Return Heterogeneity in Retirement Accounts*

Lorenzo Bretscher†
University of Lausanne & SFI & CEPR

Riccardo Sabbatucci‡
Stockholm School of Economics & SHoF

Andrea Tamoni§
Rutgers Business School

July 3, 2023

ABSTRACT

We study the performance of IRA pension plans from 2004 through 2018. We document novel evidence of large return heterogeneity across income groups in the US, and provide estimates of its impact on wealth inequality. High-income individuals substantially outperform low-income ones, and this return differential is almost three times as large in "tax-free" Roth IRAs. These returns cannot be matched by equity market returns, but are consistent with high-income individuals having exposure to private assets.

Keywords: Roth IRA; Traditional IRA; Return Heterogeneity; Wealth Inequality; Retirement Accounts

*We are grateful to Sylvain Catherine, Carlo Favero, Clemens Sialm, Stephan Siegel, Paolo Sodini, Chester S. Spatt, and Quinton White for valuable suggestions. We also received helpful comments from conference and seminar participants at Rutgers Business School, WFA 2023, and the Stockholm School of Economics.

†Department of Finance, University of Lausanne, Lausanne 1015, Switzerland. **Email:** lorenzo.bretscher@unil.ch

‡Department of Finance, Stockholm School of Economics and Swedish House of Finance (SHoF), Stockholm 111 60, Sweden. **Email:** riccardo.sabbatucci@hhs.se

§Department of Finance and Economics, Rutgers Business School, Newark 07102, NJ, USA. **Email:** atamoni@business.rutgers.edu

"Nearly half all Americans report that they have less than $10,000 in savings. Of people aged 51 to 61 – those ages just prior to retirement – 30% have less than $10,000 in savings. And close to 40% of baby boomers have less than $10,000 in savings. " — January 15, 1999

Senate Finance Committee Chairman William V. Roth, Jr.

1 Introduction

Saving for retirement constitutes a fundamental decision for the vast majority of individuals. Several variables, including age, income, and tax rates, affect the consumption-saving behavior of individuals. Figure 1 shows the distribution of pension plan assets in the US over time. Individual retirement accounts (IRAs), introduced in 1997 as part of the Taxpayer Relief Act, are the *most important* defined-contribution (DC) retirement plans for U.S. workers, with assets totaling $12.2tn, followed by 401(k) plans at $6.7tn, as of 2020. Most importantly, IRA plans account for more than 35% of the *total* value of U.S. retirement accounts (e.g., defined-benefit plus defined-contribution plans). Despite this fact, little attention has been devoted to studying IRA plans in the literature.

Several IRA plans exist, with the largest two being the "traditional" and "Roth" plans. These two plans mainly differ in how contributions and capital gains are taxed (Table 1). Contributions in traditional IRA plans are made "before tax", and withdrawals at retirement are taxed as income at the marginal tax rate. In contrast, Roth IRA contributions are "after-tax", may be withdrawn at any time without penalty, and investment returns are tax-free in retirement.[1] In other words, Roth IRA assets grow tax-free.

In this paper, we document a key, novel fact. We provide evidence of return heterogeneity in US individual retirement accounts (IRA) as a function of income. High-income individuals substantially outperform low-income ones, and this return differential is almost three times as large in "tax-free" Roth IRAs. This evidence is relevant for policy makers since any structural and persistent difference in the rate of return to IRAs plans across income groups might be linked to the observed growing

[1] If withdrawn after a five-year holding period and age over 59 and half.

wealth inequality documented in the recent academic literature.[2]

More in detail, we address three research questions: (i) is there evidence of return heterogeneity in US retirement accounts, e.g., are the returns of high-income IRA individuals different from those obtained by low-income individuals? (ii) is this return spread between high- and low-income homogeneous across IRA plans or concentrated in specific ones?; (iii) what is the impact of return heterogeneity in IRA plans for wealth inequality?

Our findings are striking and unexpected. First, we analyze the aggregate (i.e., across plans) IRA performance among different income groups. We find large heterogeneity in IRA returns across income groups. Individuals having annual income between $10,000 and $100,000 earn around 2%-3% per year on their IRA investments, a performance which is in sharp contrast to the 8% average returns earned by high-income individuals. For example, individuals with income greater than $1mn earn about 10% per year. These higher returns are not associated with greater risk-taking, since the Sharpe ratios of these high-income groups are also large, ranging from 0.351 to 0.519. The fact that the volatility of IRA returns for these high-income brackets is similar to that of the S&P500 suggests that some of these high-income individuals invest in equity portfolios or other risky asset classes that outperform the broad US equity market in terms of both absolute and risk-adjusted performance. Overall, IRA plans owned by individuals earning more than $200,000 outperform those held by low-income individuals by a factor of three over our sample period.

Second, we try to understand *which* IRA plans account for this large performance difference between high- and low-income individuals. In other words, we investigate whether the return spread between high- and low-income individuals is homogeneous across IRA types, or is mostly detectable in Roth IRA plans. We perform our analysis under three scenarios, based on assumptions about the annual flows into IRA plans by income groups. Under the baseline scenario, we find that the return spread between

[2] As a point in case, entrepreneur and investor Peter Thiel supposedly contributed private, high-growth assets at very low nominal cost into his Roth IRA account – including PayPal shares costing less than a penny each – which skyrocketed the value of his Roth IRA account to $5bn as of 2019. No tax will ever be due on this amount. See "What Peter Thiel's Roth IRA Means for Yours", WSJ, July 2, 2021

high- and low-income in Roth IRAs is 15.9%, while that in traditional IRAs is 4.7%, over the full sample. Importantly, this difference is mainly driven by high-income individuals, whose performance in Roth IRA plans is more than twice as large as that obtained in traditional IRA plans (17.9% versus 8.6%). Results under a conservative scenario, labeled "bias towards equality", are still impressive for high-income Roth IRA investors, with a return spread of 8.9% and a Sharpe ratio more than twice as large (0.51 vs. 0.19) relative to low-income Roth IRA investors. Results under the extreme scenario, labeled "bias towards inequality", depict a situation where the average return of Roth IRA high-income investors is more than 40% higher than the one obtained by low-income individuals. This evidence suggests that the two IRA plans – traditional and Roth – are exploited differently by high-income investors in terms of allocations. This also implies that the resulting performances are not just due to an "investor-type fixed effect"'[3](e.g., all high-income individuals are better investors and always outperform equally across the entire range of IRA plans). Rather, our results suggest that high-income individuals invest in top performing assets *only* through their "tax-free" Roth IRA plans.

Third, we try to understand what can explain the performance of Roth IRAs owned by high-income individuals. We find that these returns cannot be matched by investments in public assets, such as the aggregate equity market or value stocks. Even the performance of growth stocks, the best performing public asset during our sample period, is unable to explain the performance of Roth IRAs held by high-income individuals. Accounting for private investments, such as venture capital, private equity or startups, it is possible to obtain performances similar to those of high-income Roth IRA plans.

Lastly, we analyze whether and to what extent this return heterogeneity has affected households' wealth distribution and inequality using data from the Survey of Consumer

[3]This implicitly assumes that income is the main characteristic defining investor types. As a matter of fact, there might be heterogeneity within high-income individuals. As an example, high-income individuals might differ in terms of investment skills or investable assets, e.g., startups. However, while it is possible that sorting takes place within the high-income group, e.g., high-income individuals self-select themselves in different IRA account types, this is irrelevant for the purpose of our paper. In fact, documenting heterogeneity of returns across income groups, within retirement accounts, is a key finding per se, with several policy implications (e.g., optimal taxation of DC retirement options).

Finances augmented with the Forbes 400 data. More precisely, we study how IRA return heterogeneity has potentially impacted (i) the wealth share of high-income households, (ii) the standard wealth share, and (iii) the wealth of the ten percent richest households. Our findings suggest that up to 7% of the increase in the wealth share over our sample period could be attributed to the observed heterogeneity in IRA returns.

Overall, our results suggest that, on the one hand, the introduction of Roth IRAs, established to help hard-working, middle-class Americans saving for retirement, had greatly benefited high-income individuals instead. On the other hand, the existence of Roth IRAs might have spurred productive and disruptive entrepreneurship, creating incentives for high-income individuals to fund private firms, including startups.

Our paper contributes to the literature on the structure and performance of retirement plans, which largely focus on public pension funds (Novy-Marx and Rauh (2014), Andonov and Rauh (2021)) and individual 401(k) plans (Poterba and Wise (2000), Robert Clark and Mitchell (2017), Sabbatucci et al. (2023)). Differently from these papers, we study the performance of the largest US defined-contribution plans, IRAs, which has received little attention in the literature.

We also contribute to the growing literature that emphasize the importance of portfolio choice and asset price changes for wealth inequality (Gomez (2018); Kuhn, Schularick, and Steins (2020); Toledano (2020)). Indeed, the increase in wealth concentration may results in part from the surge of top incomes combined with an increase in saving rate inequality, but it may also be driven by unequal returns earned on financial investments, particularly if returns are higher for those who are already wealthier. On the theoretical side, several studies propose an exogenous process for heterogeneity in returns and calibrate this process to match the target inequality (Benhabib et al., 2011, 2017; Benhabib and Bisin, 2018; Gabaix et al., 2016; Hubmer et al., 2020). Quadrini (2000) and Kaplan et al. (2018) propose micro-founded models that generate heterogeneity in returns and satisfactorily match the wealth distribution. On the empirical front, and particularly relevant for our paper, Campbell et al. (2019) use detailed administrative data on the equity portfolios of Indian investors to show that heterogeneous investment returns account for 84 percent of the increase in inequality of

wealth held in equities. Bach, Calvet, and Sodini (2020) and Fagereng et al. (2020) document return heterogeneity in Sweden and Norway, respectively. Similarly, Fagereng et al. (2019) exploit Norwegian administrative panel data and show that wealthier households do not have higher saving rates; rather, they accumulate more wealth through capital gains. Gomez and Gouin-Bonenfant (2020) link the decline in interest rates to the rise in wealth inequality through the effects of lower interest rate on the cost of raising new capital for entrepreneurs. Greenwald et al. (2021) discuss the effects of changes in long-term real rates on total wealth, the sum of financial and human wealth. Parker et al. (2022) document that the typical middle- and upper-middle-class US investor hold about 70% of investable wealth in equity, and that this allocation changed sharply since the 2000s due to financial innovation (namely, target dated funds) and regulatory changes (namely, Passage of the Pension Protection Act of 2006).

Differently from all the papers cited above, we are the first to document return heterogeneity across incomes in the US, with a focus on IRA retirement accounts.

The remainder of the paper proceeds as follows. Section 2 describes the institutional framework and the data used in the paper, while Section 3 presents the return performance of IRA plans across income groups. Section 4 presents evidence of what investments could be explaining our findings, while Section 5 discusses the impact of IRA return heterogeneity for wealth inequality. Section 6 provides some robustness checks, and Section 7 concludes.

2 Institutional Framework and Data

Four main types of IRA plans are currently available to individuals: traditional IRA, Roth IRA, SEP IRA and SIMPLE IRA.[4] The first two plans account for the majority of the assets invested in IRA accounts, as shown in the bottom panel of Figure 1, and they differ substantially in how they treat contributions and withdrawals at retirement.

In traditional IRAs, introduced with the Employee Retirement Income Security Act of

[4]Technically, there exist two additional, smaller plans: a payroll deduction IRA, where employees establish a traditional or Roth IRA with a financial institution and authorize a payroll deduction amount for it, and a SARSEP IRA, which is no longer available.

1974 (ERISA) and made popular with the Economic Recovery Tax Act of 1981, contributions are tax-deductible[5] (i.e., contributions are made with "pre-tax" assets), all transactions and earnings within the IRA have no tax impact, but withdrawals at retirement are taxed as income.

In contrast, in Roth IRAs, named after Senator William V. Roth Jr. and introduced as part of the Taxpayer Relief Act of 1997, contributions are non-deductible and transactions within the IRA have no tax impact. The contributions may be withdrawn at any time without penalty, and no taxes on profits are due during retirement.

SEP and SIMPLE IRAs are employer-sponsored plans. The former requires an employer (typically a small business or self-employed individual) to make retirement contributions into a traditional IRA established in the employee's name, while the latter requires the employer to match employee's contributions to the plan, similarly to a 401(k).

We obtain data on IRA plans from the IRS from 2004 through 2018.[6] More precisely, we collect and aggregate IRA data on two dimensions.

Aggregate data by IRA plan. The IRS provides aggregate annual data on IRA plans' year-end fair market values, contributions, withdrawals, rollovers, and Roth conversions. We collect these data for the main four IRA plans: traditional IRA, Roth IRA, SEP, and SIMPLE IRA.

IRA data by income and plan. The IRS provides partial data on individual IRA plans by adjusted income groups.[7] More precisely, we have data on year-end fair market

[5]The deduction may be limited if the individual also has a retirement plan at work and her income exceeds a certain threshold. See here.

[6]See https://www.irs.gov/statistics. In principle, the data is available starting from 2000. However, we start our analysis in 2004 for two main reasons. First, the IRS slightly changed the informational content of the data at the end of 2002. Second, there is no data available for the year 2003, which would create a gap in our analysis.

[7]The IRS defines adjusted gross income (AGI) as gross income minus adjustments to income. Gross income includes wages, dividends, capital gains, business income, retirement distributions as well as other income. Adjustments to income include items such as educator expenses, student loan interest, alimony payments or contributions to a retirement account. While theoretically two individuals with similar gross incomes but different pension contributions could be assigned to two contiguous AGI income bins, this might only marginally affect our empirical analysis since (i) pension contributions as a fraction of gross income are usually limited in size (e.g., less than 10% per year, on average, according to the U.S. Census), and (ii) we focus on high-income individuals with AGI greater than $200,000 in most of the empirical analysis, implying that small differences in pension contributions will likely leave unchanged

values and total annual dollar contributions split into fifteen adjusted income groups. Data on withdrawals, rollovers, and Roth conversions are also available for the same fifteen income groups, but are only reported in terms of aggregate amounts across all IRA plans (i.e., not separately for each individual IRA plan).

Lastly, we use data from the Survey of Consumer Finance (SCF) and Forbes 400 list for the years 2004, 2007, 2010, 2013, 2016, and 2019 in Section 5.

2.1 Traditional versus Roth IRAs

Traditional and Roth IRAs are the most important IRA plans. Table 1 compares the two plans. Contributions in traditional IRAs are on a pre-tax basis, while they are after-tax in Roth IRAs. Moreover, while there are generally no income limits on who can contribute to traditional IRAs, only individuals earning less than $125,000 ($198,000 if filing jointly) can contribute to Roth IRAs. In both cases, only a few thousand dollars can be contributed at most on any given year.

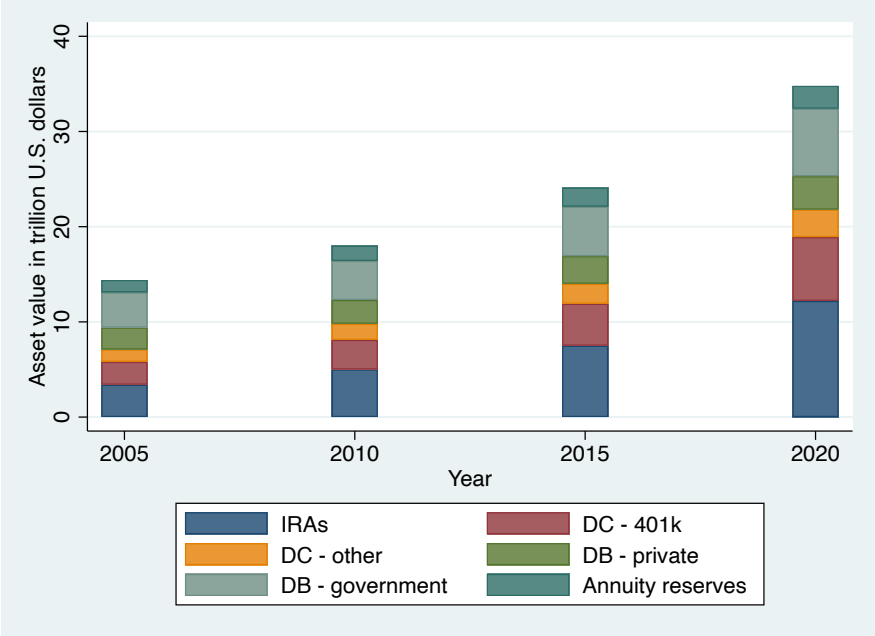
Focusing on taxes and withdrawal rules, earnings grow tax-deferred in traditional IRAs, and they are not taxed until withdrawal after age 59 and half. In contrast, any earnings in Roth IRA plans grow tax-free and can be withdrawn without taxes or penalties after, as of 2022, a five-year holding period. Additionally, contributions can be withdrawn tax-free and penalty-free at any time. This seems to suggest that Roth IRAs are best suited for individuals who expect to be in a higher tax bracket at retirement age or might face future liquidity needs (since contributed amounts can be withdrawn at anytime). Finally, in terms of required minimum distributions, traditional IRAs require individuals to start distributing assets after turning 72, while no mandatory distribution exists for Roth IRAs.

Lastly, both types of IRAs can invest in all assets, with the exceptions of collectibles, life insurances and S-corporations.[8]

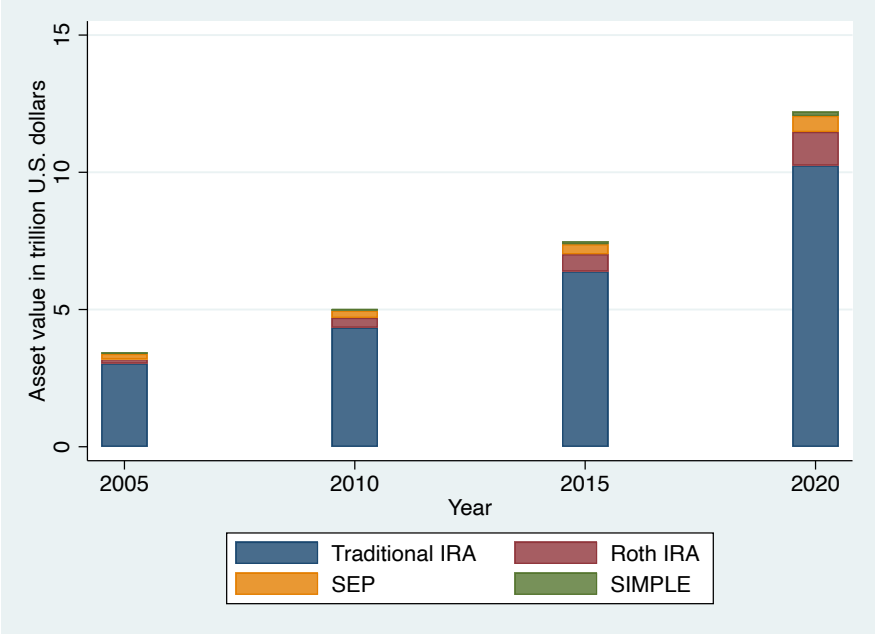
An interesting property of both IRA plans is their theoretical equivalence, in terms of

the categorization of individuals in the high-income group.

[8]Differently from, for example, 401(k) retirement accounts, IRA plans do not usually have a default investment option.



(a) Total Retirement Assets.



(b) IRA types.

Figure 1: Retirement plans in the US. The top panel plots the size of all pension plans in the United States, over time. Each bar includes individual retirement accounts (IRAs), defined contribution (DC) plans (401k plans and others), defined benefit (DB) plans (private and government), and annuity reserves. The bottom panel plots the assets of the four main types of IRA accounts over time. Data are from Statista (top panel) and the IRS (bottom panel).

cash flow dynamics, under the assumption of constant tax rates. More precisely, the final cash flows received at retirement for a pre-tax $1 invested today are equivalent if tax rates are constant and distributions happen at the same time. However, these assumptions do not hold in practice, and this might explain the different choices individual investors make regarding retirement plans and asset allocations.

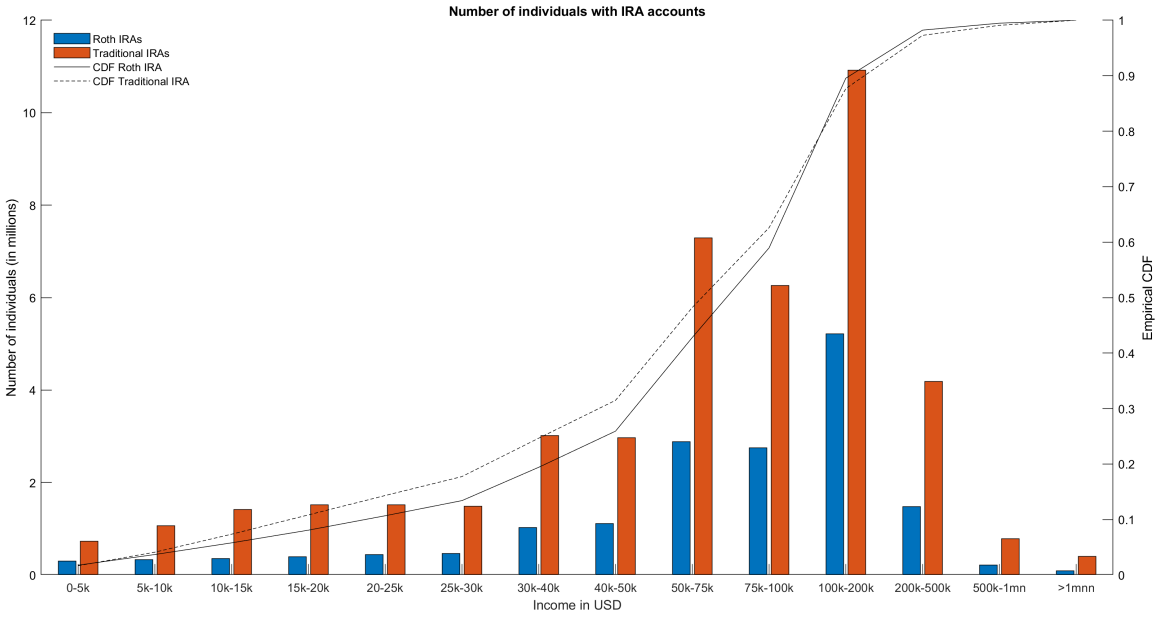
To summarize, three main factors affect the choice between traditional and Roth IRA plans. The first factor is the trade off between the level of tax rates at the time of contribution and the expected tax rates at retirement. For example, if marginal tax rates are expected to be lower in retirement, deferring income taxes through traditional IRAs might be optimal.

The second factor is the discount rates used by individuals to evaluate the present value of future taxes. More precisely, traditional IRA and Roth IRAs are equivalent in terms of future, after-taxes value of the accounts, under the assumptions discussed above. However, individuals display different utility functions and might prefer, perhaps irrationally, to pay a small amount of taxes today rather than a larger amount of nominal taxes in a few decades.

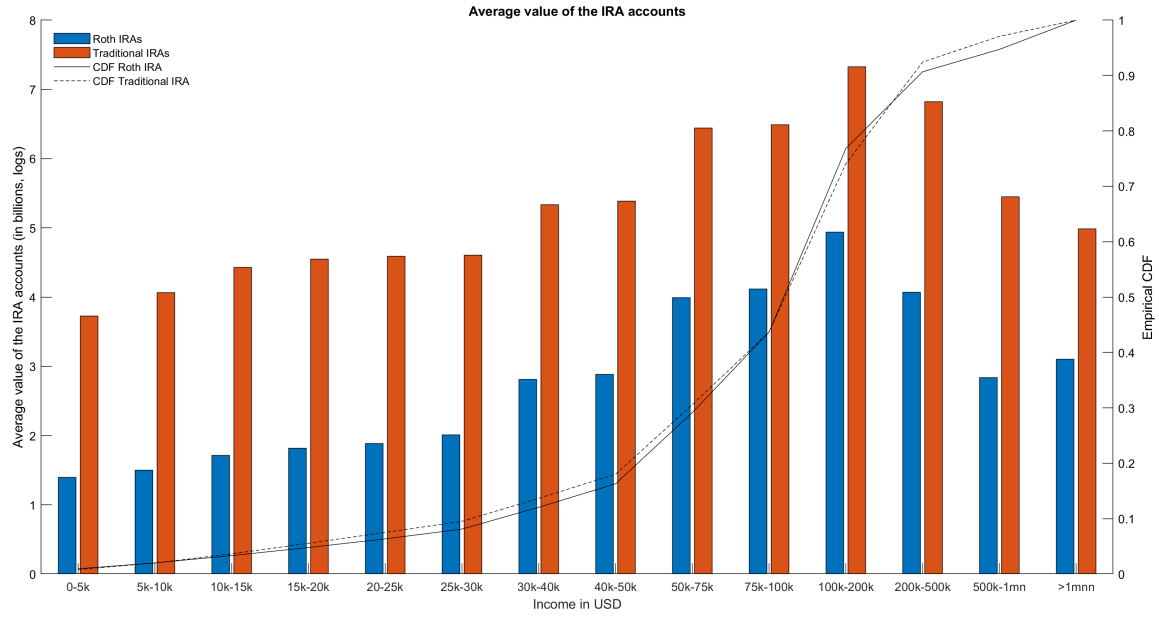
The last factor affecting the choice between traditional and Roth IRAs is the "fungibility of money" or liquidity, evidenced by the different withdrawal and penalty rules. In fact, contributions into Roth IRAs can always be withdrawn without any penalty, while this is not the case with traditional IRAs. Hence, if individuals need money to cover unexpected expenses, Roth IRA is the only account that allows withdrawals without incurring additional costs.

Figure 2 shows the distribution, across income groups, of the number of individuals owning traditional and Roth IRA plans (top panel) and their average assets (bottom panel), together with the empirical cumulative density function of the distributions for the two IRA plans.

Looking at the top panel, we notice that the total number of individuals holding the two IRA plans is different, with 44.12 million individuals having a traditional IRA plan compared to 17.26 million individuals owning a Roth IRA account. However, their distribution across income brackets is similar, as evidenced by the empirical cumulative



(a) Number of individuals



(b) Average IRA assets

Figure 2: Comparison of traditional and Roth IRA plans. This figure plots the distribution across income brackets of (i) the number of individuals owning a traditional or Roth IRA account (top), and (ii) the time series average of the asset values in those accounts (bottom). The sample period is from 2004 through 2018.

density functions. Moving to the bottom panel, we note that the average assets in traditional IRA accounts are substantially larger than the ones in Roth IRAs; however, their distribution across income groups is again very similar. We observe no stochastic dominance in either case, as the cdf of Roth IRA plans cross those of the traditional IRAs around the $100,000 threshold, implying that sorting (e.g., high-income individuals choosing different IRA plans than low-income individuals) is not a particularly relevant issue in the data.

3 The Performance of IRA Plans

Figure 1 shows that IRAs are the most important defined-contribution pension plans in the US. However, perhaps surprisingly, limited evidence on their overall performance exists to date. The IRS does not report any performance metric, since it requires assumptions on the unobserved heterogeneity of rollovers and contributions, across plans and income groups. Despite these shortcomings, understanding the returns obtained by investors in IRA plans is of paramount importance, and can provide insights on how this investment channel is related to the increase in wealth inequality documented in the recent literature.[9]

We estimate an implied rate of return, for all IRA plans and across income groups, as:

$$Ret_t = \frac{Value_t - Contributions_t - Rollovers_t + Withdrawals_t - Conversions_t}{Value_{t-1}} - 1 \quad (1)$$

where $Value_t$ is the year-end dollar value of the IRA plan, $Contributions_t$ are the annual dollar contributions into the plan, $Rollovers_t$ are the annual dollar rollovers from other

[9] An important literature has linked the skewed distribution of wealth to (a) savings rates (e.g., Krueger et al. (2016) study consumption rates out of income; Saez and Zucman (2016) compute "synthetic saving rates" by following percentile groups, rather than individuals, over time, and show that rising income and saving rate inequality is fueling wealth inequality; Bach et al. (2020) examine how the saving rate out of wealth varies across the wealth distribution; Fagereng et al. (2019) focus on the saving rate out of income; Mian et al. (2020) study the savings by different parts of the income or wealth distribution; see also De Nardi (2004) who discusses bequests that are increasing in wealth); (b) idiosyncratic returns to wealth (see Benhabib et al. (2011, 2017) for capital income risk, and Cagetti and De Nardi (2006) and Quadrini (2000) for entrepreneurial risk); (c) heterogeneity and asymmetry of returns to wealth (Fagereng et al., 2016, 2020).

pension plans, $Withdrawals_t$ are the annual dollar value of withdrawals from the IRAs, and $Conversions_t$ are the annual dollar value of conversions from traditional IRAs to Roth IRAs.

This is the usual definition of returns on the existing stock of wealth. The implicit assumption in the return calculation (1) is that the flow components (contributions, rollovers, withdrawals and conversions) happen at the end of the year, and, hence, do not affect the return of the IRA plans throughout the year.[10] Although this assumption might appear restrictive, two remarks are in order. First, at the aggregate plan level, some of these components offset each other during the year. For example, withdrawals and the sum of rollovers and conversions are largely equivalent in size over our sample period.

Second, and most importantly, the *relative* value of these flows is a small fraction of the total market value of all IRA accounts, as can be observed quite clearly in Figure 3. In particular, for traditional IRAs (Roth IRAs) this fraction equals 1.66% in 2008 and 2.25% in 2018 (7.83% in 2008 and 5.42% in 2018).[11] Moreover, rollover and withdrawal activity among Roth IRA investors is extremely rare (see Figure E.3 in Holden and Schrass (2021)). Hence, any return differential deriving from these components is of negligible magnitude.

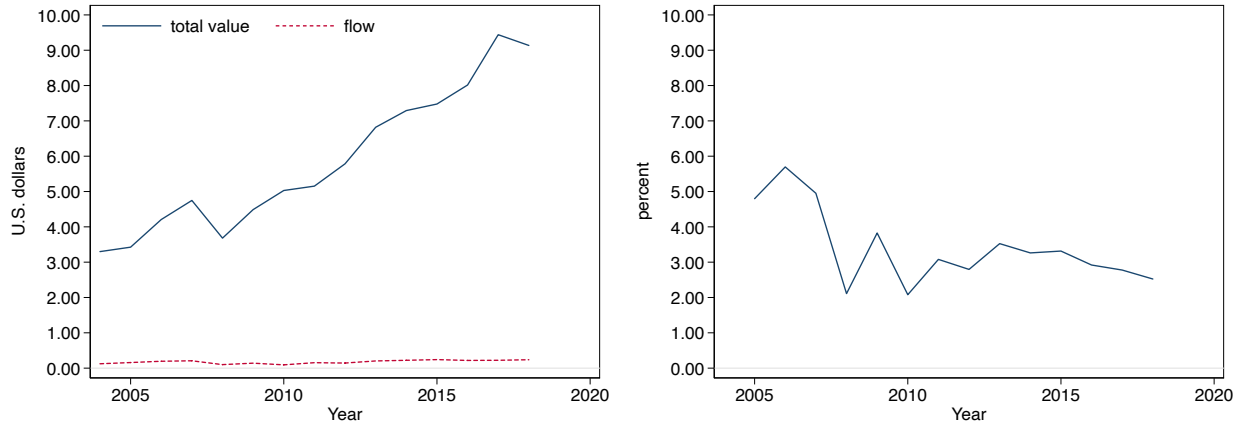
Lastly, it is worth emphasizing that households do pay attention to their IRA investments, especially Roth ones. Holden and Schrass (2021) report that more than three in ten Roth IRA investors contribute to their Roth IRAs in any given year, and that contributions are persistent over time. Most interestingly, conversion activity picked up in 2010, exactly when the income limits for Roth conversions were lifted[12] and taxpayers

[10] In Section 6.1 we report our results using an alternative return definition proposed in the literature (Fagereng et al., 2020), and show that our results are unaffected. Furthermore, this choice is analogous to the practice of using dividends not reinvested in the market in the asset pricing literature (Binsbergen and Koijen (2010) and Sabbatucci (2022)).

[11] Note that corresponding numbers for the SEP (SIMPLE) plans equal 1.27% (14.6%) in 2008 and 0.21% (8.74%) in 2018. However, SEP and SIMPLE plans are not directly comparable to traditional and Roth IRA, since the contributions into these employer-sponsored plans include both contributions made directly by the individuals as well as those made by their employers.

[12] Prior to 2010, anyone (except married taxpayers filing separately) with an annual adjusted gross income (AGI) below $100,000 could convert a traditional IRA to a Roth IRA. This income limit prevented high-income earners from participating. Following the Tax Increase Prevention and Reconciliation Act (TIPRA) of 2005, the income limit has been repealed, opening the floodgates for new Roth conversions for higher-income clients starting in 2010.

could choose to pay the taxes on 2010 conversions in 2011 and 2012.



(a) Total Value versus Flows

(b) Flows / Total value

Figure 3: Flows versus values across IRA plans. This figure shows the size of the total value of existing IRA accounts and their annual flows. Panel A plots the total value against the flow component across all IRA plans, over time. According to equation (1), flows are defined as the sum of contributions, rollovers, and conversions net of withdrawals. Panel B plots the flow component scaled by the lag of the total value of IRA plans. The sample period is from 2004 through 2018.

3.1 Findings Across Income Groups

In this section, we analyze the returns of IRA plans across different income groups.[13] To this end, we first calculate IRA plan returns for individuals within fourteen income brackets, based on the IRS classification.[14] For each income group, we observe the aggregate (across all IRA plans) contributions, rollovers, withdrawals, and conversions; thus, no assumption is needed to analyze the joint performance of IRA plans by income. Before turning to the specific numbers, we recall that traditional and Roth IRAs are by far the largest IRA plans (see Panel (b) in Figure 1). On average, their combined assets account for 94% of total IRA assets, suggesting that these aggregate IRA returns are representative of the overall IRA plans performance.

[13] Piketty (2014, p. 430) writes: "Many economic models assume that the return on capital is the same for all owners, no matter how large or small their fortunes. This is far from certain, however: it is perfectly possible that wealthier people obtain higher average returns than less wealthy people."

[14] There is also a no-income category, which we exclude from the calculations since it is relatively small in terms of individuals.

Panel A in Table 2 reports summary statistics on the IRA returns of the different income groups. Average returns across income groups follow a U-shape pattern. Individuals with annual income between $10,000 and $100,000 earn around 2%-3% per year on their IRA investments. Their risk-adjusted performance is also not impressive, with Sharpe ratios ranging from 0.122 to 0.244, half of that obtained by simply investing in the stock market. In contrast, individuals with reported annual income in excess of $100,000 perform extremely well, with average returns often above 8% and, in the case of individuals earning more than $1mn, almost 10% per year. These higher returns are not associated with greater risk-taking, as confirmed by high Sharpe ratios ranging from 0.351 to 0.519. Quite interestingly, IRA plans of high-income individuals display risk-adjusted performances that are better than the 0.413 Sharpe ratio attained by the S&P500; given that the volatility of IRA plans for these high-income brackets is similar to that of the S&P500, it follows that some of these high-income households invest in equity portfolios, or other risky asset classes, that outperform the broad US equity market in terms of both absolute and risk-adjusted performance. Section 4 discusses this point more in detail.

The top panel of Figure 4 shows that IRA plans owned by individuals earning more than $200,000 outperform those held by low-income individuals by a factor of three over our sample period, with only a short-lived underperformance during the Global Financial Crisis in 2008-2009. In the bottom panel, we collapse the income categories into two main groups, low-income and high-income, with $200,000 being the threshold. We choose this threshold to minimize the transition of individuals between one category to the other at any point in time, and to limit the mechanical increase in the value of the accounts once households move from the low- to the high-income category. In other words, with this threshold, once an individual enters the high-income category, she is likely to remain there with a high probability; furthermore, the probability of moving from the low-income to the high-income category in any given year, is also relatively small. Figure 5 provides supporting evidence for our choice. Most importantly, only 0.25% of the population moves to the high-income category, on average, every year.[15]

[15]Since IRAs, especially Roth, are relatively new, very few account holders are retirees, implying that

These two facts together imply that the impact of households moving from the low- to the high-income group is negligible in the return calculation for the two categories.

The bottom panel of Figure 4 shows that, from 2010 through 2018, the cumulative performance differential between low- and high-income groups is around 110%. This evidence can also be observed in Panel B of Table 2, which shows returns of high-income individuals more than twice as large as those of low-income individuals, and a Sharpe ratio that is 50% larger.

3.2 Findings Across IRA Plans and Income Brackets

The empirical evidence presented so far is striking and suggests that high-income individuals have outperformed both the aggregate equity market and low-income individuals in their IRA plans. In this section, we try to understand more in detail *which* IRA plans account for this large performance differential. In other words, is this return spread between high- and low-income individuals homogeneous across all IRA types, or is this performance gap concentrated, for example, in Roth IRA plans?

Answering this question is clearly important from a policy perspective, since it can shed some light on whether the increase in observed wealth inequality over the last twenty years in the United States can be related to the performance heterogeneity of "tax-free" investment pension accounts.[16] As a matter of fact, if the Roth IRAs of high-income individuals tend to outperform the market, but those of low-income individuals tend to underperform, the compounding effect of this return differential will generate an increase in total wealth inequality.[17]

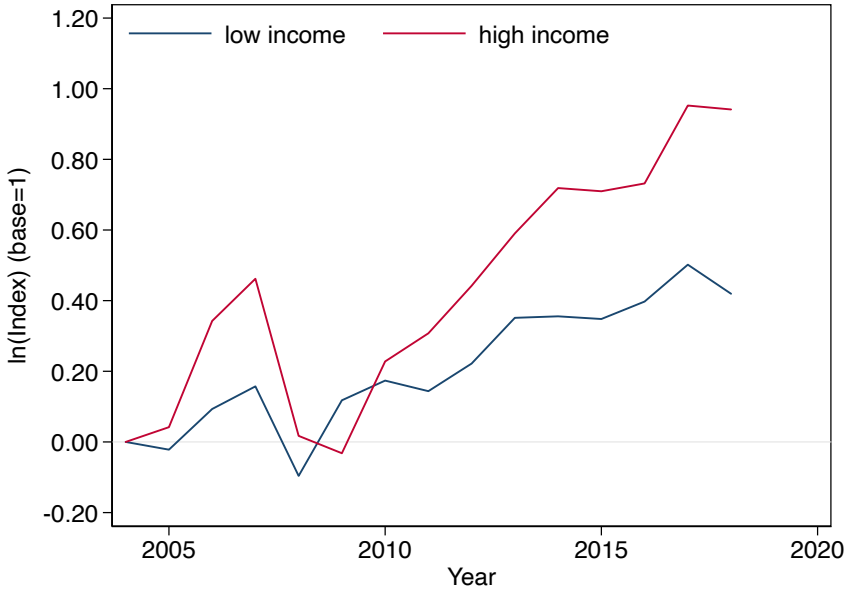
Panel C of Table 2 provides evidence of large return heterogeneity across aggregate

the increase in the number of high-income account holders is mainly due to new individuals becoming high-income.

[16]The concentration of wealth far exceeds the concentration of labor income and exhibits rapid growth both in the United States (Cagetti and De Nardi, 2008; Piketty, 2014; Saez and Zucman, 2016; Smith, Zidar, and Zwick, 2021) and around the world (Alvaredo et al., 2017).

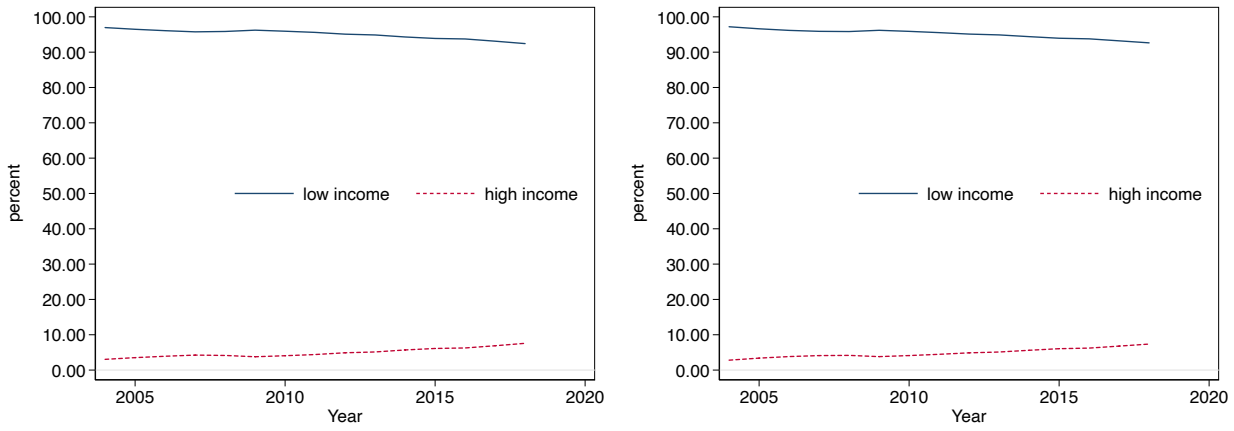
[17]Our analysis only considers the effect of IRA differential returns, abstracting from the impact of other investments, e.g., additional pension assets owned by the individuals, on wealth inequality. Of course, the total net effect on wealth inequality also depends on the correlation between income and wealth (which, in our argument, we assume to be non-negative), and the aggregate investment performance of the individuals.

(a) By Income Groups.



(b) High and Low Income.

Figure 4: Cumulative returns of IRA plans by income groups. The top panel plots the log of cumulative return indices (normalized to a value of 1 at the beginning of the sample) for all four IRA plans combined across different income groups. The bottom panel plots the IRA plans' cumulative returns for low-income (e.g., below $200,000, blue line) and high-income (red line) individuals (normalized to a value of 1 in 2004). The sample period is from 2004 through 2018.



(a) Population of taxpayers

(b) Taxpayers eligible to contribute to IRAs

Figure 5: Income distribution. This figure plots the time series of the fraction of taxpayers that we classify as either low or high income in our analysis, using $200,000 as threshold. Panel A reports the fraction of low and high income individuals in the full population of taxpayers, while Panel B reports the split only using the subset of the population of taxpayers that are eligible to make IRA contributions, every year. The sample period is from 2004 through 2018.

plans. Aggregate Roth IRA ("tax-free") returns are, on average, *lower* than traditional IRA ("taxable"), with a non-negligible difference of about 1% per year over our sample. This differential performance can also be observed by looking at the Sharpe ratios, with traditional IRA plans having a 50% higher Sharpe ratio than Roth IRA plans (0.443 vs. 0.297). Given the existence of a few, extremely large Roth IRA accounts with stellar return performances, this implies that the *average* Roth IRA holder has substantially underperformed both traditional IRA investments and the aggregate stock market.[18]

Motivated by this discussion and the evidence in Section 3.1, we next turn to a detailed analysis of the performance of traditional and Roth IRA plans *across* income groups.

3.2.1 Scenarios

The IRS does not report details on withdrawals, rollovers, and Roth conversions by income and IRA plan pairs. However, the aggregate dollar value of these flows by income brackets, jointly for all IRA plans, are available. Hence, by imposing mild assumptions on how these aggregate flows get distributed amongst different IRA plans, we are able to study the performance of each individual IRA plan (e.g., traditional and

[18] See, for example, the ProPublica report.

Roth IRA plans, separately), by income brackets.

We define a baseline scenario and two extreme ones. We start by calculating, every year, the fraction of total (i.e., across all IRA plans) dollar rollovers, withdrawals and Roth conversions within each income bracket. We also observe, every year, the IRA plan-specific total (i.e., across all income brackets) dollar rollovers, withdrawals, and Roth conversions. In the baseline scenario, we calculate, for each income bracket, the dollar flows into the different IRA plans by assuming the within-income-bracket fraction of total flows stays constant across IRA plans. As an example, the percentage of total (i.e., across all IRA plans) rollovers for individuals with annual income between \$50,000 and \$75,000 in 2015 is 8.66%, while the total (i.e., across all income-brackets) dollar rollovers into traditional (Roth) IRA plans are \$459,901,573 (\$8,377,438). We then calculate the value of rollovers for individuals with income between \$50,000 and \$75,000 and traditional IRA plans as $8.66\% \times \$459{,}901{,}573 = \$39{,}839{,}552$, and for Roth IRA as $8.66\% \times \$8{,}377{,}438 = \$725{,}706$. We do not require any assumption on contributions, since we observe them for each separate IRA plan and income bracket.

In the first extreme scenario, labeled *bias towards equality*, we depart from the baseline case and "stack the cards against" high-income investors. We do so by noting that investment returns are decreasing in contributions, rollovers, and Roth conversions but increasing in withdrawals as can be seen from equation (1).[19] Thus, we proportionally distribute the total dollar rollovers and Roth conversions *only* amongst individuals with annual income above \$75,000. At the same time, we proportionally allocate the withdrawals *exclusively* to individuals with annual income below \$75,000.

In the second extreme scenario, labeled *bias towards inequality*, we mirror the distributional assumptions of the first extreme scenario, thus favoring the high-income individuals.

Figure 6 shows the cumulative returns of the different IRA plans for low- and high-income individuals, under the baseline scenario.[20] High-income individuals outperform low-income ones in traditional IRA, SEP, and SIMPLE plans by an average

[19] Traditional IRAs and Roth IRAs are the only plans affected.

[20] The threshold is \$200,000, as in Section 3.1.

margin, across plans, of 84%. In the case of Roth IRAs, the outperformance is extraordinary: high-income individuals earned 523% more than low-income individuals over the period between 2004 and 2018.[21] The return spread between high- and low-income individuals obtained in Roth IRAs is both economically and statistically different (p-value: 0.23%) from that obtained in traditional IRA plans. We also find that the average difference between low- and high-income Roth IRA performances (panel b) is statistically significant at the 1% level (t-stat: 7.95). Columns (1)-(4) of Table 3 report, for low- and high-income, summary statistics for the performance of different IRA plans under the baseline scenario. Over the full sample, the return spread between high- and low-income individuals is 15.9% in Roth IRAs, while it is 4.7% in traditional IRAs. This difference is mainly driven by high-income individuals, whose performance in Roth IRA plans is more than twice as large as that obtained in traditional IRA plans (17.9% versus 8.6%). This evidence suggests that two IRA plans are exploited differently by high-income investors in terms of allocations. Importantly, the difference in performance is not merely driven by risk exposure, since the Sharpe ratio of high-income individuals in Roth IRAs is 0.704, larger than the 0.485 obtained by the same investors in traditional IRAs, with the latter being comparable to the Sharpe ratio of the aggregate stock market (0.413). In other words, high-income investors seem to invest in top-performing assets only through "tax-free" Roth IRA plans.

Perhaps our baseline scenario gives too much advantage to high-income individuals. Therefore, we study the performance of IRA plans under our two extreme scenarios to understand how the results are sensitive to our assumptions. Figure 7 and columns (5)-(12) in Table 3 report the results.

The dashed lines in all four panels in Figure 7 represent the time series dynamics of the returns for the high- (in red) and low- (in blue) income individuals under the *bias towards equality* scenario. We observe that high-income individuals underperform low-income individuals in traditional IRA plans, as we expect. However, quite strikingly, this is *not* the case for Roth IRA plans. In fact, even under this extremely

[21] From Panel (b), the values at the end of 2018 for the Roth IRA of high- and low-income individuals are 1.98 and 0.15, implying outperformance of $(e^{1.98}/e^{0.15} - 1) \times 100 = 523$ percentage points over this period.

(a) Traditional IRA.

(b) Roth IRA.

(c) SEP.

(d) SIMPLE.

Figure 6: Cumulative returns for high- and low-income individuals in different IRA plans under the baseline scenario. This figure plots, in each panel, the cumulative returns (normalized to a value of 1 at the beginning of the sample) for high- and low-income individuals in each of the four IRA plans. Panels (a)-(d) report results for traditional, Roth, SEP, and SIMPLE IRA plans, respectively. The sample period is from 2004 through 2018.

conservative scenario, high-income individuals earn 171% more than low-income individuals over our sample period.[22] In fact, column (8) in Table 3 shows that the Sharpe ratio of high-income individuals in Roth IRA plan is now 0.51, lower than the 0.704 obtained under the baseline scenario, but still more than double that of low-income individuals (0.19). In other words, even under the most conservative scenario, high-income individuals have a Roth IRA risk-adjusted performance that is almost *three* times as large as that of low-income individuals. This is however not the case for the other IRA plans.

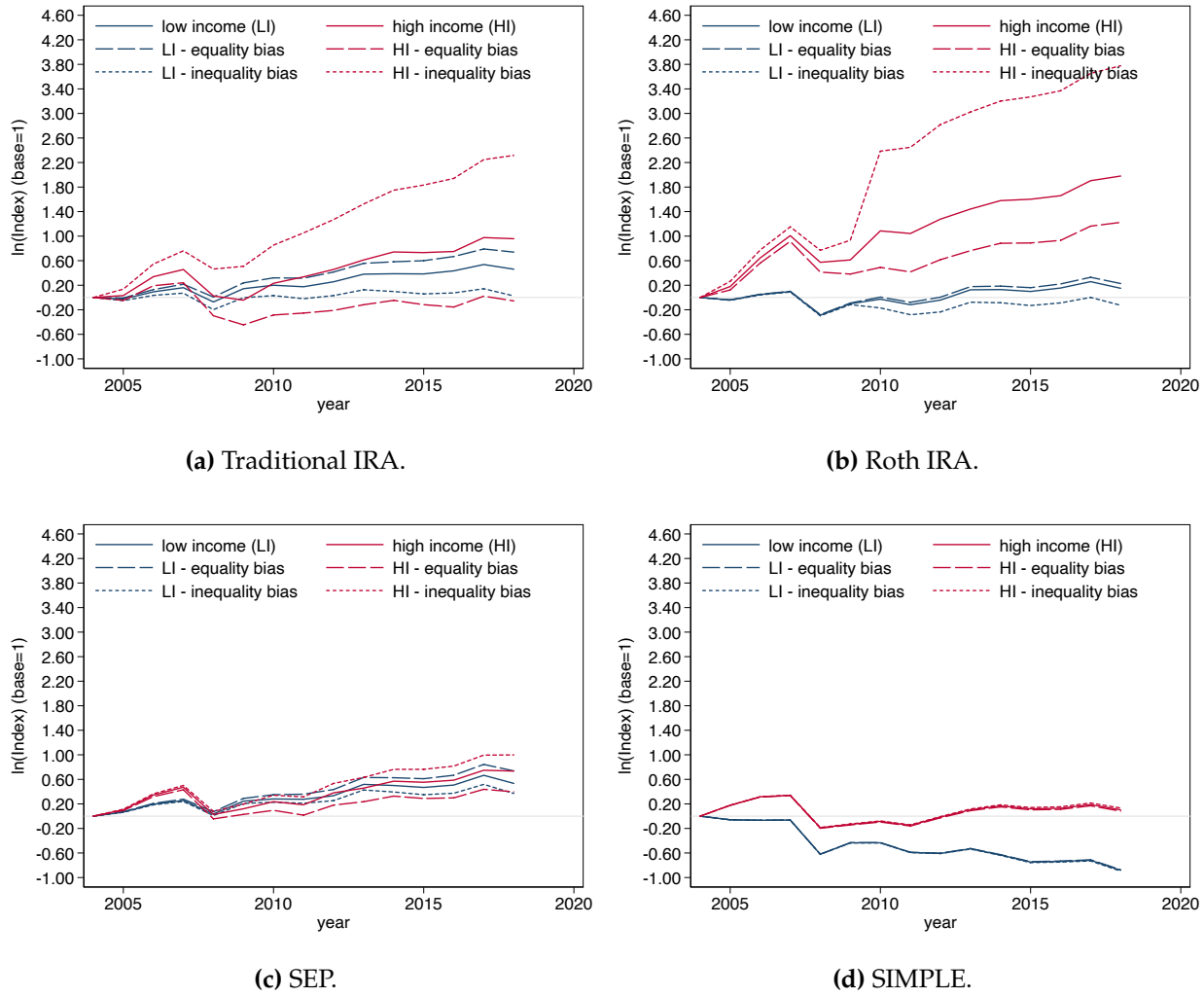
Under the *bias towards inequality* scenario, the results become, as expected, even more impressive. The performance of high-income individuals skyrockets in both traditional and Roth IRA plans, with a cumulative return difference between high- and low-income groups of about 900% and 4,200%, respectively, over the full sample. The improvement is particularly noticeable for high-income investors in traditional IRA plans, who obtain returns close to 20%, still less than half of those obtained by the same investors in Roth IRA plans (43.8%).

Interestingly, there is no meaningful difference in the SEP and SIMPLE IRA plans under either of the two extreme scenarios (see Figure 7, bottom panels).

Overall, our analysis highlights that the differential performance between high- and low-income individuals is particularly large in "tax-free" Roth IRA accounts. This evidence is indicative of how this type of pension plan has been used by high-income individuals to invest in assets that performed extremely well over the last fifteen years. Most importantly, these are not the same assets purchased by, or available to, the same high-income investors in other IRA plans, such as traditional ones. Indeed, high-income individuals investing in Roth IRA obtained returns that have been at least twice as large as those they obtained in "taxable" traditional IRA plans under *any* of the three scenarios (e.g., 17.9% vs. 8.6%, 11.5% vs. 1.1%, 43.8% vs. 19.5%). This implies that the resulting performances are not just due to an "investor-type fixed effect" (e.g., all high-income individuals are better investors and always outperform equally across the entire range of

[22] From Panel (b), the values at the end of 2018 for the Roth IRA of high- and low-income individuals under the conservative scenario are 1.22 and 0.23, implying outperformance of $(e^{1.22}/e^{0.23} - 1) \times 100 = 171$ percentage points over this period.

IRA plans). Rather, our results suggest that high-income individuals invest in top performing assets *only* through their "tax-free" Roth IRA plans.



(a) Traditional IRA.

(b) Roth IRA.

(c) SEP.

(d) SIMPLE.

Figure 7: Cumulative returns for high- and low-income individuals in different IRA plans under all three scenarios. This figure plots, in each panel, the cumulative returns (normalized to a value of 1 at the beginning of the sample) for high- and low-income individuals in each of the four IRA plans under all three scenarios: the baseline scenario (continuous red and blue lines), *bias towards equality* (blue and red long dashed lines) and the *bias towards inequality* (blue and red short dashed lines). The sample period is from 2004 through 2018.

4 What Can Explain the Performance of High-Income Individuals?

Our findings provide striking evidence that Roth IRA plans are not all equal: high-income individuals exploit them more successfully. Several plausible explanations exist, but are difficult to test since detailed holdings of IRA plans are not publicly available.

In this section, we shed some light on plausible wealth accumulation dynamics in order to understand what can drive the IRA performances and, more specifically, those of high-income individuals.

The first possible explanation is that high-income individuals tend to have a larger fraction of their wealth invested in equity or other risky assets (Bach et al., 2020; Fagereng et al., 2019, 2020). This fact, paired with the low interest rate regime and global quantitative-easing measures introduced over the last decade, helped equity markets and other risky, long-duration asset classes (e.g., real estate) more than low-risk, fixed income instruments, hence increasing the return differential amongst investor types (Gomez and Gouin-Bonenfant (2020), Greenwald et al. (2021)). However, this explanation can only partially explain our results, since, all else being equal, there should be no return differential between Roth IRA and other types of IRAs, such as traditional IRAs, *within* investor type (e.g., high-income individuals) if we assume similar asset allocations across IRA plans.

Thus, a more plausible explanation is that asset allocation is heterogeneous across IRA plans, even for the same type of investors. In other words, investors may decide to hold assets in Roth IRAs that are different from those held in traditional IRA plans, hence realizing different returns on the two plans. As an example, startup founder shares or real assets (e.g., real estate), typically owned by high-income individuals, could perhaps be held in Roth but not in traditional IRA plans.[23]

In order to test this hypothesis, the cross-sectional asset allocations of IRAs would be required. However, such detailed data is not available. Hence, we proceed heuristically and try to match the realized performance of Roth IRAs plans (Table 3) by varying

[23] See, for example, this ProPublica report.

assumptions on the nature of assets held in those accounts.

4.1 Public Investments

First, we benchmark IRA returns against two well-known, public equity investment strategies: value and growth stocks. Growth stocks constitute one of the most profitable, publicly available asset over our sample period. Indeed, growth stocks have largely outperformed value stocks over the last decade by 59%, and this might help explain the differential performance we observe across income for a given type of IRA plan, or across IRA accounts for a given income group.[24]

We proxy the investment in value stocks using the largest, most liquid ETF tracking value stocks, namely the iShares Russell 1000 Value ETF (IWD), created in May 2000, holding more than $55bn of net assets as of December 2021. Similarly, we use the iShares Russell 1000 Growth ETF (IWF) to proxy for an investment in growth stocks. The choice of these funds is straightforward: they are the most accessible, cheapest, liquid, and actually tradable assets that retail investors can hold in their pension accounts in order to get exposure to the underlying asset category (i.e., value or growth stocks).

Figure 8 plots the cumulative performance of Roth IRAs for low- and high-income individuals, together with the cumulative return performance of value and growth stocks, while Panel A in Table 4 reports some summary statistics.

A strategy investing only in value (growth) stocks would obtain an average return of 5.2% (8.9%), and a Sharpe ratio of 0.312 (0.496). This suggests that, if high-income individuals invested all of their traditional IRA wealth into growth stocks over the period 2004-2018, the performance of their traditional IRA plans could be matched. However, this is an unlikely scenario. In fact, Holden and Schrass (2021) show that, at the aggregate level, these IRA accounts also invest in bonds and global equities, not only on US growth stocks, over our sample period.

Most importantly, the return differential between growth and value stocks is not

[24]The values at the end of 2018 for a one U.S. Dollar investment in 2004 in growth and value stocks are 2.66 and 1.67, respectively. Equivalently stated, growth stocks have outperformed value stocks by $(2.66/1.67 - 1) = 0.59$.

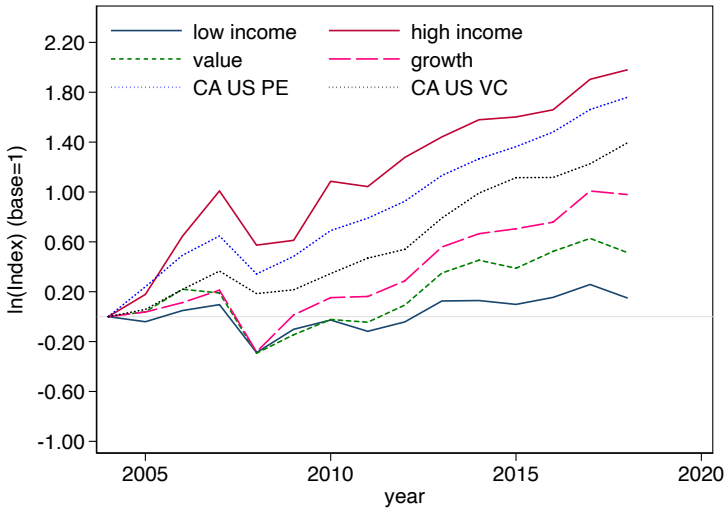


Figure 8: Cumulative returns of Roth IRAs for high- and low-income individuals, value and growth stocks, venture capital and private equity. This figure plots the cumulative returns (normalized to one at the beginning of the sample) for high- and low-income individuals in Roth IRA plans, the iShares Russell 1000 Value ETF (IWD), the iShares Russell 1000 Growth ETF (IWF), and the US venture capital and private equity returns from Cambridge Associates. The sample period is from 2004 through 2018.

large enough to explain the return heterogeneity across income groups within a given IRA plan, nor the return heterogeneity across IRA plans within the high-income group. Indeed, assuming that high- and low-income households invest their total Roth IRA wealth into growth and value assets, respectively, the return differential between growth and value stocks cannot explain the gap in Roth IRA performance across these two income groups (Figure 8).[25] Similarly, assuming high-income individuals were investing 100% of their Roth IRA wealth in growth stocks and 100% of their traditional IRA assets in value stocks, this would still not explain the almost 9% return differential between their Roth and traditional IRA plans (17.9% vs 8.6%, see Table 4).

In Table 4 we formally test the ability of growth and value stocks to capture the return performance of Roth IRA plans owned by high-income individuals. In particular, the last column of Table 4 reports a difference-in-means test between the Roth IRA performance and the growth/value strategy returns. The difference is economically large (9% and 12.7%, respectively) and statistically significant, especially considering the small sample

[25] Assuming that high-income households invest their total Roth IRA wealth in growth stock is justified by Betermier et al. (2017). An alternative justification for our assumption is the fact that tax-advantaged individuals might prefer to hold assets with call-option, lottery-style payoffs in tax-free Roth IRA.

size. Furthermore, risk can only partially explain this differential performance. Indeed, despite a higher return volatility of 25.4% versus that 17.9% of growth stocks, Roth IRAs of high-income individuals generate a risk-adjusted performance (Sharpe ratio) that is 42% larger than that obtained by growth stocks.

We next formally test whether high-income individuals are more exposed to growth stocks over our sample period, by running a regression of returns of traditional/Roth IRAs, for both low- and high-income individuals, on the Fama-French five risk factors:

$$r_{i,p,t} - rf = \alpha_i + \beta^{MKT} r_{MRP} + \beta^{SMB} r_{SMB} + \beta^{HML} r_{HML} + \beta^{RMW} r_{RMW} + \beta^{CMA} r_{CMA} + \varepsilon_{i,t} \tag{2}$$

where $i = \{\text{low-income, high-income}\}$ and p denotes the plan (e.g., traditional or Roth IRA). The risk factors are the market excess return (r_{MRP}), the small-minus-big return (r_{SMB}), the return on the high-minus-low book-to-market portfolios (r_{HML}), the return on the robust-minus-weak profitability (r_{RMW}), and the return on the conservative-minus-aggressive investment portfolios (r_{CMA}).[26]

Panel A in Table 5 reports the results of regression (2). All four return series are positively exposed to the aggregate market, with a loading less than one. This suggests that IRAs are, on average, "delevered" equity investments, and include exposure to other asset classes (Holden and Schrass, 2021). Consistent with our previous analysis, columns (3) and (4) show that high-income individuals are more exposed to large cap and growth stocks than low-income individuals, as implied by the large, negative coefficients on both SMB and HML.

Lastly, we emphasize how the benchmark equity risk model (Fama-French five factors) can explain almost 94% (83%) of the variation in Roth (traditional) IRAs of low-income individuals, but only 13.64% (42%) of the one of high-income individuals. This striking difference is consistent with high-income individuals investing in different assets classes than low-income individuals in their IRA plans. Next, we investigate this hypothesis.

[26] All risk factors can be obtained from Ken French's website.

4.2 Private Investments

Section 4.1 provides evidence that the performance of traditional IRA plans of high-income households could be approximately replicated by a portfolio of the best performing public assets (e.g., growth stocks). The puzzling result is, however, that Roth IRAs returns of high-income households cannot simply be replicated via an unlevered, long position in publicly traded assets. This can be seen in Panel A of Table 4 when comparing the mean (17.9% vs. 8.9%) and Sharpe ratios (0.704 vs. 0.496) of Roth IRAs to those of growth stocks.

One possible explanation is that high-income individuals invest in high-risk, private investments through their tax-advantaged Roth IRA accounts. In order to test this hypothesis, we collect US venture capital and private equity[27] return data from Cambridge Associates (Ang et al., 2018) over our 2004-2018 sample, and compare their performance to that of Roth IRAs owned by high-income individuals. Panel B in Table 4 reports summary statistics of those private equity return series (second and third rows). Figure 8 shows the cumulative US venture capital and private equity return time series, along with the baseline Roth IRA returns of high-income individuals.

Panel B in Table 4 shows that the returns and Sharpe ratios of private investments are larger than those of public equity. Moreover, most private business wealth is idiosyncratic (Bach et al., 2020), suggesting that individuals might obtain even better performances in their private investments. Focusing on Figure 8, the PE time series and the Roth IRA returns of high-income individuals perform similarly. In other words, if 100% of Roth IRA assets were invested in private equity, then it would be possible to almost match the Roth IRA performance. This hypothesis is formally tested in the last column of Table 4, Panel B, where we report the difference-in-means test between the Roth IRA performance and the VC/PE strategy returns. Focusing on the PE return, we confirm that the difference, 17.9%-14.3%=3.6%, is not statistically significant at standard confidence levels. This suggest that accounting for private investments, such as private equity or startups, is a promising avenue to obtain performances similar to those of high-income Roth IRA plans.

[27] We use the CA US Buyout Index.

However, a caveat is in order. So far in this section, we have assumed that high-income individuals keep their entire Roth IRA wealth fully invested in PE. A more realistic assumption is that only part of their Roth IRA assets are private investments. To this end, we compute the return and risk of private assets held in Roth IRAs required to match the Roth IRA returns of high-income individuals, under the assumption that the latter invest only a fraction ω of their wealth in private businesses, based on SCF data,[28] and the remaining share (1-ω) in the public equity market (e.g., S&P500). The last row of Table 4, Panel B, shows that — in this scenario — the implied average private equity return needed to match the Roth IRA performance should be much larger than the actual PE one from the Cambridge Associates, hovering at around 35%.[29]

Taking stock of all the evidence, however, it seems plausible that part of the Roth IRA performance might be driven by exposure to assets whose returns resemble those obtained by private investments over the 2004-2018 period. We formally test this conjecture by estimating the following regression,

$$r_{i,p,t} - rf = \alpha_i + \beta^{MKT} r_{MRP} + \beta^{PE} r_{PE} + \varepsilon_{i,t} \tag{3}$$

where $i = \{\text{low-income, high-income}\}$, and p denotes the plan (e.g., traditional or Roth IRA).

Panel B of Table 5 reports the results of regression (3). Columns (1) and (2) report results for low-income individuals, while columns (3) and (4) those for high-income individuals. Results for low-income individuals are consistent with those found in Section 4.1, showing a positive and statistically significant exposure of 0.69 (0.81) to the aggregate equity market for traditional (Roth) IRA plans, and a zero exposure to private equity returns. It is also quite interesting how the private equity returns increase the explanatory power of the returns of low-income individuals compared to the other risk

[28] We use the percentages of wealth invested in business equity from the Survey of Consumer Finances (SCF). That is, in 2004, 38.5% of the wealth of individuals with an income above the 90th percentile was invested in business equity. This fraction equals 40.8% in 2007, 37.7% in 2010, 35.5% in 2013, 38.1% in 2016, and 37% in 2019.

[29] We formally test for the difference in average private equity return and the implied Roth IRA returns on private businesses, and find that this difference is economically large at 21% and statistically significant with a p-value of 0.091.

factors in Panel A. In fact, the R^2 increases around 3% in both regressions (1) and (2), while the number of additional regressors shrinks from four to only one (e.g., private equity returns). In contrast, columns (3) and (4) highlight that private equity returns are the main driver of high-income individuals IRA returns, with coefficients greater than one and highly statistically significant, with the aggregate equity market playing only a minor role. Despite this, around 40% of the variation in Roth IRA returns of high-income individuals remains unexplained by aggregate sources of risk, suggesting the idiosyncratic risk is quite relevant for those accounts.

5 IRA Return Heterogeneity and Wealth Inequality

In Section 3 we documented large return heterogeneity in pension plan (IRA) returns as a function of household income, while in Section 4 we tried to characterize what type of investments could have explained such performance gap, highlighting the potential role of private investments. In this section, we analyze whether and to what extent this return heterogeneity has affected households' wealth distribution and inequality, by using data from the Survey of Consumer Finances (SCF). Following Bricker et al. (2019) and Smith et al. (2021), we append the Forbes 400 members to the SCF and adjust the sampling weights to account for overlap between the two.[30] We study three different scenarios, in which we generate counterfactual wealth dynamics to understand how IRA return heterogeneity impacted (i) the wealth share of high-income households, (ii) the standard wealth share, and (iii) the wealth of the ten percent richest households.

The Survey of Consumer Finances (SCF) is a representative panel dataset of US households, providing information on income and wealth, where wealth can be decomposed into individual assets and liabilities.[31] The SCF dataset is available every three years. For our analysis we rely on the data sets for the years 2004, 2007, 2010, 2013,

[30] The use of SCF for estimating top wealth traces back to Wolff (1998). Our results are similar if we use the original, unadjusted SCF dataset, and note that incorporating the Forbes data has only a modest effect on our overall top share estimates, consistent with Smith et al. (2021). See also Parker et al. (2022) who use SCF data to construct a sample of retirement investors.

[31] For our analysis, we rely on the "Summary Extract Public Data", which contains aggregated summary variables expressed in USD (inflation-adjusted to 2019).

2016, and 2019. Each individual in the panel is weighted to make the observations a representative sample of the US populations at each point in time. The SCF data is very granular, and includes, for each individual, his income and the components making up his net wealth. Importantly, it includes the fraction of total net wealth derived by pension assets, and in particular that available through IRA plans.[32]

Before presenting our analysis, we emphasize once again the relevance of IRA plans. In the SCF, IRA assets make up 51% of the total pension assets, on average, over the last 16 years. This share increased from 48% in 2004 to 52% in 2019. Moreover, pension savings account, on average over our sample, for 16% of total net wealth. This share has been increasing constantly over time from 13% in 2004 to 17% in 2019, representing the second largest subcomponent of individual net wealth.[33]

Our first scenario estimates the impact of return heterogeneity in IRA plans on the wealth dynamics of low- and high-income groups. More in detail, every period we sort all households in the SCF by income and we split them in low- and high-income groups using \$200,000 as a threshold.[34] We then aggregate their total net wealth, by group, obtaining an estimate of the realized wealth dynamics for the two income groups.[35] Having access to the subcomponents of wealth, our counterfactual analysis takes the amount of IRA assets at the beginning of the sample 2004 for the high-income population, averaging 58% of total net wealth in our sample, and imposes annual returns on those assets equal to those obtained by the low-income households estimated each period (Figure 4, bottom panel).[36] We hence obtain a counterfactual wealth dynamic for high-income households, which we compare to their actual wealth. Both

[32]The SCF dataset lists IRA and Keoghs accounts together. We assume that 10% of the IRA pension assets in the SCF are Keoghs, consistent with fraction of people being self-employed in the United States according to the U.S. Bureau of Labor Statistics. For the Forbes 400 observations, we assume the individuals are high-income, and their IRA wealth is equal to the average size of the IRA accounts of households in the SCF dataset having more than \$1bn dollar of total wealth.

[33]The largest subcomponent is primary residence housing with a net equity share of 19% of net wealth.

[34]The SCF defines income at the houeshold level, while the IRS at the individual level. However, a high-income individual for the IRS will mechanically be high-income in the SCF (household level) too, hence our counter-factual analysis for the high-income group is unaffacted by composition issues.

[35]We assume an annual IRA savings rate of 2% of income for each individual, consistent with the savings rate in the US observed in the IRA data from the IRS.

[36]We use the sample average, by income group, for the missing year 2019.

series are normalized by the sum of low- and high-income household wealth to deliver a wealth share. Our results, plotted in the top panel of Figure 9, indicate that over time wealth inequality has increased from around 0.53 to 0.64, a twenty-one percent increase. Absent any return heterogeneity in IRA accounts, the wealth shares at the end of 2019 would be slightly smaller, around 0.63. These numbers suggest that the return spread on IRA assets has widened the wealth inequality between low- and high-income households by about 2% (0.01 divided by 0.53) in our sample.

Our second empirical analysis estimates the impact of IRA return heterogeneity on wealth inequality using a more classic "wealth inequality" approach. First, we rank households in the SCF dataset based on their net wealth, and split them into the top 10% and remaining 90% wealth subgroups. We then calculate the actual wealth shares over time (e.g., wealth of the top 10% / total wealth), as it is standard in the literature (Smith et al., 2021). Our counterfactual wealth shares is constructed by shutting down the return heterogeneity channel. In other words, we assume that high-income individuals, regardless of whether they belong to the top 10% of net wealth or not, obtain returns on their IRA assets equal to those of low-income individuals. Hence both the numerator (wealth of the top 10%) and denominator (total wealth) will change, since the dollar value of IRA assets for high-income household will be recalculated in both components of the ratio. The middle panel of Figure 9 shows the evolution of the actual and counterfactual wealth over time.

Our results show an actual increase of almost 10% in the wealth shares from 0.695 to 0.77 over the last fifteen years, consistent with recent estimates from the Federal Reserve Bank of Saint Louis and Smith et al. (2021). How much of this increase is attributable to the return heterogeneity in IRA accounts? Our estimate indicates that the wealth share differential is around 0.05, meaning that approximately 7% of the increase in the wealth shares can be linked to return heterogeneity in IRA plans.

Our last analysis focuses on the wealthy individuals (namely, those owning the top 10% of wealth), and ask how much their total net wealth would be, in USD, had some of them not obtained higher returns in their IRA pension assets. The bottom panel of Figure 9 shows that the wealth of the rich households hovers around $77tn in 2019, up

from 48tn in 2004, equivalent to a 60% increase over our 15 years sample. Most interesting, around $3tn of this wealth could be linked to the return spread in IRA pension assets. To put this number in perspective, the total US equity public market was around $34tn at the end of 2019. This means that, over our sample period, the increase in total wealth of the rich US households potentially due to pension asset performance differential has been close to 10% of the total US equity market value in 2019, a non trivial amount.

Overall, our results indicate that return heterogeneity in IRA pension plans is a potential driver of wealth share dynamics over time.

6 Robustness

6.1 Alternative Return Definition and Timing of Flows

Our baseline return definition (1) implicitely assumes that flows into IRA accounts happen at the end of the year. While this assumption is relatively harmless in our setting, since net flows only represent 2% of total wealth invested in IRA accounts, and hence have a tiny impact on annual return dynamics, we repeat our analysis by using the alternative return definition of Fagereng et al. (2020), which takes into account the potential different timing of the various flows components:

$$Ret_t = \frac{Value_t - Contributions_t - Rollovers_t + Withdrawals_t - Conversions_t - Value_{t-1}}{Value_{t-1} + 0.5 \times (Contributions_t + Rollovers_t - Withdrawals_t + Conversions_t)} \tag{4}$$

Our results are robust to the choice of return definition, since in practice the difference in the annual time series is marginal. The statistics in Table 2 for our baseline return definition are reproduced in Table 6 under the alternative return definition (4). The average returns across plans for low- and high-income individuals are 3.6% and 8.2% (with annualized standard deviations of 10.9% and 17%), respectively. These numbers are indeed very similar to the numbers reported in Panel B of Table 2.

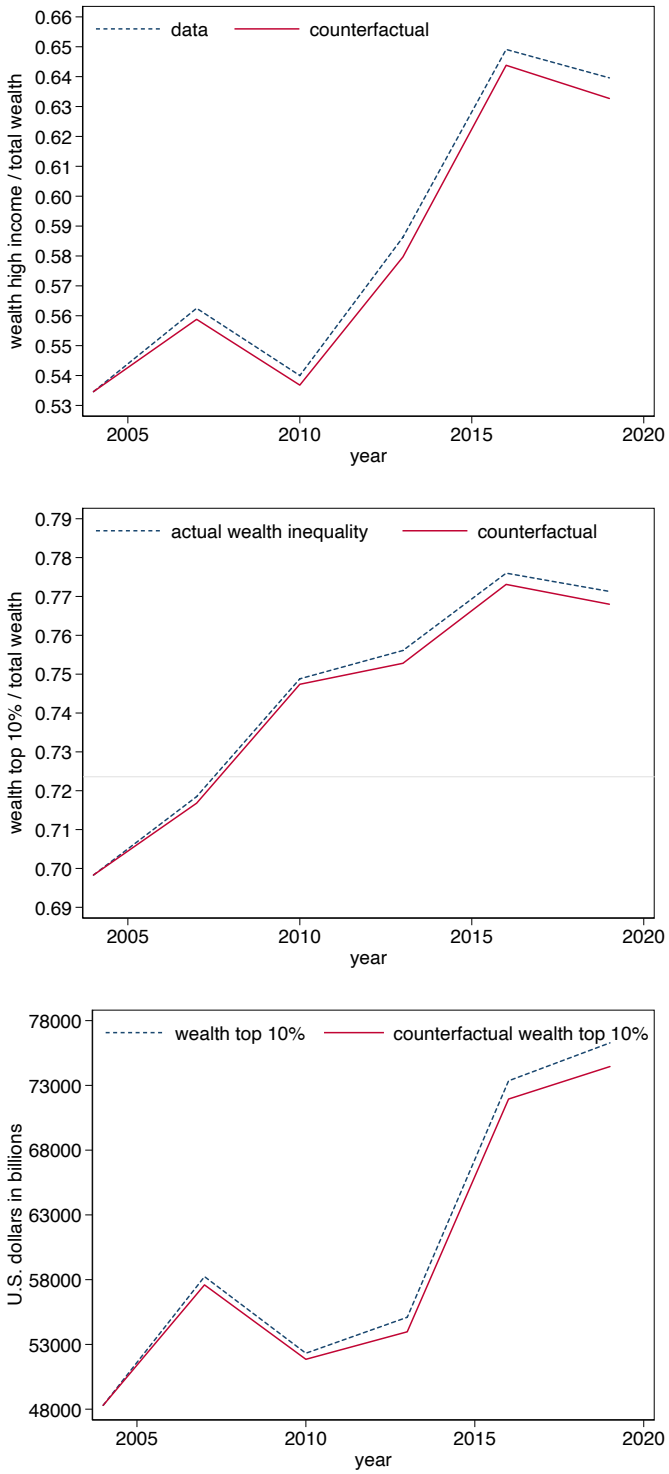


Figure 9: Wealth shares and total wealth in the USA. The top panel plots the empirical wealth share as a function of income, along with a counterfactual obtained had there been no return heterogeneity in IRA returns between low- and high-income individuals. Wealth inequality is defined as the ratio of total wealth held by high-income individuals divided by total wealth. The middle panel plots the empirical and counterfactual wealth shares over time. The counterfactual wealth share is calculated under the assumption of no return heterogeneity in IRA pension plans. The wealth share is defined as the ratio of the total wealth held by the top 10% in the total wealth distribution, divided by the total wealth. The bottom panel plots the total dollar wealth of the top 10% households in the wealth distribution, along with their counterfactual total dollar net wealth under the assumption of no return heterogeneity in retirement assets. The sample period is from 2004 through 2019.

6.2 Excluding the Global Financial Crisis

Table 7 computes the performance of different IRA plans for low- and high-income individuals under the baseline scenario, excluding the years of the Global Financial Crisis. We find that the return spread between high- and low-income individuals is 18% in Roth IRA, while it is 7.5% in traditional IRA. This difference is again mainly driven by high-income individuals, whose performance in Roth IRA plans is more than twice as large as that obtained in traditional IRA plans (19.0% versus 9.8%).

These figures are close in magnitude to those reported in Table 3 over the full sample. Overall, our conclusions are unaffected - qualitatively and quantitatively - by the exclusion of the Global Financial Crisis.

6.3 Excluding Roth IRAs Outliers

The results presented in Section 3.2 are obtained using aggregate data from all existing IRA accounts. As a consequence, it might be possible that the results concerning Roth IRAs of high-income individuals could be driven by a few wealthy outliers with extremely large realized returns. In order to gauge the sensitivity of our results to outliers, we re-run the analysis by excluding them using extremely conservative assumptions.

First, we assume there exist ten extremely large Roth IRA accounts of $5bn each. Given the fact that there was only one Roth IRA account in 2019 of around $5bn, this is an extremely conservative assumption. In fact, since Roth IRAs of high-income individuals have a total value of $270bn as of the end of 2018, this implicitly assumes that around 20% of Roth IRA assets of high-income individuals are owned by those ten individuals.

Second, we further assume that these outliers perform extremely well. More in detail, since we know that Mr. Thiel holds a substantial fraction of his Roth IRA account in both Facebook and Paypal shares, we assume that these ten outliers have an under-diversified portfolio invested half in Facebook and half in Paypal. We then use the Facebook returns from 2013 to 2018 and Paypal from 2016 to 2018 and backfill the previous returns to 2004 using their time-series annual averages of 36.63% and 36.59%,

respectively. In other words, we assume that these outliers earned outsized returns of close to 37% every year since 2004.

Lastly, we calculate the "ex-outliers" time series value of Roth IRAs of high-income individuals, from 2003 through 2018, by excluding, every year, the "implied" value of Roth IRAs owned by these outliers using the assumed return distribution.

Even under these extremely conservative assumptions regarding outliers, the average returns of Roth IRAs of high-income individuals are around 15.8% per year, approximately two percent lower than the full-sample average returns of 17.9%. The volatility of the "ex-outliers" Roth IRA returns slightly increases to 28.4%, resulting in a Sharpe ratio of 0.556, lower than the 0.704 reported in Table 3, but still four times that of low-income individuals under our baseline scenario. In summary, our results are not driven by outliers in Roth IRA plans.

7 Conclusion

In this paper, we present the first study of the performance of IRA pension plans over time, with a focus on traditional and tax-free Roth IRAs.

We first analyze the aggregate (i.e, across plans) IRA performance among different income groups, and find that average IRA returns across income groups follow a U-shape pattern, with individuals having annual income between $10,000 and $100,000 earning around 2%-3% per year on their IRA investments, while high-income individuals perform extremely well, with average returns above 8% and, in the case of individuals earning more than $1mn, almost 10% per year. These higher returns are not associated with greater risk-taking, since the Sharpe ratios of these high-income groups are also large, ranging from 0.351 to 0.519. Overall, IRA plans owned by individuals earning more than $200,000 outperform those held by low-income individuals by a factor of three over our sample period.

Second, we investigate whether the return spread between high- and low-income individuals is homogeneous across IRA types, or is mostly detectable in tax-free Roth IRA plans. Under the baseline scenario, we find that the return spread between high-

and low-income households in Roth IRAs is 15.9%, while that in traditional IRAs is 4.7%, over the full sample. Importantly, this difference is mainly driven by high-income individuals, whose performance in Roth IRA plans is more than twice as large as that obtained in traditional IRA plans (17.9% versus 8.6%). This evidence suggests that the two IRA plans – traditional and Roth – are exploited differently by high-income individuals in terms of allocations.

We then try to understand what can explain the performance of Roth IRAs owned by high-income individuals, and find that these returns cannot be matched by investments in public assets, including those with the best performance over our sample periods, namely growth stocks. However, the Roth IRA performance of high-income individuals is consistent with that of private investments, such as venture capital, private equity or startups.

Lastly, we analyze whether and to what extent this return heterogeneity has affected households' wealth distribution and inequality using data from the Survey of Consumer Finances augmented with the Forbes 400 data. Our findings suggest that up to 7% percent of the increase in the wealth share over our sample period could be attributed to the observed heterogeneity in IRA returns.

Overall, our results suggest that, on the one hand, the introduction of Roth IRAs, established to help hard-working, middle-class Americans save for retirement, had greatly benefited high-income households instead. However, Roth IRAs might have spurred productive and disruptive entrepreneurship, creating incentives for high-income individuals to fund private firms, including startups.

References

ALVAREDO, F., L. CHANCEL, T. PIKETTY, E. SAEZ, AND G. ZUCMAN (2017): "Global Inequality Dynamics: New Findings from WID.world," *American Economic Review*, 107, 404–09.

ANDONOV, A. AND J. D. RAUH (2021): "The Return Expectations of Public Pension Funds," *The Review of Financial Studies*, 35, 3777–3822.

ANG, A., B. CHEN, W. N. GOETZMANN, AND L. PHALIPPOU (2018): "Estimating Private Equity Returns from Limited Partner Cash Flows," *The Journal of Finance*, 73, 1751–1783.

BACH, L., L. E. CALVET, AND P. SODINI (2020): "Rich Pickings? Risk, Return, and Skill in Household Wealth," *American Economic Review*, 110, 2703–47.

BENHABIB, J. AND A. BISIN (2018): "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56, 1261–91.

BENHABIB, J., A. BISIN, AND M. LUO (2017): "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review*, 107, 593–97.

BENHABIB, J., A. BISIN, AND S. ZHU (2011): "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 79, 123–157.

BETERMIER, S., L. E. CALVET, AND P. SODINI (2017): "Who Are the Value and Growth Investors?" *The Journal of Finance*, 72, 5–46.

BINSBERGEN, J. H. V. AND R. S. KOIJEN (2010): "Predictive Regressions: A Present-Value Approach," *The Journal of Finance*, 65, 1439–1471.

BRICKER, J., P. HANSEN, AND A. H. VOLZ (2019): "Wealth concentration in the U.S. after augmenting the upper tail of the survey of consumer finances," *Economics Letters*, 184, 108659.

CAGETTI, M. AND M. DE NARDI (2006): "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 114, 835–870.

——— (2008): "Wealth Inequality: Data and Models," *Macroeconomic Dynamics*, 12, 285–313.

CAMPBELL, J. Y., T. RAMADORAI, AND B. RANISH (2019): "Do the Rich Get Richer in the Stock Market? Evidence from India," *American Economic Review: Insights*, 1, 225–40.

DE NARDI, M. (2004): "Wealth Inequality and Intergenerational Links," *The Review of Economic Studies*, 71, 743–768.

FAGERENG, A., L. GUISO, D. MALACRINO, AND L. PISTAFERRI (2016): "Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality," *American Economic Review*, 106, 651–55.

——— (2020): "Heterogeneity and Persistence in Returns to Wealth," *Econometrica*, 88, 115–170.

FAGERENG, A., M. B. HOLM, B. MOLL, AND G. NATVIK (2019): "Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains," NBER Working Papers 26588, National Bureau of Economic Research, Inc.

GABAIX, X., J.-M. LASRY, P.-L. LIONS, AND B. MOLL (2016): "The Dynamics of Inequality," *Econometrica*, 84, 2071–2111.

GOMEZ, M. (2018): "Asset Prices and Wealth Inequality," .

GOMEZ, M. AND E. GOUIN-BONENFANT (2020): "A Q-Theory of Inequality," *Working Paper*.

GREENWALD, D. L., M. LEOMBRONI, H. LUSTIG, AND S. VAN NIEUWERBURGH (2021): "Financial and Total Wealth Inequality with Declining Interest Rates," .

HOLDEN, S. AND D. SCHRASS (2021): "The IRA Investor Profile - Roth IRA Investors' Activity 2010-2018," Tech. rep., ICI.

HUBMER, J., P. KRUSELL, AND A. A. S. JR. (2020): "Sources of US Wealth Inequality: Past, Present, and Future," in *NBER Macroeconomics Annual 2020, volume 35*, National Bureau of Economic Research, Inc, NBER Chapters, 391–455.

KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

KRUEGER, D., K. MITMAN, AND F. PERRI (2016): "Macroeconomics and Household Heterogeneity," Elsevier, vol. 2, chap. Chapter 11, 843–921.

KUHN, M., M. SCHULARICK, AND U. I. STEINS (2020): "Income and Wealth Inequality in America, 1949–2016," *Journal of Political Economy*, 128, 3469–3519.

MIAN, A., L. STRAUB, AND A. SUFI (2020): "The Saving Glut of the Rich," Working Paper 26941, National Bureau of Economic Research.

NOVY-MARX, R. AND J. RAUH (2014): "The Revenue Demands of Public Employee Pension Promises," *American Economic Journal: Economic Policy*, 6, 193–229.

PARKER, J. A., A. SCHOAR, A. T. COLE, AND D. SIMESTER (2022): "Household Portfolios and Retirement Saving over the Life Cycle," *Working Paper*.

PIKETTY, T. (2014): *Capital in the Twenty-First Century*, The Belknap Press of Harvard University Press: Harvard University Press, 2 ed.

POTERBA, JAMES, M. S. F. V. AND D. A. WISE (2000): "Saver Behavior and 401(k) Retirement Wealth," *American Economic Review*, 90(2), 297–302.

QUADRINI, V. (2000): "Entrepreneurship, Saving, and Social Mobility," *Review of Economic Dynamics*, 3, 1–40.

ROBERT CLARK, A. L. AND O. S. MITCHELL (2017): "Financial knowledge and 401(k) investment performance: a case study," *Journal of Pension Economics and Finance*, 16, 324–347.

SABBATUCCI, R. (2022): "Are Dividends and Stock Returns Predictable? New Evidence Using Cash M & A Flows," *Working Paper*.

SABBATUCCI, R., A. TAMONI, AND S. XIAO (2023): "Stock Demand and Price Impact of 401(k) Plans," *Working Paper*.

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, 519–578.

SMITH, M., O. M. ZIDAR, AND E. ZWICK (2021): "Top Wealth in America: New Estimates and Implications for Taxing the Rich," .

TOLEDANO, C. (2020): "House Price Cycles, Wealth Inequality and Portfolio Reshuffling," Tech. rep.

WOLFF, E. N. (1998): "Recent Trends in the Size Distribution of Household Wealth," *Journal of Economic Perspectives*, 12, 131–150.

Table 1: Comparison of traditional and Roth IRA plans. This table compares traditional and Roth IRA plans in terms of contributions, taxes, withdrawals, distributions, liquidity and types of assets as of Q1 2022.

	Traditional IRAs	**Roth IRAs**
Contributions		
Source	Pre-tax basis	After-tax basis
Eligibility / Income requirements	Anyone 18 or over with earned income can contribute to a traditional IRA. There are specific income limits for how much might be tax-deductible	Those with earned income below a certain threshold. In 2021, $125,000 for single and $198,000 for married filing jointly
Maximum amount (2021)	$6,000 ($7,000 if over 50)	$6,000 ($7,000 if over 50)
Taxes		
Contributions grow	Tax-deferred	Tax-free
Tax deductability	Immediate tax benefits	No current-year tax benefits
Withdrawals		
Penalties	Earnings and contributions are penalty-free only after age 59 and half. They are taxed as current income when withdrawn	Earnings can be withdrawn penalty- and (federal) tax-free after 5 years and age 59 and half. Contributions can always be withdrawn tax- and penalty-free at any time
Distributions		
Required minimum distributions	After turning 72	None
Liquidity		
Plan liquidity	Even if invested in liquid investments, the taxes and penalties for withdrawals make traditional IRAs an illiquid asset	Since contributions can be withdrawn without any taxes or penalties, Roth IRAs may be considered liquid assets, especially if invested in short-term assets (e.g., money-market mutual funds)
Assets		
Investable assets	All except life insurance, collectibles, S-corporations	All except life insurance, collectibles, S-corporations

Table 2: Returns - summary statistics. Panel A reports summary statistics on the IRA returns (across all plans) by income groups. Panel B reports summary statistics on the IRA performance of high- (e.g. more than $200,000) and low-income individuals by collapsing the income groups reported in Panel A. Panel C reports summary statistics on the returns of the four IRA plans separately and jointly, as well as for the S&P500.

A. Income Groups

	Mean	Median	SD	Sharpe Ratio
$1 – $5,000	0.065	0.032	0.319	0.205
$5,000 – $10,000	0.039	-0.026	0.218	0.180
$10,000 – $15,000	0.027	-0.007	0.184	0.146
$15,000 – $20,000	0.025	0.005	0.182	0.137
$20,000 – $25,000	0.032	0.012	0.133	0.244
$25,000 – $30,000	0.019	0.001	0.133	0.143
$30,000 – $40,000	0.021	-0.017	0.144	0.145
$40,000 – $50,000	0.012	-0.024	0.097	0.122
$50,000 – $75,000	0.017	0.042	0.105	0.158
$75,000 – $100,000	0.025	0.040	0.122	0.203
$100,000 – $200,000	0.064	0.087	0.132	0.484
$200,000 – $500,000	0.084	0.082	0.161	0.519
$500,000 – $1,000,000	0.088	0.109	0.194	0.455
$1,000,000 or more	0.096	0.046	0.273	0.351

B. High- versus Low-Income

	Mean	Median	SD	Sharpe Ratio
less than $200,000	0.036	0.054	0.111	0.324
more than $200,000	0.085	0.105	0.174	0.489

C. By Plan

	Mean	Median	SD	Sharpe Ratio
SEP	0.054	0.077	0.119	0.450
Traditional IRA	0.049	0.061	0.111	0.443
Roth IRA	0.040	0.069	0.134	0.297
SIMPLE	-0.030	0.007	0.136	-0.219
All together	0.046	0.060	0.112	0.410
S&P 500	0.067	0.105	0.163	0.413

Table 3: Plan-specific returns under the different scenarios. This table reports summary statistics on the performance of the different IRA plans for high- and low-income individuals, under the three different scenarios. Columns (1)-(4) report results for the baseline scenario, columns (5)-(8) for the "biased towards equality" scenario, and columns (9)-(12) for the "biased towards inequality" scenario. The first (second) column in each group reports the average (median) return, while the third column reports the return volatility. The last column within each block reports the Sharpe ratio of the IRA plan. The sample period is from 2004 through 2018.

	Baseline				Biased Towards Equality				Biased Towards Inequality			
	Mean	Median	SD	SR	Mean	Median	SD	SR	Mean	Median	SD	SR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Traditional IRA												
less than $200,000	0.039	0.054	0.107	0.362	0.060	0.078	0.107	0.556	0.007	0.026	0.107	0.068
more than $200,000	0.086	0.117	0.178	0.485	0.011	0.038	0.167	0.066	0.195	0.231	0.186	1.048
B. Roth IRA												
less than $200,000	0.020	0.053	0.133	0.149	0.026	0.056	0.134	0.190	0.000	0.018	0.133	-0.001
more than $200,000	0.179	0.163	0.254	0.704	0.115	0.121	0.224	0.510	0.438	0.211	0.847	0.517
C. SEP												
less than $200,000	-0.049	-0.012	0.143	-0.346	-0.049	-0.012	0.143	-0.342	-0.050	-0.012	0.143	-0.353
more than $200,000	0.020	0.054	0.149	0.134	0.018	0.052	0.149	0.124	0.022	0.055	0.149	0.148
D. SIMPLE												
less than $200,000	0.045	0.050	0.119	0.381	0.060	0.067	0.118	0.511	0.033	0.034	0.119	0.281
more than $200,000	0.066	0.103	0.152	0.435	0.041	0.072	0.150	0.272	0.086	0.123	0.153	0.562

Table 4: Returns of high-income individuals versus public and private asset returns . This table reports summary statistics of IRA returns of high-income individuals, together with those of public and private investments. Panel A reports the baseline return estimates for Traditional and Roth IRAs of high-income individuals (rows (1)-(2)) in Table 3, and the returns of tradable growth and value strategies (rows (3)-(4)). The first (second) column in each group reports the average (median) return, while the third and fourth columns report the return volatility and Sharpe ratio, respectively. The last column reports the difference in means test (t-test) with respect to Roth IRA returns (p-values in parenthesis). Significance levels are denoted by $*$ = 10%, $**$ = 5%, and $***$ = 1%. Panel B reports the same analysis for US venture capital and private equity (buyout) return data from Cambridge Associates. The last row shows the implied return of private assets held within Roth IRA plans under the assumption that high-income households invest a fraction of their Roth IRA wealth in the S&P500 and the remaining in private assets. The sample period is from 2004 to 2018.

Panel A: Public Investments

	Mean	Median	SD	SR	Δ
	(1)	(2)	(3)	(4)	(5)
Traditional IRA (baseline)	0.086	0.117	0.178	0.489	
Roth IRA (baseline)	0.179	0.163	0.254	0.704	
Growth	0.089	0.092	0.179	0.496	0.090*
(p-value)					(0.076)
Value	0.052	0.109	0.167	0.312	0.127**
(p-value)					(0.017)

Panel B: Private Investments

	Mean	Median	SD	SR	Δ
	(1)	(2)	(3)	(4)	(5)
Roth IRA (baseline)	0.179	0.163	0.254	0.704	
Cambridge Associates VC Index	0.110	0.132	0.109	1.001	0.069
(p-value)					(0.124)
Cambridge Associates PE Index	0.143	0.148	0.132	1.081	0.036
(p-value)					(0.222)
Returns of private assets in Roth IRA (implied)	0.352	0.260	0.544	0.648	

Table 5: Drivers of IRA returns. This table reports the results from regressions of IRA returns on financial and macroeconomic explanatory variables. Panel A reports the results of regressions of IRA returns for low- and high-income individuals on the Fama-French five factors model. Panel B reports the same regressions using the US private equity returns from Cambridge Associates in addition to the aggregate equity market. The sample period is from 2004 through 2018.

Panel A: Fama-French Five Factor Model				
	Low-income		High-income	
	Traditional	Roth	Traditional	Roth
	(1)	(2)	(3)	(4)
β^{MKT}	0.636***	0.746***	0.906**	0.720
	(0.103)	(0.077)	(0.299)	(0.514)
β^{SMB}	0.061	0.087	-1.301	-0.934
	(0.261)	(0.194)	(0.759)	(1.306)
β^{HML}	-0.134	-0.091	-0.367	-0.671
	(0.158)	(0.118)	(0.460)	(0.791)
β^{RMW}	0.320	0.010	0.175	-1.060
	(0.293)	(0.219)	(0.853)	(1.468)
β^{CMA}	0.113	0.049	1.537	1.829
	(0.308)	(0.230)	(0.897)	(1.543)
constant	-0.036	-0.055***	-0.011	0.127
	(0.020)	(0.015)	(0.058)	(0.099)
N	14	14	14	14
R^2	82.48%	93.68%	42.12%	13.64%

Panel B: Private Equity Returns				
β^{MKT}	.693***	0.812***	0.003	-0.360
	(0.115)	(0.083)	(0.295)	(0.464)
β^{PE}	-0.197	-0.101	1.124***	1.880***
	(0.157)	(0.113)	(0.404)	(0.635)
constant	-0.004	-0.045***	-0.073	-0.049
	(0.017)	(0.012)	(0.045)	(0.071)
N	14	14	14	14
R^2	85.80%	95.27%	63.56%	54.71%

Table 6: Alternative return definition - summary statistics. Panel A reports summary statistics on the IRA returns (across all plans) by income groups. Panel B reports summary statistics on the IRA performance of high- (e.g. more than $200,000) and low-income individuals by collapsing the income groups reported in Panel A. Panel C reports summary statistics on the returns of the four IRA plans separately and jointly, as well as for the S&P500.

A. Income Groups

	Mean	Median	SD	Sharpe Ratio
$1 – $5,000	0.065	0.031	0.317	0.205
$5,000 – $10,000	0.039	-0.026	0.216	0.180
$10,000 – $15,000	0.027	-0.007	0.184	0.148
$15,000 – $20,000	0.025	0.005	0.182	0.137
$20,000 – $25,000	0.033	0.012	0.133	0.246
$25,000 – $30,000	0.019	0.001	0.132	0.144
$30,000 – $40,000	0.021	-0.017	0.143	0.145
$40,000 – $50,000	0.012	-0.024	0.096	0.121
$50,000 – $75,000	0.016	0.041	0.104	0.157
$75,000 – $100,000	0.024	0.040	0.120	0.203
$100,000 – $200,000	0.062	0.086	0.129	0.482
$200,000 – $500,000	0.081	0.080	0.157	0.517
$500,000 – $1,000,000	0.086	0.106	0.190	0.452
$1,000,000 or more	0.096	0.045	0.272	0.352

B. High- versus Low-Income

	Mean	Median	SD	Sharpe Ratio
less than $200,000	0.036	0.053	0.109	0.326
more than $200,000	0.082	0.102	0.170	0.484

C. By Plan

	Mean	Median	SD	Sharpe Ratio
SEP	0.053	0.075	0.118	0.446
Traditional IRA	0.048	0.060	0.110	0.441
Roth IRA	0.037	0.066	0.128	0.288
SIMPLE	-0.028	0.006	0.127	-0.224
All together	0.045	0.059	0.110	0.407
S&P 500	0.067	0.105	0.163	0.413

Table 7: Plan-specific returns when excluding the financial crisis. This table reports summary statistics on the performance of the different IRA plans for high- and low-income individuals when excluding the financial crisis. The first (second) column in each group reports the average (median) return, while the third column reports the return volatility. The last column within each block reports the Sharpe ratio of the IRA plan. The sample period is from 2004 through 2018 (excluding the year 2008).

	Mean	Median	SD	SR
	(1)	(2)	(3)	(4)
A. Traditional IRA				
less than $200,000	0.023	0.051	0.094	0.248
more than $200,000	0.098	0.124	0.179	0.546
B. Roth IRA				
less than $200,000	0.006	0.049	0.127	0.044
more than $200,000	0.190	0.179	0.261	0.727
C. SEP				
less than $200,000	0.031	0.039	0.111	0.280
more than $200,000	0.017	0.050	0.154	0.110
D. SIMPLE				
less than $200,000	-0.069	-0.016	0.128	-0.540
more than $200,000	0.017	0.050	0.154	0.110