Private Capital Markets and Inequality[*]

Ararat Gocmen[†] Clara Martínez-Toledano[‡] Vrinda Mittal[§]

May 17, 2025

Abstract

This paper studies the relationship between the growth in private capital markets and the rise in economic inequalities in the U.S over the last two decades. First, we document that the share of financing raised by early-stage private companies from U.S.-based high-net-worth individuals (HNWIs) tripled from 2004 to 2022. Second, exploiting state-level variation in exposure to the expanded federal capital gains tax exclusion on qualified small business stock (QSBS), we find that the growth in HNWIs' early-stage investments increased the average income gap between HNWIs and other income earners by 6.0%. Third, we show that this rise in income concentration appears to have been driven by HNWIs' excess returns on their early-stage investments relative to public stock market returns. Finally, using counterfactual simulations, we find that HNWIs' excess returns on these investments accounted for 11% and 5% of the growth in the top 1% share of income and wealth, respectively, from 2010 to 2019.

[*]We thank Saleem Bahaj, Ralph Koijen, Ramana Nanda, Tarun Ramadorai, Uta Schönberg, and seminar participants at the CEPR Household Finance, Lake District Corporate Finance, Northeastern Finance Conference, Helsinki Graduate School of Economics, Imperial College Business School, Institute for Private Capital, London Business School, Norwegian School of Economics, Paris School of Economics, University College London, University of Copenhagen, University of North Carolina, University of Oxford, European Economic Association Annual Congress, London Inequality Workshop, Madrid Public Economics Workshop, and Rome Junior Finance Conference for helpful comments. Scheduled to present at NBER SI Household Finance 2025, WFA, EFA, FIRS, LBS Private Capital, IPC PERC Oxford.

We acknowledge financial support from the James M. and Cathleen D. Stone Centre on Wealth Concentration, Inequality, and the Economy at the University College London Department of Economics. We are very grateful to Maitreyee Guha for excellent research assistance.

[†]University College London and Centre for Research and Analysis of Migration; address: University College London, Department of Economics, Drayton House, 30 Gordon Street, London WC1H 0AX, UK; email: ararat.gocmen.19@ucl.ac.uk

[‡]Imperial College Business School and Centre for Economic Policy Research; address: Imperial College London, South Kensington Campus, London SW7 2AZ, UK; email: c.martinez-toledano@imperial.ac.uk; website: https://sites.google.com/view/claramartinez-toledanotoledano/home.

[§]Kenan-Flagler Business School, University of North Carolina at Chapel Hill; address: University of North Carolina at Chapel Hill, McColl Building, 300 Kenan Center Drive, Chapel Hill, NC 27599, USA; email: Vrinda_Mittal@kenan-flagler.unc.edu; website: https://www.vrindamittal.com.

1 Introduction

Two important stylized facts have marked the last four decades in the U.S. On the one hand, income and wealth concentration have steadily risen (Saez and Zucman, 2016; Piketty et al., 2018). On the other hand, private capital markets have expanded, while public stock market listings have fallen (Stulz, 2020; Ewens and Farre-Mensa, 2022). The aim of this paper is to examine whether high-net-worth individuals (HNWIs) have increased their participation in private capital markets and, if so, whether their growing presence in these markets could be behind these two key macroeconomic trends.[1]

The rise in economic inequalities and the increasing participation of HNWIs in private capital markets could be strongly linked for two main reasons. First, only the wealthiest individuals are generally able to invest in private companies (Jensen et al., 2017; Mikhail, 2022), which leads private business wealth being highly concentrated at the top of the income and wealth distribution (Kopczuk and Zwick, 2020). Second, if the returns on private companies were also larger than the returns on public companies (Kartashova, 2014; Brown and Kaplan, 2019; Balloch and Richers, 2023), then, all else equal, income and wealth inequality would increase.

We focus on the U.S. for two reasons. First, U.S. companies account for about half of all the financing raised in global private capital markets (Lerner and Nanda, 2020). Second, the U.S. federal government introduced tax breaks after the 2008 financial crisis to incentivize HNWIs' investments in early-stage private companies (Polsky and Yale, 2023). These reforms represent a quasi-exogenous shock to HNWIs' participation in private capital markets, which we exploit to study how HNWIs' early-stage investments shaped the dynamics of inequality in the U.S.[2] To rationalize the effects on inequality, we further estimate the returns associated with these early-stage investments and compare them to public stock market returns. We carry out our analyses using data on private capital market activity from Pitchbook, which includes information on the financing received by private companies, the investors participating in each deal, and the changing valuation of each company across deals.

We provide three sets of results. First, we show that U.S.-based HNWIs' participation in

[1] By "HNWIs," we refer to individuals who satisfy the U.S. Securities and Exchange Commission's (SEC) definition of accredited investors: those whose combined net worth with their spouse (excluding the value of their primary residence) exceeds \$1 million, or whose combined (individual) income has exceeded \$300,000 (\$200,000) for the last two years. The SEC's website provides further details about this definition: https://www.sec.gov/resources-small-businesses/capital-raising-building-blocks/accredited-investors.

[2] Throughout this paper, we use "early-stage investments" interchangeably with "investments in startup companies." By these, we refer to investments in companies that Pitchbook categorizes as in either the pre-seed, seed, early, or late stage of their development.

U.S. private capital markets has grown considerably in recent decades. This growth has been driven mainly by their investments in early-stage companies—increasing from about \$0.4 billion in 2004 to almost \$15 billion in 2022—rather than by their investments in more mature private companies. HNWIs' early-stage investments have increased not only in absolute but also in relative terms. In particular, the share of overall financing raised by U.S. startups from U.S.-based HNWIs tripled from about 2% to 6% over the 2004-2022 period.[3] We further document that the accumulated value of HNWIs' early-stage investments since 2004 reached almost \$460 billion by 2022, which was more than double their counterfactual value had HNWIs instead invested in either the NASDAQ 100, S&P 500, or Russell 2000 public stock market indices. HNWIs have thus earned, on average, excess returns on their early-stage investments relative to public stock market returns.[4]

Second, we establish a link between the increase in early-stage investments by HNWIs and the rise in U.S. income inequality over the last two decades. To do so, we exploit the expansion of an existing federal tax exclusion on long-term capital gains from the sale of qualified small business stock (QSBS). This tax exclusion is meant to incentivize HNWIs' investments in early-stage companies, since it applies only to QSBS purchased from companies that satisfy the following three conditions. First, the issuing company must be a U.S. C-corporation; second, the company's gross assets must have never exceeded \$50 million, inclusive of the financing raised from the investor purchasing the QSBS; and, third, at least 80% of the company's assets must be actively used in qualified—mainly non-professional and non-extractive—trades and businesses. At the exclusion's original introduction in 1993, investors who held QSBS for at least five years could exclude 50% of the first \$10 million of the associated capital gains. However, in response to the 2008 financial crisis, the federal government temporarily expanded the exclusion rate to 75% in 2009, and to 100% in 2010. Finally, the 100% exclusion was made permanent in 2015.

To study how early-stage investments by HNWIs shape the dynamics of income inequality, we follow a two-step approach. In the first step, we study the effects of the QSBS reforms on HNWIs' early-stage investments, and in the second step, we estimate the effects of these investments on income inequality. In both steps, we exploit state-level variation in exposure to the federal reforms, based on the number of HNWIs who resided in each state in 2008 (i.e., the year immediately prior to these reforms). To avoid additional regulatory burdens, startups generally only raise money from HNWIs, who have home

[3] Although HNWIs account for a relatively small share of overall early-stage financing, the growth in HNWIs' early-stage investments could still have important implications for economic inequalities, given that the ownership of private companies is highly concentrated at the top of the income and wealth distribution (Kopczuk and Zwick, 2020).

[4] We estimate the returns on each company based on the history of its valuation, as recorded by Pitchbook each time the company raised new financing.

bias and invest disproportionately more in companies in the same states where they reside. Hence, the reforms should increase HNWIs investments more in states where the ex-ante number of resident HNWIs is higher. This setting has two main threats for identification. On the one hand, HNWIs may choose to settle in certain states specifically to get access to exclusive new investment opportunities (e.g., aspiring venture capitalists moving to California). On the other hand, resident HNWIs' early-stage investment portfolios may be biased toward local startups. In the latter case, if the startups in states with more resident HNWIs are exposed to different economic shocks than those in states with fewer resident HNWIs, then startup investments by HNWIs residing in one state may grow faster or slower relative to those by HNWIs residing in another state, and for reasons entirely unrelated to the reforms.

In the first step, we thus compare how the QSBS reforms affected the in-state investments in early-stage companies by resident HNWIs relative to those by other types of investors—namely, resident institutional investors, non-resident institutional investors, and non-resident HNWIs. This comparison allows us to control for interacted state-year fixed effects. We estimate that the expansion of the QSBS tax exclusion explained, on average, 20% of the overall growth in U.S.-based HNWIs' investments in early-stage companies between 2004-2008 and 2009-2022. To ensure that this result is driven by the QSBS reforms, we also perform a company-level analysis where we compare how the gap in the probability of raising financing from HWNIs between QSBS-eligible and QSBS-ineligible companies changed after the reforms. We find that the probability of raising financing from HNWIs increased, on average, by 2.1 percentage points more for QSBS-eligible relative to QSBS-ineligible companies in the post-reform period. In the second step, we use a similar state-level specification to the one that we use for investments, comparing how the QSBS reforms affected the average income gap between HNWIs and other income earners. We find that the increase in HNWIs' early-stage investments increased the average income gap between the top 0.5% and bottom 99.5% by 6.0% in the post-reform period.

Third, we document that HNWIs' excess returns on their early-stage investments (relative to public stock market returns) can be associated with the rise in U.S. income concentration over the last two decades. To that end, we first decompose our state-level income measures into three components: labor income, realized capital gains, and other capital income. We then show, using the same state-level design, that the rise in the average income gap between the top 0.5% and bottom 99.5% in the post-reform period was mainly due to a disproportionate rise in the capital gains of the top 0.5%. We further show that this rise in the average capital gains gap between the top 0.5% and the bottom 99.5% is strongly associated with an increase in HNWIs' excess returns from their early-stage investments. We also document, by means of counterfactual simulations, that these excess returns account for 11% and 5% of the overall nationwide growth in the top 1% shares of taxable income and wealth, respectively, in the post-reform period. Finally, we show that the rise in economic inequalities can lead to a further increase in HNWIs' participation in private capital markets, suggesting the existence of a feedback loop between the two phenomena.

This paper contributes to three main strands of the literature. First, we contribute to the growing theoretical and empirical literature on the dynamics of income and wealth inequality, which—in addition to savings, bequests, interest rates, and labor income—has emphasized asset prices and returns as important determinants of those dynamics (De Nardi, 2004; Jones, 2015; De Nardi and Fella, 2017; Gomez, 2017; Kuhn et al., 2017; Feiveson and Sabelhaus, 2018; Bach et al., 2020; Fagereng et al., 2020; Cioffi, 2021; Greenwald et al., 2021; Hubmer et al., 2021; Mian et al., 2021; Xavier, 2021; Bauluz et al., 2022; Meeuwis, 2022; Andersen et al., 2023; Blanchet and Martínez-Toledano, 2023; Martínez-Toledano, 2023; Nekoei and Seim, 2023; Gomez and Gouin-Bonenfant, 2024). While confirming the importance of return heterogeneity as one such determinant, we also identify a new channel to explain it—namely, the differences across the income and wealth distribution in individuals' access to and participation in private capital markets.

In this way, we further contribute to the separate literature that focuses on measuring the returns to different asset classes and providing explanations for the heterogeneity in those returns across investors. A number of theoretical papers have suggested that such heterogeneity can be driven by differences in entrepreneurial ability (Lucas, 1978), information (Peress, 2004), or sophistication (Kacperczyk et al., 2019). Several recent empirical studies have instead explored the return heterogeneity within a particular asset class, including household wealth (Bach et al., 2020; Fagereng et al., 2020; Xavier, 2021; Balloch and Richers, 2023), bank deposits (Deuflhard et al., 2019), and stocks (Calvet and Fisher, 2007; Campbell et al., 2019). Our paper is most closely related prior studies examining differences in returns between public and private companies (Moskowitz and Vissing-Jørgensen, 2002; Kartashova, 2014; Brown and Kaplan, 2019; Brown et al., 2021; Balloch and Richers, 2023), most of which have focused on documenting either the under- or outperformance of buyout funds. Our findings that early-stage private companies have outperformed public stock markets in the U.S. over the last two decades are in contrast to those of Moskowitz and Vissing-Jørgensen (2002) for the 1990s, but they are consistent with those of Kartashova (2014) and Balloch and Richers (2023) for more recent periods.

Finally, we contribute to the literature on investors in private capital markets. Most existing studies have focused on institutional investors like pension funds and endowments (Lerner and Schoar, 2004; Lerner et al., 2007; Sørensen, 2007; Robinson and Sensoy, 2013; Maurin et al., 2022; Mittal, 2024). Following the recent literature's growing interest in angel investors (Lindsey and Stein, 2020; Bach et al., 2022; Karlsen et al., 2023), we focus on HNWIs' participation in private capital markets, especially on their investments in early-stage companies. Though other papers have also studied the effects of tax breaks that incentivize investments in early-stage companies (Edwards and Todtenhaupt, 2020; Denes et al., 2023; Chen and Farre-Mensa, 2024), we are, to be best of our knowledge, the first to do so with an emphasis on the implications for income and wealth inequality.

The rest of this paper is organized as follows. Section 2 first describes the data that we use, while Section 3 documents key stylized facts. Section 4 then analyzes the effects of the QSBS reforms on HNWIs' early-stage investments, after which Section 5 discusses the implications of these investments for inequality. Section 6 finally concludes.

2 Data

2.1 Private Capital Market Activity

Pitchbook. Our main data source for private capital market activity is Pitchbook, which is a commercial data provider that collects data on financing deals, the investors and funds investing in them, and the companies invested in. From 2004 to 2022, it contains information on 441,430 deals for 201,414 U.S. companies, corresponding to 3,089,706 distinct investments by 126,440 distinct investors. Pitchbook collects this data from various sources, including press releases and regulatory filings by companies, Freedom of Information Act requests to public pension funds, and correspondence with the general and limited partners of private investment funds (Cumming and Monteiro, 2023).[5]

First, we use Pitchbook to identify the investors participating in each financing deal and measure the amount that they invested in the deal. We classify investors into HNWIs (i.e., individuals, angel groups, and family offices) and institutional investors (e.g., pension funds, endowment plans, foundations, funds of funds). We also classify private capital market deals into early-stage investments (i.e., equity investments in startup companies), private equity (i.e., equity investments in more mature companies), private debt (i.e., debt investments by non-bank entities in the form of non-bond loans), and real assets (i.e., equity investments in real estate, infrastructure, or natural resources).[6] These categories include both direct investments in companies by investors and intermediated investments by private investment funds. For intermediated investments, we attribute the investments by each fund to its limited partners, based on the amount committed to the fund by

[5] For further details about Pitchbook's data collection process see Pitchbook's website: https://pitchbook.com/research-process.

[6] We follow the standard categorization used by private investment professionals: https://www.preqin.com/academy/lesson-2-private-capital/what-is-private-capital. We prefer "early-stage" to "venture capital" investments, since the latter usually refers only to intermediated investments in startup companies.

each limited partner. Appendix A contains a detailed description of our data-cleaning procedure for Pitchbook's data on companies, deals, investments, and investors.[7]

Second, we also rely on data from Pitchbook to calculate the returns on HNWIs' early-stage investments. Although returns are often calculated at the fund level, we instead emphasize investment-level returns, given that 81.6% of U.S.-based HNWIs' early-stage investments in U.S. companies between 2004 and 2022 are direct rather than intermediated. We first use information on the changing valuation of each company across deals to calculate each company's rate of return. In particular, we compare—between any consecutive pair of deals—the company's pre-money valuation from the later deal (i.e., its valuation before accounting for the new financing that it raised as part of this deal) to its post-money valuation from the earlier deal (i.e., its valuation after accounting for the new financing that it raised as part of that deal). This comparison accounts for any shareholder dilution between deals, differentiating the growth in the value of the investments by the earlier deal's investors from the value of the new investments by the later deal's investors. Based on this approach, we can calculate the historical rate of return on each company. We can then calculate the return earned on each investment in each year, and thereby the annual returns of HNWIs on all of their early-stage investments, by letting the value of each investment in each company evolve over time according to the company's historical rate of return. There are instances where the valuation of a company is missing as part of a particular deal, or where the company did not raise financing in a particular year. We discuss the methodology that we use to impute these missing valuations, as well detail other aspects of our return methodology in Appendix B.

Finally, we use our investment-level dataset to construct a company-level panel dataset for our company-level analyses. In particular, we distinguish between QSBS-eligible and QSBS-ineligible companies by relying on the legal type, primary industry, and financing history of each company. We also use each company's financing history to identify the year in which it was founded, as well as when it first became bankrupt, went public, or was acquired by another company. We use this information for two distinct purposes. On the one hand, since our company-level analysis focuses on companies that were active at the time of the QSBS reforms, the information allows us to distinguish these companies from those that had already become inactive or were founded only after the reforms. On the other hand, the same information also makes it possible to study the effects of the reforms on bankruptcies, initial public offerings, and acquisitions. We provide further details about how we identify QSBS-eligible companies in Appendix Sections A.2 and A.3.

[7] We prefer Pitchbook's data to data from other providers (e.g., Preqin, Burgiss) because Pitchbook has more extensive data coverage, especially of early-stage investments by HNWIs. For further details, see Pitchbook's website: https://pitchbook.com/compare/pitchbook-vs-preqin.

2.2 High-Net-Worth Individuals and Inequality

Geographic Wealth Inequality Database. To conduct the state-level regression analyses, we require a measure of the total number of HNWIs residing in each U.S. state. For that, we rely on the Geographic Wealth Inequality Database (GEOWEALTH-US) built by Suss et al. (2024). This database provides estimates of the number of HNWIs residing in each state in every year from 2005 to 2022. To obtain these series, the authors first estimate the relationship between wealth and other observable characteristics on the sample of individuals who appear in the Survey of Consumer Finances (SCF). They then predict the wealth of individuals sampled in U.S. population surveys, in which those same characteristics—other than wealth—are also observable.

Based on observable income and estimated wealth, Suss et al. (2024) define HNWIs to resemble the U.S. Securities and Exchanges Commission's (SEC) definition of accredited investors.[8] This is because, in the view of the SEC, only accredited investors are sophisticated enough to invest in unregistered securities like the QSBS issued by early-stage private companies. Private investment funds and companies that raise financing from non-accredited investors are therefore required to register their securities with the SEC.[9]

In Appendix C, we validate the number of accredited investors residing in each state as measured in GEOWEALTH-US with alternative estimates from Phoenix Marketing International/MarketCast Wealth and Affluent Monitor, Forbes 400, Credit Suisse, and the SCF. Our baseline measure based on GEOWEALTH-US appears to be consistent with alternative sources, both across states and over time.

Statistics of Income. To conduct the state-level inequality analyses, we build state-level income inequality series based on the personal income tax statistics from the Statistics of Income (SOI) database provided by the U.S. Internal Revenue Service (IRS). Specifically, we use the historical data tables that provide information on a range of personal income tax items, which are aggregated by state and adjusted gross income (AGI) bracket for each year from 2004 to 2022. AGI refers to income from all sources, including labor income, investment income, business profits, and retirement income, adjusted for tax deductions.

[8] For this definition, see footnote 1. Suss et al. (2024) estimate the wealth of individuals who appear in cross-sectional rather than longitudinal population surveys. As a result, they only consider whether an individual's household income exceeds $300,000 in the current year for the income test. Furthermore, they consider the individual's wealth excluding the value of their primary residence for the wealth test.

[9] Private funds and companies can raise financing from up to 35 non-accredited investors before triggering SEC registration: https://www.sec.gov/resources-small-businesses/exempt-offerings/private-placements-rule-506b. In practice, however, this number is so low—and the amount that can be raised from them so limited—that non-accredited investors have generally been excluded from private capital markets; the only exception has been when private companies raise financing via crowdfunding (Jensen et al., 2017).

We apply the method of generalized Pareto interpolation (GPI) developed by Blanchet et al. (2022). GPI is a non-parametric approach that avoids the assumptions of a Pareto approximation, which are often violated by empirical data. For every state in each year, we construct the state-level income distribution across individuals using data on IRS tax filing units, assuming that the reported household income of couples filing jointly is shared equally between spouses. Our series are consistent with those of Sommeiller and Price (2018), who build state-level income inequality series for the U.S. using personal income tax tabulations from 1917 to 2015.

Using the information available on the composition of income for each tax bracket, we further decompose the aggregate income in that bracket into labor income, realized capital gains, and capital income (i.e., dividends, interest income, and other income from investments). We also aggregate our state-level income inequality series at the national level to implement the counterfactual simulations of U.S. income inequality in Section 5.2.

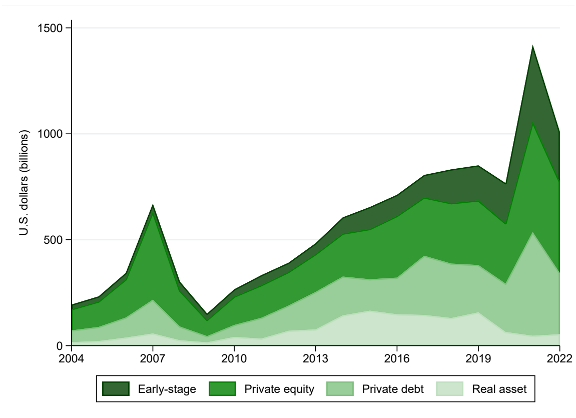
Survey of Consumer Finances. We also construct a series for the nationwide wealth distribution to implement the counterfactual simulations of U.S. wealth inequality in Section 5.2. We rely on the SCF, which provides a representative picture of the structure of the incomes, assets, and debts of U.S. households. The SCF oversamples individuals at the top of the wealth distribution, enabling a more accurate measurement of the wealth of the wealthiest individuals. The survey is updated every three years and is available between 1989 and 2022. We build the wealth distribution for every wave of the survey between 2004 and 2022 using a measure of net wealth—that is, the sum of private business wealth, public equity, real estate, interest-earning assets, and other financial and non-financial assets, minus all liabilities. To complement our analysis of HNWIs' returns on their early-stage investments based on the data from Pitchbook, we also rely on the SCF to compute their returns on both private and public equity using the methodology used by Moskowitz and Vissing-Jørgensen (2002) and Kartashova (2014).

3 Descriptive Evidence

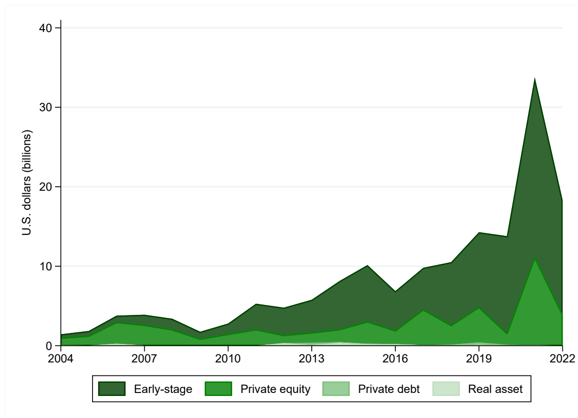
This section presents descriptive evidence on the evolution of private capital market activity in the U.S., focusing on the growing participation of HNWIs in these markets. We also provide evidence on the returns earned by HNWIs on their early-stage investments, benchmarking them with respect to public stock market returns.

Figure 1: U.S. Private Capital Market Activity

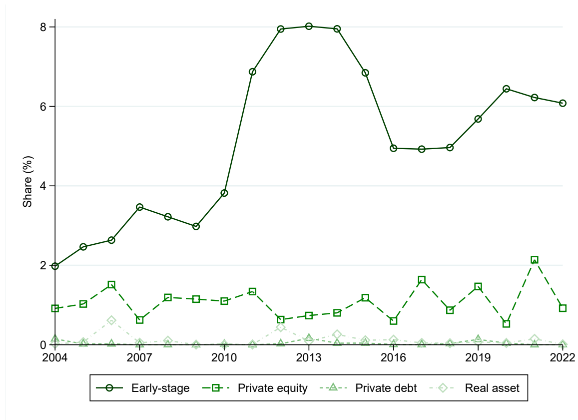
(A) Investments by All Investors



(B) Investments by U.S.-based HNWIs



(C) U.S.-based HNWIs' Shares of Investments



Source: Pitchbook.
Notes: Only investments in U.S. companies are considered. The values in Panels A and B are expressed in nominal terms. High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices. Panel B excludes 7 private equity investments by HNWIs because the individual amounts invested as part of them are outliers (in excess of $2 billion).

3.1 HNWIs' Increasing Participation in Private Capital Markets

We first describe the evolution of private capital market activity in the U.S. Figure 1, Panel A shows that the total financing raised by U.S. private companies grew from about \$190 billion in 2004 to about \$670 billion in 2007. This growth was mainly driven by private equity, private debt, and real asset financing, rather than by early-stage financing. With the onset of the 2008 global financial crisis, U.S. private capital market activity collapsed to below its initial level. However, it quickly recovered starting in 2010, alongside the broader economic recovery from the crisis. The post-crisis period was marked by the faster growth of early-stage financing, relative to that of private equity, private debt, and real asset financing. By the second half of the 2010s, U.S. startups were raising \$100-170 billion annually, compared to only \$20-45 billion during the 2000s. This trend accelerated during the COVID-19 pandemic, with the total financing raised by early-stage companies peaking in 2021, before the tightening of U.S. monetary policy in 2022 and the subsequent decline in the demand for risky assets.[10]

The rise in early-stage financing has been partly driven by the increasing participation of HNWIs in these markets. Figure 1, Panel B shows that the total amount invested annually in U.S. startups by U.S.-based HNWIs was relatively stable between 2004 and 2008. However, it grew from about \$1 billion to over \$20 billion between 2008 and 2021. In line with the patterns documented in Figure 1, Panel A for U.S. private capital markets as a whole, the total amount invested by HNWIs decreased to about \$10 billion in 2022. However, HNWIs barely increased their investments in private equity, private debt, or real asset deals over the 2008-2022 period. Early-stage investments have thus become the most important private asset class in HNWIs portfolios. The increase in HNWIs' early-stage investments has ben primarily driven by the entry of first-time investors (see Appendix Figure A3, Panel A). Despite the entry of new investors and the sharp increase in the total amount invested by HNWIs in early-stage companies, the sector composition of these investments has remained remarkably stable (see Appendix Figure A4).

The growth in HNWIs' early-stage investments has outpaced the growth of overall early-stage financing. Figure 1, Panel C shows that U.S.-based HNWIs only accounted for 2% of the total financing raised by U.S. early-stage companies in 2004. However, this share spiked to 8% during the mid-2010s, eventually settling above 6% during the early 2020s. HNWIs have therefore emerged as a new and important source of financing for

[10] U.S. companies account for about half of the total financing raised in private capital markets globally (see Figure A1, Panel B in Appendix A). U.S. private capital market activity has grown not only in absolute terms but also relative to the overall size of the U.S. economy. Appendix Figure A2 shows that private capital market investments as a share of gross domestic product have increased from about 1.5% in 2004 to 4% in 2022.

U.S. startups. Even though HNWIs account for a relatively small share of early-stage financing compared to institutional investors, private business wealth is very unequally distributed across households and accounts for a large share of the portfolio of the wealthy (see Appendix Figure A5). Thus, a drastic increase in the private business wealth of the wealthy could significantly contribute to a rise in economic inequalities in the U.S. In Sections 4 and 5, we formally explore the link between the increasing participation of HNWIs in early-stage markets and the rise in U.S. economic inequalities between 2004 and 2022. Throughout the rest of the paper, we focus on investments in early-stage companies, as this is the private asset class that has mainly driven the increasing participation of HNWIs in private capital markets.

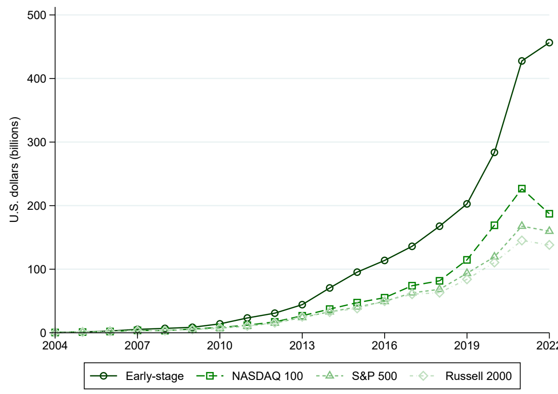
3.2 Excess Returns on Early-Stage Investments

We next present descriptive evidence on HNWIs' returns on their early-stage investments, as well as the counterfactual returns that they would have earned if they had instead invested in public stock markets. For that, we follow the methodology described in Section 2.1. Figure 2, Panel A plots the total accumulated value of U.S.-based HNWIs' investments in U.S. early-stage companies from 2004 to 2022, which equals the accumulated value of their initial investments plus their accumulated returns on these investments. It also plots the total value of HNWIs' counterfactual investments in either the NASDAQ 100, the S&P 500, or the Russell 2000. The accumulated value of HNWIs' early-stage investments by 2022 was more than double that of their counterfactual investments in any of the public indices. HNWIs thus earned excess returns on their early-stage investments, relative to the returns that were available in public stock markets.

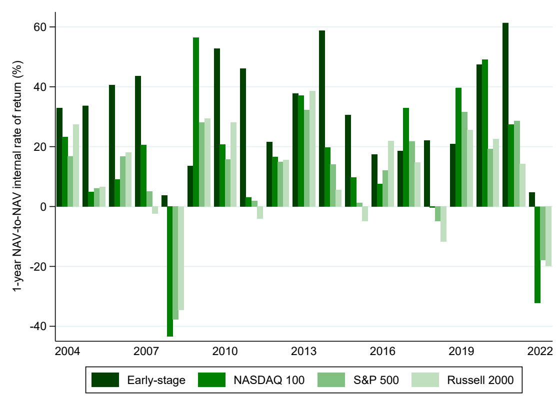
To explain the divergence between the total value of HNWIs' early-stage investments and that of their counterfactual investments, Figure 2, Panel B plots the average rate of return in every year on both sets of investments. Following Phalippou (2024), we calculate the internal rate of return (IRR) on the pooled investments at each 1-year horizon, comparing each investment's start-of-year net asset value (NAV) to its end-of-year NAV. This calculation accounts for the fact that HNWIs enter and exit investments at different points during a year, whereas a mere weighted average of investment-specific rates of return would not. In 13 out of the 19 years from 2004 to 2022, HNWIs' average rate of return on their early-stage investments exceeded the rate that they would have earned on their counterfactual investments in any of the public indices considered. Appendix Figures B1 and B2 show that this excess average rate of return on HNWIs' early-stage investments is robust to the exclusion of imputed valuations (see Section B.1), as well as to the assumption of a more pessimistic scenario where we impute bankruptcies for financially inactive companies. Furthermore, Appendix Figure B3 shows that the returns on private business equity have also outperformed those of listed equity from 2004 to 2022 using the SCF and the methodology developed by Moskowitz and Vissing-Jørgensen (2002) and Kartashova (2014).

Figure 2: Returns on Early-Stage Investments in U.S. Companies by U.S.-based HNWIs

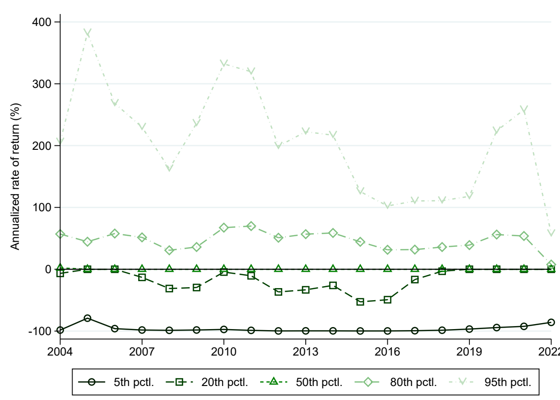
(A) Accumulated Value of Investments



(B) 1-Year NAV-to-NAV IRR



(C) Distribution of Annualized Rates of Return



Source: Pitchbook, S&P Capital IQ.
Notes: The values in Panel A are expressed in nominal terms. The rates in Panel B are based on both observed and imputed valuations, while those in Panel C are based on only observed valuations. The 1-year NAV-to-NAV IRRs in Panel B are calculated following Phalippou (2024). High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices.

These excess average returns mask substantial heterogeneity across HNWIs' early-stage investments. Figure 2, Panel C plots the 5th, 20th, 50th, 80th, and 95th percentiles of the distribution of annualized rates of return across HNWIs investments from 2004 to 2022. Since an investor can exit or enter an investment during the middle of a year, we annualize the rate of return on each investment before calculating the distribution. Whereas the investment at the 5th percentile consistently lost almost all of its value, and the median investment earned zero return, the investment at the 95th percentile more than doubled—if not tripled or even quadrupled—in value. Our finding that the median early-stage investment by HNWIs earned zero returns is consistent with existing evidence on the distribution of early-stage returns in general (Karlsen et al., 2023; Stanley and Øvrum, 2023). Furthermore, Appendix Figure B4 shows that a similar pattern of heterogeneity holds if we also consider imputed valuations rather than only observed valuations. The distribution of returns was therefore highly right-skewed: though half of these investments did not yield positive returns, the minority of investments that did so resulted in outsized capital gains. In Section 5, we formally explore whether these excess returns could be behind the recent rise of economic inequalities in the U.S. and quantify its importance.

4 Effects of Qualified Small Business Stock Reforms

This section analyzes how of much of the increasing participation of HNWIs in early-stage markets is attributable to the expansions of the U.S. federal capital gains tax exclusion on qualified small business stock (QSBS), which incentivized HNWIs to invest in early-stage companies in the aftermath of the 2008 financial crisis. We exploit these reforms—which represent a quasi-exogenous shock to HNWIs' participation in private capital markets—as a first step in our empirical analysis, before turning in Section 5 to our ultimate goal, which is to study how HNWIs' early-stage investments shaped the dynamics of inequality in the U.S. We start this section by providing institutional details about the QSBS capital gains tax reforms and then carry out both state-level and firm-level analyses to evaluate the reforms' role in increasing the participation of HNWIs in early-stage markets.

4.1 The QSBS Capital Gains Tax Exclusion

The QSBS capital gains tax exclusion was first introduced by the U.S. federal government in 1993. Set forth in Section 1202 of the U.S. Internal Revenue Code, it is a personal income tax exclusion on the capital gains realized from the sale of QSBS. To qualify for the exclusion, an investor needs to hold the QSBS for at least five years, and the amount of gain eligible for exclusion is limited to the larger of \$10 million or 10 times the acquisition value of the stock. For a company to be categorized as a qualified small business, it needs to meet the following three requirements: (1) it must be an active business that is incorporated as a U.S. C-corporation (a type of legal entity that is taxed separately from its owners); (2) it must have had gross assets of \$50 million or less at all times before and immediately after the QSBS was issued; and (3) at least 80% of the company's assets must be actively used in a qualified trade or business.[11] Companies satisfying these requirements tend to be startups in high-growth sectors (e.g., information technology) that are attractive to early-stage investors. The exemption was thus explicitly designed to incentivize investments in such startups (Polsky and Yale, 2023).[12]

Although the QSBS capital gains tax exclusion has been in place since 1993, it was only in the aftermath of the 2008 financial crisis that it become attractive from a tax savings perspective.[13] Figure 3, Panel A shows that, from 2004 to 2008, 50% of the first $10 million in capital gains realized from the sale of QSBS were *expected* to be excludable from the federal long-term capital gains tax, while the remaining 50% of gains were expected to be taxable at a fixed 28% rate (i.e., the federal long-term capital gains tax rate at the time of the exclusion's original introduction in 1993).[14] Nevertheless, since the federal tax rate on capital gains from other long-term investments was itself 15%, the expected federal tax wedge on QSBS capital gains—measured as the difference between the federal long-term capital gains tax rate and the tax rate on QSBS capital gains—was negligible, as shown in Figure 3, Panel B. Investors therefore had little to no incentive to favor QSBS investments over other investments.

With the onset of the 2008 financial crisis and the associated contraction of credit, the U.S. federal government decided to expand the QSBS tax exclusion to help small private companies raise financing. In particular, the government temporarily expanded in 2009 the excludable share of QSBS capital gains from 50% to 75% until the end of 2010, as part of the American Recovery and Reinvestment Act (ARRA). In 2010, the QSBS exclusion rate was temporarily raised further from 75% to 100% until the end of 2011, as

---

[11] Disqualified trades and businesses are determined by the I.R.S. and include companies that: (1) perform services related to health, law, engineering, architecture, accounting, actuarial science, performing arts, consulting, athletics, finance, banking, insurance, leasing, investing, or brokerage; (2) rely on an employee or owner's reputation (i.e., if it endorses products or services, uses an individual's image, or has an employee make appearances at events or on media outlets); (3) produce products, such as fossil fuels, for which percentage depletion (a type of tax deduction) can be claimed; (4) operate a hotel, motel, restaurant, or similar business; or (5) are farming businesses.

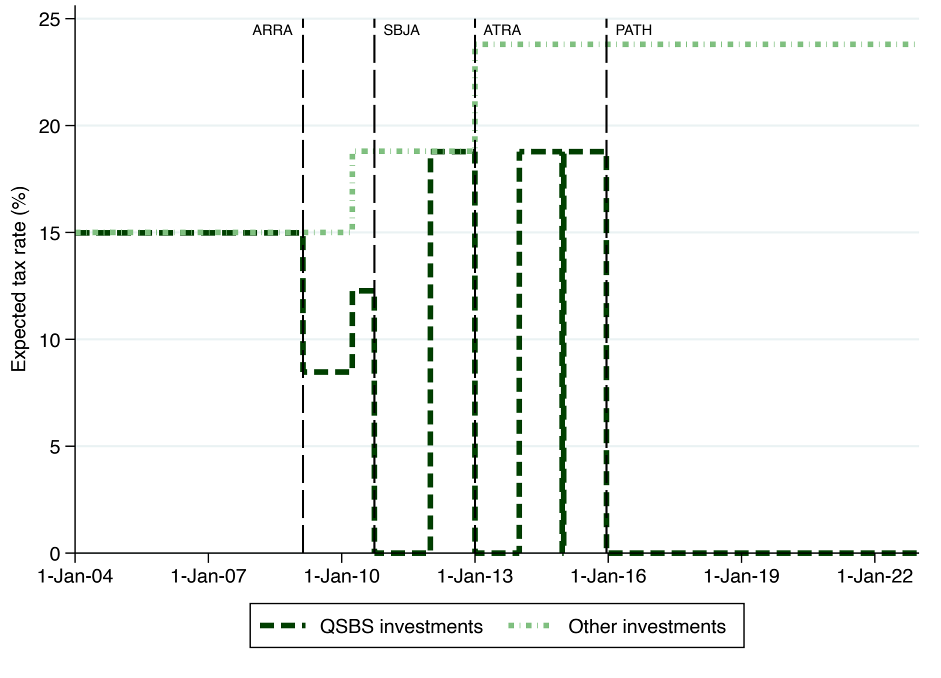
[12] The QSBS exclusion also varies across states due to differences in their state-level long-term capital tax rates or in their decisions to adopt the federal tax rules about the QSBS exclusion for their own state-level tax rules. States either fully conform with federal tax law and apply the same exclusion rate, partially conform and apply a different exclusion rate to the federal rate, or do not conform at all and fully tax QSBS capital gains at the state level. For example, California does not apply the federal exclusion to their own state-level tax rules.

[13] Figure D1 in Appendix D plots the history of the federal tax wedge on QSBS investments all the way back to its original introduction in 1993. The figure shows that cuts to the federal long-term capital gains tax rate in 1997 and 2003 eroded any incentives to favor QSBS investments over other investments.

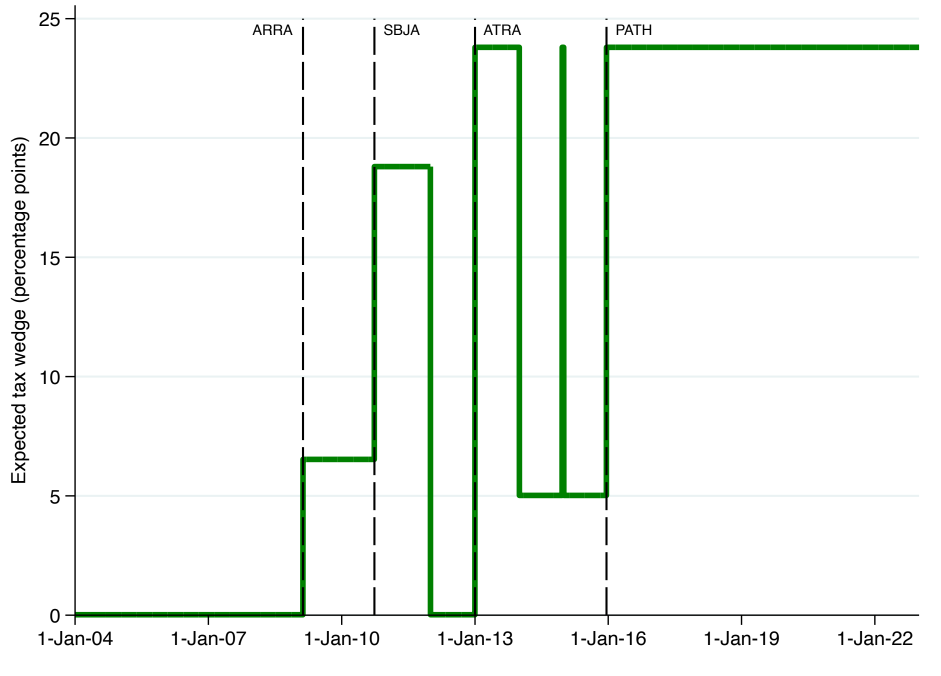
[14] We refer to the tax rates as expected rates, since benefiting from the tax exclusion requires at least a five-year gap between the investment date and the selling date, and since individuals face uncertainty about the future evolution of the long-term federal capital gains tax rate from the time of their investment decision. Thus, the tax wedge can either shrink or expand after the investment decision has been made. In contrast, the exclusion rate is determined at the time of the investment.

Figure 3: Recent History of the Federal Tax Exclusion on QSBS Capital Gains

(A) Expected Tax Rates on QSBS vs. Other Investments



(B) Expected Tax Wedge on QSBS Investments



Source: Polsky and Yale (2023).
Notes: Panel B plots the difference between the two lines in Panel A. The highlighted legislation are the American Recovery and Reinvestment Act (ARRA), the Small Business Jobs Act (SBJA), the American Tax Payer Relief Act (ATRA), and the Protecting Americans from Tax Hikes Act (PATH).

part of the Small Business Jobs Act (SBJA). These temporary expansions of the QSBS exclusion rate repeatedly expired and were retroactively extended until 2015, when the 100% exemption was made permanent as part of the Protecting Americans from Tax Hikes Act (PATH). The 100% exclusion rate has thus been in place since 2010. Given that the federal long-term capital gains tax rate has ranged from 15 to 24% since 2008, the QSBS tax exclusion has made it considerably more attractive for HNWIs to invest in early-stage companies compared to publicly listed stocks or other financial assets.

4.2 Effects on HNWIs' Early-Stage Investments

4.2.1 State-Level Analysis

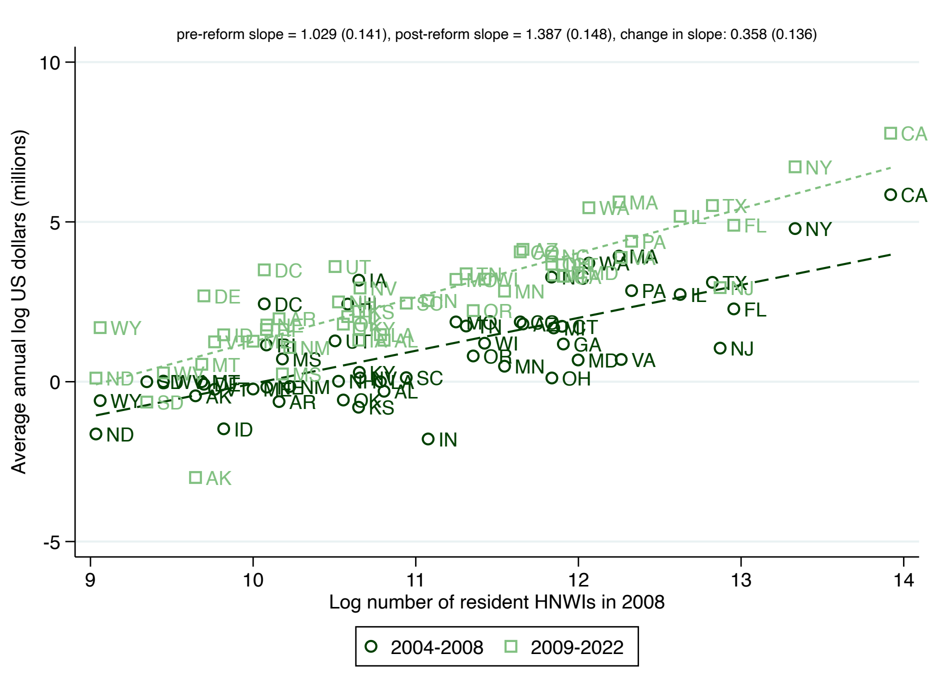
To assess the effects of the QSBS reforms on the participation of HNWIs in early-stage markets, we first carry out a state-level analysis. The reforms constitute common shocks to HNWIs residing in all U.S. states, since startups generally do not raise financing from less wealthy non-accredited investors to avoid the additional regulatory burdens of doing so (see Section 2.2). We thus rely on the number of resident accredited investors as a proxy for the number of resident HNWIs who could potentially invest in early-stage companies and on the amount invested by resident HNWIs as their actual participation. Following this intuition, the reforms should increase HNWIs' investments more in states where the ex-ante number of resident HNWIs is higher.

Figure 4, Panel A visualizes this intuition, which we later formalize in our regression analyses. For both the pre-reform (2004-2008) and post-reform (2009-2022) periods, it plots the relationship between the average annual log millions of dollars invested by resident HNWIs in U.S. startups (on the vertical axis) and the log number of resident HNWIs in 2008 (on the horizontal axis), with each pair of points representing a different U.S. state (including the District of Columbia). We find a relatively large and statistically significant increase in the slope of this relationship after the QSBS reforms, suggesting that HNWIs' participation in early-stage markets increased more in states where, there were ex-ante more resident HNWIs who could have entered these markets.

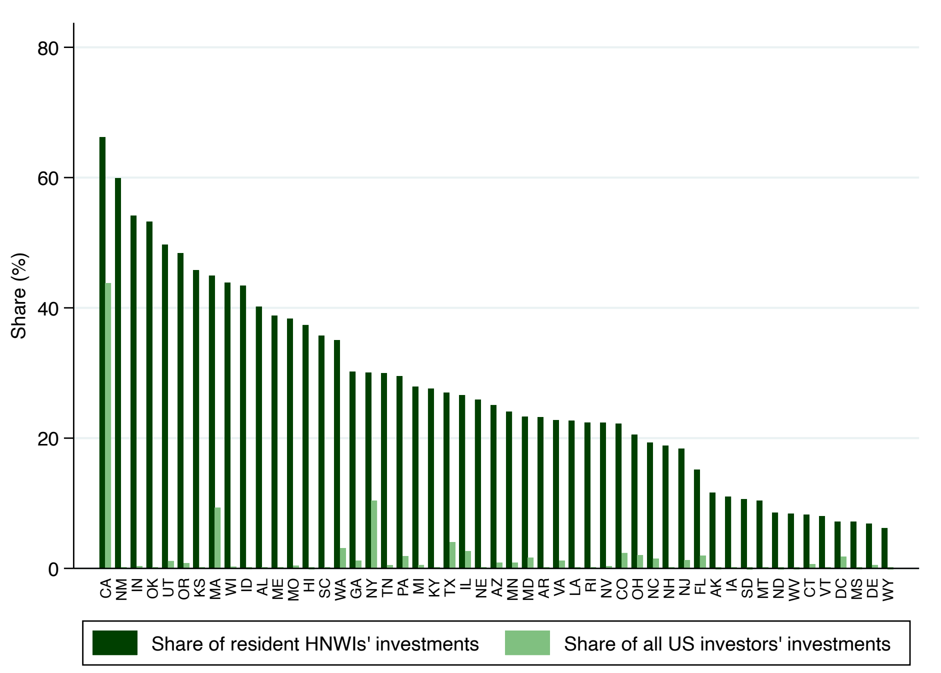
There are two main confounding factors that could also explain this increase in the slope beyond the QSBS reforms. On the one hand, HNWIs may have chosen to settle in certain states to get access to exclusive local investment opportunities (e.g., aspiring angel investors moving to California). This would especially be a threat if HNWIs exhibited home bias and thus had the tendency to invest primarily in local startups. Figure 4, Panel B shows that this was indeed the case in the U.S. from 2004 to 2022: in every state, the in-state investment share of HNWIs residing in that state exceeded the share of total

Figure 4: Motivating How to Evaluate the Effect of the Reforms
on Early-Stage Investments by HNWIs

(A) Early-Stage Investments by Resident HNWIs vs.
Number of Resident HNWIs



(B) In-State Bias of Early-Stage Investments by
Resident HNWIs: 2004-2022



Source: Pitchbook, GEOWEALTH-US.
Notes: Panel A plots the relationship between the average annual log millions of dollars invested by resident HNWIs in U.S. startups (on the vertical axis) and the log number of resident HNWIs in 2008 (on the horizontal axis), with each pair of points representing a different U.S. state (including the District of Columbia). State-year observations for which the amount invested is zero are dropped. Panel B compares the share of investments by each state's resident HNWIs invested in companies headquartered within that state to the share of investments by all U.S. investors invested in companies headquartered within that same state.

early-stage financing from all U.S. investors raised by companies headquartered within that state.[15] On the other hand, given that HNWIs exhibit such home bias, if startups in states with more resident HNWIs were exposed to different economic shocks than those in states with fewer resident HNWIs, then the startup investments by HNWIs in the first set of states may have grown faster for reasons entirely unrelated to the reforms.

To overcome these threats to identification, we analyze how the QSBS reforms affected the in-state investments in early-stage companies by resident HNWIs relative to those by other types of investors—namely, resident institutional investors, non-resident institutional investors, and non-resident HNWIs. This comparison makes it possible to control for interacted state-year fixed effects. Figure 5, Panel A shows that, since 2008, in-state investments by resident HNWIs grew more than in-state investments by other investors. Moreover, the amounts invested by resident HNWIs and by other investors increasingly diverged as newly enacted legislation ensured the permanence of the 100% QSBS exclusion.

To formalize this finding, we estimate the following regression:

$$\ln Y_{i,s,t} = \beta_t(\ln X_{s,2008} \times \mathbb{1}_{i=\text{resident HNWIs}}) + \alpha_{i,s} + \gamma_{i,t} + \delta_{s,t} + \zeta_{i,t} W_{s,t} + \epsilon_{i,s,t}, \qquad (1)$$

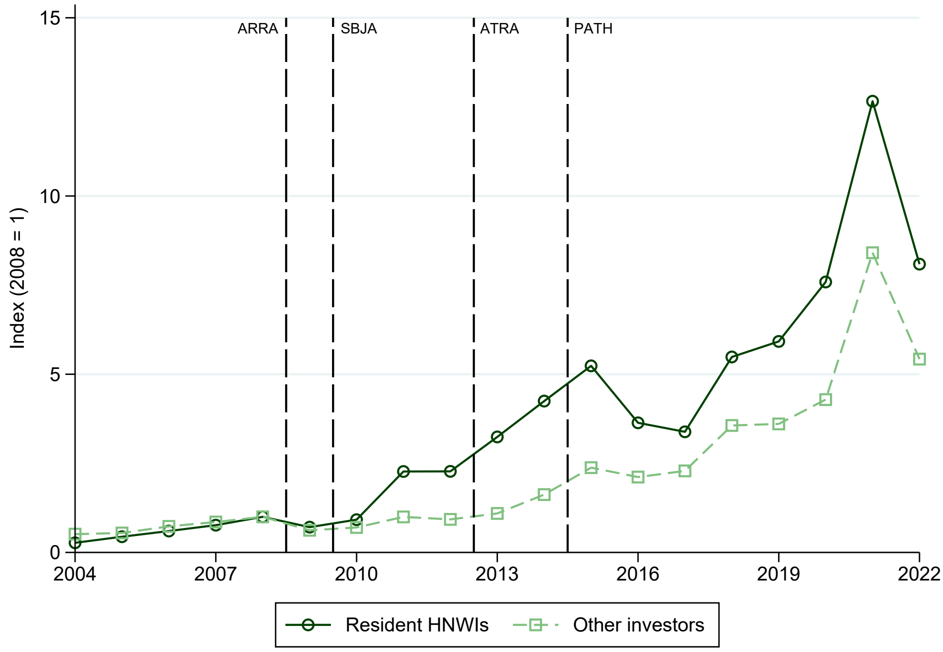
where $Y_{i,s,t}$ is the log millions of dollars invested by investors of type i in startups headquartered in state s in year t, $X_{s,t}$ stands for the log number of HNWIs residing in s in 2008, and $\mathbb{1}_{i=\text{resident HNWIs}}$ is a dummy variable equal to 1 for resident HNWIs and to 0 for resident institutions, non-resident institutions, and non-resident HNWIs. We also include investor type-state fixed effects $\alpha_{i,s}$, investor type-year fixed effects $\gamma_{i,t}$, state-year fixed effects $\delta_{s,t}$, and a vector of observable control variables $W_{s,t}$ whose effects $\zeta_{i,t}$ we allow to vary by both investor type and year.[16] If we exclude the coefficient β_{2008} when estimating Equation (1), then we can interpret the parameter of interest β_t as the change since 2008 in the elasticity of resident HNWIs' early-stage investments with respect to the number of resident HNWIs in 2008. In contrast to the change in slope plotted in Figure 4, Panel A, β_t identifies a relative change in slope, netting out the average change in slope across the other types of investors. Thus, β_t more cleanly identifies the effect of the QSBS reforms on HNWIs' participation in early-stage markets, accounting for potentially confounding shocks to local investment opportunities.

[15] Appendix Figure D2 shows that the same home bias prevails if we focus only on the pre-reform period from 2004 to 2008. Appendix Figure D3 further shows that, other than in the companies headquartered in their own state, HNWIs residing in all states also tend to invest only in companies headquartered in California (where Silicon Valley is located) and, to a lesser extent, Massachusetts.

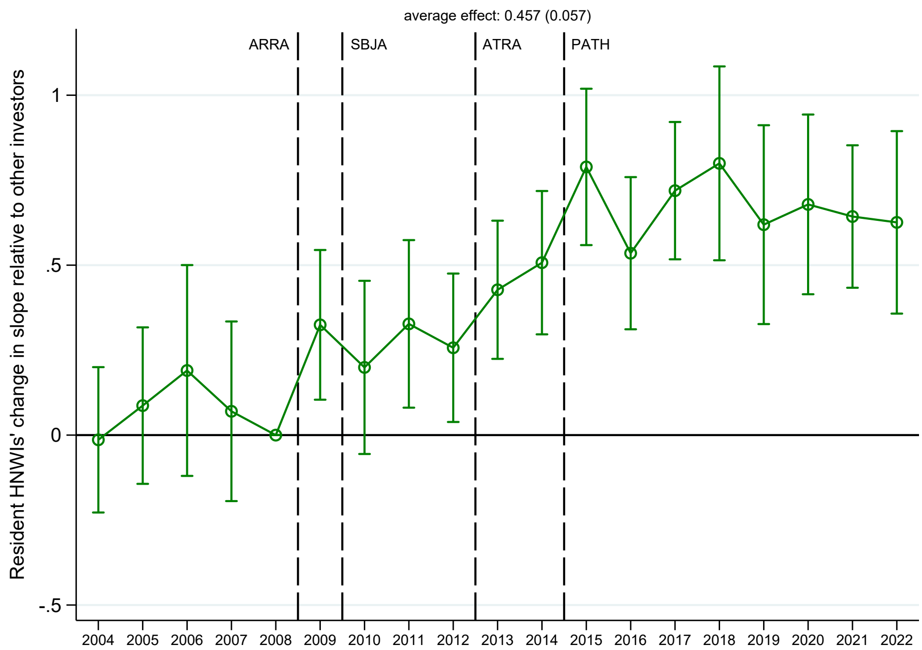
[16] The only control variable that we include in this vector is the state-level long-term capital gains tax wedge on QSBS investments for individuals residing in state s. The purpose of this control variable is to account for likely differences between resident HNWIs and other investors in the responsiveness of their investment activity to any state-specific tax reforms related to QSBS. We further assume that such state-specific reforms were exogenous to resident HNWIs' investment activity.

Figure 5: Effect of the Introduction of the Policy on Resident HNWIs' In-State Early-Stage Investments

(A) In-State Investments by Resident HNWIs vs. Other Investors



(B) Difference in Differences Estimates from Equation (1)



Source: Pitchbook, GEOWEALTH-US.
Notes: Panel A shows the evolution of in-state investments by resident HNWIs vs. other investors over time, indexed to 1 in 2008. Panel B shows the coefficient estimates β_t in Equation (1) for HNWIs, where the dependent variable measures the log millions of dollars invested. The bars represent 95% confidence intervals. Standard errors are clustered at the state level. The average effect reported in Panel B is based on a modified version of Equation (1) where β_t is replaced with $\beta_{t:t>2008}$.

Figure 5, Panel B plots our baseline estimates of β_t from Equation (1) for 2004-2022, replacing $\ln Y_{i,s,t}$ with $\ln(1 + Y_{i,s,t})$ as the outcome variable to ensure a balanced panel. The estimated coefficients for the pre-reform period are never statistically significant and exhibit no pre-trends. We find an immediate increase in resident HNWIs' early-stage investments after the initial temporary expansions of the QSBS tax exclusion in 2009-2010. We find even further increases during 2013-2014, when the 100% exclusion was repeatedly—but still only temporarily—renewed. Finally, when the full exclusion was made permanent in 2015, our estimated effect reaches its peak, remaining around this elevated level until 2022. The dynamic effects that we estimate are therefore consistent with the actual timing of the policy's introduction.[17]

Finally, we quantify the scale of the effects of the QSBS reforms relative to the overall growth in early-stage investments by HNWIs. We start by replacing the dynamic coefficient β_t from Equation (1) with the static coefficient $\beta_{t:t>2008}$ to obtain an average effect of 0.457. With approximately 6.4 million accredited investors residing in the U.S. in 2008, this estimate implies that the QSBS reforms explain $(6.4 \times 10^6)^{0.457} \approx$ \$1,300 million = \$1.3 billion of the increase in early-stage investments by HNWIs between the average pre-reform year and the average post-reform year. Since the overall increase was \$6.4 billion (from \$0.9 billion to \$7.3 billion per yer on average), the QSBS reforms account for $1.3/6.4 \approx 20\%$ of the increase in HNWIs' participation in early-stage markets after 2008.

4.2.2 Company-Level Analysis

We next study the extent to which the expanded federal tax exclusion on QSBS capital gains increased HNWIs' investments in QSBS-eligible firms in particular. To that end, we consider the sample of U.S. companies that, as of 2008, were already in existence and

[17] Our baseline estimates are robust to the use of different regression specifications. First, we find similar results when estimating Equation (1) on unbalanced panels with $\ln Y_{i,s,t}$ as the outcome (i.e., dropping observations for which $Y_{i,s,t} = 0$), or with $Y_{i,s,t}$ but using a Poisson pseudo-maximum likelihood (PPML) estimator as suggested by Chen and Roth (2024) (see Appendix Figure D4). The average effect estimated used the PPML estimator is only 0.264 if we compare the post-reform period to the whole pre-reform period. However, this average effect becomes 0.396—close to our baseline estimate of 0.457—if we compare the post-reform period only to 2008. Second, Appendix Figure D5 further shows that the inclusion of the state-level long-term capital gains tax wedge on QSBS investments as a control variable meaningfully alters only the PPML estimates (Panel C) but not the ordinary-least-squares estimates with either $\ln(1 + Y_{i,s,t})$ (Panel A) or $\ln Y_{i,s,t}$ (Panel B) as the outcome. This suggests that state-specific reforms related to QSBS may affect the extensive margin of resident HNWIs' participation in early-stage markets (i.e., do any resident HNWIs participate in the market?) but not its intensive margin (i.e., how much do participating HNWIs actually participate?). Third, we also show that the increase in the triple-difference parameters identified in Equation (1) is driven by an increase in early-stage investments by resident HNWIs, rather than by a decrease in investments by other investors—exactly as we would have expected, given the policy's design (see Appendix Figure D6). Finally, we show that our baseline estimates are also robust to dropping California and the eight other states containing cities that are major U.S. technology hubs (see Appendix Figure D7).

that had never before been bankrupt, publicly listed, or acquired by another company. We distinguish between treated and control companies based on their QSBS eligibility. Specifically, the treated companies are those that satisfy the three necessary conditions to be a qualified small business: (1) they are legally structured as tax-paying C-corporations; (2) they operate primarily in a QSBS-eligible industry; and (3) they have raised no more than $50 million in financing as of or before 2008.[18] In contrast, the controls are other U.S. companies that were active, private, and independent as of 2008 but that failed to satisfy at least one of these three conditions.

We identify the effects of the QSBS reforms on treated companies, relative to control companies, using the following regression:

$$Y_{i,t} = \beta_t \text{QSBS}_i + \alpha_i + \gamma_{\text{corp}(i),t} + \delta_{\text{ind}(i),t} + \zeta_{\text{size}(i),t} + u_{i,t} \tag{2}$$

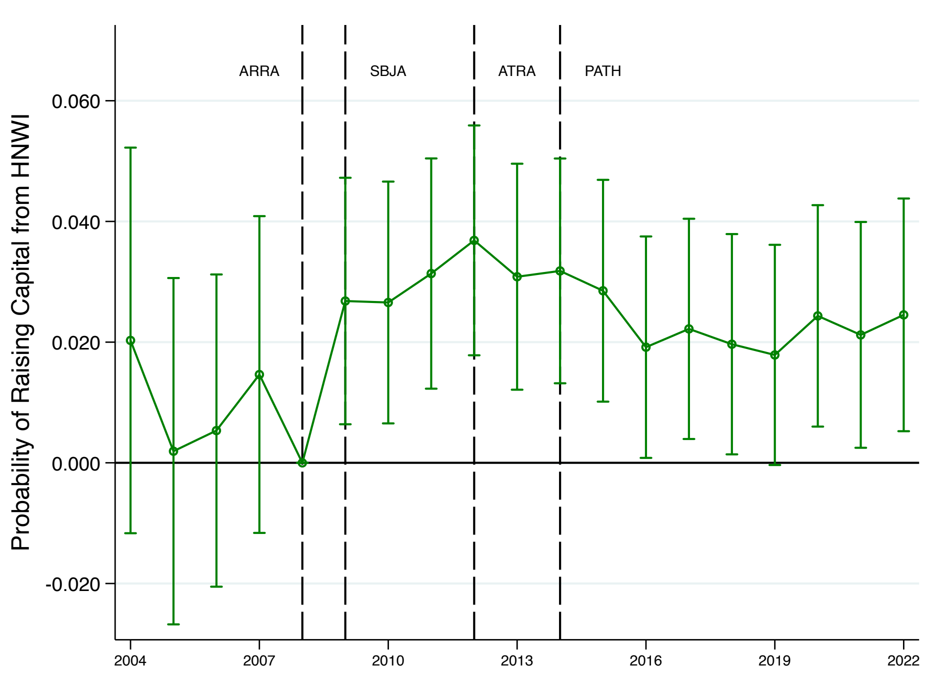
where $Y_{i,t}$ is the probability of company i raising financing in year t from at least one U.S.-based HNWI, and QSBS_i is a dummy variable equal to 1 for treated companies and 0 for control companies. We also include company fixed effects α_i, corporate structure-year fixed effects $\gamma_{\text{corp}(i),t}$, industry-year fixed effects $\delta_{\text{ind}(i),t}$, and size-year fixed effects $\zeta_{\text{size}(i),t}$, where corp(i) is a dummy variable for whether company i is a C-corporation, ind(i) is a categorical variable indicating one of 215 industries, and size(i) is a dummy variable for whether the company had raised no more than $50 million in financing as of or before 2008. If we exclude the coefficient β_{2008} when estimating Equation (2), then we can interpret the parameter of interest β_t as the percentage-point change since 2008 in the probability that QSBS-eligible U.S. companies raised money from at least one U.S.-based HNWI, relative to QSBS-ineligible U.S. companies.

Figure 6 plots our estimates of Equation (2). The estimated coefficients for the years 2004-2007 are statistically insignificant, suggesting the absence of pre-trends. Immediately after the first reform was introduced in 2009, the probability that QSBS-eligible companies raised financing from HNWIs jumped by approximately 3 percentage points, with this effect steadily declining over time to 2 percentage points. The average effect in the entire post-reform period was 2.1 percentage points, with this estimate statistically significant at the 1% level. This increase in QSBS-eligible companies' probability of raising financing from HNWIs is consistent with the QSBS reforms driving the state-level results documented in Section 4. Finally, Table D1 reports the reforms' effects on other company-level outcomes. We show that treated companies became more likely to remain private and less likely to go bankrupt, consistent with the reforms' goal to relax the financing constraints faced by

[18] We use capital raised as a proxy for gross assets—which is the measure with which the tax code determines QSBS-eligible firms—since Pitchbook only has information on gross assets for a very small sample of companies.

high-growth startups in the aftermath of the 2008 financial crisis.

Figure 6: Effect of the QSBS Reforms on QSBS-Eligible Companies' Probability of Raising Financing from HNWIs



Source: Pitchbook.
Notes: The regression is based on 336,228 company-year observations. The figure shows difference in differences coefficient estimates β_t in Equation (2). The y variable is the probability of a company raising capital from HNWIs. The bars represent 95% confidence intervals. Standard errors are clustered at the company level. High-net-worth individuals (HNWIs) refer to U.S.-based investors categorized by Pitchbook as individuals, angel groups, and family offices.

5 Excess Returns and Inequality

5.1 Effects on Income Inequality

This section studies the implications of the QSBS reforms on state-level income inequality. We rely on the taxable income distribution that we construct for every state in every year using the Statistics of Income from the IRS. In particular, we decompose the distribution into 103 income groups, with each of the first ninety-nine groups covering a percentile, while the top percentile is further split into four groups covering the 99th to 99.5th, 99.5th to 99.9th, 99.9th to 99.99th, and 99.99th to 100th percentiles. We split the top percentile because average income differs drastically across these four income groups. We then calculate the average taxable income $Y_{g,s,t}$ (in thousands of dollars) of the individuals belonging to each income group $g \in \{1, \dots, 99, 99.5, 99.9, 99.99, 100\}$ of the distribution in state s in year t. Ultimately, we run the following regression:[19]

$$\ln Y_{g,s,t} = \alpha_{g,s} + \beta_t(\ln X_{s,2008} \times \mathbb{1}_{g>99.5}) + \gamma_{g,t} + \delta_{s,t} + \zeta_{G(g),t} W_{s,t} + \epsilon_{g,s,t}, \tag{3}$$

where $X_{s,t}$ is the number of accredited investors residing in state s in 2008; $\mathbb{1}_{g>99.5}$ is a dummy variable equal to 1 for income groups $g > 99.5$ in the top 0.5% of the state-level income distribution, and 0 otherwise; $W_{s,t}$ is a vector of time-varying state-level controls whose dynamic effects $\zeta_{G(g),t}$ are the same for all income groups g within decile G of the distribution, except for the top income groups $g > 99.5$ that are assigned to their own G; and $\alpha_{g,s}$, $\gamma_{g,t}$, $\delta_{g,t}$ are interacted group-state, group-year, and state-year fixed effects, respectively. We can therefore interpret the parameter of interest β_t as the effect of the reforms on the average (log) income gap between the top 0.5% and bottom 99.5% of the state-level income distribution, which we choose to further interpret as the income gap between HNWIs and other income earners.[20]

Figure 7, Panel A plots our baseline estimates of Equation (3), replacing $\ln Y_{g,s,t}$ with the asinh$Y_{g,s,t}$ transformation to ensure a balanced panel, given that average income is negative for 1% of the observations.[21] We find that the average income gap between HNWIs and other income earners grew more after the reforms in those states with more ex-ante resident accredited investors than in those with fewer accredited investors.[22] This rise in income inequality is consistent with the findings that we presented previously—namely, that resident HNWIs' investments in local startups increased disproportionately more in those same states (Section 4.1), and that these investments yielded excess returns relative to the returns that were available in public stock markets (Section 3.2).

The increase in income inequality peaked five years after the QSBS reforms were introduced. This is consistent with the fact that investors needed to hold their QSBS investments for at least five years before selling them in order to benefit from the capital gains tax exemption (Section 4.1). To corroborate that this increase in income inequality was driven

[19] As we discuss in Section 4.1, only the first \$10 million of the long-term capital gains from each QSBS investment are exempted from the personal income tax. However, as we discuss in section 5.2, QSBS investments that yield at least \$10 million in capital gains tend to yield much more than just \$10 million. We therefore expect the majority of QSBS capital gains to be included in the measures of taxable income that we consider.

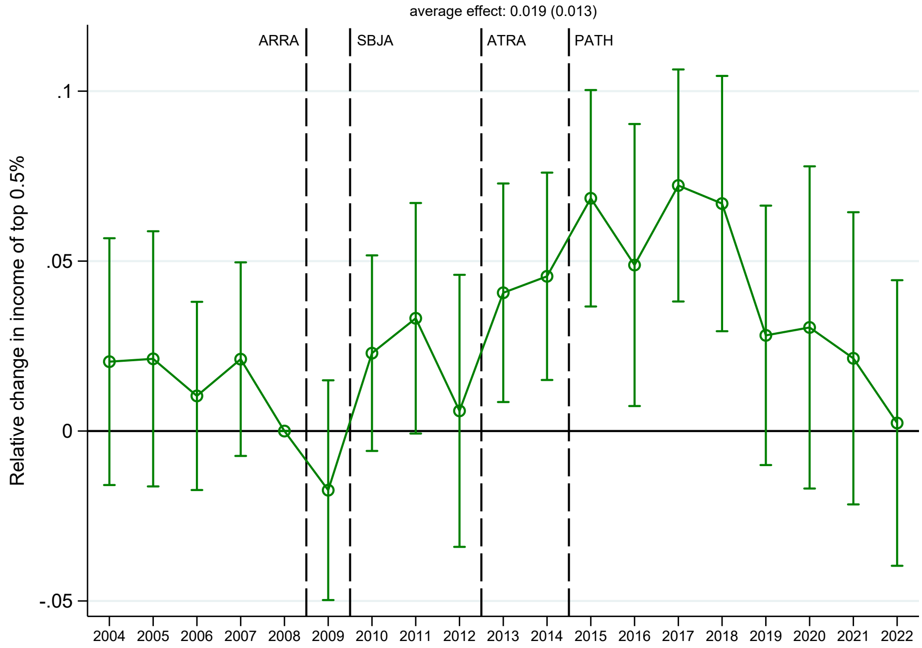
[20] Our decision to split the state-level income distribution at the 99.5th percentile is not arbitrary. It is motivated by the fact that this is the largest top income group for which in every state and year the average income of individuals exceeded \$200,000—the individual income threshold for qualifying as an accredited investor (see Section 2.2). Hence, we are certain that accredited investors belong to that income group across all states and years.

[21] Appendix Figure E1 shows that our baseline estimates using the asinh $Y_{g,s,t}$ transformation are robust to the alternative transformation $\ln Y_{r,s,t}$ based on unbalanced panels, as well as to asinh$(1000 \times Y_{r,s,t})$.

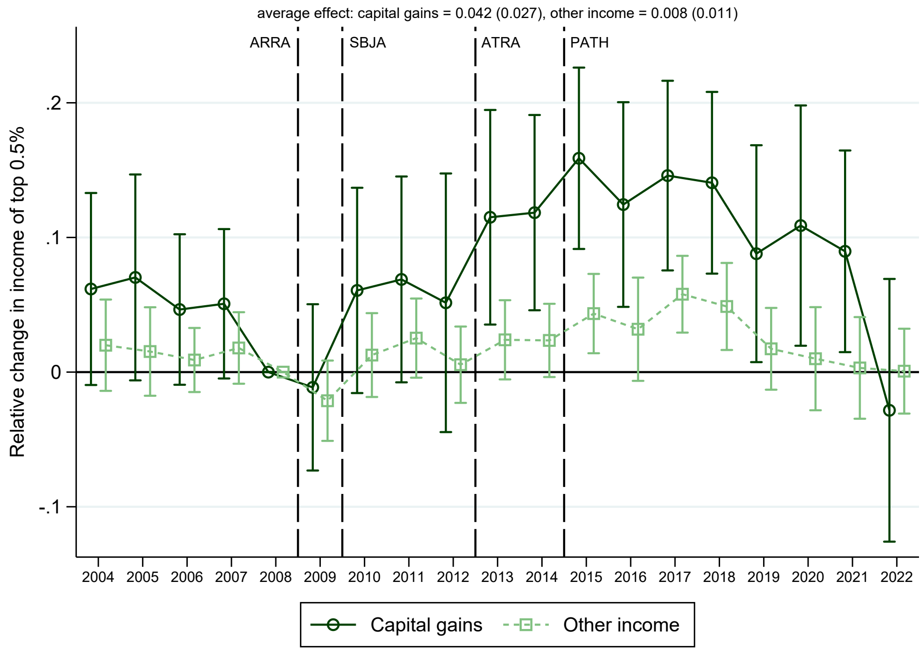
[22] We further show that this increase in income inequality was driven by an increase in the income of the top 0.5%, rather than by a decrease in that of the bottom 99.5% (see Appendix Figure E2, Panel A).

Figure 7: Income Gap between the Top 0.5% and Bottom 99.5% of the State-Level Income Distribution after the QSBS Reforms

(A) Total Income



(B) Capital Gains vs. Other Income



Source: SOI Tax Stats, GEOWEALTH-US.
Notes: This figure shows coefficient estimates β_t from Equation (3), which shows the evolution of income of the top 0.5% in a given state over time post the QSBS reforms. The regression controls for time-varying state-level controls, interacted income group-state, income group-year, and state-year fixed effects. Panel A shows total income, and Panel B shows capital gains and other income. Other income includes labor income and capital income. Bars represent 95% confidence intervals, and standard errors are clustered at the state level. The average effect reported in the figures is based on a modified version of Equation (3) where β_t is replaced with $\beta_{t:t>2008}$.

by an increase in HNWIs' capital gains, we decompose the average taxable income for each income group into capital gains and all other types of income—namely, labor and capital income. Figure 7, Panel B shows that the gap between the capital gains of the top 0.5% and the bottom 99.5% increased much faster after the reforms than the gap for the other income component.[23] This increase in the inequality of capital gains was driven by an increase in the gains of the top 0.5%, rather than by a decrease in those of the bottom 99.5% (see Appendix Figure E2, Panel B).[24]

For further reassurance, we also compare this increase in the inequality of capital gains to the growth in residents HNWIs' returns on their early-stage investments. Specifically, we replace the term $\beta_t(\ln X_{s,2008} \times \mathbb{1}_{g>99.5})$ in Equation (3) with $\beta \ln R_{s,t}$, where $R_{s,t}$ is the average returns earned in year t by HNWIs residing in state s from their accumulated investments in local startups.[25] Using the transformations asinh $Y_{g,s,t}$ and asinh $R_{s,t}$ to ensure a balanced panel, we estimate $\beta = 0.059$ (significant at the 1% level). This suggests that, for every 10% increase in HNWIs' early-stage returns, the gap between the capital gains of HNWIs and other income earners increased by about 0.6%.

Finally, we quantify by how much the QSBS reforms increased the average income gap between HNWIs and other income earners. For that, we compare our estimate of the reforms' average effect on resident HNWIs' investments in local startups (Figure 5, Panel B) to our estimate of its effect on state-level income inequality (Figure 7, Panel A). We find that the income gap between the top 0.5% and bottom 99.5% increased by $0.019/0.457 \approx 4.2\%$ for every 100% increase in HNWIs' early-stage investments. Since the reforms increased these investments by about \$1.3 billion per year relative to the pre-reform average of \$0.9 billion (see Section 4.1), they also increased this income gap by $1.3/0.9 \times 4.2\% \approx 6.0\%$.

---

[23] We distinguish between the estimated effects on labor and capital income in Appendix Figure E3.

[24] In Appendix Figures E4 and E5, we show that our baseline estimates of the effect of the policy on the capital gains of the top 0.5% are sensitive to the specific transformation of the outcome—namely, asinh $Y_{r,s,t}$, $\ln Y_{r,s,t}$, or asinh$(1000 \times Y_{r,s,t})$—that we choose to use. This is because a higher share of observations were negative for capital gains (4.9%) than for overall taxable income (1.0%), especially in the first years after the QSBS reforms. These years coincide with the immediate aftermath of the 2008 financial crisis, explaining why a higher share of the observations for capital gains were negative.

We prefer the transformation asinh $Y_{r,s,t}$ for three reasons. First, the high share of negative observations makes the $\ln Y_{r,s,t}$ transformation inappropriate. Second, when we use the asinh$(1000 \times Y_{r,s,t})$ transformation, the difference-in-difference estimates for the bottom 99.5% are implausibly negative, with these estimates driving the implausibly positive estimates of the triple-difference coefficients. Finally, the IRS Statistics of Income report the aggregate income in each range of the state-level income distribution in thousands of dollars, making it more appropriate to measure the outcome variable in this unit.

[25] We calculate these average returns as the thousands of dollars of returns earned by resident HNWIs on their accumulated investments in local startups (as measured from Pitchbook), divided by the number of residents in the top 0.5% of the state-level income distribution (as measured from the Statistics of Income).

5.2 Counterfactual Simulations of Income and Wealth Inequality

The reduced-form analyses in the previous subsection make it possible to assess the implications of the growing participation of HNWIs in private capital markets on state-level income inequality, but they do not help us understand the overall effect on U.S. income nor wealth inequality. This is the reason why in this section we run counterfactual simulations to quantify how the rise in the partipation of HNWIs in private capital markets has shaped overall U.S. income and wealth inequality between 2004 and 2019.

To carry the counterfactual simulations for income inequality, we rely on the taxable income inequality series estimated based on the Statistics of Income published by the U.S. IRS. The methodology used to build the series is based on Blanchet et al. (2022), and is explained in Section 2.2. We focus the counterfactual analysis on the top 1% income group, as this is the usual indicator used to examine the dynamics of income concentration. We run three different counterfactual simulations for the post-reform 2010-2019 period using the private and counterfactual public gains derived from Pitchbook and described in Section 3.2.[26] First, we re-estimate the taxable income inequality series excluding taxable private capital gains. We distribute the private capital gains proportionally on an annual basis so as to match the total capital gains distribution of accredited investors who are full or partial owners of a C-corporation or a partnership in the SCF. This group of investors is highly concentrated at the very top, since only the top 1% income group accounts for approximately 98% of their capital gains over the 2001-2022 period. Second, we re-estimate the taxable income inequality series replacing taxable private capital gains by the counterfactual gains had these money been invested in the S&P500. Note that because there is no exemption in the tax code for investing in public stocks, these capital gains are 100% taxable. Finally, we re-estimate the taxable income inequality series replacing the taxable private capital gains by total private capital gains, that is, the sum of taxable and non-taxable capital gains.

Table 1 compares the differences in growth rates of the top 1% taxable income share under the baseline and the different counterfactual scenarios. We have three main takeaways. First, private capital gains from early-stage investing account for 15% of overall growth in the top 1% taxable income share over 2010-2019. Second, had U.S. HNWIs instead invested in the S&P500, the top 1% taxable share would have grown by 11% less than it actually did. These two results are consistent with the fact that private capital gains are highly concentrated at the top of the income distribution and the return premium over public markets we document in Section 3.2. Third, if all QSBS capital gains had been

[26] We start the simulations in 2010, since it is the first available post-reform year for which there is a wave of the SCF available.

taxable, the top 1% taxable income share would have grown by 4% more than it actually did, as top income holders would have obtained even higher realized gains.[27]

Table 1: Growth rate in taxable income share of top 1% under various scenarios

Period	Baseline	W/o taxed priv	W/o taxed priv, W/ pub	W/o taxed priv, W/ priv
2001-2019	15.27%	14.20%	14.51%	15.58%
Baseline (=100)		93	95	102
2010-2019	6.60%	5.63%	5.90%	6.85%
Baseline (=100)		85	89	104

Source: SOI Tax Stats, Pitchbook.
Notes: This table summarizes the growth rates in the top 1% taxable income share between 2001-2019 and between 2010-2022 for the baseline series.

To carry the counterfactual simulations for wealth inequality, we rely on the wealth inequality series estimated based on the U.S. SCF provided by the Federal Reserve Board. We also focus the counterfactual analysis on the top 1% wealth group, and run two different counterfactual simulations for the post-reform 2010-2019 period using the private and counterfactual public gains derived from Pitchbook and described in Section 3.2.[28] First, we re-estimate the wealth inequality series excluding cumulated private capital gains. We distribute the cumulated private capital gains proportionally on an annual basis so as to match the total distribution of business wealth among top 1% wealth holders who are full or partial owners of a C-corporation or a partnership in the SCF. Second, we re-estimate the wealth inequality series replacing the cumulated private capital gains by the counterfactual cumulated gains had these money been invested in the S&P500.

Table 2 compares the differences in growth rates of the top 1% wealth share under the baseline and the different counterfactual scenarios. We have two main takeaways. First, private capital gains from early-stage investing account for 6% of overall growth in the top 1% wealth share over 2010-2019. Second, had U.S. HNWIs instead invested in the S&P500, the top 1% wealth share would have grown by 5% less than it actually did. These findings are consistent with the fact that private business wealth is highly concentrated at the top of the wealth distribution and the return premium over public markets we document in Section 3.2.[29]

[27] Figure E6 compares the evolution of the three different counterfactual income inequality series to the baseline scenario over the period 2010-2019.

[28] We start the simulations in 2010, since it is the first available post-reform year for which there is a wave of the SCF available.

[29] Figure E7 compares the evolution of the two different counterfactual wealth inequality series to the baseline scenario over the period 2010-2019.

Table 2: Growth rate in wealth share of top 1% under various scenarios

Period	Baseline	W/o private	W/o private, W/ public
2001-2019	15.61%	14.92%	15.08%
Baseline (=100)		96	97
2010-2019	9.27%	8.69%	8.81%
Baseline (=100)		94	95

Source: Survey of Consumer Finances, Pitchbook.
Notes: This table summarizes the growth rates in the top 1% wealth share between 2001-2019 and between 2010-2022 for the baseline series.

5.3 Feedback Loop between Private Markets and Inequality

To what extent is the relationship between HNWIs' increasing participation in private capital markets and rising inequality self-reinforcing? We briefly consider this question in this last section, where we evaluate the effect of previous entrepreneurial success on a HNWI's later activity as an investor in other early-stage companies.

Specifically, we first identify investments by HWNIs to capitalize the companies of which they were founders themselves. We then calculate the distribution of these founders' lifetime rates of return on their capitalization investments, grouping the investments into three categories: those that yielded returns in the bottom 75% of this distribution, those ranked between the 75th and 90th percentiles, and those in the top 10%. Lastly, we estimate the effect of ranking in the top 10% of this distribution (relative to ranking either in the bottom 75% or in between the 75th and 90th percentiles) on the log amount invested by these former founders as part of their later early-stage investments, controlling for year fixed effects. We find that former founders who are in the top 10% of successful entrepreneurs invest almost 60% more as part of their later early-stage investments in comparison to those ranked in the bottom 75%. Even compared to those ranked in between the 75th and 90th percentiles, the most successful former founders invest 35% more. These results suggest the existence of a feedback loop between rising private capital market activity and rising economic inequalities.[30]

6 Conclusion

This paper studies the interplay between the growth in private capital markets, the shrinking in public markets, and the rise in income and wealth concentration over the last two decades in the U.S. For that, we rely on novel data sources, and exploit an exemption

[30] We will further develop this section in a future version of this paper.

from capital gains tax for investments in early-stage companies introduced during the financial crisis as a quasi-experimental shock increasing the participation of HNWIs in private capital markets.

We obtain three main findings. First, we document that the share of financing raised by early-stage companies from U.S.-based high-net-worth individuals (HNWIs) tripled from 2004 to 2022. Second, exploiting state-level variation in exposure to the expanded federal capital gains tax exemption on qualified small business stock (QSBS), we find that HNWIs' growing participation in private capital markets increased the income gap between HNWIs and other income earners by 6.0%. Finally, using counterfactual simulations, we find that HNWIs' excess returns on these investments accounted for 11% and 5% of the growth in the top 1% share of income and wealth, respectively, from 2010 to 2019. The rise in economic inequalities further rises private capital markets activity, suggesting the existence of a feedback loop among the two.

Taken together, our paper reveals that private capital market dynamics may have non-negligible distributional implications due to the differences in portfolio composition and in returns across asset classes across the income and wealth distribution. Our analyses are based on reduced-form approaches and partial equilibrium counterfactual simulations. Further research is needed to quantify the distributional implications of changing private capital markets taking a general equilibrium approach.

References

Andersen, Asger Lau et al. (2023). "Monetary Policy and Inequality". In: *The Journal of Finance* 78.5, pp. 2945–2989.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini (2020). "Rich Pickings? Risk, Return, and Skill in Household Wealth". In: *American Economic Review* 110.9, pp. 2703–2747.

Bach, Laurent et al. (2022). "Who Becomes a Business Angel". Unpublished.

Balloch, Cynthia Mei and Julian Richers (2023). *Asset Allocation and Returns in the Portfolios of the Wealthy.* Tech. rep. DP 885. London School of Economics Financial Markets Group.

Bauluz, Luis, Filip Novokmet, and Moritz Schularick (2022). *The Anatomy of the Global Saving Glut.* Tech. rep. 9732. CESifo.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2022). "Generalized Pareto Curves: Theory and Applications". In: *Review of Income and Wealth* 68.1, pp. 263–288.

Blanchet, Thomas and Clara Martínez-Toledano (2023). "Wealth inequality dynamics in Europe and the United States: Understanding the determinants". In: *Journal of Monetary Economics* 133, pp. 25–43.

Brown, Gregory et al. (2021). "Can investors time their exposure to private equity?" In: *Journal of Financial Economics* 139.2, pp. 561–577.

Brown, Gregory W. and Steven N. Kaplan (2019). "Have Private Equity Returns Really Declined?" In: *The Journal of Private Equity* 22.4, pp. 11–18.

Calvet, Laurent E. and Adlai J. Fisher (2007). "Multifrequency news and stock returns". In: *Journal of Financial Economics* 86.1, pp. 178–212.

Campbell, John Y., Tarun Ramadorai, and Benjamin Ranish (2019). "Do the Rich Get Richer in the Stock Market? Evidence from India". In: *American Economic Review: Insights* 1.2, pp. 225–240.

Chen, Jiafeng and Jonathan Roth (2024). "Logs with Zeros? Some Problems and Solutions". In: *The Quarterly Journal of Economics* 139.2, pp. 891–936.

Chen, Jun and Joan Farre-Mensa (2024). "The Effects of the Qualified Small Business Stock Exemption on Entrepreneurship and Innovation". Unpublished.

Cioffi, Riccardo A. (2021). "Heterogeneous Risk Exposure and the Dynamics of Wealth Inequality". Unpublished.

Cumming, Douglas and Pedro Monteiro (2023). "Sovereign wealth fund investment in venture capital, private equity, and real asset funds". In: *Journal of International Business Policy* 6, pp. 330–355.

De Nardi, Mariacristina (2004). "Wealth Inequality and Intergenerational Links". In: *The Review of Economic Studies* 71.3, pp. 743–768.

De Nardi, Mariacristina and Giulio Fella (2017). "Saving and wealth inequality". In: *Review of Economic Dynamics* 26, pp. 280–300.

Denes, Matthew et al. (2023). "Investor Tax Credits and Entrepreneurship: Evidence from U.S. States". In: *Journal of Finance* 78.5, pp. 2621–2671.

Deuflhard, Florian, Dimitris Georgarakos, and Roman Inderst (2019). "Financial Literacy and Savings Account Returns". In: *Journal of the European Economic Association* 17.1, pp. 131–164.

Edwards, Alexander and Maximilian Todtenhaupt (2020). "Capital gains taxation and funding for start-ups". In: *Journal of Financial Economics* 138.2, pp. 549–571.

Ewens, Michael and Joan Farre-Mensa (2022). "Private or Public Equity? The Evolving Entrepreneurial Finance Landscape". In: *Annual Review of Financial Economics* 14, pp. 271–293.

Fagereng, Andreas et al. (2020). "Heterogeneity and Persistence in Returns to Wealth". In: *Econometrica* 88.1, pp. 115–170.

Feiveson, Laura and John Sabelhaus (2018). *How Does Intergenerational Wealth Transmission Affect Wealth Concentration?* Tech. rep. Board of Governors of the Federal Reserve System.

Gomez, Matthieu (2017). "Asset prices and wealth inequality". Unpublished.

Gomez, Matthieu and Émilien Gouin-Bonenfant (2024). "Wealth Inequality in a Low Rate Environment". In: *Econometrica* 92.1, pp. 201–246.

Greenwald, Daniel L. et al. (2021). *Financial and Total Wealth Inequality with Declining Interest Rates.* Tech. rep. 28613. National Bureau of Economic Research.

Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr (2021). "Sources of US Wealth Inequality: Past, Present, and Future". In: *NBER Macroeconomics Annual* 35.1, pp. 391–455.

Jensen, Marlin R.H., Beverly B. Marshall, and Jr. Jahera John S. (2017). "Can Non-Accredited Investors Find and Invest in the Next Unicorn?" In: *Alternative Investment Analyst Review* 6.1, pp. 76–85.

Jones, Charles I. (2015). "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality". In: *Journal of Economic Perspectives* 29.1, pp. 29–46.

Kacperczyk, Marcin, Jaromir Nosal, and Luminita Stevens (2019). "Investor sophistication and capital income inequality". In: *Journal of Monetary Economics* 107, pp. 18–31.

Karlsen, Johan et al. (2023). "The Better Angels of our Nature?" Unpublished.

Kartashova, Katya (2014). "Private Equity Premium Puzzle Revisited". In: *American Economic Review* 104.10, pp. 3297–3334.

Kopczuk, Wojciech and Eric Zwick (2020). "Business Incomes at the Top". In: *Journal of Economic Perspectives* 34.4, pp. 27–51.

Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins (2017). "Income and Wealth Inequality in America, 1949–2016". In: *Journal of Political Economy* 128.9, pp. 3469–3519.

Lerner, Josh and Ramana Nanda (2020). “Venture Capital’s Role in Financing Innovation: What We Know and How Much We Still Need to Learn”. In: *Journal of Economic Perspectives* 34 (3), pp. 237–261.

Lerner, Josh and Antoinette Schoar (2004). “The illiquidity puzzle: theory and evidence from private equity”. In: *Journal of Financial Economics* 72 (1), pp. 3–40.

Lerner, Josh, Antoinette Schoar, and Wan Wongsunwai (2007). “Smart Institutions, Foolish Choices: The Limited Partner Performance Puzzle”. In: *The Journal of Finance* 62.2, pp. 731–764.

Lindsey, Laura A. and Luke S.D. Stein (2020). “Angels, Entrepreneurship, and Employment Dynamics: Evidence from Investor Accreditation Rules”. Working paper.

Lucas, Robert E. (1978). “On the Size Distribution of Business Firms”. In: *The Bell Journal of Economics* 9.2, pp. 508–523.

Martínez-Toledano, Clara (2023). “House Price Cycles, Wealth Inequality and Portfolio Reshuffling”. Unpublished.

Maurin, Vincent, David T. Robinson, and Per Strömberg (2022). “A Theory of Liquidity in Private Equity”. In: *Management Science* 69.10, pp. 5740–5771.

Meeuwis, Maarten (2022). “Wealth Fluctuations and Risk Preferences: Evidence from U.S. Investor Portfolios”. Unpublished.

Mian, Atif R., Ludwig Straub, and Amir Sufi (2021). *The Saving Glut of the Rich*. Tech. rep. 26941. National Bureau of Economic Research.

Mikhail, Diana (2022). “The Composition of Limited Partners in Private Equity Funds”. PhD thesis. Carnegie Mellon University.

Mittal, Vrinda (2024). “Desperate Capital Breeds Productivity Loss: Evidence from Public Pension Investments in Private Equity”. Unpublished.

Moskowitz, Tobias J and Annette Vissing-Jørgensen (2002). “The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?” In: *American Economic Review* 92.4, pp. 745–778.

Nekoei, Arash and David Seim (2023). “How Do Inheritances Shape Wealth Inequality? Theory and Evidence from Sweden”. In: *The Review of Economic Studies* 90.1, pp. 463–498.

Peress, Joel (2004). “Wealth, Information Acquisition, and Portfolio Choice”. In: *The Review of Financial Studies* 17.3, pp. 879–914.

Phalippou, Ludovic (2024). “The Tyranny of IRR”. In: *SSRN*.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). “Distributional National Accounts: Methods and Estimates for the United States”. In: *The Quarterly Journal of Economics* 133.2, pp. 553–609.

Polsky, Gregg and Ethan Yale (2023). “A Critical Evaluation of the Qualified Small Business Stock Exclusion”. In: *University of Georgia School of Law Research Paper Series*.

Robinson, David T. and Berk A. Sensoy (2013). "Do Private Equity Fund Managers Earn Their Fees? Compensation, Ownership, and Cash Flow Performance". In: *The Review of Financial Studies* 26 (11), pp. 2760–2797.

Saez, Emmanuel and Gabriel Zucman (2016). "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data". In: *The Quarterly Journal of Economics* 131.2, pp. 519–578.

— (2022). *Top Wealth in America: A Reexamination*. Tech. rep. 30396. National Bureau of Economic Research.

Sommeiller, Estelle and Mark Price (2018). *The new gilded age: Income inequality in the U.S. by state, metropolitan area, and county*. Tech. rep. Economic Policy Institute.

Sørensen, Morten (2007). "How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital". In: *The Journal of Finance* 62 (6), pp. 2725–2762.

Stanley, Edward and Matias Øvrum (2023). "Morgan Stanley Research". In: *Morgan Stanley Research*.

Stulz, René M (2020). "Public versus private equity". In: *Oxford Review of Economic Policy* 36.2, pp. 275–290.

Suss, Joel, Tom Kemeny, and Dylan S. Connor (2024). "GEOWEALTH-US: Spatial wealth inequality data for the United States, 1960–2020". In: *Scientific Data* 11.

Xavier, Inês (2021). "Wealth Inequality in the US: the Role of Heterogeneous Returns". Unpublished.

Appendix A: Pitchbook Data

In this appendix, we provide a detailed description of our data-cleaning procedure for Pitchbook's data on companies, deals, investments, and investors. In Section A.7, we also produce additional figures related to the findings in Section 3.1 of the main text.

A.1 Overall Pitchbook Data

The data we purchased from Pitchbook includes datasets on companies that have ever raised certain types of financing from private capital markets; the specific deals as part of which they raised those and other types of financing; and the investors who participated in those deals, whether directly or via private investment funds.

We merge together two versions of the data. One version (updated as of 9 August 2024) contains information on companies that have ever raised either private equity, venture capital, or pre-venture financing or that are less than two years old. The other version (updated as of 23 November 2023) also contains information on companies that have only ever either raised debt financing, been acquired by another company, or been publicly traded.[31] When a company appears in both versions of the data, we consider their information only from the most recently updated version.[32]

A.2 Companies Data

We observe 937,903 companies, 62.9% of which appear in the more recently updated version of the data. For 85.4% of all companies, we observe the year in which they were founded; for the remaining companies, we will need to impute their missing year founded based on their deal history (see Section A.3). We also observe that 40.0% of all companies are headquartered in the United States, further observing the state (including the District of Columbia) in which they are headquartered for all but 1.1% of these U.S. companies.

To identify U.S. companies that are C-corporations, we parse their legal name; we observe this for 78.7% of these companies, using their trade name whenever their legal name is missing. We can identify limited partnerships ("LP"), limited liability partnerships

[31] Pitchbook also tracks companies that are two years old or older but that have never raised any of these types of financing. However, these companies are missing from our versions of the data.

[32] Our access to each version of the data is based on a separate license. The first license allows us to update the data at regular intervals but contains information on only a subset of the companies that have ever raised financing from private capital markets. In contrast, the second license does not allow for updates but contains information on the whole universe of such companies.

("LLP"), and limited liability limited partnerships ("LLLP"), none of which can be taxed as C-corporations. Although we can also identify limited companies ("LC" or "Ltd"), limited liability companies ("LLC"), professional limited liability companies ("PLLC"), and professional corporations ("PC"), these can be—but are not necessarily—taxed as C-corporations. We therefore classify only—and all—of the remaining corporations ("Corp" or "Inc") as C-corporations.[33] By this classification, 49.7% of the U.S. companies in the data are C-corporations.

To next identify those U.S. companies that are active primarily in qualified trades or businesses, we consider each company's primary industry code.[34] We classify a company as active primarily in a disqualified trade or business if its primary industry code is related to either accounting services, consulting services, engineering services, financial services, healthcare services, legal services, athletics, hospitality, performing arts, agriculture, or natural resources.[35] By this classification, 68.7% of the U.S. companies in the data are active primarily in qualified trades or businesses.

Finally, as an initial attempt to identify when—if at all—the aggregate gross assets of each U.S. company first exceeded $50 million, we consider its financial history. We observe 970,345 quarterly financial statements for 82,443 U.S. companies, which represent 22.0% of all the U.S. companies in the data. We calculate gross assets as the sum of "cash and cash equivalents" and "net property, plant, and equipment," observing both balance sheet items in 36.0% of these financial statements.[36]

For only 4,348 U.S. companies do we either observe their gross assets in 2008 or observe their gross assets exceeding $50 million before 2008; for 29.3% of these companies, gross assets never exceeded $50 million in or before 2008. Given this small number of U.S. companies whose gross assets history we observe before the expansions of the QSBS tax exclusion in 2009-2010, we will need to consider a proxy for gross assets that can be constructed for a larger number of companies (see Section A.3).

---

[33] Companies seeking financing in capital markets are unlikely to be taxed as S-corporations, since these can have at most 100 shareholders (Polsky and Yale, 2023).

[34] This is missing for only 0.1% of U.S. companies. In this rare case, we assume that a company missing its primary industry code is active primarily in a disqualified trade or business.

[35] Section 1202 of the Internal Revenue Code requires that at least 80% of the company's assets be used in the active conduct of one or more qualified trades or businesses (Polsky and Yale, 2023). However, since we cannot observe how much of its assets a company actually uses in each trade or business in which it is active, we consider only its primary industry code.

[36] In addition to cash, gross assets include "the fair market value of property contributed to the corporation measured at the time of the contribution" and "the adjusted basis of property other than contributed property" (Polsky and Yale, 2023). While we expect the "adjusted basis" component to be captured in the "net property, plant, and equipment" measure that we use when calculating gross assets, it is unlikely that this measure accurately captures the "fair market value" component.

Given our classifications of C-corporations and of qualified trades or businesses, as well as our calculation of gross assets, we can classify 18.9% of U.S. companies as eligible issuers of QSBS as of the end of 2008, before the first expansion of the QSBS tax exclusion in 2009. However, the companies that we can classify as either QSBS-eligible or QSBS-ineligible are limited to only those same 4,348 whose pre-2009 gross assets history we observe.

A.3 Deals Data

We now consider each company's deal history. For all but 1.0% of the 937,903 companies in the data, we observe at least one deal as part of which they raised financing from capital markets. Among those companies for which we observe at least one deal, the median company (in terms of the number of deals) raised financing twice.

We observe 2,116,217 deals, of which we drop 2.2% that were never (or have not yet been) completed.[37] Of the remaining deals, we drop a further 8.5% whose completion date we do not observe, since all of our analyses will require exact knowledge of the timing of deals.

We distinguish between different categories of deals. First, using Pitchbook's description of each company as well as its primary industry code, we identify real asset deals that involve purchases of real estate, infrastructure, and natural resources. Then, using the deal types maintained by Pitchbook, we classify each of the remaining deals as either a debt, private equity, early-stage, or other deal. In the private equity category, we include all buyouts, growth/expansion investments, and private investments in public equity ("PIPEs"). In the early-state category, we instead include all capitalization, crowdfunding, grant, angel, seed, accelerator/incubator, early-stage, and late-stage investments. Of the 1,894,673 completed deals whose completion date we observe, 2.3% are real asset deals, 8.1% are debt deals, 14.8% are private equity deals, and 40.7% are early-stage deals.

For 50.2% of all of these deals, we observe their size—that is, the amount of financing that a company raised as part of the deal. Since this may have included both equity and debt financing, we can further decompose the deal into its separate equity and debt components. To that end, we observe the amount of equity and debt financing that the company raised for 28.2% and 8.4% of these deals, respectively.

Of the 21,436 deals for which we observe the overall size of a deal as well as both its equity and debt components, 93.9% satisfy (to the nearest thousand U.S. dollars) the identity that the size of the deal equals the sum of these two components. Thus, we can use this identity to impute either the equity or debt component of a deal whenever only one of

[37] These deals could have failed or been canceled after being announced initially. Alternatively, despite already being announced, they could still be in the process of being completed.

them is missing but the overall size of the deal is not missing. We can likewise impute the size of a deal whenever it is missing but neither the equity nor debt component of the deal is missing. In this way, we impute the equity component for 15.8% of deals, the debt component for 0.9%, and the overall size for 2.2%.[38]

For the 14.6% of companies whose year founded is missing (see Section A.2), we can impute it as the earliest year in which the company ever completed a deal.[39] In this way, we impute the year founded for 93.5% of the companies missing it.

To next calculate a proxy for each company's gross assets (see Section A.2), we calculate the total amount of financing raised by the company up to the completion date of each deal. When calculating financing raised, we consider only those deals in which we expect the new financing to have increased the amount of gross assets on its balance sheet.[40] We can calculate financing raised as of the end of 2008 for 28,600 U.S. companies, 80.1% of which had not raised financing in excess of $50 million in total. The number of companies whose pre-2009 financing history we observe is therefore more than six times as large as the number of companies whose pre-2009 gross assets history we observe.

Given our calculation of financing raised as a proxy for gross assets, as well as our classifications of C-corporations and of qualified trades or businesses, we can classify 46.9% of U.S. companies as eligible issuers of QSBS as of the end of 2008. To verify the accuracy of this proxy-based classification of QSBS-eligible companies, we can compare it to our initial classification of those companies whose gross assets we observe in the data (see Section A.2). Of the 3,424 companies classifiable according to both classifications, the proxy-based one correctly classifies 76.9% of the QSBS-eligible companies and 81.2% of the QSBS-ineligible ones. If we restrict this comparison to only the 338 companies founded between 2001 and 2008, the proxy-based classification becomes even more accurate, correctly classifying 82.9% of the QSBS-eligible companies and 93.4% of the QSBS-ineligible ones. Given its relative accuracy and broader applicability, we will consider only this proxy-based classification in all of our analyses.

Finally, we can also use the deal history of each company to identify when it first went bankrupt, was traded publicly, or was acquired by another company. This is because Pitchbook tracks bankruptcies, public offerings, and acquisitions as deals. Of the 375,339

---

[38] Whenever Pitchbook records the one non-missing component of a deal as exceeding its overall size, or whenever both components of the deal are missing, we impute the deal's missing component(s) by instead considering the equity or debt investments by specific investors as part it (see Section A.5). For example, whenever we observe neither any debt component of a non-debt deal nor any debt investments by specific investors as part of it, we set the deal's missing equity component equal to its overall size.

[39] When we do observe the year founded, the imputed year founded precedes it in less than 0.1% of cases.

[40] For example, this excludes buyouts and debt refinancings.

U.S. companies in the data, 10.5% eventually go bankrupt, 4.4% are eventually traded publicly, and 50.7% are eventually acquired. We can use the earliest date of each of these different types of events to identify which companies were still active, private, and independent as of the end of 2008. We can also use these dates to observe exactly how long each of these companies remained active, private, or independent after the expansions of the QSBS tax exclusion in 2009-2010.

A.4 Data on Participants in Capitalization Deals

We now identify the participants in a particular type of early-stage deals: the capitalization deals as part of which a company's founders, their family and friends, and other investors provide the company with its startup capital. Pitchbook's data model for deals of this type differs from its data models for other deals, for which it directly lists the investments made by specific investors (see Section A.5). For capitalization deals, we will instead need to extract the names of the participating investors from Pitchbook's description of each deal, which is missing for only 0.3% of these deals. We will then need to further map each named investor to a unique one of the investor IDs maintained by Pitchbook.

Of the 2,116,217 deals in the data (including the ones that we drop; see Section A.3), only 0.2% are capitalization deals. Nevertheless, since the investors participating in these deals at the very start of a company's life will experience outsized increases in the value of their investments if the company succeeds and grows, it will be important to our analysis of returns (see Appendix B) that we correctly identify these initial investors.

From the string describing each capitalization deal, we extract each substring of consecutively capitalized words, interpreting it as the name of an investor.[41] We then extract further information about whether the company's "founder(s)" or their "family" and "friend(s)" participated in the deal, treating both the group of unnamed founders and the group of unnamed family and friends as additional investors. The result of this procedure is a list of 7,105 investments by specific investors across 3,797 capitalization deals.

We next identify each company's founders and other board members. For 47.5% of the 937,903 companies in the data, we can identify at least one founder. Furthermore, of the 7,105 capitalization investments that we observe, we can attribute 52.1% to board members who share their name with a specific investor and 49.7% specifically to founders.

For those investors that we cannot match to board members, we attempt to match each of them by name to one of the 397,735 investors for which Pitchbook maintains an investor

[41] We first clean the text to ensure that only the names of investors are capitalized.

ID.[42] We can successfully match 236 additional investors to unique IDs, but we are left with no information about any of the 376 investors remaining unmatched.

Finally, we consider the investments by unnamed investors. We first drop the 1,250 investments by unnamed founders in deals for which we have already matched at least one named investor to a named founder. We then translate the remaining 402 investments by unnamed founders into 994 investments by the named founders of those companies.[43] We are left with 6,447 investments in capitalization deals, 71.1% of which are by founders, while 16.8% are by their unnamed family and friends.

A.5 Investments Data

We now consider investments by specific investors as part of deals of all types. Pitchbook's data model for equity investments differs from its data model for debt investments, so will we have to consider each data model separately. Moreover, a single deal can involve both equity and debt investments; in such a case, we will have to consolidate these two different types of investments, assigning the investments of each type to the corresponding equity or debt component of the deal (see Section A.3).

Of the 2,116,217 deals in the data (including the ones that we drop; see Section A.3), 68.1% involve at least one equity investment. We observe 2,705,601 equity investments, 21.7% of which are intermediated by private investment funds, while the majority are made directly by investors without the use of intermediaries.[44] Furthermore, for 17.0% of equity investments, we observe the exact amount invested by a specific investor.

For the remaining equity investments, we must impute the amount invested by each investor. To do this, we first sum the observed amounts invested across all of the equity investments in the deal. We then subtract this sum from the observed amount invested in total as part of the deal's equity component, distributing the remainder equally across all of the equity investments whose exact amount invested we do not observe.[45]

---

[42] If there is no investor ID associated with an investor's name, we instead attempt to match it to one of the 3,173,982 people for which Pitchbook maintains a person ID.

[43] We retain only the 58 investments by unnamed founders as part of the capitalization deals for those companies with no named founders.

[44] This is based on whether Pitchbooks names a fund (or multiple funds) through which an investment is made, such that we can ultimately attribute the investment to the fund's limited partners (see Section A.6). There may be cases when the named investor is a fund manager but the fund through which it may have invested is unnamed. In this case, we can attribute the investment only to the fund manager, but not to any of the limited partners of the unnamed fund. This attribution is not necessarily inaccurate, since the fund manager is the true investor in cases when it invested on its own behalf as a general partner.

[45] Whenever the summed amount exceeds the size of the deal's equity component, we first calculate

7.7% of deals also involve at least one debt investment. Across 228,321 credit facilities, we observe 554,155 debt investments, 7.6% of which are intermediated by private investment funds.[46] Moreover, we observe the exact amount lent by a specific lender for 20.5% of debt investments, and the exact amount lent in total as part of a specific credit facility for 76.7% of facilities. Whenever either of these amounts is missing, we impute it in the same way that we do the missing amount invested by a specific equity investor.

We next reconcile these equity and debt investments by specific investors with the lists of participants in capitalization deals that we previously identified (see Section A.4). This is necessary because, in rare cases, Pitchbook actually records participation in these capitalization deals as such investments. To avoid double-counting, we drop 2.0% of the 6,447 investments in capitalization deals for which we observe either an equity or debt investment by the same investor in the same deal. We then recategorize all the remaining investments in capitalization deals as equity investments, imputing each investment's share of the corresponding deal's equity component. We are left with a total of 3,266,077 investments by specific investors, 83.0% of which are equity investments. For an additional 547,030 deals, we observe no specific investors.

Whenever a deal has both equity and debt components (see Section A.3), we split the deal into two. For 38,512 deals, we observe at least one equity investment and at least one debt investment by specific investors. Additionally, we observe 2,342 debt deals with an equity component but no equity investments; 30,585 non-debt deals with a debt component but no debt investments; 1,535 debt deals with only equity investments; and 3,392 non-debt deals with only debt investments. Splitting each of these deals into two, we observe a total of 3,850,961 investments across 2,192,583 deals.[47]

Finally, we distinguish private from traditional debt. The latter includes bank loans and bonds, while the former includes all other debt investments by non-bank lenders.[48] Splitting each deal involving both types of debt into two, we are left with 2,213,522 deals.

Of the 1,988,903 completed deals whose completion date we observe, we drop 40.7% that

---

the share of the equity investments whose exact amount invested we do not observe, distributing it equally across this first set of investments. We then distribute the remaining share across the other equity investments whose exact amount invested we do observe, distributing it according to each investment's share of the summed amount invested across this second set of investments.

[46] Only 20.4% of these intermediated debt investments are made through funds for which the fund ID maintained by Pitchbook appears in our versions of the data. For the remaining intermediated debt investments, we are left with no information about the limited partners of the intermediating funds. Thus, we can attribute them only to the fund managers of these funds.

[47] We reassign all equity investments in debt deals to duplicated deals that we recategorize as private equity, and all debt investments in non-debt deals to duplicated deals that we recategorize as debt.

[48] We categorize all debt investments by lenders of an unknown type as traditional debt.

are not private capital market deals.[49] Of the 1,178,794 remaining, 65.4% are early-stage deals, 24.4% are private equity deals, 6.6% are private debt deals, and 3.7% are real asset deals. Across all of these deals, we observe 2,468,987 investments.

A.6 Investors Data

We now consider information about the specific investors who made these private capital market investments. We will first identify the limited partners of private investment funds, attributing any investments intermediated by these funds entirely to their limited partners. We will then identify all investments (whether intermediated or direct) that we can ultimately attribute to high-net-worth individuals (HNWIs).

We observe 44,198 private investment funds, 47.7% of which are headquartered in the U.S. For 48.1% of all funds, we observe at least one limited partner who has committed capital to it. Furthermore, among the 158,248 relationships that we observe between funds and their limited partners, the limited partners are HNWIs in 3.4% of cases.[50]

We next determine how much of each fund's investments can be attributed to each of its limited partners. We observe the exact amount committed by the limited partner for 47.5% of the 165,646 commitments in the data.[51] Whenever this amount is missing, we impute it in the same way that we do the missing amount invested by a specific investor as part of a deal (see Section A.5).[52]

We then attribute the 356,076 private capital market investments intermediated by funds for which we observe at least one limited partner to their limited partners. Specifically, we replace these investments by the funds' fund managers with the corresponding 4,446,362 investments by their limited partners. Given the 2,112,911 direct investments that we also observe, we are left with a total of 6,559,273 investments.[53]

Finally, we consider similar information about the specific investors who invested directly

[49] These include traditional debt deals, mergers/acquisitions, and IPOs, among other deal types.

[50] In 21.7% of these cases, the limited partners are family offices or wealth management firms that invested on the behalf of HNWIs.

[51] This exceeds the number of relationships that we observe between funds and their limited partners, since the same limited partner can commit to the same fund multiple times.

[52] To do this, we compare the sum of the observed amounts committed across the fund's limited partners to the observed size of the fund overall. In the rare cases when we do not observe the size of the fund, we impute it based on the fund's target size as set by the fund manager(s) at inception.

[53] In terms of the size of each investment, the intermediated investments tend be smaller than the direct ones. For example, the size of the median intermediated early-stage investment is about \$0.115 million, while that of the median direct early-stage investment is about \$0.583 million.

in companies. After doing so, we find that, across all direct and intermediated investments, 63.1% were made by investors headquartered or residing in the U.S.[54] Furthermore, 6.7% of all investments were made by HNWIs, while 78.7% of these investments by HNWIs were made as part of early-stage deals.[55] We lastly identify investments by companies' founders, including but beyond those that they made as part of capitalization deals (see Section A.4); overall, 1.5% of all of the early-stage investments made by HNWIs were by founders who financed their own startups.

A.7 Additional Figures on Investments

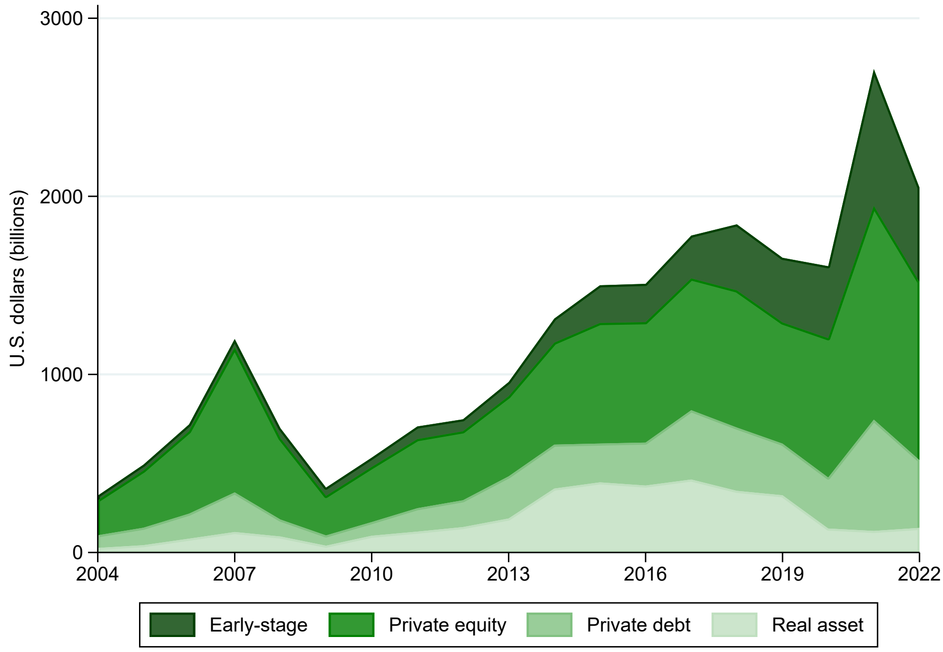
This appendix contains additional figures related to Section 3.1 of the main text.

[54] For all the founders of a company and their family and friends who participated in a capitalization deal for the company (see section A.4) but for which we have no additional information, we assume that they reside in the same location as the company's headquarters.

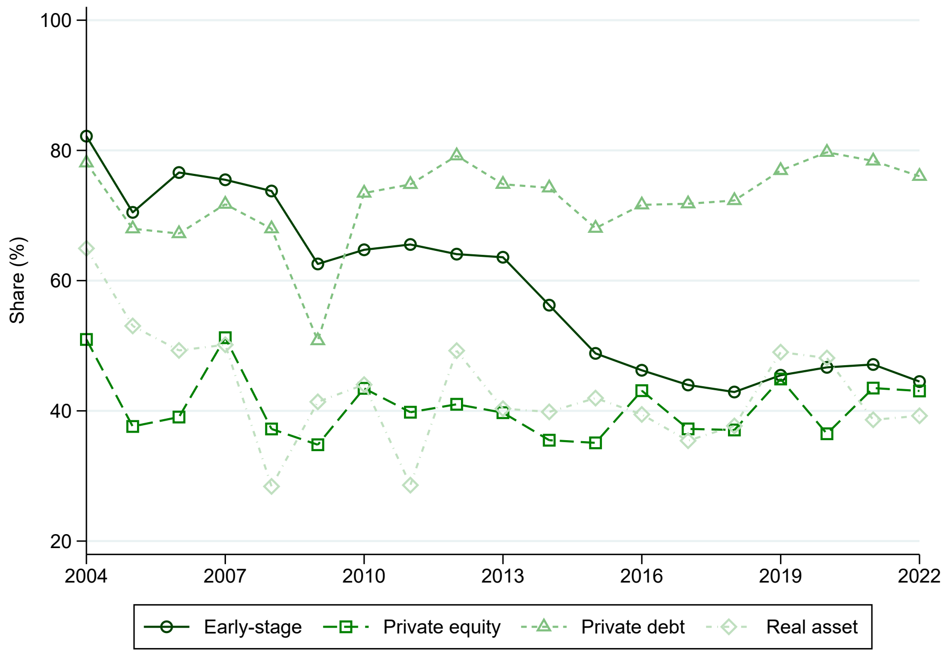
[55] In addition to HNWIs themselves, their family offices, and their wealth managers, we also include angel groups in this classification. These refer to groups of HNWIs who meet together to coordinate their angel investments in startup companies.

Figure A1: Global Private Capital Market Activity

(A) Investments in All Companies

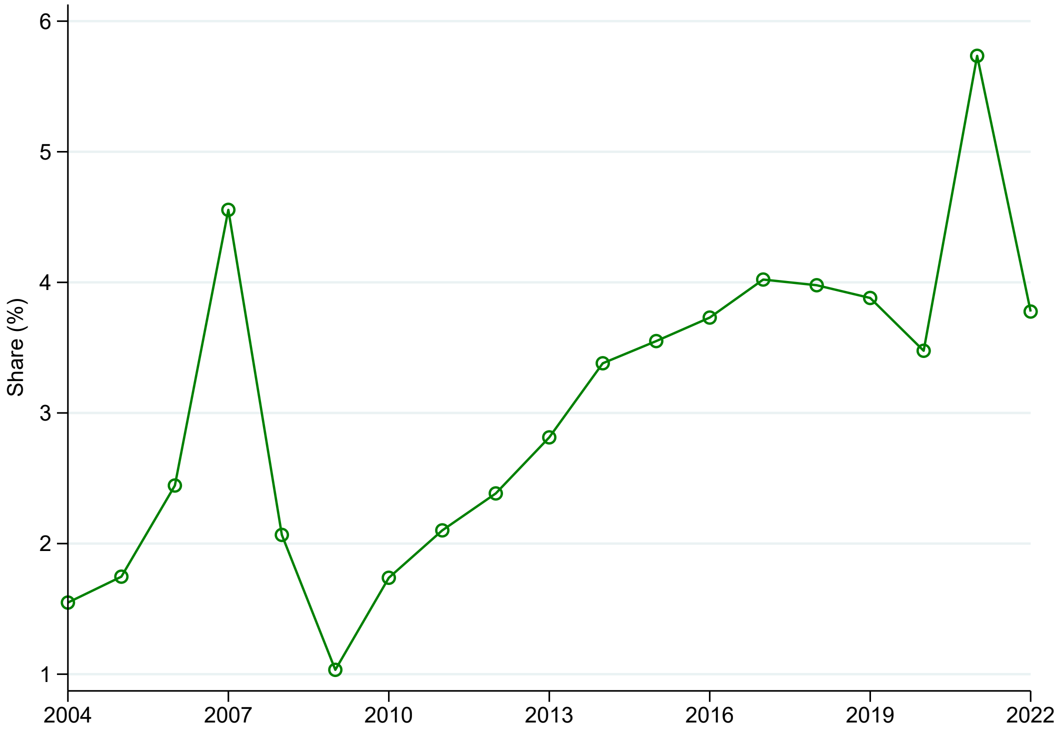


(B) U.S. Companies' Shares of Investments



Source: Pitchbook.
Notes: The values in Panel A are expressed in nominal terms.

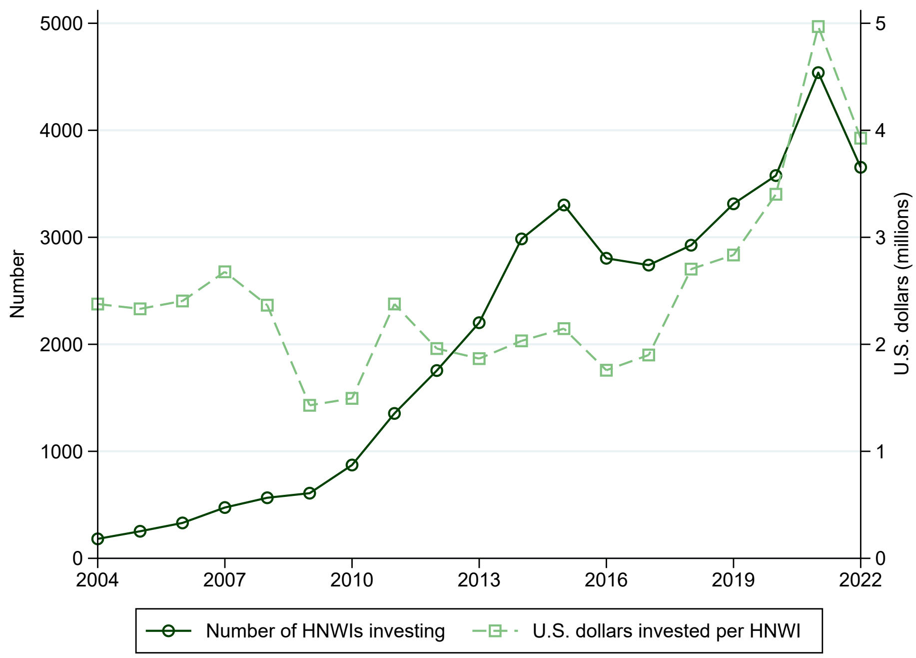
Figure A2: U.S. Private Capital Market Activity as a Share of U.S. GDP



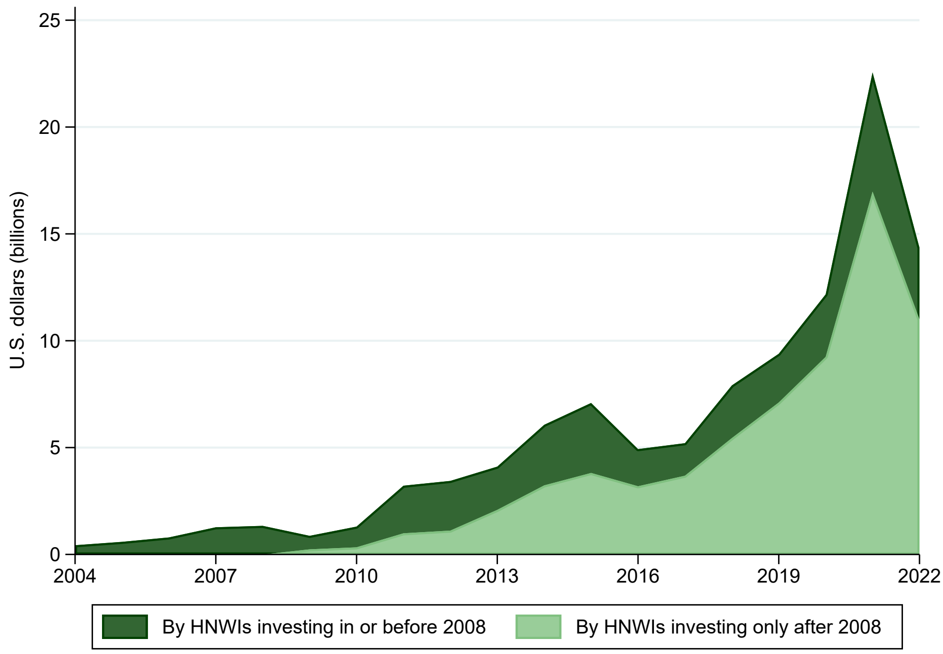
Source: Pitchbook, Bureau of Economic Analysis.
Notes: When calculating the share, private capital market investments and gross domestic product (GDP) are both measured in nominal terms.

Figure A3: HNWIs' Participation in Early-Stage Markets

(A) Number of HNWIs with Early-Stage Investments

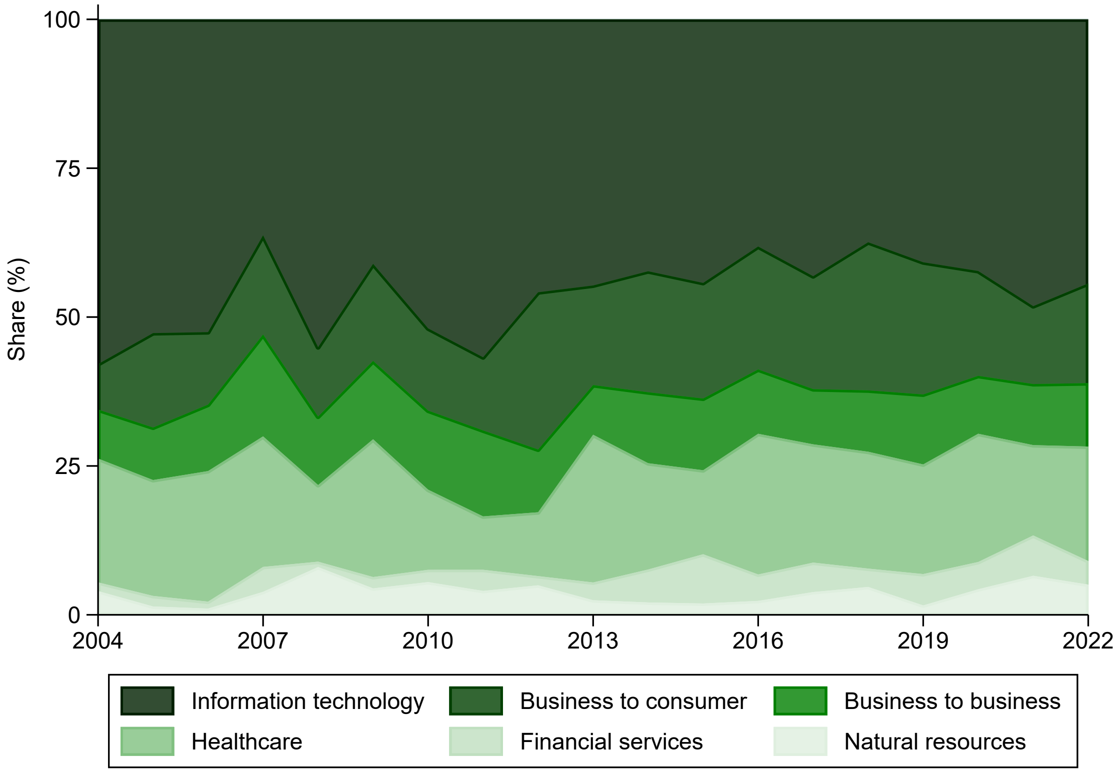


(B) Early-Stage Investments by HNWIs



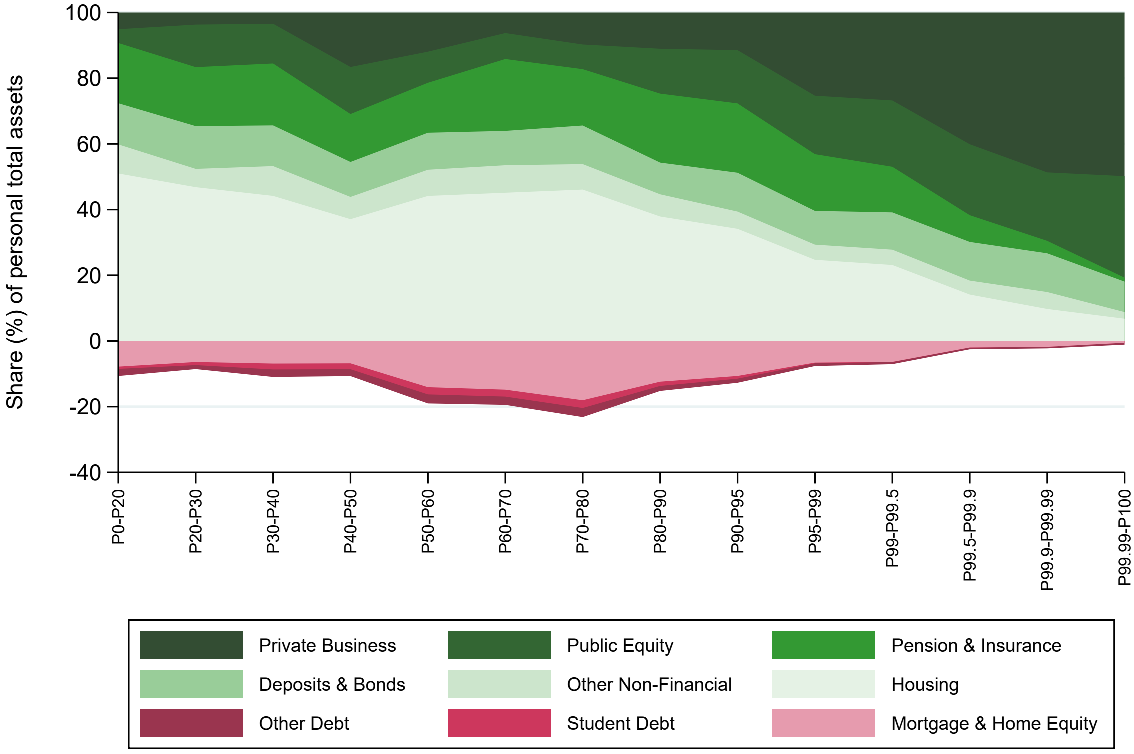
Source: Pitchbook.
Notes: Only early-stage investments in U.S. companies by U.S.-based HNWIs are considered. The values in Panel B are expressed in nominal terms. High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices.

Figure A4: Sector Composition of HNWIs' Early-Stage Investments



Source: Pitchbook.
Notes: Only early-stage investments in U.S. companies by U.S.-based HNWIs are considered. High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices. The natural resources sector encompasses Pitchbook's own distinct sector categories for energy and for materials and resources.

Figure A5: U.S. Asset Composition by Income Level in 2022



Source: SCF.
Notes: This figure plots the asset composition of wealth along the income distribution in the U.S., based on household-level information from the 2022 wave of the Survey of Consumer Finances (SCF).

Appendix B: Return Methodology

In this appendix, we provide a detailed description of our methodology to calculate HNWIs' returns on their early-stage investments, given Pitchbook's data on how the valuation of each company has changed across deals. In Section B.3, we produce additional figures related to Section 3.2 of the main text.

B.1 Valuations Data

We consider Pitchbook's data on the valuation history of each company. Whenever a company raises financing as part of a new deal, its investors and its management must agree on its new valuation. By comparing its new pre-money valuation (before accounting for the financing that it raised as part of the new deal) to its previous post-money valuation (after accounting for the financing that it raised as part of its previous deal), we will be able to calculate the annual returns that its existing investors earned on their investments (see Section B.2) in a way that accounts for any dilution of their shares between deals.

Since our analysis of returns will start from 2004, we consider the 1,774,304 deals that were completed after 2003. Of these, we have complete information about the valuation of the company raising financing—that is, about both its pre- and post-money valuations as part of a deal, as well as the size of the equity component of that deal (see Appendix Section A.3)—in only 12.0% of cases.

96.7% of deals with complete valuation information satisfy (to the nearest thousands U.S. dollars) the identity that the equity component of a deal equals the difference between the company's post- and pre-money valuations. However, for 14.6% of these deals, Pitchbook only estimates the company's valuation (rather than observing it). When constructing each company's valuation history, we will consider only the 175,381 deals whose valuation information neither violates this identity nor is estimated by Pitchbook.

For the remaining deals, we will need to impute the missing valuations of the company receiving financing. For our baseline analysis of returns (see Section B.2), the only imputation that we apply is to the 88,915 deals that identify when a company went permanently bankrupt, in which case we impute the company's valuation as \$0.01.[56]

Of the 264,296 deals (in our baseline analysis) for which we either observe or impute the valuation of the company raising financing, we drop 117 that were completed on a date

[56] In rare cases (where we do not apply this imputation), a company went bankrupt only temporarily before later raising financing and being attributed a positive valuation as part of a new deal.

on which multiple deals were completed for that same company.[57] We also drop 60,773 deals for companies for which we observe only one such deal, since we need to observe the valuation of a company at least twice (and on two different dates).[58]

For the 203,406 remaining deals that we observe across 108,572 companies, we first calculate the number of days between each consecutive pair of deals for each company, as well as the percent change between the company's previous post-money valuation and its new pre-money valuation. We then convert this percent change into a daily compounded rate.[59] We lastly use this rate to construct the history of the daily rate of return on each company from 2004 to 2024, applying a rate of zero to all dates preceding the completion date of the company's first deal and following the completion date of its last deal.

We next construct an alternative, pessimistic valuation history for each company. In particular, we assume that, in addition to the bankruptcies that we actually observe, every company that has raised no new financing since the end of 2021 went bankrupt three years after the completion date of its last deal.

Finally, to be able to compare the returns earned by HNWIs on their early-stage investments to counterfactual returns that they would have earned had they instead invested in publicly listed stocks (see B.2), we also consider a complete and more sophisticated imputation of the valuation of each company in every year since their founding (and until their eventual bankruptcy, if it ever occurs).[60] Using the valuations of companies whose valuation we observe at least twice (which thereby allows us to control for company fixed effects), we estimate a regression of their log valuation on observable company-level characteristics, interacted with year fixed effects.[61] We then use the estimated parameters from this regression to predict the change in each company's valuation over time.[62]

---

[57] We keep only the last deal (in terms of its deal number) completed on that date for the company.

[58] To ensure that we capture the losses of investors on their investments in companies that eventually go bankrupt but whose valuation we never observe, we first impute the valuation of each such company (as part of every one of its deals) as a fixed $1 million.

[59] If D is the number of days between consecutive deals for the company and R is the percent change between its previous post-money valuation and its new pre-money valuation, then the daily compounded rate is $r = \exp\left(\frac{\ln(1+R)}{D}\right) - 1$. This follows from $(1+r)^D = 1 + R \implies \ln(1+r) = \frac{\ln(1+R)}{D}$.

[60] If we observe multiple deals for a company during a given year, we impute the company's valuation as part of every deal for which its valuation is otherwise missing during that year.

[61] We include triple-interacted sector-by-stage-by-year fixed effects in this regression. We classify companies into three aggregated sectors (with each company's sector fixed over time): business-to-business and business-to-consumer products and services; information technology; and all other industries (including energy, financial services, healthcare, and materials and resources). We also classify companies into eight stages of maturity based on the number of times that they have raised financing (with each company's stage changing over time): one, two, three, four, five, six, seven, and eight times or more.

[62] For each deal as part of which we observe a company's valuation, we first calculate the difference between the company's valuation and the value predicted by the regression. We then add this difference

B.2 Calculation of Investment-Level Returns

We now calculate the returns earned by HNWIs on their early-stage investments, using the valuation history that we constructed for each company (see Section B.1). After calculating these returns, we will also calculate the counterfactual returns that they would have earned had they instead invested in publicly listed stocks.

We first consider the returns earned by U.S.-based HNWIs on their 92,004 early-stage investments in U.S. companies from 2004 to 2022.[63] For each investment, we calculate the change in its value since the date on which the HNWI to which it is attributable initially made the investment; we do this until the date on which the HNWI eventually exited the investment, which we observe for 35.5% of these investments.[64] We then calculate the return (in U.S. dollars) earned by the HNWI in each year in which they held the investment, also calculating the corresponding rate of return that it yielded.[65]

We can calculate these returns based on each of the three distinct methodologies that we used to construct the valuation history for each company: our baseline methodology that considers only valuations and bankruptcies that we can actually observe, a pessimistic alternative that assumes additional bankruptcies, and the complete imputation of every company's valuation in every year (see Section B.1). While we can calculate the annual returns on all 92,004 early-stage investments by HNWIs using the last methodology, we can do so for only 64.4% of these investments in our baseline analysis.[66]

We next consider how to aggregate returns across investments. Most simply, we can calculate the accumulated value of all investments as of the end of each year. However, we can also calculate the pooled internal rate of return across all investments.[67]

---

to the predicted value for each year in which we do not observe the company's valuation, taking the difference from the most recent deal for which we do observe its valuation.

[63] 81.7% of these investments were direct, while the rest were intermediated by funds (see Appendix Section A.6). It is for this reason that we focus on investment-level returns rather than fund-level returns.

[64] We assume that, after the HNWI exited the investment, they held their proceeds from the sale in the form of cash, such that there was no further accumulation in the investment's value. Since we make the same assumption for the exited counterfactual investments in publicly listed stocks, this assumption does not affect our analysis of the differences between HNWIs' actual and counterfactual returns.

[65] The annual rate of return of an investment will differ from its annualized rate of return if the HNWI held the investment for less than the whole year—that is, if they either entered or exited the investment during the middle of the year. Specifically, if the annual rate of return on an investment is R, the number of days that the investment was held during the year is D, and the number of days during that year is $d \in \{365, 366\}$, then the annualized rate of return is $r = \left[\exp\left(\frac{\ln(1+R)}{D}\right)\right]^d - 1$.

[66] This is because, for each of the companies that raised financing from the remaining investments, we observe less than two valuations across time (see Section B.1).

[67] We treat each investment's initial net asset value (NAV) during each year as a negative cash flow and its final NAV as a positive one. We then aggregate the cash flows for each date, apply the formula to

Finally, we can calculate the counterfactual returns that HNWIs would have earned on their early-stage investments had they instead invested in publicly listed stocks. To do this, rather than let the value of each investment in a company evolve according to its actual valuation history (see Section B.1), we can let it evolve according to the history of one of three major public stock market indices: the NASDAQ 100, S&P 500, or Russell 2000.[68] We then compare these counterfactual returns to the actual returns that HNWIs earned on their early-stage investments.

B.3 Additional Figures on Returns

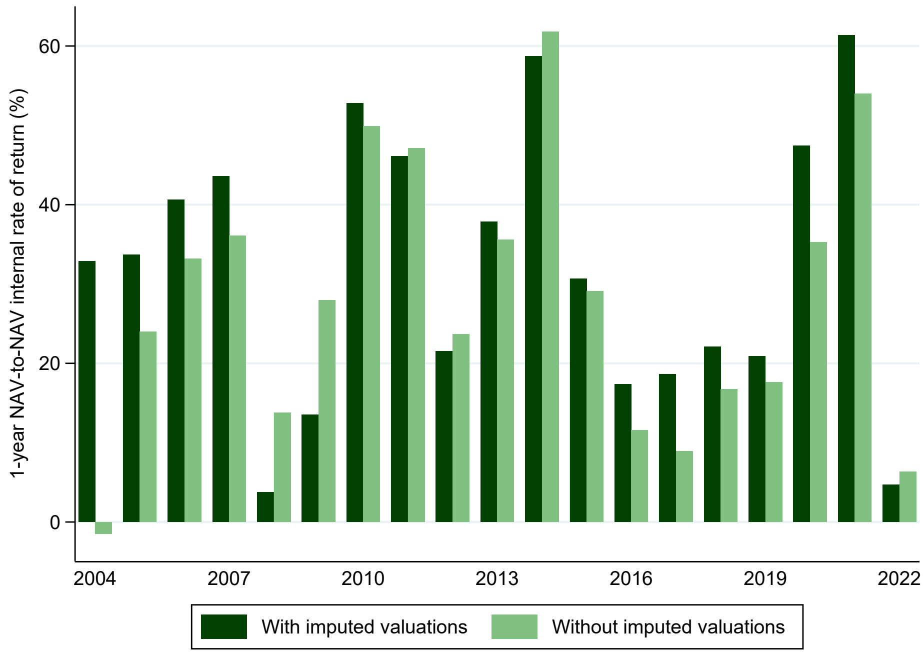
This appendix contains additional figures related to Section 3.2 of the main text.

calculate the daily internal rate of return, and convert it into an annualized rate (Phalippou, 2024).

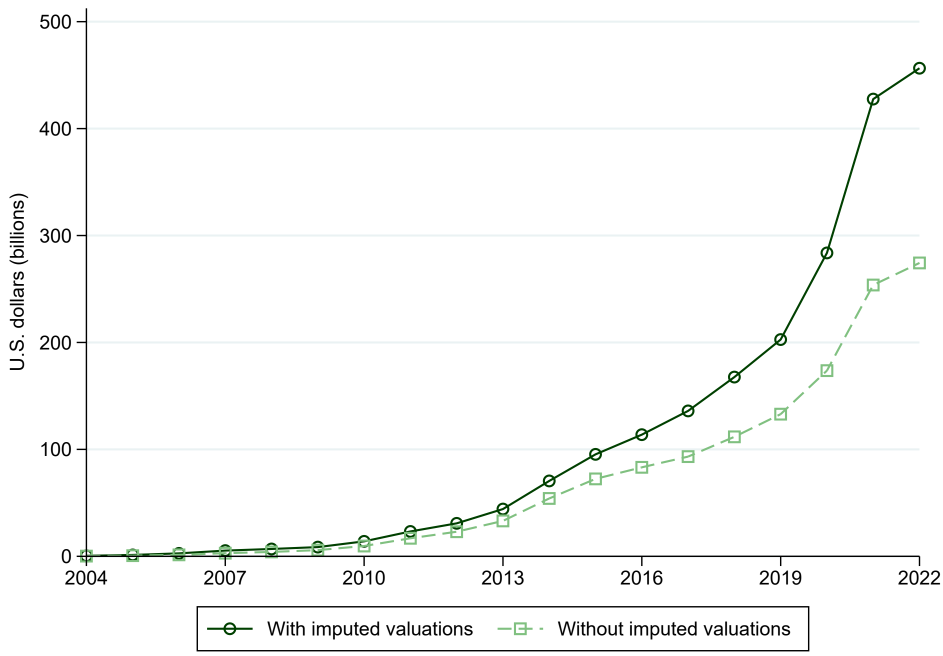
[68] We consider the Total Return versions of these indices, as recorded in data from S&P Capital IQ. We therefore assume that the dividends paid by the publicly listed companies in each index are invested immediately back into the index. In contrast, we assume that startup companies pay no dividends.

Figure B1: Returns on Early-Stage Investments in U.S. Companies by U.S.-based HWNIs: with and without Imputed Valuations

(A) 1-Year NAV-to-NAV Internal Rate of Return on Pooled Investments



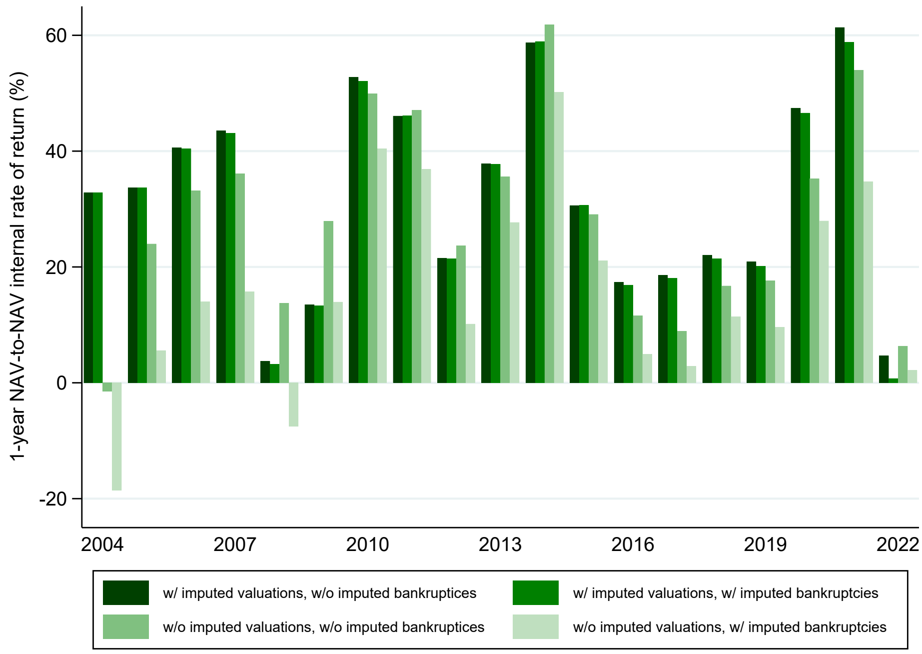
(B) Accumulated Value of Investments



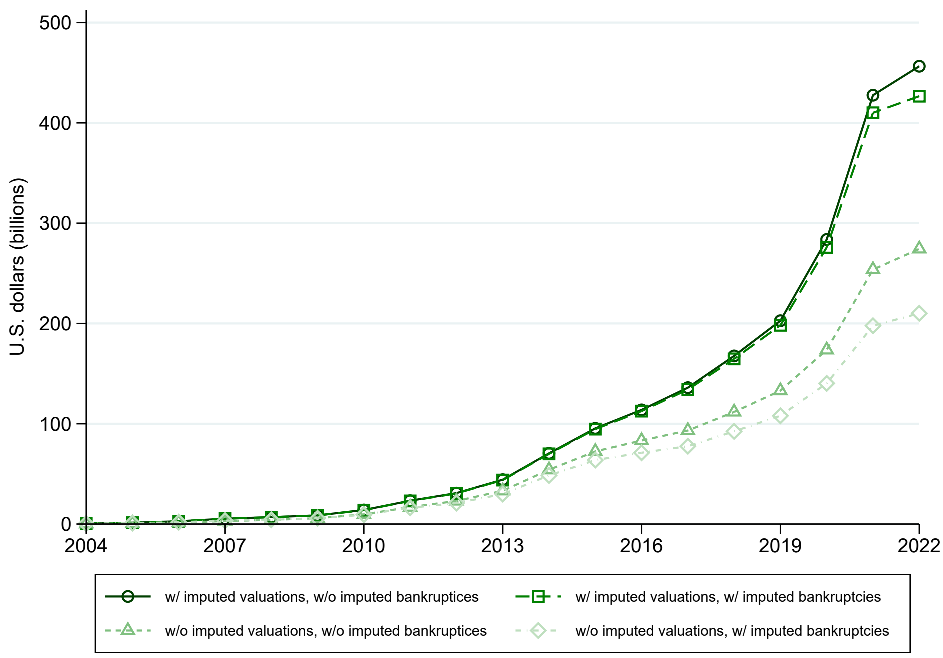
Source: Pitchbook.
Notes: The rate in Panel A is calculated as in Phalippou (2024). The values in Panel B are expressed in nominal terms. High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices.

Figure B2: Returns on Early-Stage Investments in U.S. Companies by U.S.-based HWNIs: with and without Imputed Valuations or Bankruptcies

(A) 1-Year NAV-to-NAV Internal Rate of Return on Pooled Investments

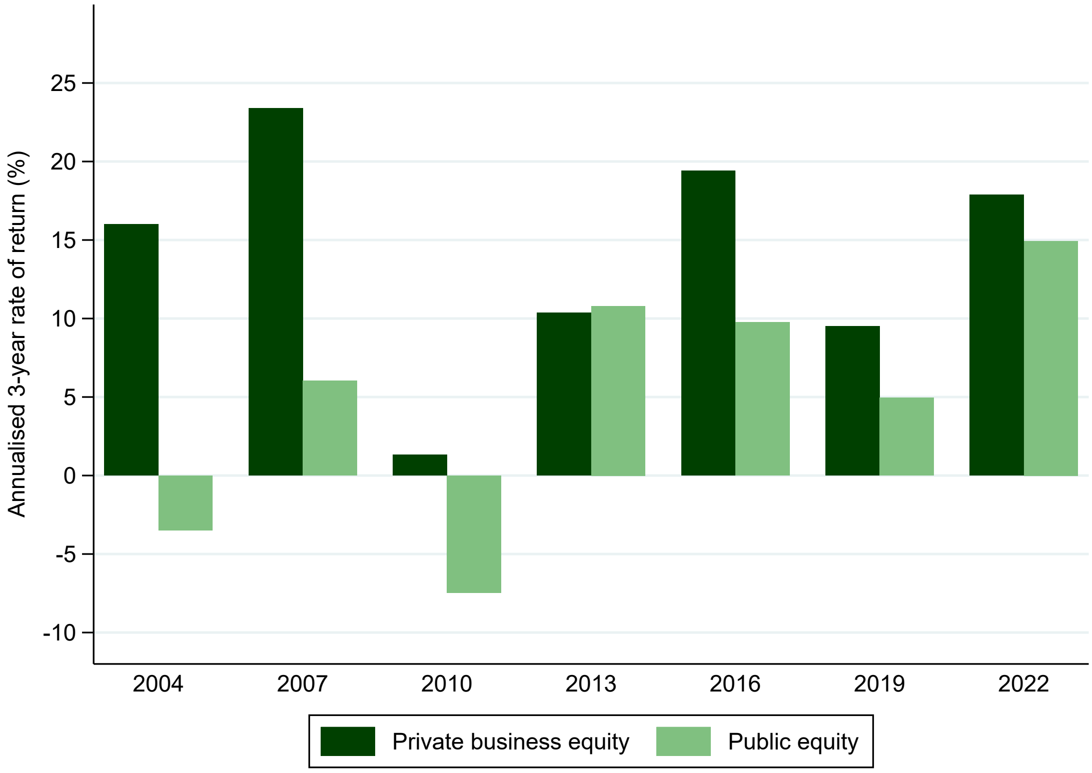


(B) Accumulated Value of Investments



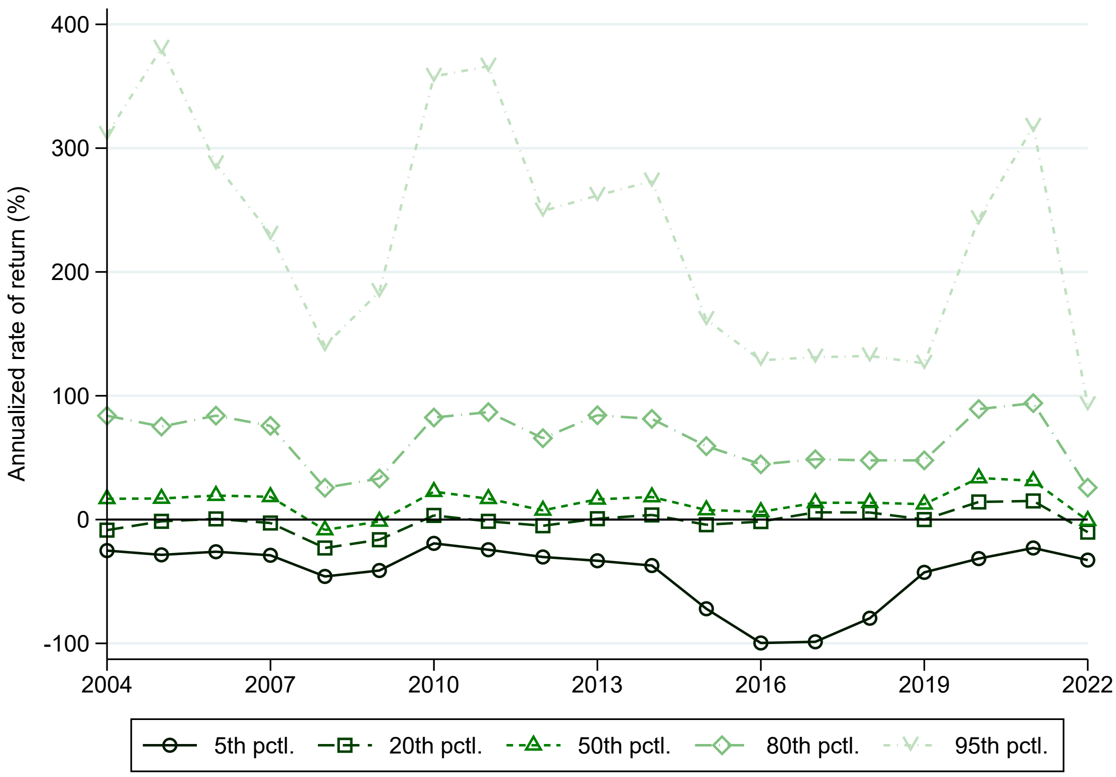
Source: Pitchbook.
Notes: The rate in Panel A is calculated as in Phalippou (2024). The values in Panel B are expressed in nominal terms. High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices.

Figure B3: Annualized 3-year Rate of Return on All Private Business Equity and All Public Equity



Source: SCF.
Notes: The annualized 3-year rate of return is estimated following the exact same methodology as in Moskowitz and Vissing-Jørgensen (2002) and Kartashova (2014).

Figure B4: Distribution of Annualized Rates of Return: with Imputed Valuations



Source: Pitchbook.
Notes: The rates are based on both observed and imputed valuations. High-net-worth individuals (HNWIs) refer to investors categorized by Pitchbook as individuals, angel groups, and family offices.

Appendix C: Data on High-Net-Worth Individuals

This appendix validates the state-level measure of the number of resident HNWIs that we use in the regression analyses of Sections 4.2.1 and 5.1 in the main text. As described in Section 2.2, we rely on the measure from the GEOWEALTH-US dataset built by Suss et al. (2024), who define HNWIs so as to resemble the SEC's legal definition of accredited investors—that is, those whose net worth (excluding the value of their primary residence) exceeds $1 million, or whose household income exceeds $300,000. We validate this baseline measure by comparing it with alternative estimates of the number of HNWIs residing in the U.S. provided by the Phoenix/MarketCast Wealth and Affluent Monitor, the Forbes 400, the Credit Suisse/UBS Global Wealth Report, and the Survey of Consumer Finances.

Figure C1, Panel A depicts the correlation between the state-level average of the log number of resident HNWIs from the GEOWEALTH-US over the period 2006-2019 and the analogous measure from the Phoenix/MarketCast Wealth and Affluent Monitor. The latter measure is based on the estimated number of individuals with $1 million or more in investable assets residing in each U.S. state, which Phoenix/MarketCast constructs by combining information from the Survey of Consumer Finances with data from Nielsen-Claritas. The correlation between the two measures is a quite high 0.9. The differences between the two measures are likely driven by the fact that Suss et al. (2024) estimate the number of resident accredited investors, while Phoenix/MarketCast estimate the number of resident millionaires (in terms of investable assets).

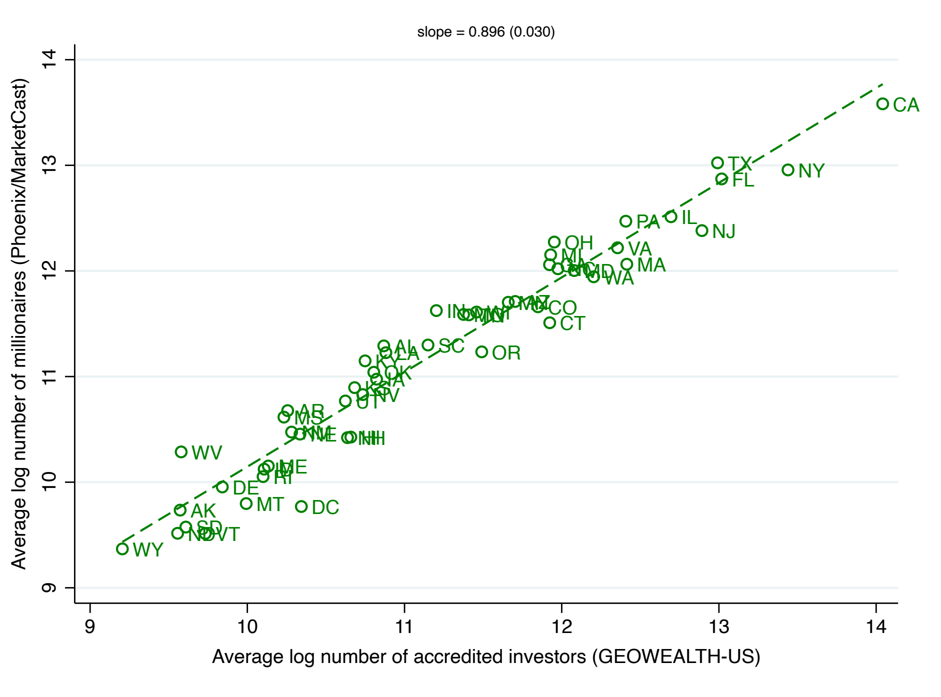
Figure C1, Panel B further depicts the correlation between the same state-level average from the GEOWEALTH-US and an analogous measure based on the Forbes 400 list of the richest Americans, which has been published in every year since 1982. We use the digitized and harmonized series from Saez and Zucman (2022). The correlation between the two measures is again quite high at 0.8. The differences between the two measures are likely driven by the fact that the Forbes 400 list only includes billionaires.

Figure C2, Panel A compares the evolution of the total number of accredited investors residing in the U.S. from the GEOWEALTH-US over the period 2006-2019 with the analogous measure from the Credit Suisse/UBS Global Wealth Report. The time series correlation between the two measures is a quite high 0.5. The differences between the two measures are likely driven by the fact that Credit Suisse/UBS estimates the number of millionaires, rather than the number of accredited investors. Furthermore, the former estimates are based on household units, while the latter are based on individual units.

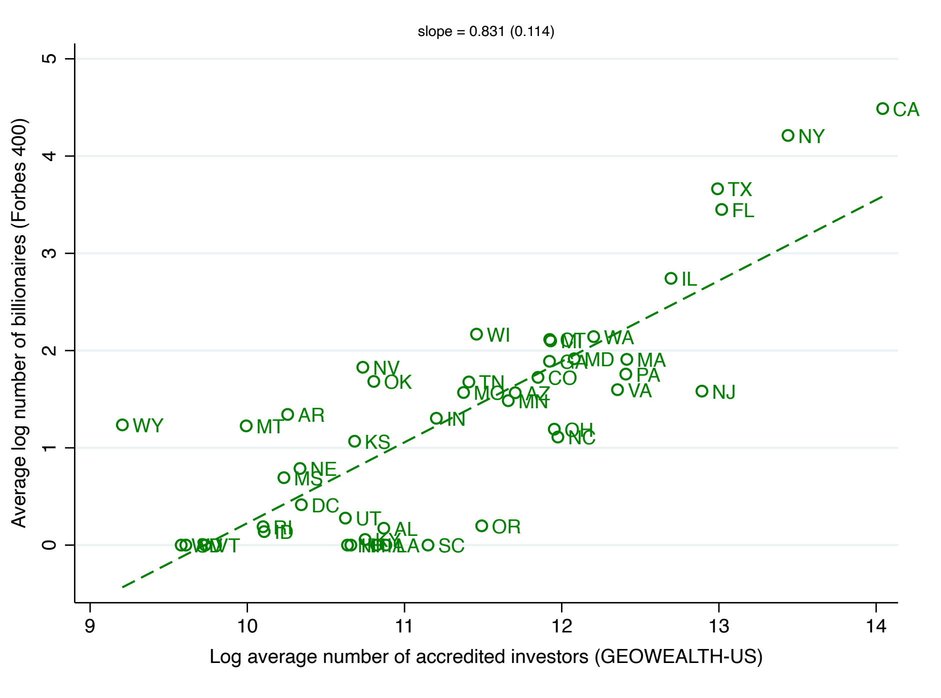
Finally, Figure C2, Panel B compares the same time series for the 2006-2022 period based on the GEOWEALTH-US with an analogous one based on the SCF.

Figure C1: HNWIs Residing in U.S. States, 2006-2019

(A) Phoenix/MarketCast Wealth and Affluent Monitor vs. GEOWEALTH-US



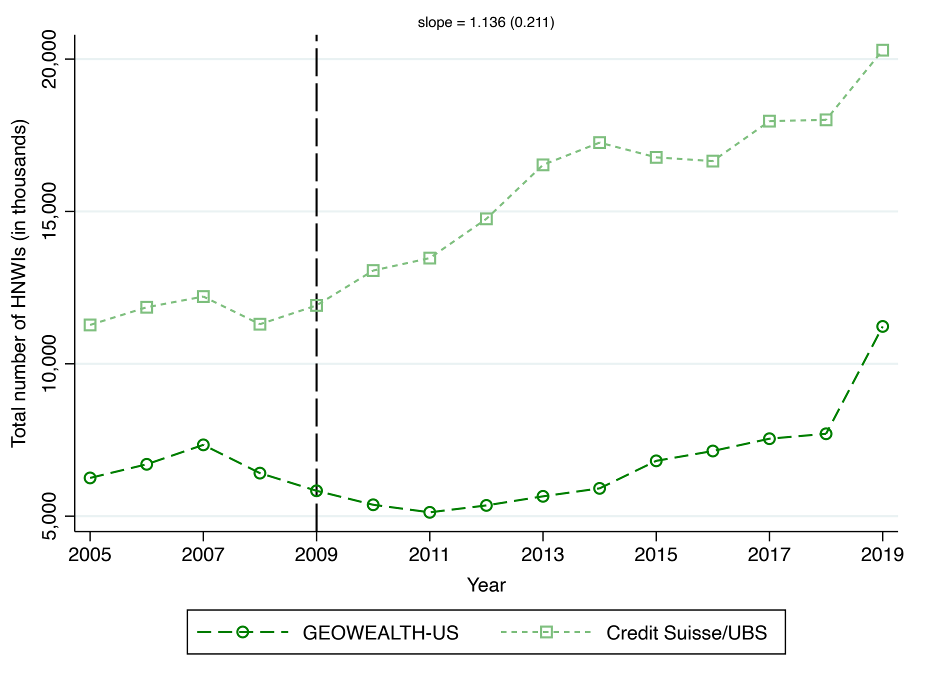
(B) Forbes 400 vs. GEOWEALTH-US



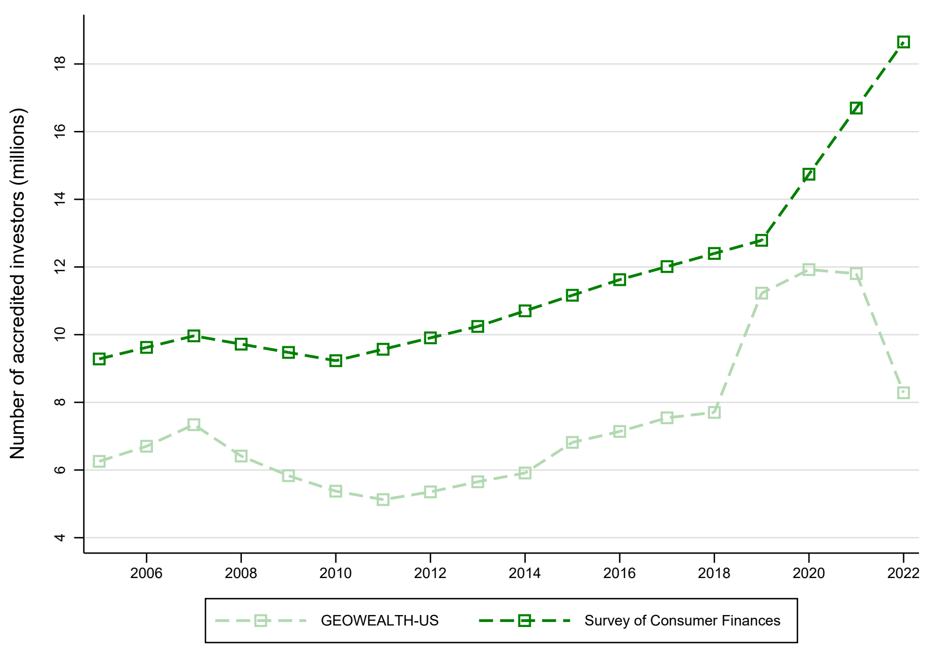
Source: Suss et al. (2024), Phoenix/MarketCast, Forbes.
Notes: Suss et al. (2024) refers to the data from the GEOWEALTH-US database. The data from Phoenix/MarketCast comes from its Wealth and Affluent Monitor, while the data from Forbes comes from the Forbes 400.

Figure C2: HNWIs Residing in U.S. States, 2006-2019

(A) Credit Suisse/UBS Global Wealth Report vs. GEOWEALTH-US



(B) SCF vs. GEOWEALTH-US



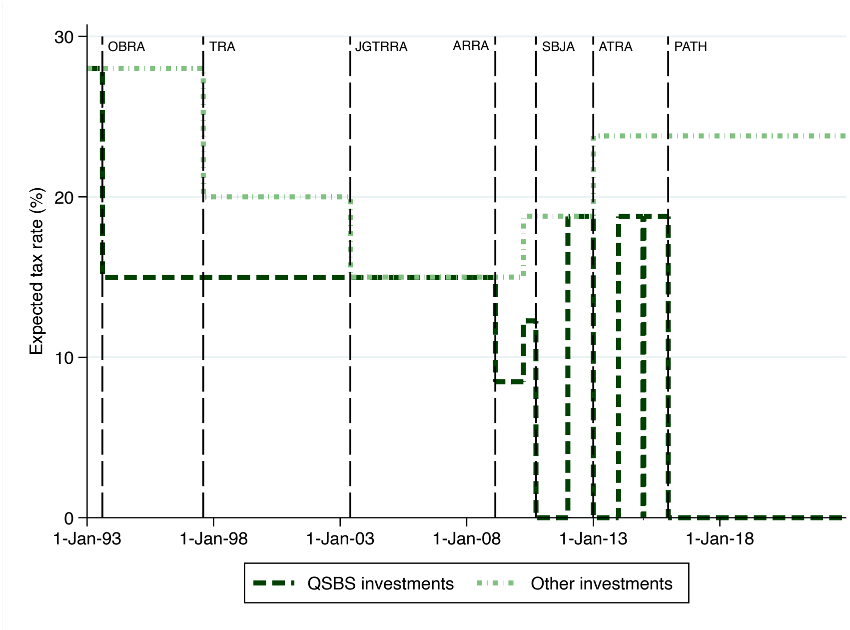
Source: Suss et al. (2024), Credit Suisse/UBS, SCF.
Notes: Suss et al. (2024) refers to the data from the GEOWEALTH-US database. The data from Credit Suisse/UBS comes from its Global Wealth Report.

Appendix D: Additional Results on Investment Effects

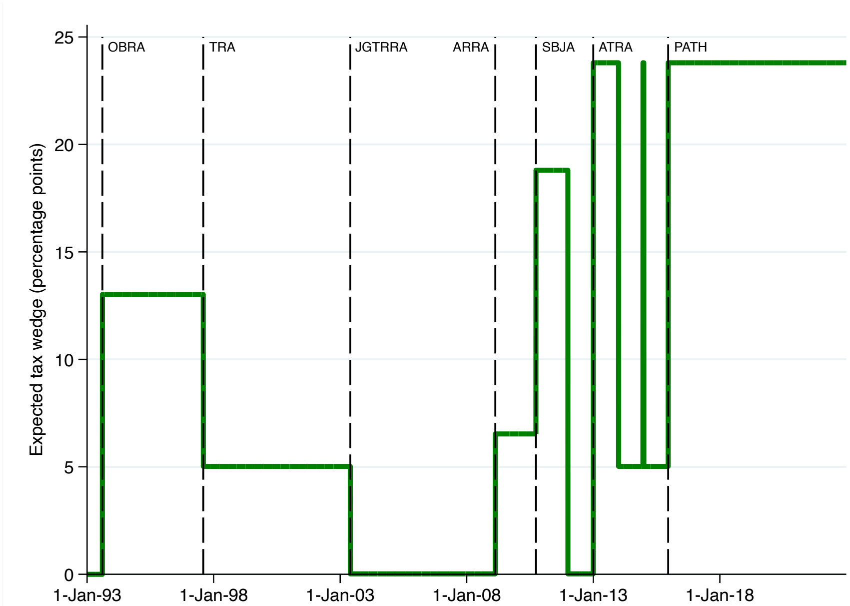
This appendix contains additional results related to Section 4 of the main text.

Figure D1: History of the Federal Tax Exemption on QSBS Capital Gains

(A) Expected Tax Rates on QSBS vs. Other Investments

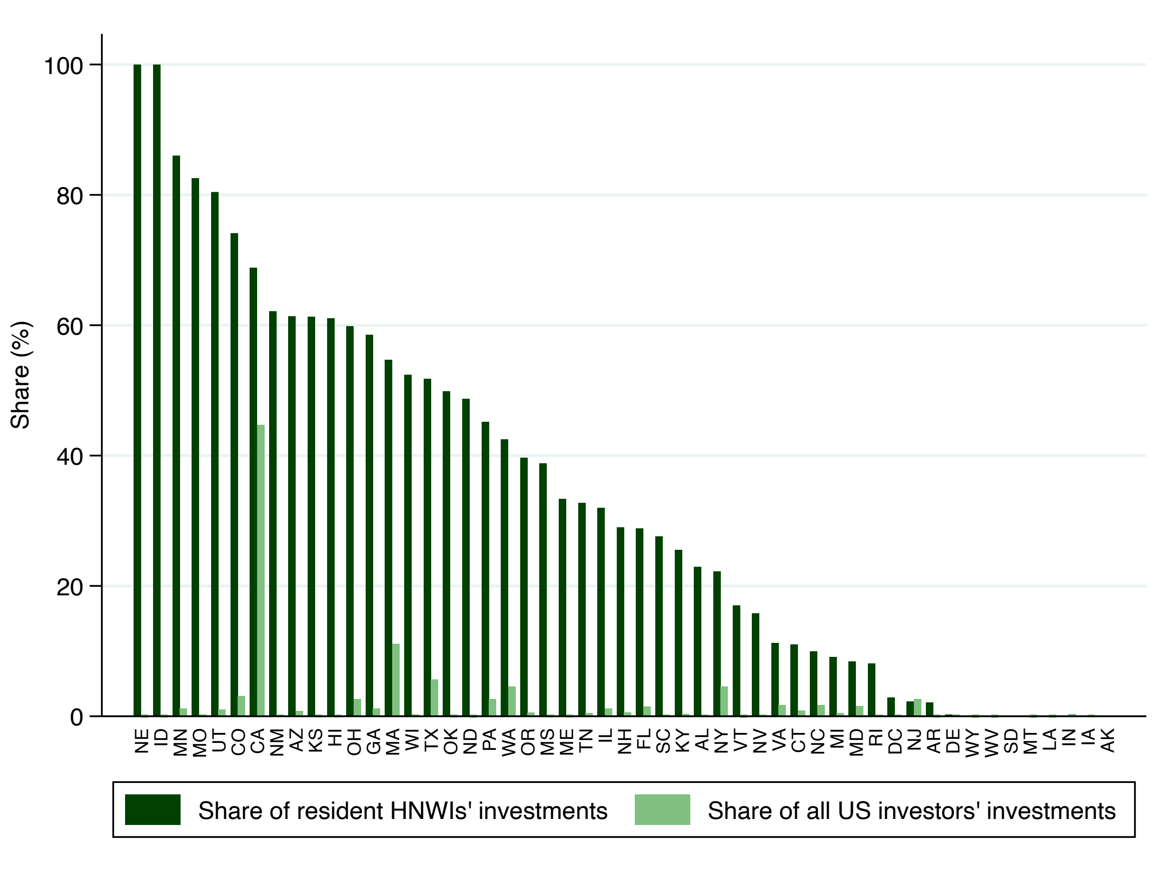


(B) Expected Tax Wedge on QSBS Investments



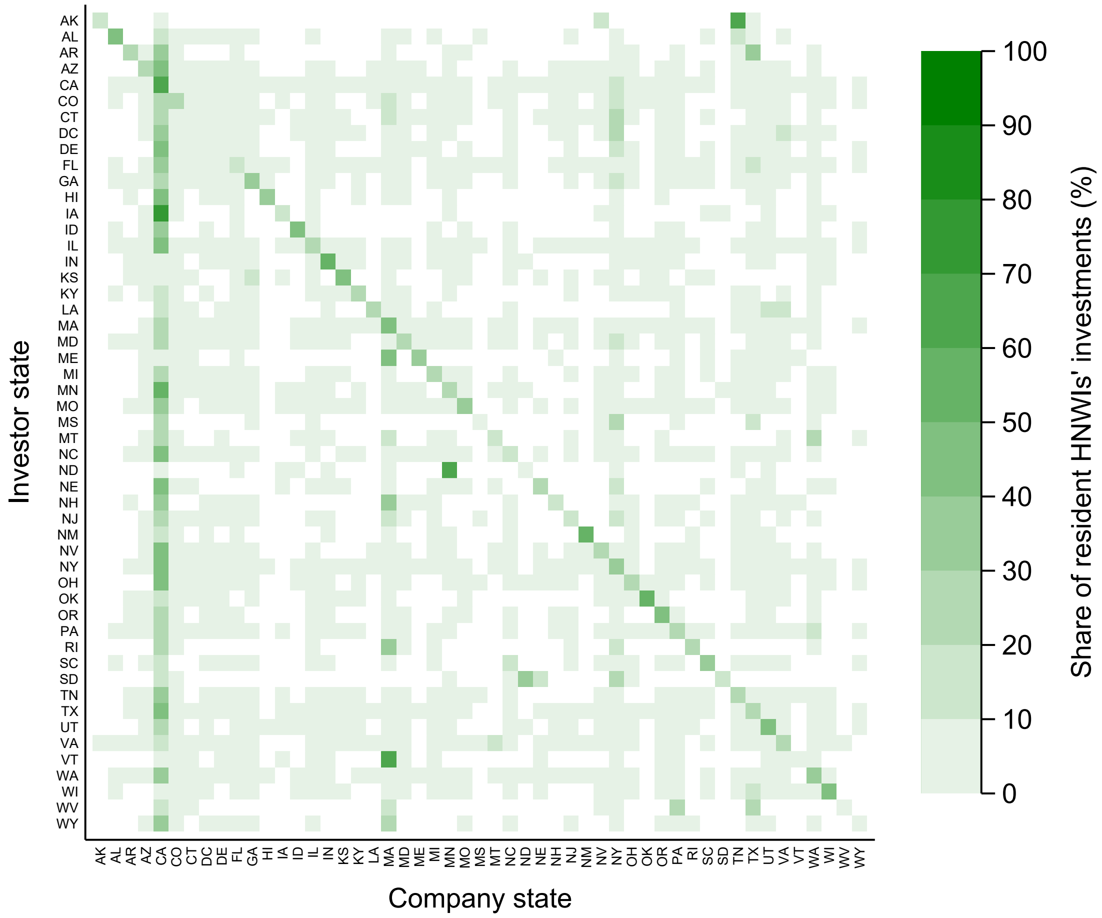
Source: Polsky and Yale (2023)).
Notes: Panel B plots the difference between the two lines in Panel A. The highlighted legislation include the Omnibus Budget Reconciliation Act (OBRA), the Taxpayer Relief Act (TRA), the Jobs and Growth Tax Relief Reconciliation Act (JBGTRRA), the American Recovery and Reinvestment Act (ARRA), the Small Business Jobs Act (SBJA), the American Tax Payer Relief Act (ATRA), and the Protecting Americans from Tax Hikes Act (PATH).

Figure D2: In-State Bias of Early-Stage Investments
by Resident HNWIs: 2004-2008



Source: Pitchbook.
Notes: The plot compares the share of investments by each state's resident HNWIs invested in companies headquartered within that state to the share of investments by all U.S. investors invested in companies headquartered within that same state.

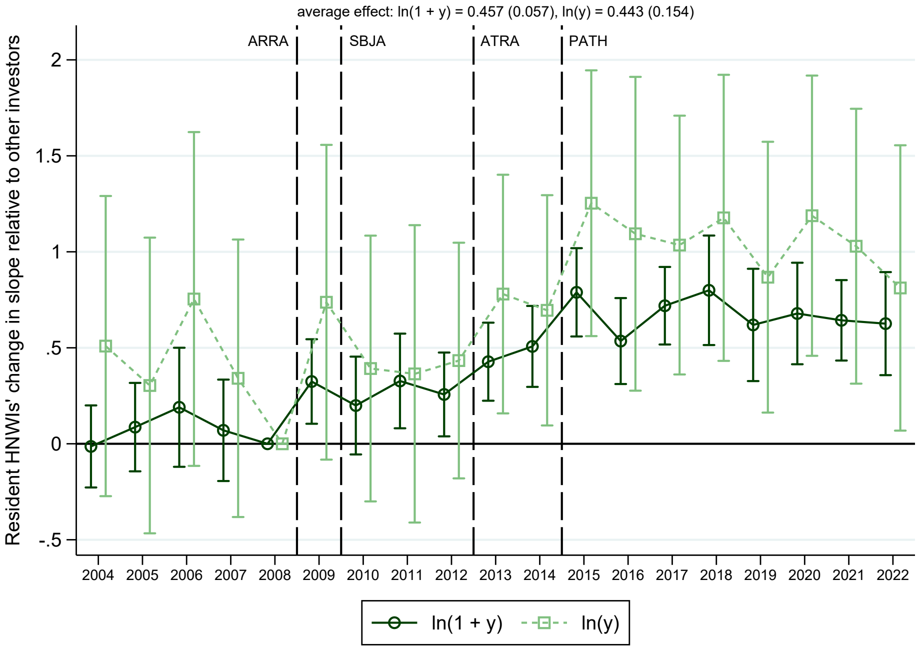
Figure D3: Distribution of Early-Stage Investments by Resident HNWIS across States: 2004-2022



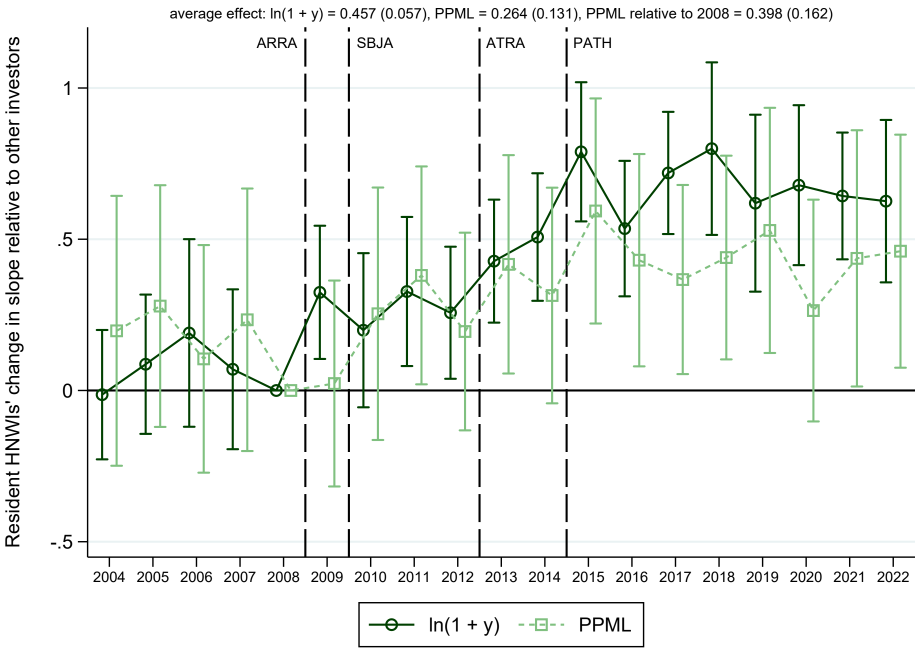
Source: Pitchbook.
Notes: The plot reports the share of investments by each (investor) state's resident HNWIs invested in companies headquartered each (company) state.

Figure D4: Robustness of Estimates of Equation (1) to Alternative Estimators

(A) $\ln(1 + Y_{i,s,t})$ vs. $\ln Y_{i,s,t}$



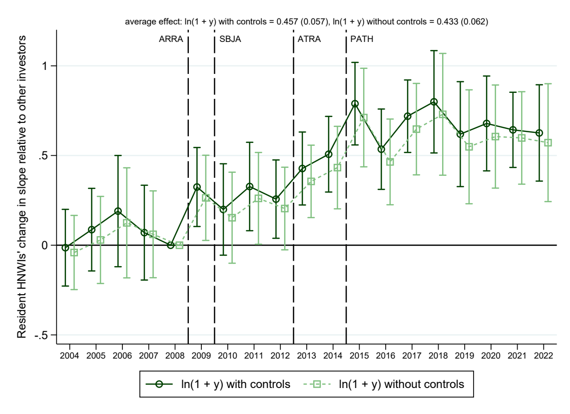
(B) $\ln(1 + Y_{i,s,t})$ vs. PPML



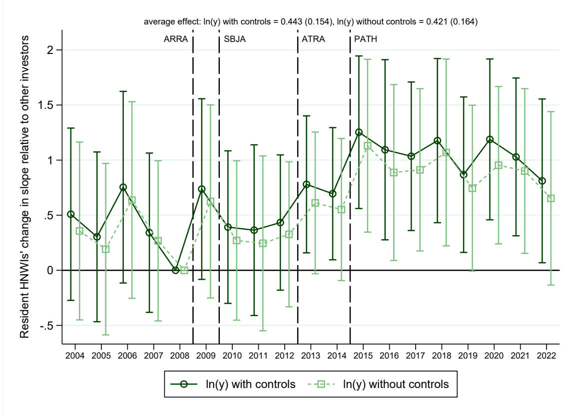
Source: Pitchbook, GEOWEALTH-US.
Notes: The average effects reported are based on a modified version of Equation (1) where β_t is replaced with $\beta_{t:t>2008}$ (or with both $\beta_{t:t<2008}$ and $\beta_{t:t>2008}$). The regression with $\ln(1 + Y_{i,s,t})$ as the outcome is based on 3,876 state-year observations. In Panel A, the regression with $\ln(Y_{i,s,t})$ as the outcome is based on 3,296 observations. In Panel B, the regression is estimated using a Poisson pseudo-maximum likelihood (PPML) estimator with $Y_{i,s,t}$ as the outcome and is based on 3,812 observations.

Figure D5: Robustness of Estimates of Equation (1) to Inclusion of Controls

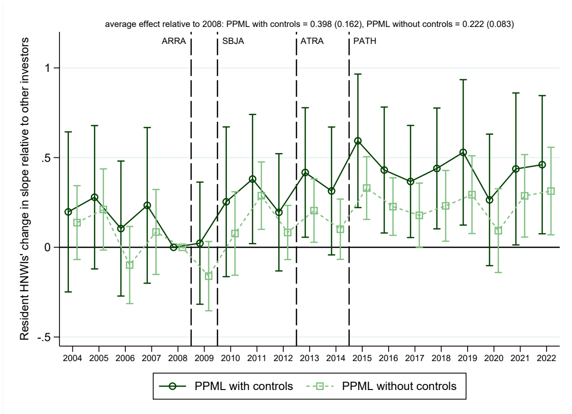
(A) $\ln(1 + Y_{i,s,t})$



(B) $\ln Y_{i,s,t}$

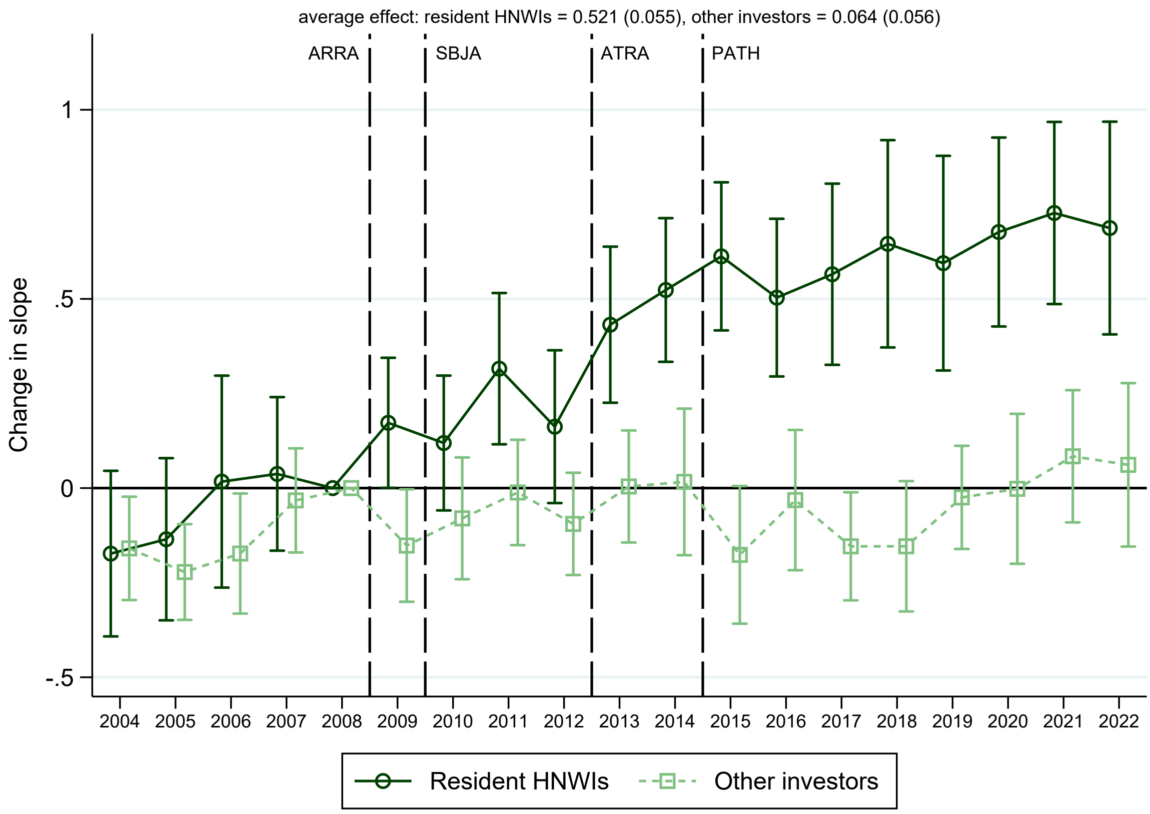


(C) PPML



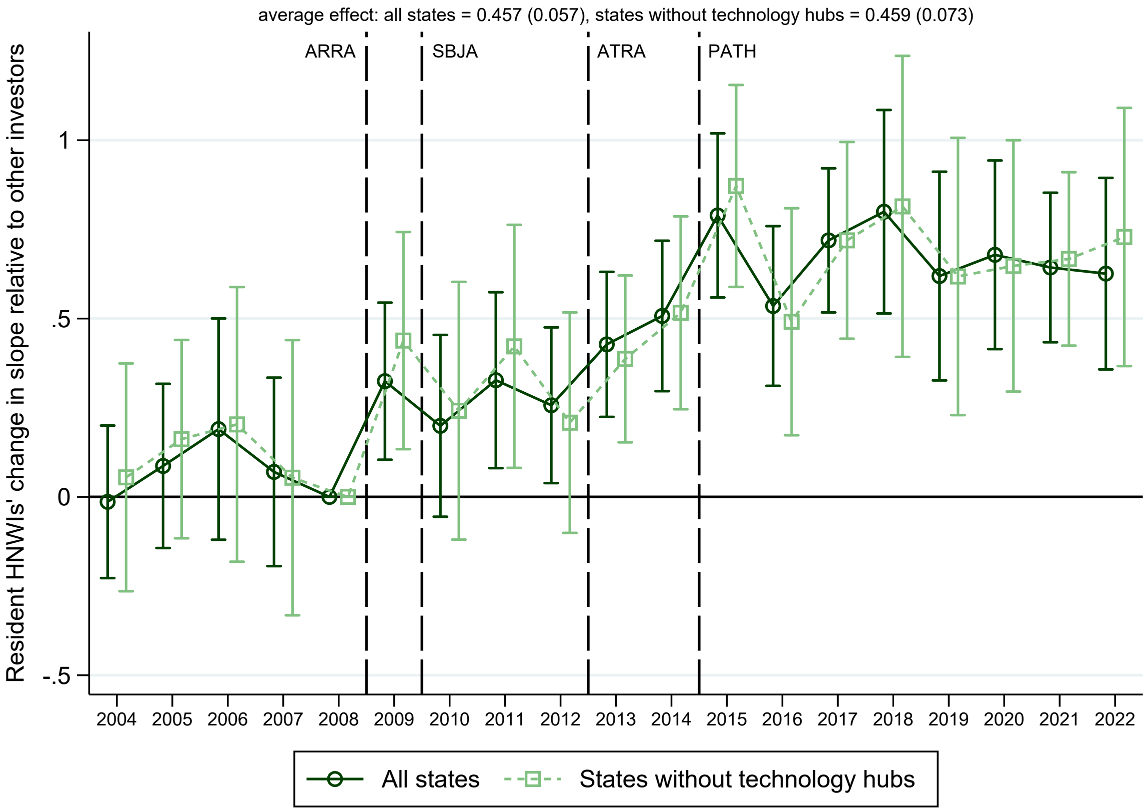
Source: Pitchbook, GEOWEALTH-US.
Notes: The average effects reported are based on a modified version of Equation (1) where β_t is replaced with $\beta_{t:t>2008}$ (or with both $\beta_{t:t<2008}$ and $\beta_{t:t>2008}$). In Panel A, the regression with $\ln(1 + Y_{i,s,t})$ as the outcome is based on 3,876 observations. In Panel B, the regression with $\ln(Y_{i,s,t})$ as the outcome is based on 3,296 observations. In Panel C, the regression is estimated using a Poisson pseudo-maximum likelihood (PPML) estimator with $Y_{i,s,t}$ as the outcome and is based on 3,812 observations. The controls include only the local long-term capital gains tax wedge on QSBS investments for individuals residing in state s.

Figure D6: Difference-in-Difference Estimates Underlying Estimates of Equation (1)



Source: Pitchbook, GEOWEALTH-US.
Notes: The difference-in-difference estimates are based on a modified version of Equation (1) where $\mathbb{1}_{i=\text{resident HNWIs}}$ and $\delta_{s,t}$ are dropped. The average effects reported are then based on a further modification that replaces β_t with $\beta_{t:t>2008}$.

Figure D7: Robustness of Estimates of Equation (1) to Subsamples of States



Source: Pitchbook, GEOWEALTH-US.
Notes: The regression is based on 3,192 state-year observations. The average effects reported are based on a modified version of Equation (1) where β_t is replaced with $\beta_{t:t>2008}$. The states without technology hubs exclude only those 9 states (California, Colorado, District of Columbia, Georgia, Illinois, Massachusetts, New York, Texas, and Washington) that contain a city listed as a technology hub on the website of the U.S. technology networking company Built In: https://builtin.com/tech-hubs.

Table D1: Other company-level outcomes

	Probability of HNWI investment	Log of HNWI investments	Probability of staying private	Probability of bankruptcy
Treated x Post	0.021***	0.013*	0.030***	0.001
	(0.006)	(0.007)	(0.009)	(0.006)
Company FE	×	×	×	×
Year FE	×	×	×	×
CCorp x Year FE	×	×	×	×
Industry x Year FE	×	×	×	×
Size x Year FE	×	×	×	×
Observations	336228	336228	336228	336228
R squared	0.223	0.190	0.781	0.434

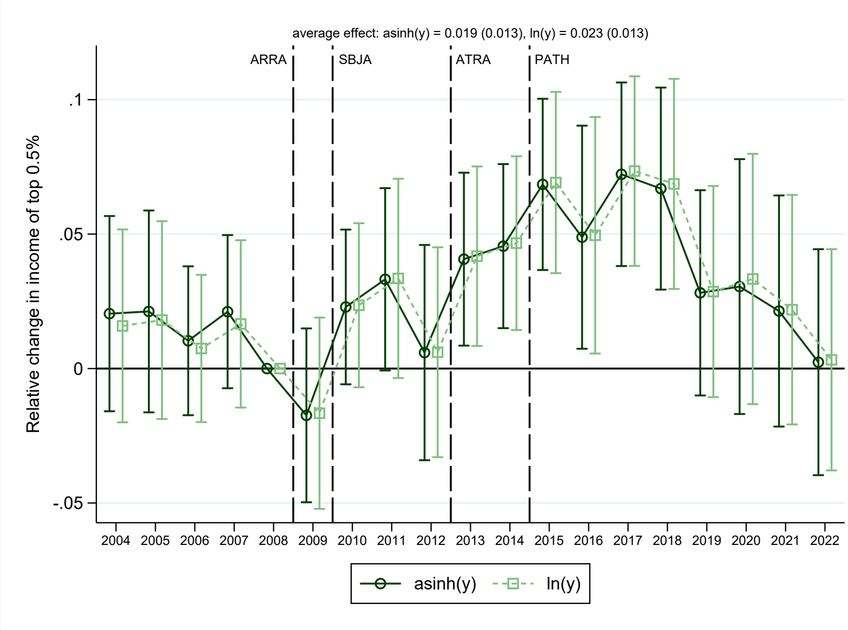
Source: Pitchbook.
Notes: The average effect reported is based on a modified version of Equation (2) where β_t is replaced with $\beta_{t:t>2008}$. High-net-worth individuals (HNWIs) refer to U.S.-based investors categorized by Pitchbook as individuals, angel groups, and family offices. ***, **, and * indicate statistical significance at the 99%, 95%, and 90% significance levels, respectively.

Appendix E: Additional Results on Inequality

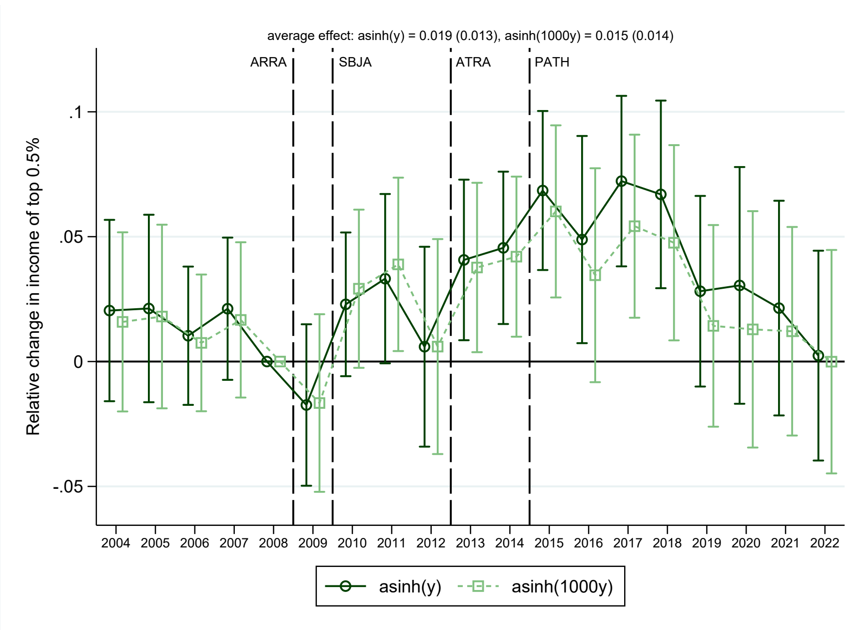
This appendix contains additional results related to Section 5 of the main text.

Figure E1: Robustness of Estimates of Equation (3) to Alternative Estimators

(A) asinh $Y_{r,s,t}$ vs. $\ln Y_{r,s,t}$



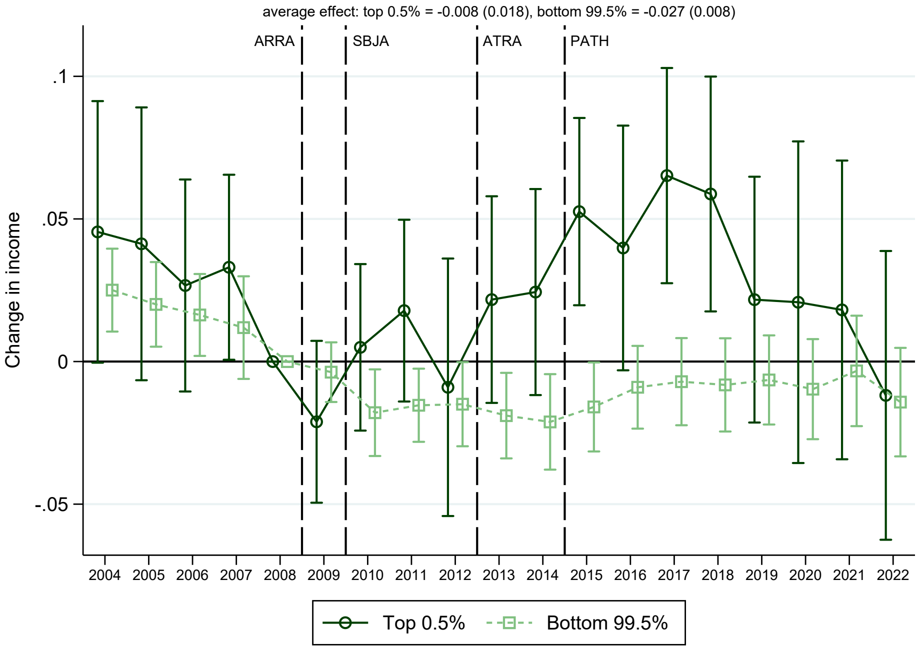
(B) asinh $Y_{r,s,t}$ vs. $\text{asinh}(1000 \times Y_{r,s,t})$



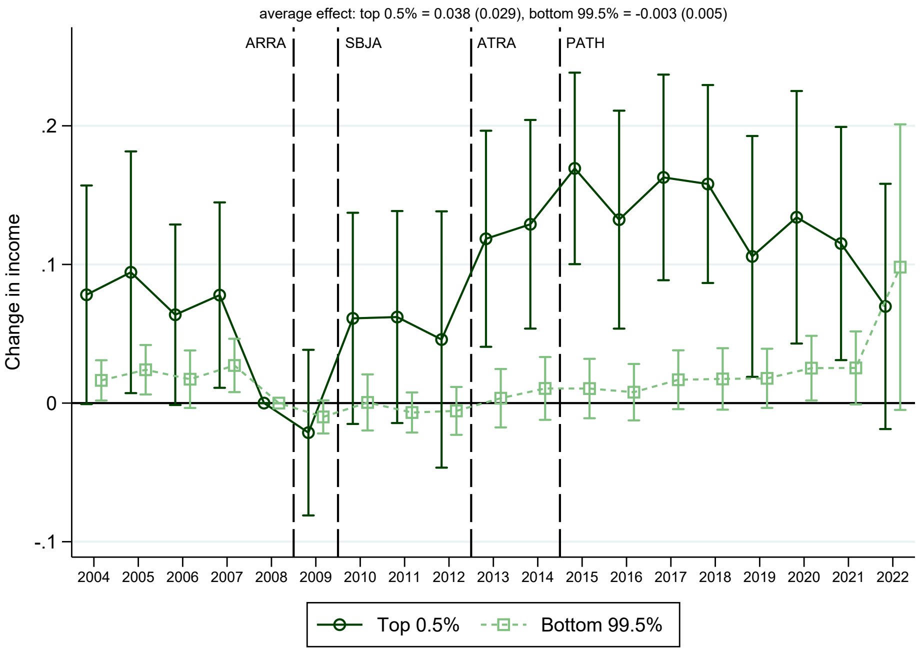
Source: SOI Tax Stats, GEOWEALTH-US.
Notes: The regression is based on 99,807 state-year observations. The average effect reported is based on a modified version of Equation (3) where β_t is replaced with $\beta_{t:t>2008}$.

Figure E2: Difference-in-Difference Estimates Underlying Equation (3)

(A) Total Income

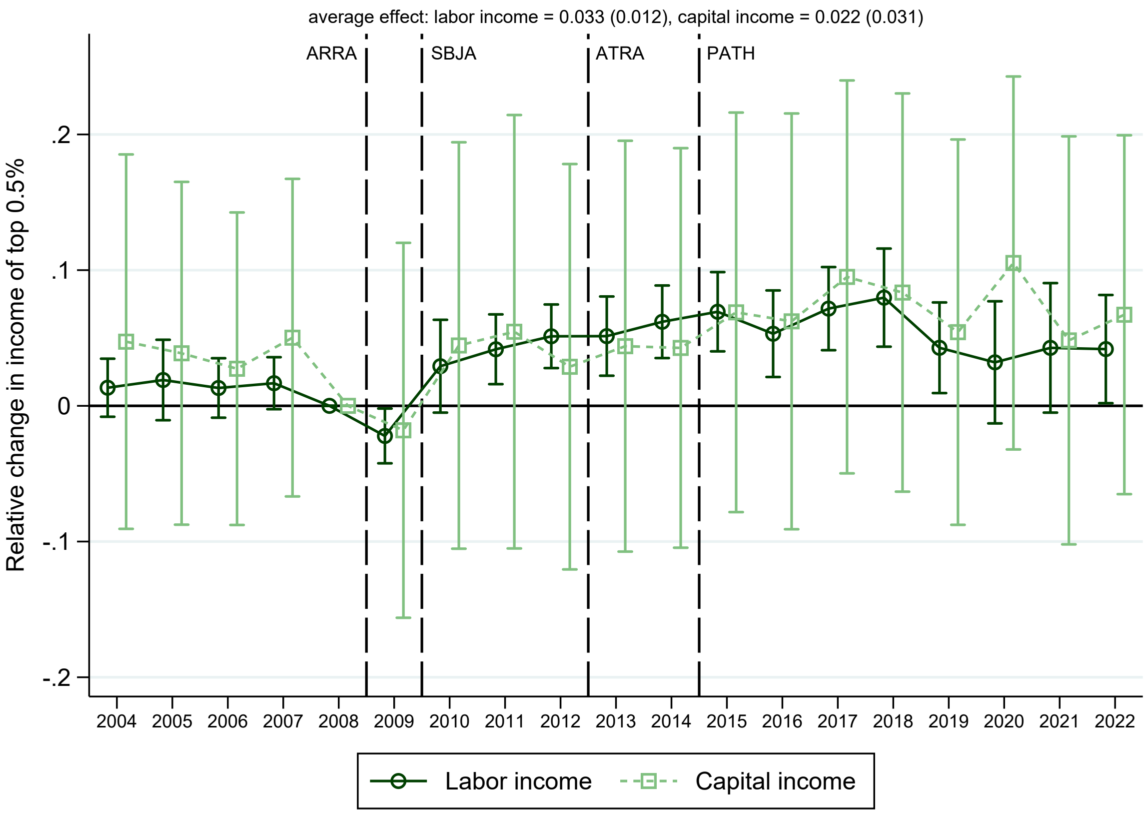


(B) Capital Gains



Source: SOI Tax Stats, GEOWEALTH-US.
Notes: The regression is based on 99,807 state-year observations. The average effect reported is based on a modified version of Equation (3) where β_t is replaced with $\beta_{t:t>2008}$.

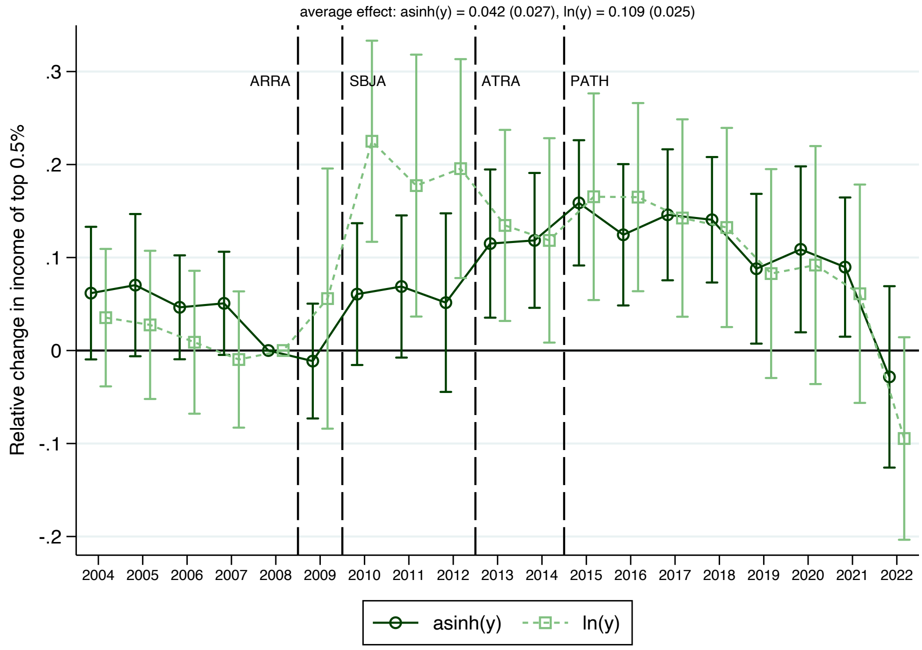
Figure E3: Decomposition of Other Income



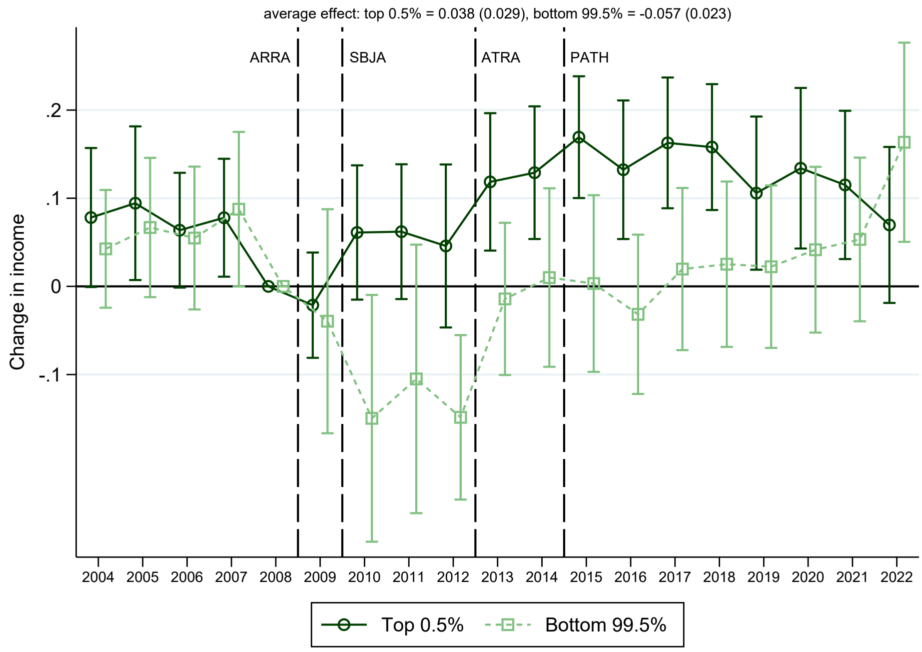
Source: SOI Tax Stats, GEOWEALTH-US.
Notes: The regression is based on 99,807 state-year observations. The average effect reported is based on a modified version of Equation (3) where β_t is replaced with $\beta_{t:t>2008}$.

Figure E4: Robustness of Effects on Capital Gains: asinh $Y_{r,s,t}$ vs. $\ln Y_{r,s,t}$

(A) Triple-Difference Estimates



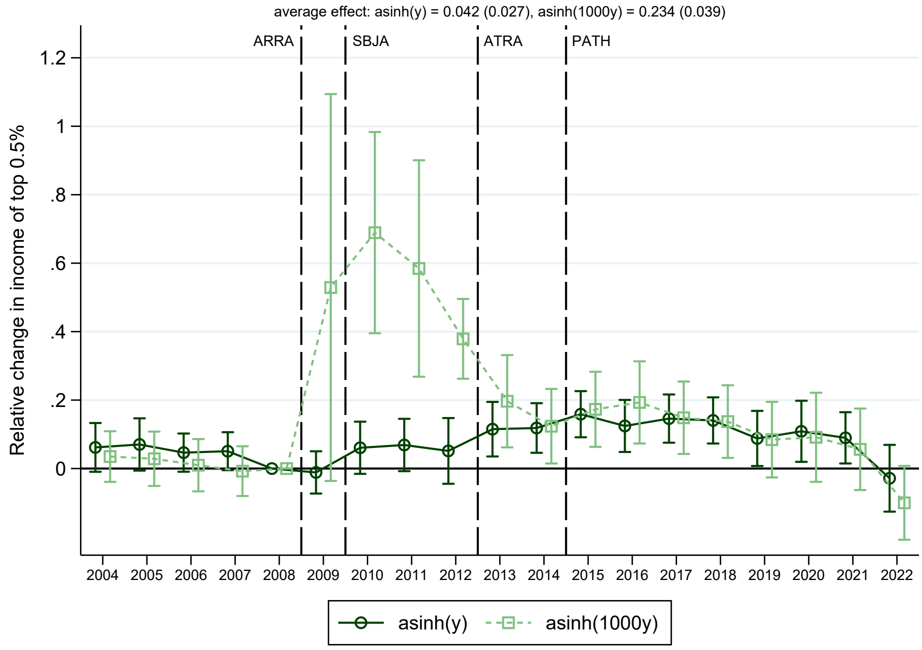
(B) Difference-in-Difference Estimates: $\ln Y_{r,s,t}$



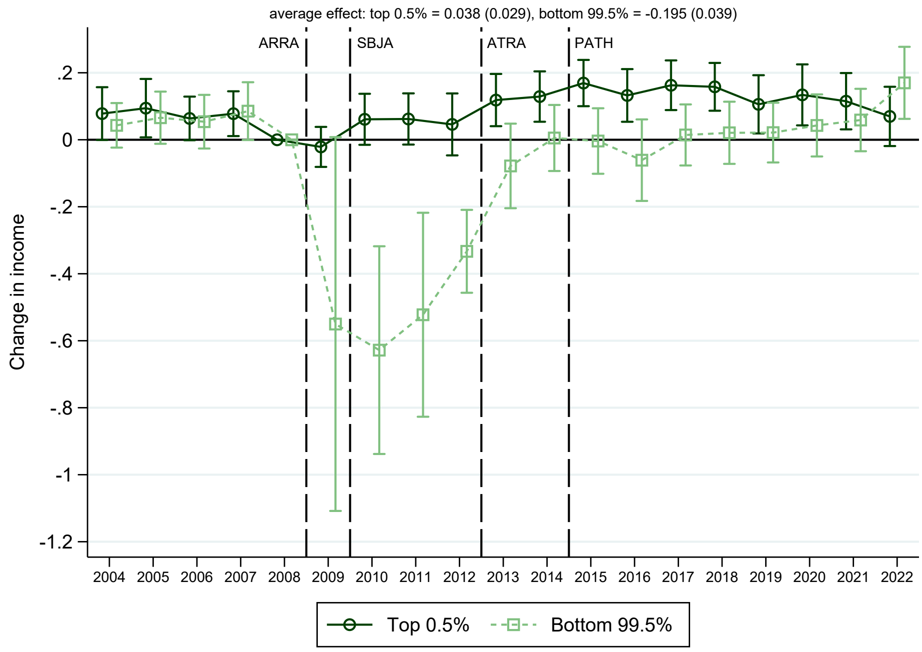
Source: SOI Tax Stats, GEOWEALTH-US.
Notes: The regression is based on 99,807 state-year observations. The average effect reported is based on a modified version of Equation (3) where β_t is replaced with $\beta_{t:t>2008}$.

Figure E5: Robustness of Effects on Capital Gains: asinh $Y_{r,s,t}$ vs. asinh$(1000 \times Y_{r,s,t})$

(A) Triple-Difference Estimates

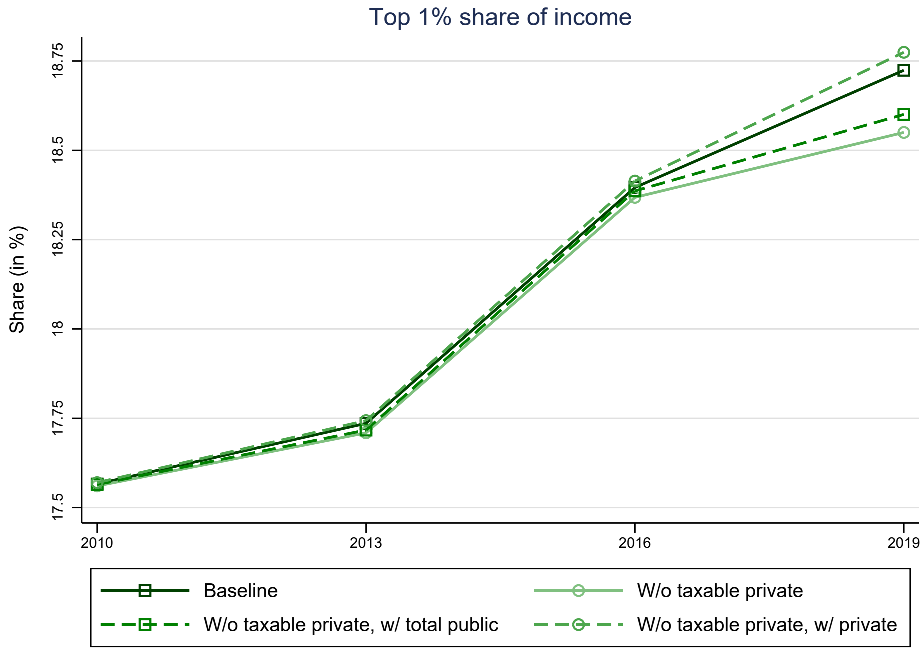


(B) Difference-in-Difference Estimates: asinh$(1000 \times Y_{r,s,t})$



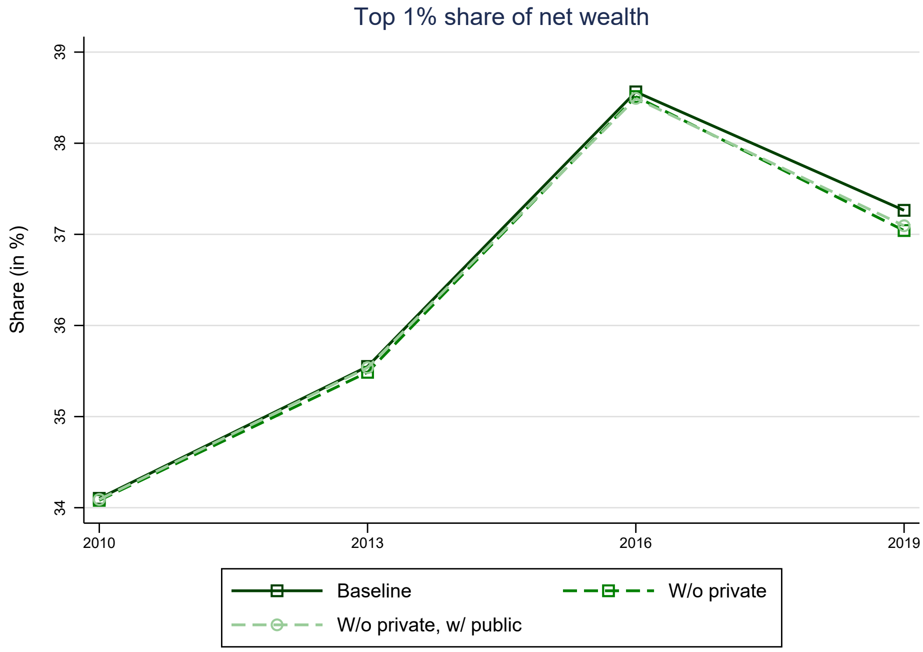
Source: SOI Tax Stats, GEOWEALTH-US.
Notes: The regression is based on 99,807 state-year observations. The average effect reported is based on a modified version of Equation (3) where β_t is replaced with $\beta_{t:t>2008}$.

Figure E6: Income Inequality Counterfactuals



Source: SOI Tax Stats, Pitchbook.
Notes: This figure compares the baseline top 1% taxable income share to the counterfactual top 1% taxable income share over the period 2010-2019 under three different scenarios: no taxable private capital gains; no taxable private capital gains but with counterfactual public capital gains; and no taxable private capital gains, but with total (taxable and non-taxable) private capital gains.

Figure E7: Wealth Inequality Counterfacuals



Source: Survey of Consumer Finances, Pitchbook.
Notes: This figure compares the baseline top 1% wealth share to the counterfactual top 1% wealth share over the period 2010-2019 under two different scenarios: no taxable private capital gains; and no taxable private capital gains but with counterfactual public capital gains.