Wodrich, N. (2020). *Estimating the top tail of the family wealth distribution in Canada.* Office of the Parliamentary Budget Officer.

Zucman, G. (2019). Global wealth inequality. *Annual Review of Economics 11*, 109–138.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Chat GPT in order to improve language and readability. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Online Appendix not intended for publication

A The relaxed lasso method

The relaxed LASSO was initially introduced by Meinshausen (2007) and can be described as an algorithm that minimizes the following penalty function

$$\begin{aligned}\sum_{i=1}^{n}\Big(P_{it}-\phi_0-\phi_1\delta_1x_{1,i}-\phi_2\delta_2x_{2,i}-\cdots-\phi_k\delta_kx_{k,i}\Big)^2\\+\lambda\pi\left(\gamma\sum_{j=1}^{k}\delta_j|\phi_j|+(1-\gamma)\sum_{j=1}^{k}\delta_j\phi_j^2\right)\end{aligned} \tag{7}$$

In the relaxed LASSO method, the goal is to find the optimal values of the parameters λ, π, and γ that minimize equation 5. The method first identifies a set of active regressors by estimating a LASSO, where some coefficients are selected ($\delta_j = 1$), making some coefficients exactly zero and effectively selecting the most relevant variables for predicting pension funds balances.

Next, the coefficients of these active regressors are further shrunk using the regularization parameter $\lambda\pi \leq 1$. This regularization term encourages sparsity in the model, indicating the importance of any variable in predicting pension funds balances.

The relaxed LASSO considers different combinations of values for the parameters λ, π, and γ. The main objective is to determine the optimal values of these parameters so that the squared error in predicting pension funds balances, as represented in Equation 5, is minimized.

B The estimation of the Pareto tail

B.1 The Pareto distribution

The approach assumes that the top tail of the wealth distribution is well characterized by the Pareto distribution, which is defined by ($\forall w > w_{min}$ and $\alpha > 0$)

$$F(w|\alpha, w_{min}) = 1 - \left(\frac{w_{min}}{w}\right)^{\alpha} \tag{8}$$

$$f(w|\alpha, w_{min}) = \frac{\alpha w_{min}^{\alpha}}{w^{\alpha+1}} \tag{9}$$

where w represents the net wealth of a household. The parameters w_{min} and α correspond to the threshold and shape parameters, respectively. A lower value of α indicates a fatter tail and a more concentrated wealth distribution.

Assuming a given w_{min}, one can approximate the distribution of wealthy households using the Pareto distribution, once α is estimated.[14] Several estimators for α are suggested in the literature. In our study, we follow the approach proposed by Vermeulen (2016) and Vermeulen (2018), who develops a pseudo maximum likelihood and weighted regression estimator for α, defined by

$$\ln\left(\left(i - \frac{1}{2}\right)\frac{\overline{N_{fi}}}{\bar{N}}\right) = C - \alpha \ln(w_i). \tag{10}$$

Here, w_i represents the wealth of household i, where households are ordered in decreasing wealth, with $i = 1$ denoting the richest household, who belongs to the auxiliary database. The poorest household in the tail is denoted by $i = n$, representing the household in the 1% percentile of wealth from the household survey database. $\bar{N} = \frac{\sum_{j=1}^{n} N_j}{n}$ is the average survey weight of all observations, where N_j is the survey weight of the household $j \in \{1, ..., n\}$. Finally, $\overline{N_{fi}} = \frac{\sum_{j=1}^{i} N_j}{i}$ is the average weight of the first i observations, and C is a constant.

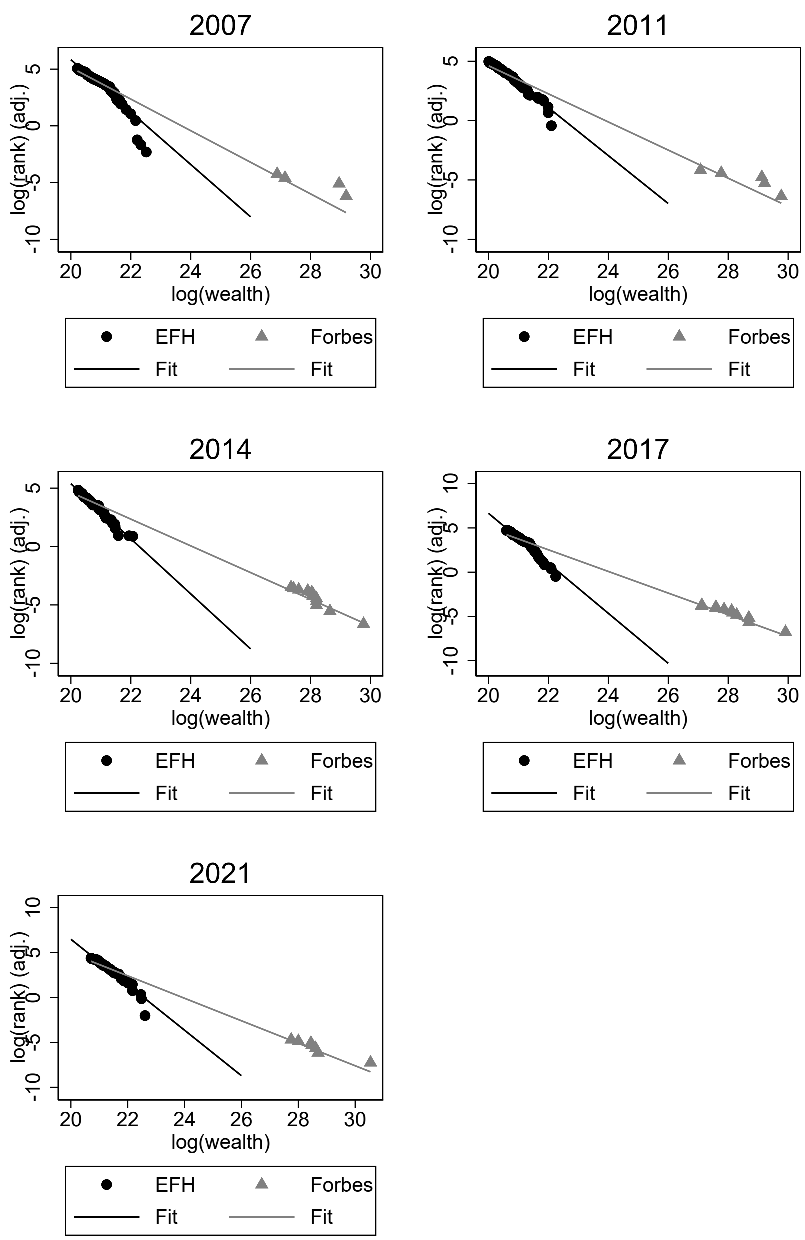
B.2 Empirical upper tail adjustment for Chile

We applied Vermeulen's methodology using two specifications: one including pension funds and one without them. In both cases we make estimations to impute financial assets, in the same way as pension funds where imputed. The reason we impute financial assets is that, as seen in Table D.2, the proportion of households manifesting financial wealth is quite unstable between years, so the

[14] There is no consensus in the literature on how to estimate w_{min}, which needs to be sufficiently large to ensure that the observations adhere to the Pareto law. This implies that only the top tail of the distribution is well approximated by the Pareto distribution. For simplicity, let's assume that w_{min} is defined as the wealth of the first household that falls within the top 1% of the wealth distribution based on the household survey.

discussion about item non-response can be extended to those assets as well, although we don't have exact data to be sure. However, financial assets are imputed until they match the proportion of households that have financial assets in 2017, given that we don't have administrative data on how many households actually have those kinds of assets. We use alternative specifications that don't impute financial assets or that impute financial assets and cash separately, as will be discussed in the robustness checks.

Figure B.1: Adjusted tail wealth distribution



Notes. Own elaboration based on country's household survey and Forbes data. For every year, the graph represents the results of the estimation of Equation 10.

After that, we use the Vermeulen methodology to correct for differential unit non-response. For simplicity, in all of our estimations, we set w_{min} as the wealth of the first household that falls within the top 1% of the wealth distribution of the EFH and estimate α with Equation 10 using

OLS. However, other researchers, such as Waltl and Chakraborty (2022), advocate for alternative methods, such as quantile regression, which we do not utilize. As illustrated in Figure B.1, the limited number of Forbes data points could be considered outliers, potentially resulting in an underestimation of the true parameter.

C Baseline estimates of aggregate wealth

In the following subsections, we explore a comprehensive analysis of each component depicted in Figure 1a, elucidating the methodologies employed in constructing these elements of total wealth.

C.1 Estimating the value of Housing

Thanks to a combination of administrative data from censuses, tax data and edification permits, we estimate the total value of housing at the national level between 2003 and 2019. To do so, we mainly rely on three variables: the total number of dwellings, the average value of a square meter and the average size of a dwelling. Our estimates distinguish formal and informal dwellings as well as four different regions (north, south, center and the metropolitan area). The following sub-sections explain how we build these estimates with more detail.

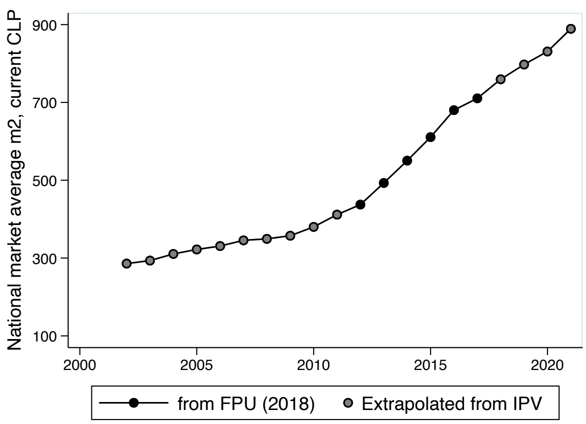
C.1.1 Formal housing

The best estimates of the market value of Chilean housing are provided by Flores et al. (2018). The authors use exhaustive administrative data –gathered for tax purposes– on the individual characteristics of both the national housing stock and the universe of market transactions. They use recorded sales and their characteristics to attribute a yearly market value to the housing stock, using the hedonic pricing method based on characteristics such as location, construction materials, or antiquity among others. However, their estimates are only available for the period 2012-2017; for other years, we use different sources to build similar estimates of our own.

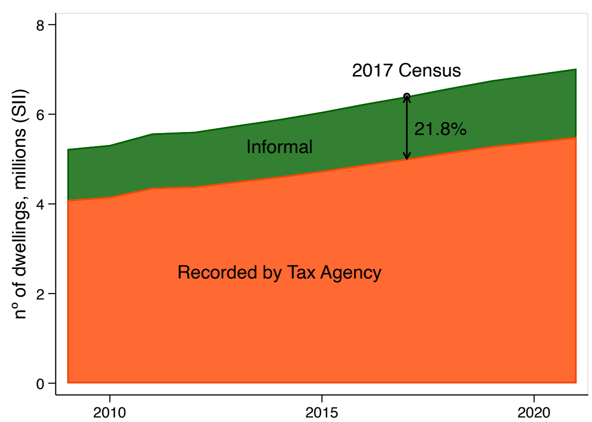
The Chilean ministry of housing and urbanism (MINVU in Spanish) reports the total number of dwellings in the country, using the same data source than Flores et al. (2018), but covering a longer period (2009-Present). We combine this information, with the housing price index reported by the Chilean Central Bank, which is also built upon exhaustive administrative data on market transactions (see figure C.1a).[15] The index reports the evolution of the average value of a square meter, from 2002, in four geographical zones, on a yearly basis. Since the evolution of average sizes of dwellings are not directly reported, we use data on new constructions permits, reported by the National Institute of Statistics, to extend the series estimates on average dwelling sizes (this is still to be implemented, for now use a constant average dwelling size).

C.1.2 Informal housing

The national census of 2017 provides a snapshot of the total amount of dwellings in the country, including both those registered by the tax agency and those that are not. We thus estimate the informal stock of housing residually. That year, 6.4 million dwellings where counted in the census, while the tax agency only counted 5 million. That is, 21.8% of dwellings are not registered in tax records. We assume that ratio to be constant, so that informal housing evolves at the same rythm than formal housing, which is reported on a yearly basis by the tax agency (see figure C.1b).



(a) The average price of a m^2



(b) Total dwellings

Figure C.1: Input variables to estimate the value of housing

C.2 Financial wealth

The estimation of national financial wealth is relatively straightforward, as the Chilean national accounts already include estimations of the net aggregates of financial wealth received by households. Within the category of assets, we differentiate between various components, such as pension

[15]We refer to the *Indice de Precio de Viviendas* - IPV in Spanish, which is presented in BCCh (2014).

funds, private firms' stocks, savings in cash, and other financial assets. Additionally, the national accounts provide information on the aggregate debt reported by households. By analyzing these components, we can accurately determine the national financial wealth for the specified years.

Pension funds are the most important component of the category of financial assets. In 2019, they represent roughly the total income of the economy. Generally, the total value of stocks is approximately half the value of pension funds, and cash is relatively less important. Total debt has grown over the period, from -34% of the national income in 2009 to -47% in 2019

One important consideration is that in 2020, due to the coronavirus pandemic recession, a law allowing the withdrawal of pension funds was accepted to reduce its effects on aggregate demand. This explains a significant reduction in total pension funds, from more than 10% of the national income. The recession itself may also explain other reductions in the market value of different assets.

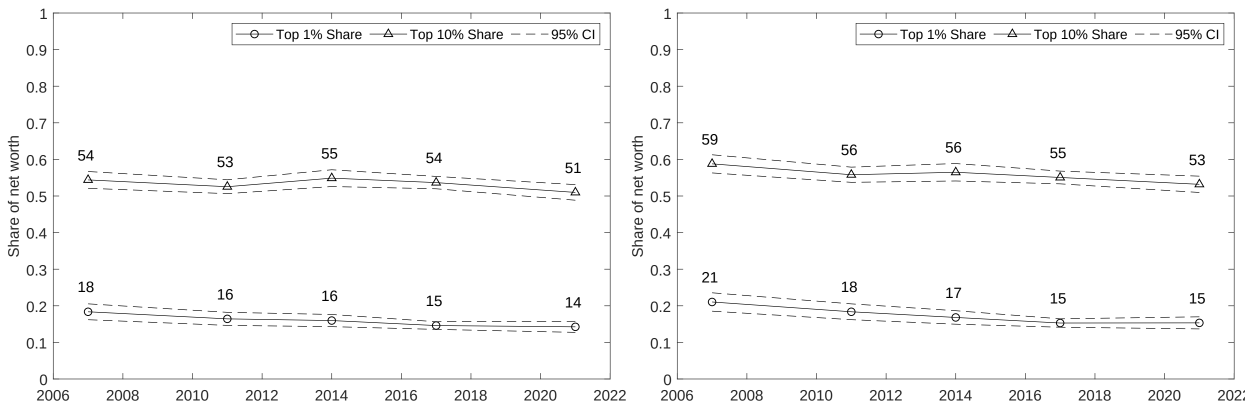
C.3 Agricultural wealth

To estimate the value of agricultural land in the country not included in the information reported by the Internal Revenue Service (*Servicio de Impuestos Internos*, SII) regarding territorial tax, we use information about the total dwellings registered by the tax agency. While these numbers are precise in terms of quantities, fiscal prices are significantly lower than market prices. Therefore, our estimations assume that this underestimation is proportional to that observed in the value of housing. In other words, to adjust fiscal prices to market prices, we divide the total value of housing described in the previous subsections by the total value of dwellings estimated by the *SII*. As a result, we apply a correction factor multiplying the fiscal value by a factor between 1.5 and 2.4, depending on the year. Thus, we estimate that agricultural wealth represents between 14% and 21% of the GDP.

D Direct estimates from the Chilean wealth survey

figure D.1 indicates a gradual decline in wealth inequality over time, with slightly higher inequality observed when pension funds are excluded from the calculation.

Figure D.1: Private wealth top 1% and 10% shares in household surveys, 2007-2021

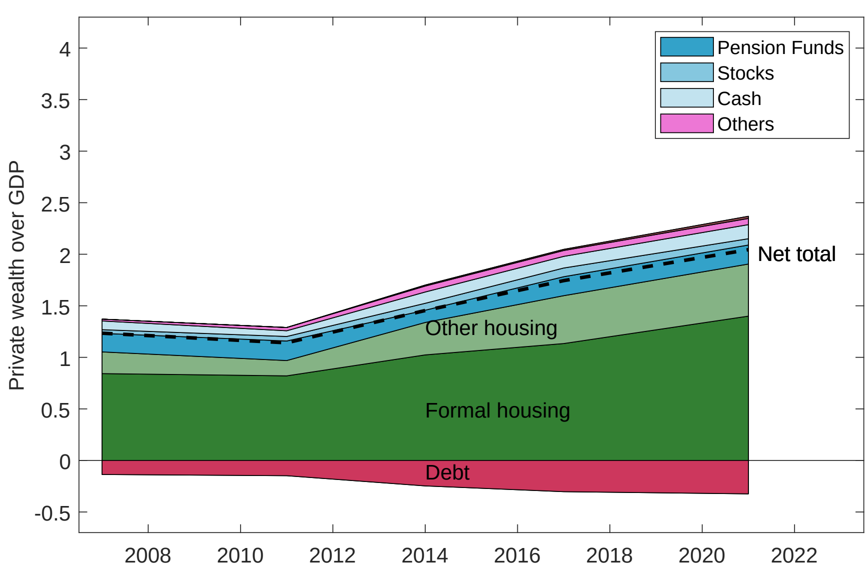


(a) Including pension funds

(b) Excluding pension funds

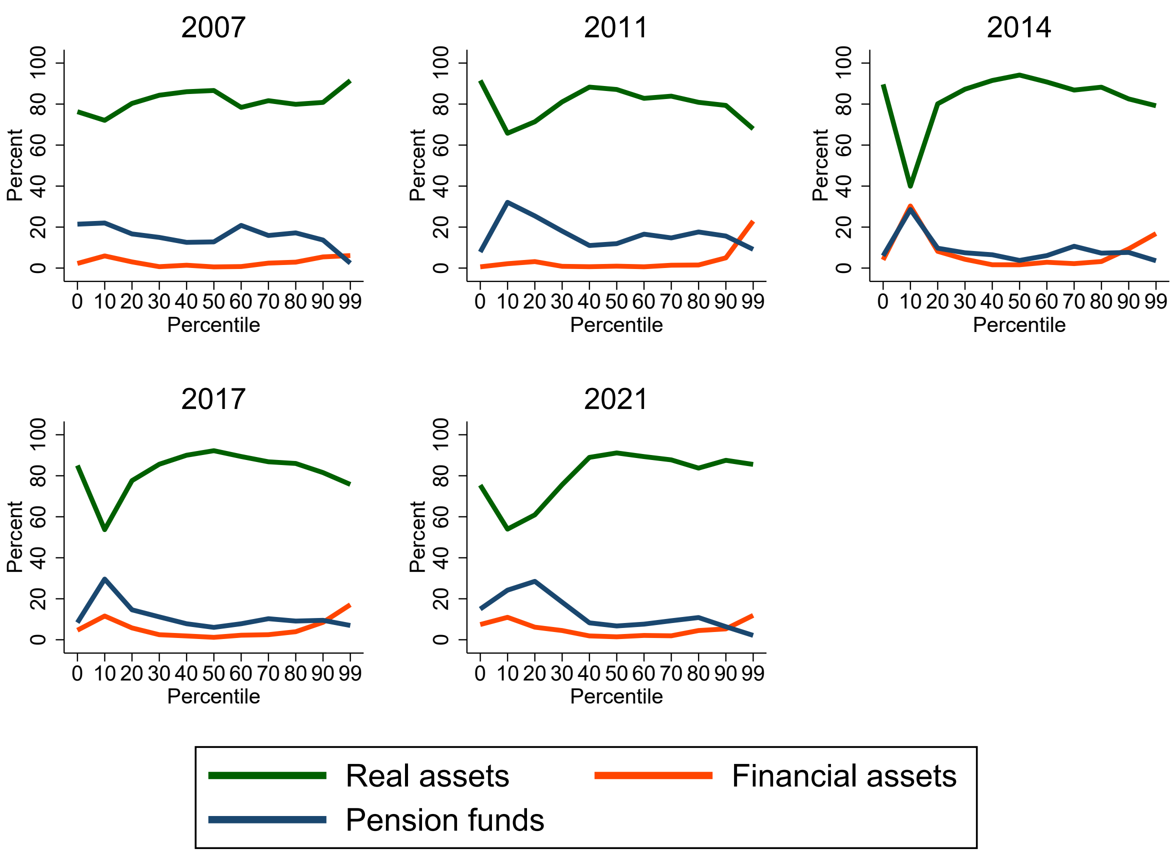
Notes. Own elaboration based on Chilean wealth survey (*Encuesta Financiera de Hogares*).

Figure D.2: Private wealth in household surveys, 2007-2021



Notes. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*). The market value of total wealth, net of liabilities, is roughly equal to 2 times GDP in 2021.

Figure D.3: Change in portfolio structure in the survey



Notes. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*). The figure shows financial assets, pension funds and real assets as share of total assets over net worth. Shares are calculated separately for each net worth survey decile.

Table D.1: Households owning pension funds in surveys

Year	Households (millions)	Declares funds	Population with funds (administrative data)
2007	3.85	33.0%	47.91%
2011	4.23	41.7%	51.56%
2014	4.70	15.7%	54.54%
2017	4.87	25.0%	55.81%
2021	5.62	44.2%	48.60%

Notes. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*) (columns 2 and 3) and Superintendency of Pensions and World Bank (column 4). In 2014, out of a total of 4,701,109 interviewed individuals from different households, 736,629 individuals (representing 15.7%) reported having pension funds greater than zero. In contrast, data from the Superintendency show that 9,746,467 individuals are affiliated with the pension system, which, according to World Bank population data, represents 55.1% of the total population.

Table D.2: Households declaring positive values in different asset classes

Year	Population (M)	Pension Funds	Financial Assets (excl. cash)	Cash	Financial Assets
2007	3.85	1.27 (33%)	0.57 (15%)	0 (0%)	0.57 (15%)
2011	4.23	1.77 (42%)	0.36 (9%)	0 (0%)	0.36 (9%)
2014	4.70	0.74 (16%)	1.22 (26%)	0.62 (13%)	1.54 (33%)
2017	4.87	1.22 (25%)	1.65 (34%)	0.92 (19%)	2.06 (42%)
2021	5.62	2.48 (44%)	2.08 (37%)	1.71 (30%)	3 (53%)

Notes. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*). All columns are measured in millions of households. Lecture: In 2007, out of a total of 3,847,952 households surveyed, 574,404 households (representing 14.93%) reported having financial assets greater than zero.

E Comparison of our wealth estimates

We now compare our estimates to those from other studies, both for Chile and for other countries. As shown in Figure E.1, when we include private pension funds, the wealth inequality estimate aligns closely with estimates from administrative data in Carranza et al. (2025). Their single estimate provides the most comprehensive and most direct assessment of the wealth distribution in the country, yet it is only available for one year, preventing it from describing a time trend. This alignment serves as a solid validation for our results. Our estimates align somewhat closely to another source, the Credit Suisse Global Wealth Reports. However, although the level is close on average, trends seem different. From the Credit Suisse's guidelines, we know their estimates for Chile are based on household financial surveys, the Forbes billionaires lists, and national accounts. On top of these common sources, our estimates include pension fund estimates from administrative data and we distinguish private pension fund assets from other financial assets when scaling the

Table D.3: Mean Household wealth (in dollars)

Year	Total	Top 1%	Top 10%	P50-P90	P0-P50
2007	\$122.693	\$4.466.464	\$854.996	\$81.798	\$8.890
2011	\$141.112	\$5.307.639	\$962.059	\$99.592	\$10.025
2014	\$136.623	\$5.529.753	\$948.345	\$93.688	\$8.515
2017	\$183.535	\$6.770.381	\$1.216.935	\$130.078	\$18.772
2021	\$161.199	\$5.539.431	\$1.059.839	\$119.080	\$14.893

Notes. Adjusted series include the right-tail adjustment using Vermeulen (2018) methodology. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*), Forbes list of the year, and relaxed LASSO imputation proposed in this article. Chilean pesos are transformed to dollars using the exchange rate of the respective year.

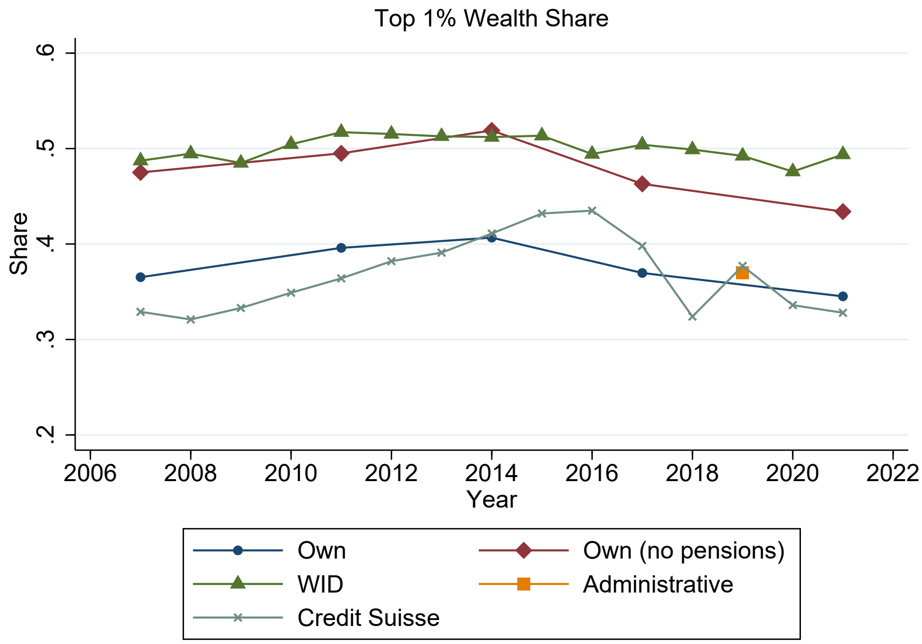
distribution to Balance sheets' levels. Additionally, given the lack of an official estimate of national non-financial wealth, we use an indirect estimate of national housing assets, its main component[16]. These differences in methods should explain discrepancies with Credit Suisse's estimates, which are available for every year, unlike ours because we chose to avoid extrapolating years without survey.

When excluding private pension funds from the wealth definition, our estimates are close in level to those of the World Inequality Database, particularly between 2007 and 2014, with estimates hovering around 50%. However, the relationship diverges after that period. By construction, the WID series for Chile are constructed based on a machine learning algorithm that considers both the distribution of income and the macroeconomic evolution of financial assets, filling the informational gaps based on information from other countries. This renders trend-analysis more difficult based on their data. Although in theory the definition of wealth in Distributional Wealth Accounts, includes private pension funds, the general algorithm used to estimate the wealth distribution of countries with scarce wealth data –such as Chile– is based on data from other countries with robust statistics, which are mostly countries with public pensions funds (Alvaredo et al., 2016). This could explain why the WID estimates match more closely to our estimates excluding private pension funds, rather than the ones with matching definitions.

Figure E.2 presents the evolution of wealth inequality in comparison with other countries. We include typical comparisons such as the USA and Western European countries. New Zealand, a country Chile often looks to emulate, is also included. South Africa and Colombia are included

[16]Contrary to financial assets, non-financial assets are not reported in the Chilean balance sheets. We constructed such estimates for the whole period based on a study from the Central bank that uses cadastral data, the administrative record of all estate transactions and hedonic pricing to estimate the market value of total housing assets. We extend such series based on construction permits and official estimates of housing price indices (see appendix C for more details)

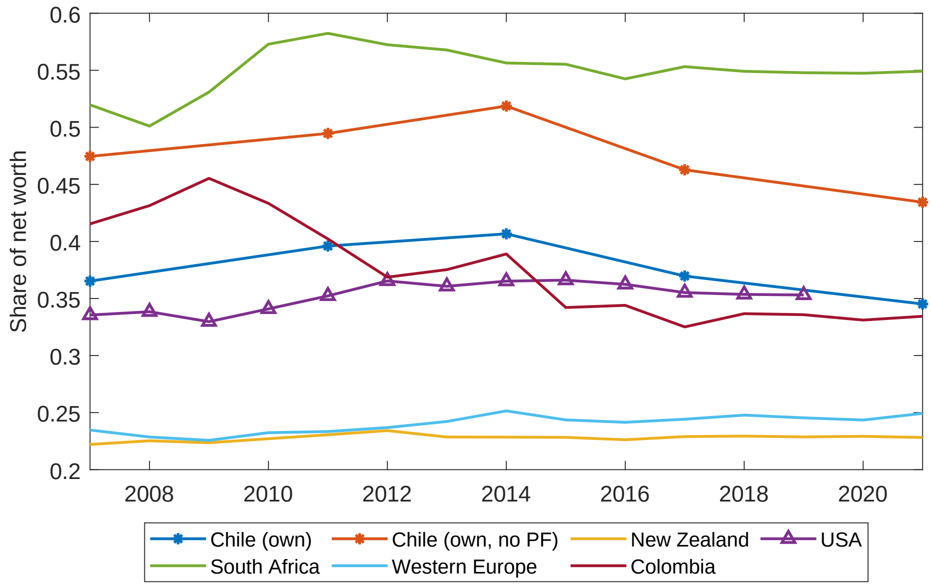
Figure E.1: Wealth Share of Top 1% Wealth Holders, 2006-2021



Notes. Own elaboration based on wid.world, the country's household survey, Forbes list of the year, Vermeulen (2018) methodology and relaxed LASSO imputation proposed in this article. The figure compares the top 1 and top 10 percent wealth share in Chile to between different specifications.

in representation of developing countries with recent state-of-the-art estimates. In our benchmark estimates, wealth inequality in Chile is significantly higher than in Western Europe, while falling below the level observed in South Africa, closer to the US or Colombia. When excluding private pension funds, the inequality in Chile appears comparable to that of South Africa. This suggests that pension funds play a crucial role in shaping the overall wealth distribution in Chile, and their inclusion or exclusion in the analysis can have a substantial impact on the measured inequality.

Figure E.2: Wealth Share of Top 1% Wealth Holders, 2006-2021



Notes. Own elaboration based on wid.world, the country's household survey, Forbes list of the year, Vermeulen (2018) methodology and relaxed LASSO imputation proposed in this article. The figure compares the top 1 percent wealth share in Chile to that of other countries.

F Robustness exercises

Table F.1: Wealth Share (%) of Top 1% Wealth Holders, 2007-2021

Year	LASSO	Random Forests	Elastic Net	Ridge Regression
2007	32.5	33.9	33.8	33.8
	(2.1)	(2)	(2)	(2)
2011	39.1	39.6	39.2	39.2
	(2.5)	(2.5)	(2.5)	(2.5)
2014	41.4	40.7	39.4	39.4
	(1.8)	(1.8)	(1.8)	(1.8)
2017	36.1	36.6	35.8	35.8
	(1.6)	(1.6)	(1.6)	(1.6)
2021	33.6	33.6	32.7	32.7
	(2)	(2)	(2)	(2)

Note: Own elaboration based on the country's household survey, Forbes list of the year, Vermeulen (2018) methodology, and the respective machine learning imputation method indicated in the corresponding column.

Table F.2: Wealth Share (%) of Top 10% Wealth Holders, 2007-2021

Year	LASSO	Random Forests	Elastic Net	Ridge Regression
2007	63.2	64.2	63.4	63.4
	(1.3)	(1.3)	(1.3)	(1.3)
2011	65.8	66.7	65.5	65.5
	(1.5)	(1.5)	(1.5)	(1.5)
2014	68.2	68.2	65.1	65.1
	(1.2)	(1.2)	(1.2)	(1.2)
2017	65.9	66.4	64.2	64.2
	(1)	(1)	(1)	(1)
2021	64	64.3	62.4	62.4
	(1.3)	(1.3)	(1.3)	(1.3)

Note: Own elaboration based on the country's household survey, Forbes list of the year, Vermeulen (2018) methodology, and the respective machine learning imputation method indicated in the corresponding column.

Table F.3: Wealth Share (%) of Top 1% Wealth Holders, 2007-2021

Year	Only Pension Funds imputation		Imputation of Pension Funds + Financial Assets + Cash	
	With Pension Funds	Without Pension Funds	With Pension Funds	Without Pension Funds
2007	36.7	50.4	35.1	45.5
	(2.1)	(2.4)	(2.2)	(2.5)
2011	41.8	54.1	39.1	47.6
	(2.5)	(2.9)	(2.5)	(3.1)
2014	42.1	52.3	40.8	50.7
	(1.8)	(2.2)	(1.8)	(2.3)
2017	37	46.3	37	45.6
	(1.5)	(1.9)	(1.5)	(1.9)
2021	34.5	43.4	35.2	43.3
	(2)	(2.3)	(1.9)	(2.3)

Note: Own elaboration based on the country's household survey, Forbes list of the year, Vermeulen (2018) methodology, and the relaxed LASSO imputation proposed in this article. Standard errors in parenthesis.

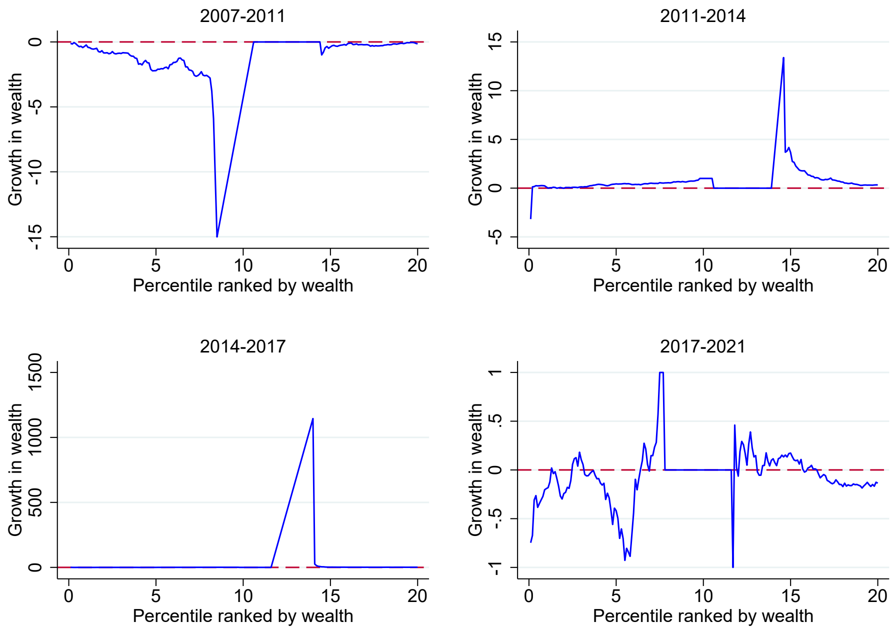
Table F.4: Wealth Share (%) of Top 10% Wealth Holders, 2007-2021

Year	Only Pension Funds imputation		Imputation of Pension Funds + Financial Assets + Cash	
	With Pension Funds	Without Pension Funds	With Pension Funds	Without Pension Funds
2007	69.8	81.6	67.2	77.1
	(1.3)	(1.1)	(1.3)	(1.2)
2011	69	80.2	66.9	75.3
	(1.4)	(1.4)	(1.5)	(1.6)
2014	70.5	80.6	69.9	79.7
	(1.2)	(1.2)	(1.1)	(1.2)
2017	66.5	75.5	66.3	75
	(1)	(1)	(1)	(1.1)
2021	65.8	72.8	65.5	72
	(1.3)	(1.3)	(1.2)	(1.3)

Note: Own elaboration based on the country's household survey, Forbes list of the year, Vermeulen (2018) methodology, and the relaxed LASSO imputation proposed in this article. Standard errors in parenthesis.

G Additional Figures and Tables

Figure G.1: Growth Incidence Curve, percentiles 0 to 15



Notes. Adjusted series include the right-tail adjustment using Vermeulen (2018) methodology. Own elaboration based on the Chilean wealth survey (*Encuesta Financiera de Hogares*), Forbes list of the year, and relaxed LASSO imputation proposed in this article. Y-axis is measured as a rate, so 1 represents a growth of 100%.

Table G.1: Wealth Share of Percentile of Wealth Holders, 2017-2021

Year	Percentile	Wealth	Real	Financial	Pensions	Liabilities
2017	0-50	5.1%	12.7%	2.0%	8.2%	33.8%
2017	50-90	28.4%	36.0%	8.5%	36.6%	35.4%
2017	90-100	66.5%	51.3%	89.5%	55.1%	30.8%
2021	0-50	4.6%	11.1%	3.0%	5.9%	28.0%
2021	50-90	29.6%	37.3%	15.5%	34.4%	37.7%
2021	90-100	65.8%	51.6%	81.5%	59.6%	34.3%

Notes. Own elaboration based on country's household survey, Forbes list of the year, Vermeulen (2018) methodology and relaxed LASSO imputation proposed in this article.

Contents