Heterogeneity and Persistence in Returns to Wealth*

Andreas Fagereng[†] Luigi Guiso[‡] Davide Malacrino[§] Luigi Pistaferri[¶]

First Version: December 2015 This version: November 2016

Abstract: We provide a systematic analysis of the properties of individual returns to wealth using twenty years of population data from Norway's administrative tax records. We document a number of novel results. First, in a given cross-section, individuals earn markedly different returns on their assets, with a difference of 500 basis points between the 10th and the 90th percentile. Second, heterogeneity in returns does not arise merely from differences in the allocation of wealth between safe and risky assets: returns are heterogeneous even within asset classes. Third, returns are positively correlated with wealth. Fourth, returns have an individual permanent component that accounts for 60% of the explained variation. Fifth, for wealth below the 95th percentile, the individual permanent component accounts for the bulk of the correlation between returns and wealth; the correlation at the top reflects both compensation for risk and the correlation of wealth with the individual permanent component. Finally, the permanent component of the return to wealth is also (mildly) correlated across generations. We discuss the implications of these findings for several strands of the wealth inequality debate.

Keywords: Wealth inequality, returns to wealth, heterogeneity, intergenerational mobility.
JEL codes: D31, D91, E21, E24, G11.

*We thank Alberto Bisin, Fatih Guvenen, Gueorgui Kambourov, Daniele Massacci, Giovanni Mastrobuoni, Benjamin Moll, Andrea Pozzi, Ali Shourideh, and Anton Tsoy for very fruitful discussions. We are grateful for useful comments from seminar participants at the 2016 ASSA meetings, the NBER Summer Institute, Bocconi University, EIEF, European University Institute, University of Chicago, UC Berkeley, Fudan University, HEC Montreal, UC Santa Cruz, MIT Sloan, Northwestern, Kellogg, NY FED, Santa Clara University, Statistics Norway, the 2016 QSPS meeting at Utah State, and the 2016 Cowles Macro/Labor conference at Yale. Funding from the Research Council of Norway is gratefully acknowledged.

[†] Statistics Norway, faa@ssb.no
[‡] Einaudi Institute for Economics and Finance (EIEF) and CEPR, guiso@tin.it
[§] Stanford University, davidem@stanford.edu
[¶] Stanford University, NBER and CEPR, pista@stanford.edu

1 Introduction

Over time and across countries, the wealth distribution appears to be extremely skewed and with a long right tail: a small fraction of the population owns a large share of the economy's wealth. In the United States (US), for example, the top 0.1% hold about 20% of the economy's net worth. Moreover, tail inequality seems to have tripled in little more than three decades (Saez and Zucman, 2016).

What produces the long tail of the wealth distribution and its extreme skewness is the subject of intense research. A traditional strand of literature started by Aiyagari (1994) (reviewed in Section 2) has focused on the role played by idiosyncratic and uninsurable labor income (i.e., human capital) risk (see Castaneda, Díaz-Giménez, and Ríos-Rull (1998); Huggett (1996)), or, more generally, heterogeneity in human capital (e.g., Castaneda et al. (2003)). The success of these models in reproducing the amount of wealth concentration observed in the data is mixed (see De Nardi, 2016), and their ability to explain rapid changes in wealth inequality dubious. A recent wave of papers has shifted attention from heterogeneity in returns to human capital to heterogeneity in returns to financial and physical capital (see Benhabib, Bisin, and Zhu (2011), Benhabib and Bisin (2016), and Gabaix, Lasry, Lions, and Moll (2015)). These papers show that models in which individuals are endowed with idiosyncratic returns to wealth that persist over time and (to some extent) across generations can generate a steady state distribution of wealth with a thick right tail that reproduces very closely what is observed in reality. Moreover, persistent heterogeneity in returns (*"type dependence"* in Gabaix, Lasry, Lions, and Moll (2015)-terminology), coupled with a positive correlation of returns with wealth ("*size dependence*"), can potentially explain rapid increases in tail inequality similar to those observed in the US over the last three decades.

There is scant evidence, however, on the qualitative and quantitative importance of the features emphasized by this more recent literature. How much heterogeneity in returns to wealth is there in the data? Do returns to wealth persist over time within a generation, as required by the Benhabib, Bisin, and Zhu (2011) model? Do they persist across generations, and if so, by how much? Are returns and their heterogeneity correlated with wealth, as required by the model of Gabaix, Lasry, Lions, and Moll (2015) to explain the rapid transitions in tail inequality? More generally, what are the empirical properties of the returns to wealth? This paper provides answers to these questions. Addressing them has so far been difficult due to data limitations: available survey data are plagued with measurement error and low response rates at the top of the wealth distribution, and they contain either limited or no longitudinal information.

We overcome these problems using two decades of administrative tax records of capital income and wealth stocks for all taxpayers in Norway. Several properties of these data make them well suited to addressing the above questions. First, measurement error and underreporting of wealth information are unlikely to be a problem, because wealth data are mostly collected through third parties (i.e., information provided by financial intermediaries). Second, the data have universal coverage, implying that there is exhaustive information about the assets owned and incomes earned by *all* individuals, including those at the very top of the wealth distribution. Furthermore, besides information on financial assets, we have data on wealth held in private businesses. These two features are critical for a study of our sort, because leaving out the wealthy or the wealth in private businesses (which happens to be highly concentrated among the wealthy) could seriously understate the extent of heterogeneity in returns to wealth, particularly if returns and the extent of heterogeneity are correlated with wealth. Most importantly, the data have an extraordinarily long panel dimension, covering 20 years – from 1993 to 2013 – and various business cycles. This allows us to study within-person persistence in returns. Finally, because over a 20-year period (some) generations overlap and because we can identify parents and children, we can also study intergenerational persistence in returns to wealth.

We find that returns to wealth exhibit substantial heterogeneity. For example, in the last year of our sample (2013), the (value-weighted) average return on overall wealth is 3.7%, but it varies considerably across households (standard deviation 6.1%). Furthermore, heterogeneity in returns is not simply a reflection of differences in portfolio allocations between risky and safe assets, and thereby compensation for risk-taking that mirrors heterogeneity in risk tolerance. Even conditioning on the share of risky assets in a portfolio, heterogeneity in returns is large and increases with the level of wealth. This result is confirmed even when looking at individuals with no private business wealth. Another remarkable finding is that asset returns increase with wealth. In 2013, the difference between the median return for people in the 90th and 10th percentiles of the wealth distribution is 180 basis points. The correlation between returns and wealth does not merely reflect risk-taking. We find that risk-adjusted measures of excess returns (the Sharpe ratio) increase with the level of wealth at the point of entry in the sample, before any investment decisions are taken.

In any given year, heterogeneity in returns to wealth may arise from differences in observables (e.g., in risk-taking), from idiosyncratic transitory variations (good or bad luck), or from a persistent unobservable component in returns to wealth. The latter is the critical component in the new literature on wealth inequality. To separate these components, we estimate a panel data statistical model for the returns to wealth that includes an individual fixed effect. To account for heterogeneity explained by observable factors, we control for the level of wealth (capturing investment-size effects on returns), the share of wealth in various types of risky assets (measuring compensation for risk), as well as for time effects and demographics. The individual fixed effect measures the component of unobserved heterogeneity that persists over time. We find that observable characteristics alone explain roughly 12% of the variability in returns to wealth. Adding individual fixed effects more than doubles the explained variability to 27%. The distribution of these fixed effects is itself quite dispersed, with a standard deviation of 2.8 percentage points and a 90th to 10th percentile difference of 6.4 percentage points.

We use our statistical model to identify the drivers of the positive correlation between wealth rank and average returns. We find that for wealth below the 95th percentile the correlation between average returns and wealth rank is largely due to a positive correlation between wealth and the fixed effects (individuals with permanently higher returns are wealthier). For wealth above the 95th percentile, it is largely driven by compensation for higher risk exposure among the wealthy, with a more limited contribution (one third) from the fixed effects.

We also study intergenerational persistence in asset returns. We find that both the return to wealth and its fixed component are correlated intergenerationally, although there is strong mean reversion. Interestingly, the association between a child's asset return and its parent's asset return, while positive for a wide range of the distribution, turns negative when the parent's return is above the 80th percentile. In other words, children of individuals who were able to achieve very high returns from wealth have returns that, while still above average, revert more quickly to the mean.

As far as we know, this is the first paper to provide systematic evidence on individual returns to wealth over the entire wealth distribution and to characterize their properties. Bach et al. (2015) perform an exercise close to ours in spirit, but our paper differs from theirs in several respects. First, their main focus is the extent and nature of the correlation between returns and wealth at the top of the wealth distribution; we are interested in studying the properties of the returns to wealth over the whole range of the wealth distribution. Second, we have access to longer panel data than they do, allowing us to study persistence in returns. Third, we can study heterogeneity and persistence in returns to wealth over and above the intra-generational dimension. Indeed, our paper is the first to provide systematic evidence of persistence in returns within and across generations. These two features are critical for explaining the long thick tail in the wealth distribution. We also provide evidence that the persistent component of returns is correlated with wealth and so is the degree of heterogeneity - two features of the data that reasonably calibrated models of wealth inequality should be able to accommodate. We also find that heterogeneity in returns varies over time. While heterogeneity in returns matters to explain the level of wealth inequality at the top, variation in heterogeneity over time matters to explain variation in wealth inequality over time. With the exception of Gabaix, Lasry, Lions, and Moll (2015), most papers have focused on explaining the distribution of wealth (or income) at a point in time, assuming the economy is in steady state. This theoretical debate lags behind the empirical one, which has shifted from measuring the extent of inequality at a point in time to documenting significant dynamics in inequality, either in income (Piketty and Saez (2003)) or wealth (e.g., Saez and Zucman (2016)).

The rest of the paper proceeds as follows. In Section 2, we review the literature. In Section 3, we present our data and discuss how we measure returns to wealth. Section 4 documents stylized facts about returns to wealth. In Section 5, we discuss our empirical model of individual returns, show how we identify persistent heterogeneity and present results about its extent. In Section 6, we discuss the drivers of the correlation between returns and wealth distinguishing between the role of observable factors, such as compensation for risk and unobserved heterogeneity. Section 7 documents intergenerational persistence. Section 8 relates our results to calibrated models of wealth inequality with returns heterogeneity. Section 9 concludes and discusses some implications of our findings.

2 Heterogeneity in returns and the distribution of wealth

In the absence of sources of heterogeneity in saving propensities or sources of income other than labor, the distribution of wealth should inherit the properties of the distribution of earnings. Hence, if the distribution of labor income has a fat tail, the wealth distribution should mirror it. Yet wealth seems to be more unequally distributed than income, and realistic calibrations of heterogeneity in earnings that produce significant wealth inequality (as in Castaneda, Díaz-Giménez, and Ríos-Rull (2003) and Kindermann and Krueger (2014)) do not seem to be able to reproduce the fatter tail in the distribution of wealth. For instance, while the calibrated model of Kindermann and Krueger (2014) comes close to matching the distribution of wealth in the US, it requires the top 0.25% of income earners to earn between 400 and 600 times more than the median earner. As Benhabib and Bisin (2016) note, this is very far from what is observed in the data - where the ratio of the income of the top 0.1% percent to the median is only around 33. A similar argument applies to Castaneda et al. (2003).

One route, followed by Krusell and Smith (1998), has been to complement Bewley-Aiyagari models of earnings heterogeneity with heterogeneity in thriftiness, allowing individuals to differ in time discounting. Differences in thriftiness, together with heterogeneity in earnings, can considerably improve the match between the wealth distribution generated by the model and that in the data. Discount rate heterogeneity has a certain appeal because of its intuitive realism. On the other hand, discount rates are hard to observe and their heterogeneity is thus difficult to assess. Hence, it is necessary to impose and accept the heterogeneity that is needed to match the distribution of wealth without being able to validate it. Furthermore, discount rate heterogeneity seems to miss one important feature of the data: the high incidence of entrepreneurs at the top of the wealth distribution. Entrepreneurship is usually associated with higher risk tolerance and idiosyncratic risk (entrepreneurs tend to hold very high stakes in their own company - see e.g., Heaton and Lucas (2000); Vissing-Jorgensen and Moskowitz (2002)), rather than with higher than average discount rates. An alternative route followed in an attempt to match the thick tail in the distribution of wealth has been to explicitly allow for entrepreneurship and idiosyncratic returns to investment, as in Quadrini (2000) and Cagetti and De Nardi (2009; 2006). These papers show that a model that incorporates individual-specific technologies – i.e., entrepreneurs - can generate more wealth inequality than that produced by Bewley-Aiyagari models of earnings heterogeneity. In these models, the driving factor that enables matching of the observed wealth inequality is given by potentially high rates of return from entrepreneurial investment, coupled with borrowing constraints (which induce a selection of entrepreneurs among wealthy people to start with). Models of entrepreneurial idiosyncratic risk-taking have been developed more recently by Aoki and Nirei (2015), and by Benhabib and Bisin (2016) using a more reduced form approach. See De Nardi (2016) for an exhaustive critical appraisal of the literature.

While idiosyncratic returns from entrepreneurship are one source of heterogeneity in returns to wealth that can help to explain wealth concentration, heterogeneity in returns to wealth can also arise from other sources. For example, Guvenen (2009) introduces return differentials by allowing all households to trade in a risk-free bond, but restricts access to the stock market to only one group of agents. This model captures limited stock market participation and generates heterogeneity in returns to wealth between stockholders and non-stockholders. Guvenen (2007) shows that a calibrated version of this model can reproduce the differences in wealth holdings observed between stockholders and non-stockholders in the US.[1]

More recently, the heterogeneous stochastic returns approach to explaining wealth concentration at the top has been systematically developed and sharpened by Benhabib, Bisin, and various coauthors in a series of contributions. Rather than focusing on the specific source of returns heterogeneity, they take the latter as given and study instead the consequences of its presence for the right tail of the wealth distribution. In one key contribution, Benhabib et al. (2011) consider an overlapping generation model where households differ both in returns to human capital and in returns to wealth. Each household is endowed at birth with a rate of return to wealth and a return to human capital, drawn from independent distributions. Hence, there is persistence in returns to wealth (and human capital) within a generation. In addition, returns persist across generations and are independent of wealth. They show that, in this model, the stationary distribution of wealth has a closed form solution and is Pareto with a thick right tail. More importantly, it is the heterogeneity in returns and their intergenerational persistence that drive the thickness in the right tail of the wealth distribution, rather than the heterogeneity in returns to human capital. In other words, if return heterogeneity explains the upper tail of the wealth distribution, then the stochastic properties of labor income risk have no effect on the thickness of the tail of the wealth distribution (see their Theorem 1). The latter is instead increasing with the degree of heterogeneity in asset returns. And because the wealthy are on average those endowed with a high rate of return, their model endogenously generates a positive correlation between the individual persistent component of returns and the individual location in the distribution of wealth. Benhabib and Bisin (2016) review the theoretical and empirical debate about the drivers of wealth inequality, highlighting the specific role of returns heterogeneity. To quantitatively assess how far heterogeneity in returns to wealth can go in explaining the distribution of wealth and the degree of concentration in the tail (as well as the patterns of mobility in the wealth distribution), Benhabib et al. (2015) calibrate their overlapping generations model to US data. Besides heterogeneity in returns to wealth, the model allows also for heterogeneity in human capital and in savings rates due to a bequest motive that varies with wealth. Benhabib et al. (2015) estimate the distribution of returns to wealth and its intergenerational persistence to match several moments of the US wealth distribution and the degree of intergenerational wealth mobility. They estimate average returns to wealth of 3.4%, with a cross-sectional standard deviation of 2.7%; intergenerational persistence in returns to wealth is positive

[1] Guvenen (2009) discusses the implications of his model of returns heterogeneity and models of discount heterogeneity as in Krusell and Smith (1998).

but modest. This amount of persistent heterogeneity plays a key role in matching the tails: indeed, the top 1% wealth share predicted by the model is almost identical to the equivalent moment in the data (33.6% in the data, 34.1% in the simulated model). Shutting down this channel alone (by forcing returns to wealth to be the same across individuals) produces much smaller top wealth shares and wealth shares at the bottom of the distribution that are abnormally inflated. Hence, returns heterogeneity appears to be a key factor for matching the empirical wealth distribution.

Gabaix, Lasry, Lions, and Moll (2015) are not only interested in the amount of wealth concentration in the steady state, but also in the speed of the transition across steady states. They show that, while the Benhabib et al. model can explain the long thick tail of the wealth distribution, it cannot explain the speed of changes in tail inequality that we observe in the data. They suggest that one way to capture the latter is to allow for some "s*ize dependence*" - a positive correlation of returns with wealth in addition to *"type dependence"* (persistent heterogeneity in returns).

Despite their theoretical appeal, explanations of the level and the dynamics of wealth inequality and concentration based on a more sophisticated process for the returns to wealth suffer from some of the same problems as models that rely on heterogeneity in discount rates. How reasonable are the findings of heterogeneity and persistence in Benhabib et al. (2015)? Is there a correlation between wealth and returns to wealth that is compatible with the speed of tail inequality observed in the data? Unlike individual discount rates, however, individual returns on wealth have the great advantage that they can be observed. Yet, data requirements are substantial: what needs to be documented is that returns to wealth have an individual component; that this component persists across individuals of the same generation; that it correlates with wealth; and that it shows some intergenerational persistence. Documenting these facts requires much more than just observability. More generally, returns to wealth may show features that a calibrated exercise should account for. The goal of this paper is to provide a systematic characterization of these properties.

3 Data sources and variable definitions

Our analysis employs several administrative registries provided by Statistics Norway, which we link through unique identifiers for individuals and households. In this section, we discuss the broad features of these data; more details are provided in the Internet Appendix.[2] We start by using a rich longitudinal database that covers every Norwegian resident from 1967

[2] Available on authors' home pages.

to 2013. For each year, it provides relevant demographic information (gender, age, marital status, educational attainment) and geographical identifiers. For the period 1993-2013 - the period we focus on here - we can link this database with tax records containing individual information about asset holdings and liabilities (such as financial assets, private businesses, real estate, and debt), as well as a detailed account of the individual's sources of income (from labor and capital). The value of asset holdings and liabilities is measured as of December 31 of each year. While tax records typically include information about income, they rarely (if ever) contain information about wealth. In Norway, this happens because of a wealth tax that requires taxpayers to report their asset holdings in their tax filings.

The data we assemble have several, noteworthy advantages over those available for most other countries, particularly for the purpose of our study. First, our income and wealth data cover all individuals in the population who are subject to income and wealth tax, including people at the very top of the wealth distribution. Given the extreme concentration of wealth at the top, this is a key feature of the data.[3] In particular, steady-state wealth inequality and the speed of transition to a new steady state are likely to be sensitive to even a small correlation between returns and wealth. Moreover, the degree of correlation and heterogeneity may be higher (as we document in Section 6) at high levels of wealth. These features can only be captured if the data include people at the very top of the wealth distribution. Second, in our data set, most components of income and wealth are reported by a third party (e.g., employers, banks, and financial intermediaries) and recorded without any top- or bottom-coding. Thus, the data do not suffer from the standard measurement errors that plague household surveys, where individuals self-report income and asset components (as for instance in the US Survey of Consumer Finances) and confidentiality considerations lead to censorship of asset holdings.[4] Third, the Norwegian data have a very long panel dimension,

[3] Wealth is highly concentrated in Norway. In 2012, the top 0.1% owned about 10% of all net worth in the economy. In 2000, before US wealth concentration started to drift up (Saez and Zucman, 2016), the top 0.1% share was similar to the US, 14% vs. 16%.

[4] Clearly, if some assets are held abroad and not reported to the tax authority this will tend to understate wealth concentration since it is plausible that these assets are disproportionately held by the wealthy (Zucman, 2014). Using information on Norwegian taxpayers who disclosed assets held offshore following an amnesty in the early 2000's, Alstadsæter et al. (2015) show that the beneficiaries of the amnesty are the very wealthy. Out of 1419 individuals who disclosed assets offshore, essentially none is below the 99th percentile and 50% are among the wealthiest 400. The chances of having assets offshore increases sharply with wealth (at home) but is never larger than 12% (Zucman, 2016), suggesting that many wealthy may have no wealth offshore. Alstadsæter et al. (2015) argue that accounting for hidden wealth can increase the top 0.1% wealth share by 4 percentages points. For our purposes, the issue is whether the existence of wealth offshore tends to distort our measure of gross (of tax) returns on wealth. If wealth is held abroad to avoid domestic taxation, our estimates of gross returns should be little affected. If it is held abroad mostly to profit from more rewarding investment opportunities not available at home, then ours are conservative estimates of the heterogeneity in

which is indispensable to identify persistent heterogeneity in returns. Because the data cover the whole relevant population, they are free from attrition, except the (unavoidable) one arising from mortality and emigration. Fourth, unique identifiers allow us to match parents with their children. Together with the long panel dimension of the data, this is key to studying intergenerational persistence in returns to wealth. Finally, our data include information not only on listed stocks but also on private business holdings. Because private business holders have large stakes in their firm, this feature is important for pinning down the extent of heterogeneity in returns. And because, as we will document, stakes in private businesses strongly increase with wealth, this feature is also important for understanding the correlation between wealth and returns. Besides these unambiguous merits, our data also have some shortcomings: one, not surprisingly, is the measurement of the value of private businesses; another is the calculation of capital gains. We discuss them below and suggest remedies.

In our main analysis, we focus on returns to financial assets, which include bank deposits, bonds, mutual funds, money market funds, stocks of listed companies, and shares in non-listed companies - i.e., private businesses.[5] Below, we briefly describe the administrative tax records for wealth and income and how we construct our measure of wealth returns. Details of the mapping between the capital income tax component and the specific asset category are provided in the Internet Appendix.

returns and their correlation with wealth. If we drop people in the top 0.5% or 1% of the wealth distribution - where all wealth offshore seems to be sitting - our results are unaffected (see Section 5.3).

[5] The main components of wealth that are left out of our analysis are housing and private pension wealth (and their related returns). Differently from all other forms of wealth, private pension funds are not subject to the wealth tax and hence do not appear in the tax records. However, contributions to private pension funds are capped to USD 1,500 annually and this wealth component is negligible (in 2013, households' deposits into private pension accounts amounted to less than 0.1% of the total deposits into financial accounts). As for housing, we exclude it for two reasons. First, a practical one: housing wealth data before 2010 are incomplete. Second, a conceptual one: returns on owner-occupied housing, which are the main component of housing wealth for the bulk of the population, are given by the services they provide. Thus, the returns on owner-occupied housing would have to be imputed. This would introduce measurement error and most likely overstate wealth returns heterogeneity. Because housing returns are essentially uncorrelated with stock returns (Curcuru et al. (2009)), our estimates provide a conservative measure of returns heterogeneity. On the other hand, leaving housing returns out of the picture is unlikely to bias the correlation between returns to wealth and the level of wealth. In fact, for the period 2010-2013 (when housing data are complete and accurate), the correlation between financial wealth and total wealth (financial wealth + housing wealth - debt) ranges from 0.98 to 0.99.

3.1 Administrative wealth and capital income records

Norwegian households are subject to both an income tax and a wealth tax.[6] Each year, people are required to report their incomes and to provide complete information about wealth holdings to the tax authorities. Tax record data are available on an annual basis from 1993.[7] The collection of tax information is mostly done through third parties. In particular, employers must send information on earned labor income both to their employees and to the tax authorities; financial intermediaries where individuals hold financial accounts (such as banks, brokers, insurance companies, etc.) do the same for the value of the assets owned by the individual as well as for the income earned on these assets. For traded assets, the value reported is the market value. The fact that financial institutions supply information about their customer's financial assets directly to the tax authority greatly reduces the scope for tax evasion, and non-reporting or under-reporting of asset holdings is therefore likely to be negligible.[8]

3.2 Wealth aggregates and returns to wealth

For our analysis, we group assets into two broad categories, safe and risky assets (w^s and w^m, respectively), and map them in relation to the corresponding values of capital income from the tax returns. We define the stock of safe assets as the sum of cash, bank deposits, treasuries, money market and bond mutual funds, bonds and outstanding claims, and receivables. The stock of risky assets is defined as the sum of the market value of listed stocks, $w^{m,l}$ (held directly, $w^{d,l}$, or indirectly through mutual funds, $w^{i,l}$) and the value of shares in private businesses and other unlisted shares, $w^{m,u}$.

While listed stocks are reported at market value, private business wealth is the value of the shares in the private business that entrepreneurs report to the tax authority to comply

[6] Wealth in excess of an exemption threshold is taxed at a flat rate of around 1% during our sample period. The exemption threshold has been increasing over time and was in the later years around NOK 1.5 million for a married couple (and half that for a single person). Importantly, households' assets are reported and recorded even if they fall short of this threshold. Certain assets are valued at a discount in certain years when calculating taxable wealth. For instance, stocks were valued at 85% of market value in 2007. We adjust these discounted values back to market values before constructing household wealth.

[7] The individuals in a household are taxed jointly (i.e., married couples) for the purpose of wealth taxation, and separately for income tax purposes.

[8] For the last ten years of our sample period a separate shareholder registry includes information on financial wealth at the level of the single financial instrument owned by the investor. These data are analogous to those for Sweden available for the years from 1999 to 2007 and used by Calvet, Campbell, and Sodini (2007) and by Bach et al. (2015). Since our goal is to measure persistence in returns, we use the much longer registry containing the more aggregate measure of asset holdings.

with the wealth tax - what we label the "assessed" value. This value does not necessarily correspond to the "market" value of these shares, i.e., the realization price if they were to be sold in the market. Indeed, it excludes the net present value calculation of the firm or goodwill. The value of unlisted stocks held by the individual taxpayer is obtained as the product of the equity share held in the firm and the assessed value of the firm. (. Needless to say, the firm may have an incentive to report an assessed value below the "true" market value. On the other hand, the tax authority has the opposite incentive and uses control routines designed to identify firms that under-report their value. Consistent with this, the (log) assessed value is strongly correlated with the firm (log) book value (correlation 0.88, Figure IA.1 in the Internet Appendix) and, in more than 50% of cases, the assessed value exceeds the book value (which may be inconsistent with the goal of minimizing the tax bill). Medium- to large-sized firms (with a turnover above NOK 5 million, or USD 500k) are required to have their balance sheet reports audited by a professional auditing firm, reducing the scope for accounting misstatements.

Total wealth is:

$$w_{it} = w_{it}^{s} + w_{it}^{m,l} + w_{it}^{m,u}$$

As for capital income y_{it}, it includes income earned on safe assets i_{it} (the sum of interest income on bank deposits and the like, other interest income, interest on loans to companies and the yield from insurance policies), dividends (from both public equity and private businesses, d_{it}), and realized capital gains and losses from all equity (g_{it}). These figures are net of any commissions paid to intermediaries. Because dividends and capital gains/losses on listed and private firms are reported jointly for tax purposes, we cannot compute separately the return from public equity and private businesses. We hence observe:

$$y_{it} = i_{it} + d_{it} + g_{it}$$

Figure 1 shows the composition of the individual portfolio (i.e., shares of wealth in safe assets, listed stocks held either directly or indirectly through mutual funds, and the share in private businesses) for people in different parts of the wealth distribution. The lower panel of the figure zooms on the top of the distribution. Safe assets clearly dominate the asset allocation of people below median wealth. Public equity (especially through mutual funds) gains weight among people above the median and below the top 1%. The share in private business strongly increases with wealth above the 95th percentile and carries very large weight, close to 90%, for the top 0.01%.

3.3 Measuring returns to wealth

Consider an individual who invests her wealth $w_{it} = \sum_j w_{it}^j$ in various financial instruments $j = 1, ..., J$, each paying an annual return r_t^j. Suppose that the individual's portfolio is passive throughout the period, so that the investments deliver an aggregate income flow $y_{it} = \sum_j r_t^j w_{it}^j$. The individual's weighted average return to wealth could thus be estimated as:

$$r_{it} = \frac{y_{it}}{w_{it}} = \sum_j \omega_{it}^j r_{it}^j \tag{1}$$

where ω_{it}^j is the share of wealth invested in asset j.[9]

Despite the richness of the data, our measure of return to wealth has to account for three limitations. First, we only observe snapshots of people's assets at the end of each period, while observing the flow of income from capital throughout the period. Second, as mentioned above, the value of private businesses does not necessarily correspond to their market value. Finally, we only observe capital gains or losses when they are realized (i.e., when assets are sold), not when they accrue economically.

We account for these three limitations using different adjustment procedures. Consider the first problem. If assets are traded during the year, the income from capital will only reflect the part earned over the holding period before (after) the assets sales (purchases). This issue is most obvious in the case in which beginning-of-period wealth $w_{it} = 0$ but $y_{it} > 0$ due to saving taking place during the period. To account for this problem, we define returns

[9]We use realized returns to compute average returns to wealth. An alternative would be to rely on an asset pricing model, such as the CAPM, and attribute to an individual holding (say) a given stock the expected return predicted by the model using time series of stock returns. This is the method used by Bach et al. (2015). Its main advantage is that it increases the precision of the estimated mean returns as one can rely on long time series of market returns. This is particularly valuable when one has short time series of realized individual returns as in Bach et al. (2015), somehow less in our case given the long panel dimension of our data. Furthermore, the method has its drawbacks. First, the higher precision comes at the cost of imposing a pricing model, typically a CAPM and its (not undisputed) underlying assumptions (e.g., ability to borrow at a risk free rate, absence of trading frictions etc.). Second, (expected) returns attributed to an individual in a given year are affected by returns realized in future years. Third, because individuals holding a given asset are imputed the same average return independently of the holding period of the asset, differences in returns due to differences in ability to time the market are not captured by this method. The method is biased towards attributing systematic differences in returns across individuals to differences in exposure to systematic risk. The realized returns approach that we use is model-free and reflects all sources of heterogeneity across individuals relevant for generating returns to wealth.

as the ratio of income from capital and the average stock of wealth at the beginning *and* end of year, i.e.:

$$r_{it}^{A} = \frac{y_{it}}{(w_{it} + w_{it+1})/2} \tag{2}$$

We use this adjustment both when we compute the returns on safe assets, $r_{it}^{s,A} = \frac{i_{it}}{(w_{it}^{s}+w_{it+1}^{s})/2}$, and when we measure returns on risky assets, $r_{it}^{m,A} = \frac{d_{it}+g_{it}}{(w_{it}^{m}+w_{it+1}^{m})/2}$. Expression (2) will be our baseline measure of returns to wealth. The results are very similar if we weight beginning and end-of-period wealth differently rather than equally.

Our sample selection is also designed to reduce errors in the computation of returns. First, we drop people with less than USD 500 in financial wealth (about NOK 3000). These are typically transaction accounts with highly volatile beginning- and end-of-period reported stocks that tend to introduce large errors in computed returns.[10] Second, we trim the distribution of returns in each year at the top and bottom 0.5%. These are conservative corrections that, if anything, reduce the extent of return heterogeneity. Finally, we focus on the Norwegian population aged between 20 and 75 (although none of our conclusions are affected if we consider a younger or older sample). We focus on this age range to make sure that the financial decision maker is the holder of the assets and, thus, that we correctly identify his/her return fixed effect.

Consider now the second limitation. Our measure of wealth from risky assets is the sum of market-valued wealth $w_{it}^{m,l}$ and the assessed-value of private business holdings $w_{it}^{m,u}$:

$$w_{it}^{m} = w_{it}^{m,l} + w_{it}^{m,u}$$

Neglecting for the time being unrealized capital gains/losses, our measure of returns to wealth (2) is overstated if private business owners understate the value of the firm relative to what they would get if they were to sell it. There is no simple way to correct for this problem.[11] Thus, to check whether our results depend on private equity, we consider an alternative measure that excludes private equity owners, defined as:

[10]For example, an individual with a (close to) zero balance (say USD 150) at the beginning of the year and a (close to) zero balance at the end of year (say USD 150), perhaps because of above average December expenditures, and average balances during the year of USD 3,500 (NOK 30,000), would report capital income of USD 70 if the interest rate is 2%. But the return computed according to (2) would be 70/150=47%. This overstatement is less likely to happen for large accounts.

[11]In principle, one could use imputation methods based on market-to-book multipliers among listed firms and apply them to similar non-listed firms. The most serious problem is to find "similar" non-listed firms,

$$r_{it}^{B} = \frac{y_{it}}{(w_{it}^{s} + w_{it}^{m,l} + w_{it+1}^{s} + w_{it+1}^{m,l})/2} \tag{3}$$

The third potential limitation of our data is that we observe capital gains/losses when they are realized, rather than as they accrue year by year. As we show in the Internet Appendix this is not a serious issue if we are interested in measuring the average returns to wealth over the life cycle of an individual and if we observe enough realizations of the capital gains.[12]

We follow a more direct route to deal with unrealized capital gains. Still focusing on a sample that excludes private equity owners, we assume that capital gains on listed shares reflect the increase in value of the stock market, and assign the stock market's aggregate capital gains to investors on the basis of their beginning-of-period total stock market wealth. Define $M_t = \sum_{j=1}^{J} P_{jt} q_j$ the aggregate stock market value, where P_{jt} is the price of stock j and q_j its quantity; let the aggregate capital gain be $G_t = \sum_{j=1}^{J} \Delta P_{jt+1} q_j$. The accrued capital gain/loss from stock holding can hence be estimated as:

$$g_{it}^{a} = \frac{w_{it}^{m,l}}{M_t} G_t$$

And our final return measure is thus:

$$r_{it}^{C} = \frac{g_{it}^{a} + y_{it} - g_{it}}{(w_{it}^{s} + w_{it}^{m,l} + w_{it+1}^{s} + w_{it+1}^{m,l})/2} \tag{4}$$

Of course, the main disadvantage of this measure is that it assumes that the composition of people's stock market portfolio is the same, which mechanically reduces the extent of heterogeneity in returns.

From now on, we focus mostly on our baseline measure (equation (2)), which has the advantage of being based on information directly available from the tax records. In Section

given that listed firms are few and observationally different from non-listed firms. This procedure would probably yield serious measurement error, making it hard to separate true heterogeneity in returns from (possibly systematic) imputation errors.

[12]Let P_t be the market price of a stock. We can show that the average return over a holding period of T years of a stock that is sold at T is the same whether the average return is computed using the annual (gross) return $1 + r(t) = R(t) = \frac{y_t}{P_t} + \frac{P_{t+1}}{P_t}$, with capital gains computed on an accrual basis , or when the annual (gross) return is $R(t) = \frac{y_t}{P_t}$ if $t < T$ and $R(t) = R(T) = \frac{y_T}{P_T} + \frac{P_{T+1}}{P_1}$ if $t = T$, as in our data (see IA).

5.3, we show that our main findings are not sensitive to adopting the alternative measures of returns (3) and (4), labeled return B and return C, respectively.

All returns statistics we report are at the individual, not the household level. In this way, we account for the fact that while households form and dissolve, individuals can be observed as they cycle through different marital arrangements. When individuals are single, the above formula applies without modifications. When individuals are married, we assume that spouses share household wealth and capital income equally. This is consistent with Norwegian laws requiring family assets to be split equally between spouses in the event of divorce. In this case, we first compute the return to household wealth, and then assign this return and the per-capita household wealth to each spouse.

3.4 Descriptive statistics

Table 1 shows summary statistics for our data. For the sake of simplicity, we report statistics for the last year in our estimation sample (2013) and, for comparison, summary statistics for 1995 in the Internet Appendix (Table IA.1). Overall, our 2013 sample includes more than 3 million individuals. In Panel A, we report some basic demographic characteristics. The sample is well balanced between males and females, and with respect to marital status (50% are married). About 80% of the individuals in the sample have at least a high school degree. Finally, 12% of individuals have a degree (college or high school) with a major in economics or business, which may be indicative of possessing above-average financial sophistication. Panel B contains statistics describing wealth levels and composition. In 2013, 45% of Norwegian households had some risky assets in their portfolio. One in nine owned shares in a private business. Conditioning on having some assets invested in risky instruments, households invested on average 29% of their assets in those risky instruments. There is more concentration among private business owners. Conditioning on having private business wealth, 44% is held in the private business itself. The last five rows of Panel B provide information on wealth levels. Total financial assets are on average about USD 87,000. As expected, the distribution is extremely skewed, with a median of about USD 21,000, while the 90th percentile is more than USD 149,000.

The last panel of Table 1 reports summary statistics for the returns. In 2013, the average return on overall wealth was 3% (median 2%), and the standard deviation 4.9%. The average return on risky assets (5.8%) substantially exceeded that on safe assets (2.5%). Statistics for the whole period 1995-2013 are qualitatively similar, although, quantitatively, the differences are enhanced by weighting the returns by assets values. For example, the average returns are

3.2% and 4.8%, respectively, in the unweighted and value-weighted case (Table 1, panel C). Similarly, the average returns from risky assets are 3.5% and 6.9% in the two cases. As we will see, the larger difference in the value-weighted case is explained by the positive correlation between returns and wealth levels.

4 Stylized facts about returns to wealth

In this section, we establish a number of stylized facts about individual returns to wealth. In the next section, we provide a formal framework for modeling returns to wealth that will help to shed light on these stylized facts.

4.1 Returns to wealth are heterogeneous

Figure 2 shows the cross-sectional distribution of average returns to wealth in 2013 (the last year of our sample) for two groups: all households (top panel) and risky asset holders (bottom panel). We overlap the distribution of returns for our baseline measure (equation 2), and for measure C (equation 4), which imputes accrued capital gains for the sample that excludes private equity holders. The figures make clear that individuals earn markedly different returns (standard deviation 4.9%, Table 1, panel C). The median return is 2%, 100 basis points lower than the mean, implying a significantly right-skewed cross-sectional distribution of returns. The difference between the median return at the 90th and the 10th percentiles is about 200 basis points. When we account for unrealized capital gains, we naturally have longer tails and a greater incidence of negative returns, suggesting that most investors hold onto poorly-performing assets. Returns are more heterogeneous among risky asset holders.

But how much return heterogeneity should we expect? As a benchmark, consider a standard Merton-Samuelson framework in which all investors have access to the same investment opportunities. In this model, investors' optimal share of risky assets ω_{it} is a function of market expected excess returns, $E(r_t^m - r_t^s)$, the variance of risky assets σ_t^2, and investor risk aversion γ_i:

$$\omega_{it} = \frac{E(r_t^m - r_t^s)}{\gamma_i \sigma_t^2}$$

It follows that the individual realized return to total wealth is a weighted average of the risk-free rate and the market return:

$$r_{it} = r_t^s + \omega_{it}(r_t^m - r_t^s) \tag{5}$$

Heterogeneity in returns is induced by differences in risk aversion and thus in (compensated) risk-taking measured by the risky share.[13] Equation (5) suggests that conditioning on having the *same* share of risky assets in a portfolio, total returns on wealth should be similar across investors. That is, the cross-sectional standard deviation of returns, given ω_{it}, should be close to zero. In Figure 3, we again use data for 2013. We allocate individuals to different bins defined by the share of their wealth held in risky assets (from 0 to 1, in 0.01 increments), and within each bin, we compute the cross-sectional standard deviation of the individual returns (the top line in the figure). Not only is the standard deviation non-zero, but it also increases dramatically with the share of risky assets held in the portfolio. Interestingly, even at $\omega_{it} = 0$ (where individuals own only safe assets), the standard deviation of returns is positive. Thus, while the allocation of wealth (between risky and safe assets) does affect the extent of heterogeneity in the overall return to wealth, it is by no means the only driver (as we shall see more clearly in a formal controlled regression, discussed in Section 5). Note that some of the heterogeneity in Figure 3 may come from holding a private business with very idiosyncratic returns and possibly some measurement error. We hence repeat the exercise, focusing only on investors who do not own *any* shares in private businesses, i.e., individuals who only invest in safe assets and stock of listed companies (our return measure B in equation (3)). The evidence is similar, although, as expected, the extent of heterogeneity is lower. Also as expected, this shows that there is much more risk involved in holding private business wealth (see among others, Carroll (2000), Vissing-Jorgensen and Moskowitz (2002) and Kartashova (2014)).

Heterogeneity in returns is present in all years and its extent varies over time. Figure 4 plots the cross-sectional mean, median, and standard deviation of returns on wealth for all sample years. Heterogeneity varies markedly over time with a cross-sectional standard deviation of returns ranging between 0.08 in 2005 and just above 0.04 in 2009. Figure 5 shows the patterns for returns on safe and risky assets. Heterogeneity in returns to risky assets is much higher, much more volatile, and much less correlated with average returns than heterogeneity on returns to safe assets.

[13] Heterogeneity may also come from human capital, as in Viceira (2001). This is irrelevant for our argument, since in these models any extra "channel" affects only the share invested in risky assets, not the return earned on each asset class.

4.2 Returns covary with the level of wealth

The second stylized fact about returns to wealth is that they are strongly positively correlated with the level of wealth. Figure 6, Panel A, plots the median return to wealth for households in different percentiles of the wealth distribution using data for 2013. The differences in returns across wealth levels are large.[14] Moving from the 10th to the 90th percentile of the wealth distribution the median return almost quadruples - from 0.7% to 2.6% [15] - suggesting that the correlation between returns and wealth holdings can potentially have large effects on wealth inequality.[16]

Note that returns decline at the top 1% of the distribution. As the red (crossed) line shows, this is entirely accounted for by private business owners (who are over-represented in the top percentiles of the distribution, see Figure 1). It is plausible that private businesses apply dividend policies that are less generous (or more liable to tax avoidance strategies) than those of listed companies, resulting in lower realized returns. For example, they do not need to distribute dividends for signaling purposes.[17]

Panel B of Figure 6 shows that the positive correlation between returns and wealth holds for both risky and safe assets (and, again, the slight decline at the very top is entirely accounted for by private equity holders). This rules out that the correlation between returns

[14]One worry is that the positive correlation between returns on wealth may be spurious because the way we measure returns may overstate the returns of the wealthy if the latter exhibit a higher propensity to save out of wealth, implying that a higher than average proportion of capital income over a year derives from savings over the year rather than initial wealth. To illustrate, suppose that y_{it} is the sum of capital income out of initial wealth ($\sum_j r_t^j w_{it}^j$) and capital income out of savings added during the year ($\sum_j r_t^j s_{it}^j f_{it}^j$), where f_{it}^j is the fraction of year the extra savings in asset j remained invested and $s_{it} = \sum_j s_{it}^j$. Assume for simplicity $f_{it}^j = 1$ for all j. If returns are independent of wealth, one can show that $sign(\frac{dr_{it}}{dw_{it}}) = sign(\frac{d(s_{it}/w_{it})}{dw_{it}})$. Hence, if the propensity to save out of wealth s_{it}/w_{it} increases with wealth, one can find a positive association between computed returns and wealth even when there is none. To check whether this is a serious concern, we construct a measure of savings out of non-capital income as $s_{it} = w_{it+1} - w_{it} - y_{it}$ and study how the propensity to save out of wealth changes with wealth. We find no evidence that it rises with wealth, while finding some evidence that, in fact, it declines with wealth. Hence, if there is any bias in the correlation between returns and wealth it is likely to be *downward*.

[15]Notice that because our measure of returns includes capital gains/losses at realization and because dividends are not distributed on all stocks, median return on risky assets in a given year can be substantially lower than the cross sectional average returns on risky assets, and even lower than returns on safe assets. Average returns on risky assets are higher, positively correlated with wealth and higher than returns on safe assets.

[16]As noticed by Piketty (2014), "It is perfectly possible that wealthier people obtain higher average returns than less wealthy people.... It is easy to see that such a mechanism can automatically lead to a radical divergence in the distribution of capital".

[17]The drop in returns at the top 1% is present only after 2005, following a reform that made distributed dividends taxable. Before 2006, when dividends were tax-exempt, the relationship between returns and wealth levels is monotonically increasing throughout the distribution, including the top percentiles.

and wealth only arises because of participation costs in risky assets markets.

The correlation between returns and wealth is not specific to a given year. It appears to be a defining feature of the data, although its size does vary over time. To summarize these features in a simple way, Figure 7 plots the median returns for households at selected percentiles of the wealth distribution over the 20-year period for which we have data. It shows very clearly that households in higher percentiles of the wealth distribution enjoy higher returns in any given year of our sample; it also shows that the difference in returns between high and low wealth levels varies considerably across the sample.

In general, a correlation between returns and wealth may arise for several reasons. In Section 6, we discuss in detail various channels of influence. One simple explanation is that wealthier households have higher risk exposure. To check whether risk-taking is the only force behind the correlation documented in Figure 6, we compute a measure of the Sharpe ratio at the individual level, using the 20 years in which the individual is potentially observed in our data. The individual Sharpe ratio is defined as:

$$S_i = \frac{\frac{\sum_{t=1}^{T_i} \tilde{r_{it}}}{T_i}}{\sqrt{\frac{\sum_{t=1}^{T_i} \tilde{r_{it}}^2}{T_i} - \left(\frac{\sum_{t=1}^{T_i} \tilde{r_{it}}}{T_i}\right)^2}} \tag{6}$$

where $\tilde{r_{it}} = (r_{it} - r_t^s)$ is the deviation of the individual return to wealth from the return on the safe asset (the annualized real 3-month rate on Norwegian T-bills).

In Figure 8, we plot the average Sharpe ratio for each percentile of the wealth distribution in 1995 (the first year for which we have data). Clearly, wealthier individuals reap higher returns for *a given* amount of risk. Focusing on a sample of individuals who never own private equity leaves the picture unaltered, although it does reveal that risk-adjusted returns are slightly lower for this group, a non-surprising feature in light of the amount of non-diversified risk that private business investment entails, which will be confirmed in controlled regressions (Section 6). The same holds if we use the alternative measures of returns (IA, Figure IA.7).

Finally, we note that the extent of heterogeneity also covaries with wealth. To document this, we compute the cross-sectional standard deviation of returns for each percentile of the wealth distribution. In 2013, heterogeneity is relatively high at low levels of wealth and fairly flat between the 20th and the 70th percentile, when it starts increasing more sharply, resulting in a U-shaped relation between the cross-sectional standard deviation of returns and wealth. While the high heterogeneity in returns at the bottom is not a feature of all years, the correlation with wealth at the top is (Figure IA.3).

5 Modeling and estimating returns to wealth

In this section, we provide a formal statistical model of individual returns, estimate it and use the results to characterize the properties of the returns. In particular, we ask whether the heterogeneity that we have documented is just a reflection of idiosyncratic realizations that are quickly reversed or whether individuals differ persistently in the returns they earn on their wealth. In other words, we investigate whether individual returns to wealth have a permanent component. Persistence in returns, as argued by Benhabib et al. (2011, 2015), is essential for heterogeneity to be able to explain the fat tail of the wealth distribution as well as the fast transitions in wealth concentration at the top (Gabaix, Lasry, Lions, and Moll (2015)).

5.1 A statistical model of returns to wealth

We specify a linear panel data regression model for wealth returns:

$$r_{igt} = X_{igt}^{'}\beta + u_{igt} \tag{7}$$

where r_{igt} denotes the return to wealth for individual i belonging to generation g in year t. X_{igt} is a vector of controls meant to capture predictable variation in returns due to individual observables. To control for differences in returns induced by riskier asset allocations, the vector X_{igt} includes the lagged share of wealth invested in risky assets. In a world where individuals are fully diversified, and thus invest in the same portfolio of risky securities with return r_t^m (the return on the market portfolio), and have access to the same returns on safe assets r_t^s, the portfolio return would be: $r_{igt} = r_t^s + \omega_{igt}(r_t^m - r_t^s)$, where ω_{igt} is the share of individual i's wealth invested in the market portfolio. Hence, a regression of returns on time dummies, the individual risky assets share ω_{it}, and their interaction would absorb all the existing variation. If some individuals can also invest in private businesses, as in Quadrini (2000), Cagetti and De Nardi (2009; 2006) and Aoki and Nirei (2015), the return on wealth can be written as $r_{igt} = r_t^s + \omega_{igt}^{m,l}(r_t^m - r_t^s) + \omega_{igt}^{m,u}(r_{igt}^{m,u} - r_t^s)$, where $\omega_{igt}^{m,l}$ and $\omega_{igt}^{m,u}$ denote the share in listed stocks and private equity, respectively, and $r_{igt}^{m,u}$ is the individual-specific return on private businesses. In this case, time effects and the two portfolio shares will not exhaust variation in returns, which now have an individual-specific component. Accordingly, in equation (7), we control separately for the share of wealth invested in risky assets overall (i.e., $\omega_{igt}^{m,l} + \omega_{igt}^{m,u}$) and in private businesses.

The correlation between returns and wealth documented in Section 4.2 may arise because of fixed entry costs in risky assets that preclude participation by low wealth households. This is indeed consistent with extensive literature on limited participation costs (surveyed in Guiso and Sodini (2013)) and emphasized by Guvenen (2009) in the context of the wealth inequality debate). Moreover, there are important economies of scale in wealth management that may result in lower fees or directly in higher returns as the size of the investment increases. In addition, recent work by Kacperczyk et al. (2014) (building on earlier ideas by Arrow (1987)) suggests that wealthy investors are more "sophisticated" than retail investors, for example because they have access to better information about where the market is heading, and hence reap higher returns on average (for a given risk). To capture any *direct* correlation between returns and wealth due to wealth size effects, we add to the specification a full set of dummies for the individual wealth percentiles computed using lagged wealth values (to avoid spurious correlations arising from the wealth accumulation equation, which implies that next year's wealth is positively correlated with current returns). We enrich the vector X'_{igt} with year fixed effects and their interaction with risky shares (to capture aggregate variation in returns) and age dummies (to pick life cycle effects in returns).

While the role played by observable characteristics is important, the focus on the error term u_{igt} is even more so. We model the error term u_{igt} as the sum of an individual fixed effect and an idiosyncratic component, which may possibly exhibit serial correlation. Hence:

$$u_{igt} = f_{ig} + e_{igt}$$

The fixed effects f_{ig} capture persistent differences across people in average returns that are not compensation for risk or a reflection of systematic differences in wealth levels. These may arise from differences in the ability to manage the portfolio or to identify and access alternative investment opportunities - including persistent differences in private businesses' productivity not already captured by the wealth level controls. The error term e_{igt} measures unsystematic idiosyncratic variation in returns reflecting "good or bad luck". This representation allows as to decompose idiosyncratic heterogeneity in returns to wealth as $var(u_{igt}) = var(f_{ig}) + var(e_{igt})$.

Because we observe several generations in our data, we can study intergenerational persistence in the fixed heterogeneity of returns by estimating:

$$f_{ig} = \rho f_{ig-1} + \eta_{ig}$$

Thus, our statistical model is able to isolate the type of heterogeneity in returns - persistent heterogeneity not due to differences in risk-taking and investment scale - whose properties (cross-sectional variance and intergenerational persistence) can in theory explain the thickness in the distribution of wealth as shown by Benhabib et al. (2011). The aforementioned variance decomposition into $var(f_{ig})$ and $var(e_{igt})$, together with intergenerational persistence in f_{ig}, plays a key role in the design of optimal capital income taxation (Shourideh (2014)).

5.2 Estimation results

Table 2 shows the results of the regression (7). The dependent variable is our baseline measure of returns on total wealth in year t (equation (2), expressed in percentage points). The first column shows estimates from a pooled OLS regression, without the fixed effects but adding a number of individual characteristics, some of them time invariant, to gain some intuition on the role played by covariates. Besides the controls already cited (the lagged shares in risky assets and private business out of total wealth, lagged wealth percentiles, time and age dummies), heterogeneity in wealth returns is captured by a set of demographics (gender, municipality fixed effects, number of years of education, a dummy for economics or business education, employment, and marital status dummies). The main sample comprises more than 50 million observations.

The estimates show that males have – *ceteris paribus* – a lower average return on wealth, but the effect is economically negligible (2.8 basis points). Returns are correlated with general education and with specific education in economics or business. An additional year of formal schooling raises returns by 3.4 basis points (i.e., completing a college degree results in about a 13.6 basis points higher average return), while having taken an economics or business education is associated with 11 basis points higher returns. Because education is a permanent characteristic, its effect cumulates over time. A systematic difference in returns of 25 basis points enjoyed by economics college graduates (the sum of the effect of completing college education and majoring in economics or business) can produce a difference in wealth at retirement of 10.5% over a working life of 40 years. This effect comes in addition to any effect that education may have on returns to wealth by twisting the portfolio allocation towards riskier and more remunerative assets (e.g., by raising the stock of human capital and inducing a greater exposure to equity shares, as in Merton (1971)). This finding is consistent with Bianchi (2015) and von Gaudecker (2015), who find a positive effect of a measure of financial literacy on the return to investments among French and Dutch investors, respectively, but with reference to a specific asset. It also supports the results of Lusardi et al. (2015), who study the effect of financial knowledge on returns to wealth and assets at retirement within a life cycle model calibrated on US data.

Not surprisingly, portfolio shares in risky assets and in private businesses have both a positive and large effect on the return to wealth, with the effect of the share invested in private businesses being significantly larger than the effect of the share in risky assets overall, as implied by calibrated portfolio models that allow for investment in private businesses (e.g., Heaton and Lucas (2001)). Increasing the share in listed stocks by 30 percentage points (about the move from the risky share of a non-participant in the stock market to that of the average participant) increases the return to wealth by roughly 20 basis points. Increasing the share in private businesses by the same amount is associated with a much larger increase in returns on wealth of 188 basis points. This finding is consistent with the idea that, because private business wealth is highly concentrated, it yields a premium to compensate for idiosyncratic risk. This runs contrary to Vissing-Jorgensen and Moskowitz (2002), who, using data from the US SCF, find no evidence that private businesses earn a premium relative to public equity; but it is consistent with the results of Kartashova (2014) who documents the existence of a private equity premium using the same survey, but extending the sample to the more recent waves. Overall, these estimates suggest that part of the observable heterogeneity in returns reflects compensation for the risk involved in investing in listed stocks or for the idiosyncratic risk of owning private businesses. Estimated time fixed effects, though not shown, are always significant, as are age dummies and wealth percentile dummies. However, the direct contribution of the wealth dummies to the average returns is modest and flat up to median wealth, and increasing moderately for wealth above median. Hence direct wealth size effects on returns are small.

Overall, observable characteristics explain only about 8% of the variance in individual returns to wealth. This limited fit (or the larger role of unobservable heterogeneity) is remarkable because, as noted, the canonical portfolio model with fully diversified risky portfolios would imply that, controlling for time variation in returns, all heterogeneity in returns should be explained by differences in the risky shares.

The second column modifies the specification by replacing the risky shares with their interaction with time dummies. This more flexible specification captures differential effects of the risky share on individual returns as the aggregate component of returns varies. In addition, the interaction between the share in private equity and the time dummies captures variation in individual returns due to tax-induced changes in incentives to distribute corporate dividends following the 2006 tax reform.[18]The fit of the model improves (the R^2 increases from

[18]As noted above, in Norway until 2005 distributed dividends were essentially exempt from tax (except for a one-time 11% tax in 2001) while capital gains were taxed at the same 28% rate as retained profits. A reform passed in 2006, but anticipated since at least 2001, introduced taxation of distributed dividends at

0.079 to 0.117), but the size and significance of the other effects are otherwise unchanged.

The third column adds the individual fixed effects to the specification in column 1.[19] As usual, the effect of time-invariant characteristics (such as gender or education) is no longer identified. The effect of listed and non-listed asset shares is now identified among individuals who change portfolio composition over time. The effect of the share in listed stocks is now larger and that of private equity smaller: the effect of a 30 percentage points increase in the share of listed stocks results in a 31 basis point increase in the return to wealth, while an equal increase in the share in private business is associated with an increase in returns of 134 basis points. The key result is that the individual fixed effects improve the fit substantially: compared to column (1), the R^2 of the regression triples, implying that returns have an important persistent individual component.[20] The fourth column uses the specification in column 2, allowing for interactions between the time effects and the risky shares. With this flexible specification and the individual fixed effects, the model can explain slightly more than a quarter of the variance in individual returns to wealth. The last column uses a finer control for risk-taking by controlling separately for the share of stocks held indirectly through mutual funds, the share of directly held stocks, and that in private businesses. Investment in single stocks carries a higher premium than investment through mutual funds, possibly because the selected stocks are highly correlated with the stock market.[21] Using these richer controls for risk-taking leaves results qualitatively unchanged.

From $u_{igt} = f_{ig} + e_{igt}$, additional persistence in returns may in principle come from e_{igt}. To check whether this is the case, we look at the auto-covariance structure of the residuals in

a flat 28% rate (the initial tax rate applied on earned income), at least for the part of returns on equity exceeding a risk free return of 3%. The corporate tax rate was kept at 28%, although dividends and capital gains received by corporations were made tax exempt (see, Alstadsæter and Fjærli (2009)). The interaction between the time effects and the share of wealth in private equity captures the fact that private business owners may have timed the distribution of dividends in response to changes in tax incentives.

[19]Because the model includes age and time effects, the individual fixed effects also capture cohort effects, posing a well known identification problem arising from the linear relation between age, time and year of birth. We deal with this issue by using the Deaton and Paxson (1994) restriction and impose that time effects sum to zero once the variables have been detrended. Since our data cover several years, we are able to separate trend and cycle, and thus feel reasonably confident about the decomposition of age, time and cohort effect based on this restriction (Deaton (1997)). Notice that the increase in the fit when adding the individual fixed effects is not due to fixed effects capturing (mostly) cohort effects. In fact, the latter are captured by the age dummies in the specification in columns 1 and 2, but the fit is modest.

[20]Some of the R^2 increase arises mechanically from the addition of the fixed effects themselves. To check the importance of this, we created a "fake" number of IDs equal to the number of individuals in our sample. The equivalent of the regression of column (2) with the fake IDs gives an R^2 of 0.15 (instead of 0.117); while in column (4) the R^2 is 0.27. Hence, the "mechanical" part can explain only about a quarter of the overall increase in returns predictability.

[21]This is consistent with Bach et al. (2015), who find that wealthy individuals tilt their portfolio towards single stocks that are highly correlated with the stock market, obtaining higher average returns.

first difference computed from the specification in column (4), i.e. $E(\Delta u_{igt} \Delta u_{igt-s})$ for $s \geq 0$ (since taking first differences of the residuals removes the fixed effect, i.e., $\Delta u_{igt} = \Delta e_{igt}$). We find that these moments are minuscule and economically indistinguishable from zero for $s \geq 2$, consistent with e_{igt} being serially uncorrelated (see Figure IA.4 in the IA).

5.3 Robustness

Table 3 shows estimation results when we drop the private equity holders (return measure B, equation 3), and when we use a measure of returns that includes unrealized capital gains (return measure C, equation 4) for the sample that excludes private business holders. In both cases, we report the OLS and the fixed effects regressions. The results are qualitatively unchanged: the sign and size of the covariates coefficients - gender, years of education, the dummy for economics or business education (see columns 2 and 3) - are the same as when using the baseline measure. Furthermore, as in Table 2, the fixed effects prove to be critical: adding them substantially increases the fit of the regressions. Most importantly for our purposes, the individual fixed effects obtained using our baseline measure of returns or using the alternative measures are strongly correlated. The rank correlation between the baseline fixed effects and those using measure B of returns is 0.94 and that between the baseline and measure C fixed effects is 0.76 (see Table I2.4 and Figure IA.5), ensuring that the properties of persistent heterogeneity that we identify and discuss below are robust to the way returns to wealth are measured. The last two columns show results when we drop people belonging to the top 0.5% of the wealth distribution from our baseline specification to account for distortions from wealth held offshore. We find no effect on our estimates. Since the fixed effects of individuals who are present in the sample for a few yers may be imprecisely estimated, the last column runs the regressions on individuals who have been observed for at least 15 years. Again, estimates are unaffected.

5.4 Persistent heterogeneity

Figure 9 plots the empirical distribution of the individual fixed effects (from the estimates in column 4, Table 2), measured as the deviation from the overall mean of returns over the sample period. The distribution has a long right tail (a skewness coefficient of 1.87)[22] and is quite dispersed, with a standard deviation of 2.8 percentage points and a 90th-10th percentile

[22] For visual clarity we collapse the frequency mass of fixed effects above the 99.5 and below the 0.5 percentile of the distribution.

difference of 6.4 percentage points. It also shows considerable excess kurtosis (15.5, Table 4, col. 1).

One interesting question is whether the persistent component of wealth returns is associated with observable characteristics that, a priori, can be deemed economically relevant. Figure 10 plots the distribution of estimated fixed effects for business owners and non-owners (first panel); top vs. bottom wealth groups (second panel); individuals with high vs. low education (third panel); and people with and without an economics or business degree (last panel). Because the first two characteristics (being a business owner and being at the top of the wealth distribution) may vary over time, the non-owners and those in the bottom wealth groups are defined using indicators for "never being a business owner " and "never being in the top 10% of the distribution". In all cases, there is substantial heterogeneity in estimated fixed effects within each group. Group differences are also economically significant. Business owners exhibit a distribution of persistent returns that is much more spread out and shifted to the right (standard deviation of 3.27 compared to 2.56 for non-business owners). This is consistent with owners of private businesses facing more heterogeneous investment opportunities and higher returns on capital. Returns are heterogeneous both among the wealthy and among people at the bottom of the wealth distribution. But the distribution of the permanent component of returns is more spread out and returns are on average higher among the wealthy, with differences in the mean and spread becoming larger at the very top of the wealth distribution (we delve deeper into the relation between wealth and returns in the next section). Individuals with more schooling have a less dispersed distribution of persistent returns to wealth, while those with a degree in economics or business face both more dispersion in persistent returns and a distribution that is more shifted to the right, which is consistent with a positive correlation with education. Table 4 shows summary statistics for the distribution of the return fixed effects for the total sample and for various population subgroups (using our baseline measure). Columns 6 and 7 of Table 4 show summary statistics for the distribution of the fixed effects when we drop the private equity holders from the estimation (return measure B, equation 3), and when we use a measure of returns that includes unrealized capital gains (return measure C, equation 4). In both cases, the estimated standard deviation is very similar to the baseline, but both distributions are somewhat more skewed and exhibit more kurtosis. The last column shows statistics when the sample includes individuals that are present for at least 15 years. Not surprisingly, heterogeneity in returns fixed effects is somewhat reduced (standard deviation 2.5 compared to 2.8 in the benchmark specification) but the general properties are very similar to the benchmark specification.

5.5 *Variance decomposition*

Our error term representation allows as to decompose idiosyncratic variation in returns to wealth as $var(u_{igt}) = var(f_{ig}) + var(e_{igt})$. As shown by Shourideh (2014), the relative importance of $var(f_{ig})$ and $var(e_{igt})$ drives the optimal taxation of capital income, particularly its progressivity. In Table 5, we report the estimated variances of the two components for different samples and specifications. In our baseline specifications (Table 2, col. 4), our estimates imply that $var(f_{ig})/(var(f_{ig}) + var(e_{igt})) = 0.24$, i.e., persistent differences in returns across individuals account for 24% of the residual variance in returns, where both $var(f_{ig})$ and $var(e_{it})$ have been computed by pooling estimated residuals for all years.

6 Returns and wealth: the role of observed and unobserved heterogeneity

In this section, we use the estimates in Table 2, column 4, to investigate what drives the correlation between returns and wealth firstly documented in Section 4.2. We compute the components of the mean predicted return for each percentile of the wealth distribution (pooling across all years): $E(r_{igt}|P_w) = E(X'_{igt}\beta|P_w) + E(f_{ig}|P_w) + E(e_{igt}|P_w)$, where $E(z_{it}|P_w)$ denotes the mean of variable z_{it} conditional on wealth percentile P_w; X_{igt} is the vector of observables in the regression (7), f_{ig} the fixed effect, and e_{igt} the estimated residual. A correlation between average returns and wealth can arise because wealth is correlated with some of the observed determinants (including any direct effect of wealth) or because it is correlated with the unobserved heterogeneity.

Figure 11 (top panel) shows the elements of the decomposition pooling data for all years. Average returns increase with wealth, and at a higher speed at the very top of the distribution. Observed heterogeneity plays an important role in explaining the *level* of average returns at each wealth percentile, but its contribution is *declining* over a very wide range of the wealth distribution, up to the 90th percentile. Hence, observables cannot explain the positive correlation between returns and wealth for levels of wealth below the 90th percentile. However, the role of observables becomes key to explaining the correlation between returns and wealth at the very top. In particular, the bottom panel of Figure 11 shows that a major role is played by the share of wealth held in risky assets, whose contribution increases very rapidly with wealth. This implies that part of the correlation between wealth and returns at the top reflects compensation for risk, as also argued by Bach et al. (2015). In contrast, the role of return fixed effects is to shape the correlation between wealth and returns throughout the distribution, not just at the top. The relation is roughly linear up to the very top percentiles of the wealth distribution, with an increase in the average return of 28 basis points for every 10 percentile increase in wealth. Simple calculations illustrate the separate importance of observables vs. fixed effects in driving the correlations between wealth and returns. The average return to wealth increases by 210 basis points as wealth increases from the 50th to the 95th percentile; 81% of this increase (171 basis points) is explained by the increase in the average fixed effects and much less (50 basis points) by compensation for (more) risk-taking. On the other hand, 67% of the 248 basis points increase in the average return that occurs as wealth climbs from the 95th to the 100th percentile reflects compensation for greater exposure to risk, and only 37% (91 basis points) the increased values of individual fixed effect. What is left is very little and it is explained by variation in the other observables and in the residual. This decomposition is very similar if we use the specification in Table 2 column 5, as well as when using the estimates with the alternative definitions of returns. However, dropping the private business holders from the sample results in a much lower correlation of returns with wealth at the top, a smaller contribution by risk-taking to explaining the correlation and a larger role played by the individual fixed effects.

To show from a different perspective that the relation between returns and wealth is only partially a reflection of higher risk-taking (and related compensation for it) at high wealth levels, Table 6 shows regressions of the individual return Sharpe ratio, computed as in equation (6), on wealth measured in 1995 (at the beginning of our sample period) and other observable characteristics. Controlling only for initial wealth (column 1), risk-adjusted returns are strongly increasing with the individual wealth percentile. Adding individual controls such as age and its square, education and its square, and a set of dummies for the share of wealth invested in private businesses causes small changes in the relation with wealth. Interestingly, more educated individuals display higher Sharpe ratios, as do individuals with a business or economics degree, suggesting that ability plays a role in explaining differences in risk-adjusted returns. Compared to those with investments in public equity (the excluded category), holders of private businesses attain lower Sharpe ratios (particularly those with intermediate exposure to it). This feature reflects poorer diversification of idiosyncratic risk among private business holders that is only partially compensated by higher monetary returns or that is compensated by non-monetary, and thus unmeasured, returns. The broad findings are unchanged when we exclude private equity holders (third column).

In sum, the decomposition results in Figure 11 and the estimates in Table 6 suggest that the positive correlation between returns and wealth is primarily driven by a positive correlation between wealth and persistent unobserved heterogeneity, possibly capturing compensation for ability to generate returns. Compensation for risk plays a role, but its importance varies over the spectrum of the distribution. Below the 95th percentile, the correlation between returns and wealth is almost entirely due to unobserved heterogeneity; above it, 2/3 of the correlation reflects compensation for greater exposure to risk, while the rest can be traced to unobserved heterogeneity.

7 Intergenerational persistence in returns to wealth

Because the Norwegian data contain both an individual identifier and a family identifier, it is possible to link individuals across generations. To focus on a sharper case, we look at fathers and children (sons and daughters). Our regression analysis provides us with an estimate of individual fixed effects for almost 2 million father-child pairs. This allows us to test whether wealth returns are correlated across generations, and whether such correlation is explained by the persistent component or by observable characteristics that may be shared by both generations.

We start by ranking parents according to their financial wealth, the return to it, and the persistent component of their returns (fixed effect). In principle, it would be best to relate parents' variables and children's variables when they are of the same age. Unfortunately, our panel is not long enough to meet this requirement. To control for the fact that parents and children are observed when they are at different points of their life cycles, we compute rank percentiles of the relevant distribution with respect to the birth cohort the individuals (father and children) belong to. Next, for each percentile of the parents' variable of interest (wealth, returns, or return fixed effect), we compute the average percentile occupied by their child in the distribution of the same relevant variable in the same year (again, relative to their year of birth cohort).

Figure 12 plots the rank correlation between the wealth percentile of the parents and that of the child (top left panel), between the returns percentiles (top right panel), and between the permanent components of these returns (bottom panel). The intergenerational rank correlation is very similar when using actual returns and when using the persistent component of returns (fixed effects). In the first case, a linear regression of the father's percentile rank onto the average child's percentile rank has a coefficient of 0.085 with a standard error of 0.002, while in the second, the coefficient is 0.10 with a standard error of 0.004. This suggests that most of the intergenerational correlation in returns to wealth is a reflection of the individual persistent component. Interestingly, there are important non-linearities: the relation turns negative at the very top of the parents' permanent returns. Children of extraordinary parents in terms of returns to wealth over their life cycle quickly revert to the mean.

By comparison, a regression of the father's wealth percentile rank on the average wealth percentile rank of the child has a coefficient of 0.29 (s.e. 0.006) (the dashed line in the graph).[23] Thus, the intergenerational correlation in returns is three times weaker than that in wealth. Furthermore, while the intergenerational correlation in returns weakens or even turns negative at the top of parents' returns, the opposite is true for the correlation in wealth across generations, which becomes stronger at the very top of the parents' wealth distribution. For the very wealthy, the pattern of intergenerational correlation in returns facilitates social mobility, while that in wealth impedes it.

Some of the intergenerational correlation in returns may come from parents and children sharing a private business (or family firm). It is also possible that children imitate the investment strategies of their parents, or that they inherit traits from their parents that matter for returns (such as preferences for risk or investment talent). However, given the positive correlation between returns and wealth, all or part of the intergenerational correlation in returns documented in Figure 12 may simply reflect the intergenerational correlation in wealth or aggregate shocks to returns. The positive correlation between the child's and the father's return fixed effects (Figure 12, second panel) rules out the second possibility, but not the first. To deal with this, we report controlled regressions of children's returns on fathers' returns. We show the results in Table 7 using children's and fathers' return percentiles; the results are similar if we use returns directly. The first column has no controls; as already shown in Figure 12, the slope coefficient is small. All the other regressions include both children's and fathers' wealth percentile dummies. Adding wealth controls and age dummies lowers the slope of the intergenerational relation, but it remains positive and significant. The results are unaffected when individual controls are added (third column). Including individual fixed effects in the last column flattens the relation even further, but considerably raises the fit (the R^2 increases from 0.06 to 0.36), which is consistent with the intergenerational correlation being driven by the permanent component of returns. The results are confirmed when private business owners are dropped from the sample and when using the alternative definitions of

[23] While the literature on intergenerational income mobility is vast (see for instance Chetty et al. (2014)), that on wealth has been limited due to wealth information being less frequently available to researchers, Charles and Hurst (2003) being an exception. More recently, a growing number of papers study intergenerational mobility of wealth using Scandinavian data, see for instance Boserup, Kopczuk, and Kreiner (2014); Adermon, Lindahl, and Waldenstrom (2015); Black, Devereux, Lundborg, and Majlesi (2015); Fagereng, Mogstad, and Rønning (2015). None of these papers study intergenerational correlation in returns to wealth.

returns (Internet Appendix Table IA.4). Intergenerational persistence is also detected if we use Sharpe ratios of fathers and children (Table 8), confirming that it is risk-adjusted returns that correlate across generations.[24]

Overall, our data suggest substantial persistence and heterogeneity in returns within a generation but mild persistence across generations. This result is similar to that found by Benhabib et al. (2015) (although their estimate is imprecise). In their calibration exercise, only mild intergenerational persistence in returns is required to match the wealth concentration data. In our case, with a substantial amount of statistical power, we find in the data an economically small but statistically significant degree of persistence.

8 Implications for models of wealth inequality and returns heterogeneity

Papers on wealth inequality in the spirit of Benhabib et al. (2011) face the problem that the key driver of wealth concentration at the top - the moments of the distribution of the persistent component of returns and the degree of intergenerational correlation - are typically unknown. Our estimates provide ready-to-use estimates that can be used to calibrate these models. Given the distribution of returns, these models imply a positive relation between the average fixed effects and the wealth percentile, which can be recovered from the calibrated model. Our decomposition in Section 6 enables estimation of this relation in the data. Table 4 shows correlation coefficients between average fixed effects and wealth percentiles, and the slope parameter of regressions of the average fixed effects on wealth percentiles. Using these data and the estimates of intergenerational persistence in Section 7, a summary characterization of the distribution of the return fixed effects (ignoring moments higher than the second) is $f_{ig} \sim (mean = 3.2\% + 0.028(P_{iw} - 50), Sd = 2.8\%)$ and $f_{ig} = const + 0.05 f_{ig-1}$, where P_{iw} is the wealth percentile of individual i and 3.2% is the average return on wealth over the sample period. This characterization is qualitatively consistent with the idea that those with a greater ability to generate persistently higher returns, measured by the fixed effects, will end up accumulating more wealth - the mechanism emphasized by Benhabib et al. (2011).[25] Alternatively, one can choose the value of the parameters of the distribution of

[24]Table IA.5 in the Internet Appendix shows the transition matrix when we allocate individuals (fathers and children) according to their returns fixed effect in quintiles (relative to their year of birth cohort). There is similar persistence across different parts of the distribution. A child born to a parent in the top quintile has a 24 percent probability of also being in the top quintile (relative to individuals of his age), and a 17 percent probability of slipping into the bottom quintile.

[25]In the IA we also study how fixed effect heterogeneity varies by wealth percentile. We find a J-shaped pattern. However, the rapid increase at the top is entirely accounted for by private equity holders. If we drop them, heterogeneity in fixed effect even declines above the 90th percentile (Figure IA.6).

individual persistent returns (mean, standard deviation, and intergenerational correlation) to match the moments of the wealth distribution as done by Benhabib et al. (2015) for the US, which can then be confronted with our data-based finding. Benhabib et al. (2015) estimate average returns to wealth of 3.4% with a cross-sectional standard deviation of 2.7% and a tiny intergenerational persistence; these parameters are very close to our estimates based on the Norwegian data. They also find that the slope of the relation between the wealth percentile and the corresponding (average) individual permanent return to wealth is about 0.01, somewhat flatter than the one we estimate but within the same broad range. The remarkable consistency between our data-based evidence and the calibration-based evidence of Benhabib et al. (2015) suggests that models that rely on returns heterogeneity are likely to succeed in explaining the concentration of wealth.

9 Discussion and Conclusions

The properties of the returns to wealth that we have documented in this paper have potentially far-reaching implications for several other strands of the current debate on wealth inequality, besides the one discussed in Section 8. Here, we discuss three and highlight some new lines of research that our findings call for.

Measurement of wealth trends Saez and Zucman (2016) have revived the debate around the medium-term dynamics of the shares of wealth at the very top of the distribution. Lacking time series of comprehensive data on wealth holdings for the US similar to those available for Norway, they use tax records of income from capital to obtain underlying wealth figures and trends in top wealth shares. Wealth is imputed by capitalizing the capital income components using the average rate of return of the corresponding component. The capitalization methods may overstate the amount of wealth concentration if returns are heterogeneous within asset classes and if returns correlate with the level of wealth - two features that our paper documents. Moreover, trends in wealth concentration and inequality may depend on whether the extent of return heterogeneity and the correlation between wealth and returns change over time (which is another feature of the data). In Fagereng et al. (2016b), we use the Norwegian data to contrast inequality measures based on actual wealth with measures obtained from imputed wealth using the capitalization method, and document that heterogeneity of returns can in principle generate significant deviations between measures of inequality based on imputed and actual wealth, although in practice the bias appears contained.

Inequality in income and inequality in wealth Some countries with low levels of income inequality display levels of wealth inequality that are similar to those of countries with much higher levels of income inequality. For example, in Denmark and Switzerland the income share of the top 10% is around 25%, much lower than the corresponding 40% share in the US.[26] However, the top 10% wealth share is 76% in Denmark and 71% in Switzerland, which is even higher than the figure for the US (70%, Davies et al., 2011). The comparison between the US and Norway is even more striking: before wealth concentration starts drifting in the US, using comparable definitions over the years 1993-2000, the top 0.1% income share in Norway is on average around 3% and the top 0.1% wealth share 12.5% on average; on the other hand, over the same period the top 0.1% income share in the US is 7.8% - more than twice that in Norway, while the average top 0.1% wealth share is as large as in Norway (13.6%).[27] Heterogeneity in returns to wealth may solve the puzzle of why a country with much lower concentration of income at the top than another country may nevertheless have similar or even higher wealth concentration at the top. Surveying the theories of skewed wealth distributions, Benhabib and Bisin (2016) revisit and put in a novel perspective two theorems, one by Grey (1994) and another by Kesten (1973). Grey's theorem asserts that, in an economy with homogeneous returns to wealth and heterogeneous income, the wealth distribution inherits the properties of the income distribution, including the thickness of its tails. Kesten's theorem asserts that, under certain conditions, heterogeneity in returns to wealth can generate a thick-tailed and skewed wealth distribution even when the distribution of returns is neither skewed nor fat-tailed, and without requiring income heterogeneity. Models that rely on heterogeneity in returns to explain wealth inequality rely on the latter property. These two theorems imply that the tail of the wealth distribution is determined either by the tail of the earning distribution or by the stochastic properties of returns, not both. This is relevant to solving the above puzzle. If returns heterogeneity determines the tail, as implied in Benhabib et al. (2015), provided the degree of heterogeneity in returns is similar across countries (not an unreasonable requirement in light of the evidence in Section 8), one can observe marked differences in income concentration and still see a similar level of concentration of wealth at the top.

[26]See World bank: World Development Indicators: http://data.worldbank.org/ and the World Wealth and Income Database: http://www.wid.world/#Database.

[27]Top income shares for the US and Norway include capital gains and are taken from the Wealth and Income Database: http://www.wid.world/#Database, see also Aaberge et al. (2016); the US top wealth shares are taken from Saez and Zucman (2016), Figure 6B. For Norway, we compute top wealth shares from the registry data using definitions that are as close as possible to those of Saez and Zucman (2016).

Taxation of capital income and taxation of wealth Our findings also relate to the emerging literature on capital income and wealth taxation. In models with heterogeneous returns, taxing income from capital and taxing capital can have important efficiency implications, as shown by Guvenen et al. (2015). In fact, holding tax revenue constant, replacing a capital income tax with a wealth tax tends to widen the after-tax heterogeneity in returns. Intuitively, taxing capital income disproportionately reduces the after-tax return of individuals with high rates of return; hence, moving to a wealth tax system redistributes the burden of taxation from high-return to low-return individuals. This may give rise to efficiency gains through two channels: because capital is reallocated to high-return individuals, and because the higher return of high-return individuals can motivate the accumulation of higher savings. The importance of these efficiency gains from tax reallocations critically depends on the nature of the heterogeneity: whether it is persistent and its extent. Our results inform both dimensions; the extent of measured persistent heterogeneity suggests that the efficiency concerns of capital income taxation raised by Guvenen et al. (2015) are of practical relevance. Furthermore, when returns also have a transitory idiosyncratic component in addition to the permanent one, the relative importance of the two sources of cross-sectional heterogeneity are relevant to the progressivity of capital income taxation (Shourideh, 2014). Our variance decomposition (Table IA.3) provides information that can be used to empirically assess how far the actual taxation of capital income is from the optimal level.

Other amplifying mechanisms for wealth inequality Besides uncovering and measuring permanent heterogeneity and persistence in returns across generations, whose role in theoretical models of wealth inequality is only now starting to be fully appreciated, our data reveal features that have so far been neglected in models that emphasize returns heterogeneity. Persistent heterogeneity in returns is positively correlated with wealth, particularly at the top and when entrepreneurs are considered. It falls when entrepreneurs are excluded. This feature suggests that the process generating heterogeneity in returns may require separate models of returns from public equity and from private equity in order to be better able to understand the drivers of wealth inequality. In closely related work (Fagereng, Guiso, Malacrino, and Pistaferri, 2016a) we also document persistence in returns across marital statuses, because people also sort on the basis of pre-marital returns to wealth and because the pre-marriage returns of both spouses affect the return to wealth of the family. We are unaware of any model that accounts for mating by returns to wealth and allocation of responsibility for wealth management within the family. Yet, they are potentially relevant to heterogeneity in returns to wealth, and thus for wealth concentration.

More generally, the effects on wealth inequality and optimal taxation of the properties of the stochastic process of returns on wealth are mediated by people's reactions to these properties, which in turn depend on specific model parameters. The identification of the latter in a life-cycle households model that explicitly allows for returns heterogeneity in human and non-human capital, as well as in key preference parameters, can make it possible to empirically quantify the relative importance of the sources of wealth inequality. The estimation of such a model is the next step in our research agenda.

References

Aaberge, R., A. B. Atkinson, and J. Modalsli (2016). The ins and outs of top income mobility. Working paper.

Adermon, A., M. Lindahl, and D. Waldenstrom (2015). Intergenerational wealth mobility and the role of inheritance: Evidence from multiple generations. Working paper.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics 109*(3), 659–84.

Alstadsæter, A. and E. Fjærli (2009). Neutral taxation of shareholder income? Corporate responses to an announced dividend tax. *International Tax and Public Finance 16*(4), 571–604.

Alstadsæter, A., N. Johannesen, and G. Zucman (2015). Who owns the wealth in the tax heavens. Mimeo.

Aoki, S. and M. Nirei (2015). Zipf's law, Pareto's law, and the evolution of top incomes in the US. *American Economic Journal: Macroeconomics, forthcoming*.

Arrow, K. J. (1987). The demand for information and the distribution of income. *Probability in the Engineering and Informational Sciences 1*(01), 3–13.

Bach, L., L. E. Calvet, and P. Sodini (2015). Rich pickings? Risk, return, and skill in the portfolios of the wealthy. Working paper.

Benhabib, J. and A. Bisin (2016). Skewed wealth distributions: Theory and empirics. Working paper.

Benhabib, J., A. Bisin, and M. Luo (2015). Wealth distribution and social mobility in the U.S.: A quantitative approach. Working paper.

Benhabib, J., A. Bisin, and S. Zhu (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica 79*(1), 123–157.

Bianchi, M. (2015). Financial literacy and portfolio dynamics. Working paper.

Black, S. E., P. J. Devereux, P. Lundborg, and K. Majlesi (2015). Poor little rich kids? The determinants of the intergenerational transmission of wealth. Working paper.

Boserup, S. H., W. Kopczuk, and C. T. Kreiner (2014). Stability and persistence of intergenerational wealth formation: Evidence from Danish wealth records of three generations. Working paper.

Cagetti, M. and M. De Nardi (2006). Entrepreneurship, frictions, and wealth. *Journal of Political Economy 114*(5), 835–870.

Cagetti, M. and M. De Nardi (2009). Estate taxation, entrepreneurship, and wealth. *The American Economic Review 99*(1), 85–111.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2007, October). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy 115*(5), 707–747.

Carroll, C. D. (2000). Portfolios of the rich. In L. Guiso, M. Haliassos, and T. Jappelli (Eds.), *Household portfolios*. MIT press.

Castaneda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (1998). Exploring the income distribution business cycle dynamics. *Journal of Monetary economics 42*(1), 93–130.

Castaneda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003). Accounting for the US earnings and wealth inequality. *Journal of Political Economy 111*(4), 818–857.

Charles, K. K. and E. Hurst (2003). The correlation of wealth across generations. *Journal of Political Economy 111*(6), 1155–1182.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Curcuru, S., J. Heaton, D. Lucas, and D. Moore (2009). Heterogeneity and portfolio choice: Theory and evidence,. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: Elsevier Science.

Davies, J. B., S. Sandström, A. Shorrocks, and E. N. Wolff (2011). The level and distribution of global household wealth. *The Economic Journal 121*(551), 223–254.

De Nardi, M. C. (2016). Savings and Wealth Inequality. Prenary talk at the 2016 Society of Economic Dynamics, Toulouse, `http://users.nber.org/~denardim/research/SEDkeynote_05_nopause.pdf`.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. World Bank Publications.

Deaton, A. and C. Paxson (1994). Saving, growth, and aging in Taiwan. In D. Wise (Ed.), *Studies in the Economics of Aging*, pp. 331–362. University of Chicago Press.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016a). Assortative mating in returns to wealth. Unpublished manuscript.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016b). Heterogeneity in returns to wealth and the measurement of wealth inequality. *The American Economic Review 106*(5), 651–655.

Fagereng, A., M. Mogstad, and M. Rønning (2015). Why do wealthy parents have wealthy children? Working paper.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2015). The dynamics of inequality.

Econometrica, forthcoming.

Grey, D. (1994). Regular variation in the tail behaviour of solutions of random difference equations. *The Annals of Applied Probability*, 169–183.

Guiso, L., M. Haliassos, and T. Jappelli (2002). *Household portfolios*. MIT Press.

Guiso, L. and P. Sodini (2013). Household Finance. An Emerging Field. In G. M. Constantinides, M. Harris, and R. M. Stulz (Eds.), *Handbook of the Economics of Finance, volume 2, part B, 1397-1532,*. Elsevier Science, North Holland.

Guvenen, F. (2007). Do stockholders share risk more effectively than nonstockholders? *The Review of Economics and Statistics 89*(2), 275–288.

Guvenen, F. (2009). A parsimonious macroeconomic model for asset pricing. *Econometrica 77*(6), 1711–1750.

Guvenen, F., G. Kambourov, B. Kuruscu, and D. Chen (2015). Use it or lose it: Efficiency gains from wealth taxation. Working paper.

Heaton, J. and D. Lucas (2000). Portfolio choice and asset prices: The importance of entrepreneurial risk. *The Journal of Finance 55*(3), 1163–1198.

Heaton, J. and D. Lucas (2001). Capital structure, hurdle rates, and portfolio choice interactions in an entrepreneurial firm. Working paper, University of Chicago.

Huggett, M. (1996). Wealth distribution in life-cycle economies. *Journal of Monetary Economics 38*(3), 469–494.

Kacperczyk, M., J. B. Nosal, and L. Stevens (2014). Investor sophistication and capital income inequality. Working paper.

Kartashova, K. (2014). Private equity premium puzzle revisited. *The American Economic Review 104*(10), 3297–3334.

Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Mathematica 131*(1), 207–248.

Kindermann, F. and D. Krueger (2014). High marginal tax rates on the top 1%? Lessons from a life cycle model with idiosyncratic income risk. Working paper.

Krusell, P. and A. A. Smith (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy 106*(5), 867–896.

Lusardi, A., P.-C. Michaud, and O. S. Mitchell (2015). Optimal financial knowledge and wealth inequality. *Journal of Political Economy, forthcoming.*

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory 413*, 373–413.

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

Piketty, T. and E. Saez (2003). Income inequality in the United States, 1913–1998. *The Quarterly Journal of Economics 118*(1), 1–41.

Quadrini, V. (2000). Entrepreneurship, saving, and social mobility. *Review of Economic Dynamics 3*(1), 1–40.

Saez, E. and G. Zucman (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics 131*(2), 519–578.

Shourideh, A. (2014). Optimal taxation of wealthy individuals. Working paper.

Viceira, L. M. (2001, April). Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income. *The Journal of Finance 56*(2), 433–470.

Vissing-Jorgensen, A. and T. J. Moskowitz (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review 92*(4), 745–778.

von Gaudecker, H.-M. (2015). How does household portfolio diversification vary with financial literacy and financial advice? *The Journal of Finance 70*(2), 489–507.

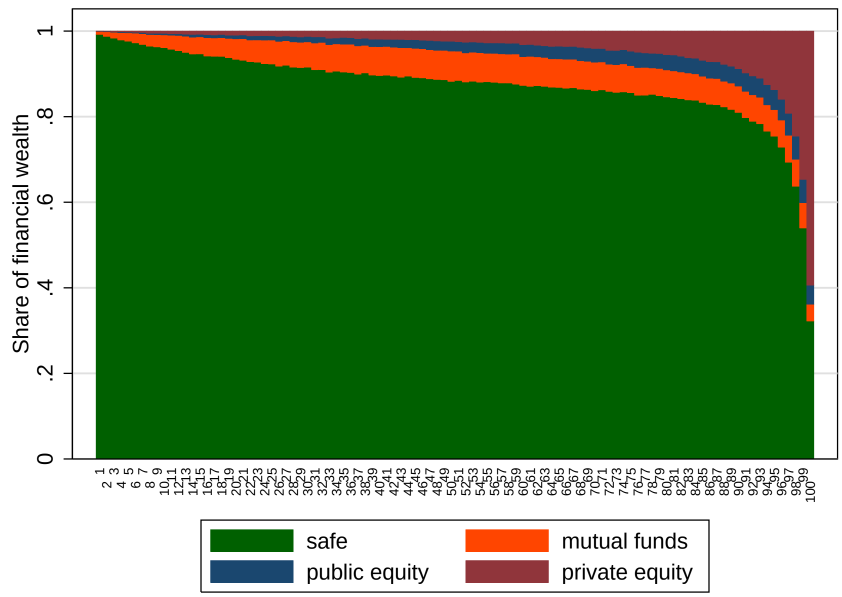
Zucman, G. (2014). Taxing across borders: Tracking personal wealth and corporate profits. *Journal of Economic Perspectives 28*(4), 121–148.

Zucman, G. (2016). The Hidden Wealth of Nations. Department of Economics and Centre for Macroeconomics public lecture, London School of Economics 2016, `http://www.lse.ac.uk/publicEvents/events/2016/06/20160630t1830vHKT.aspx`.

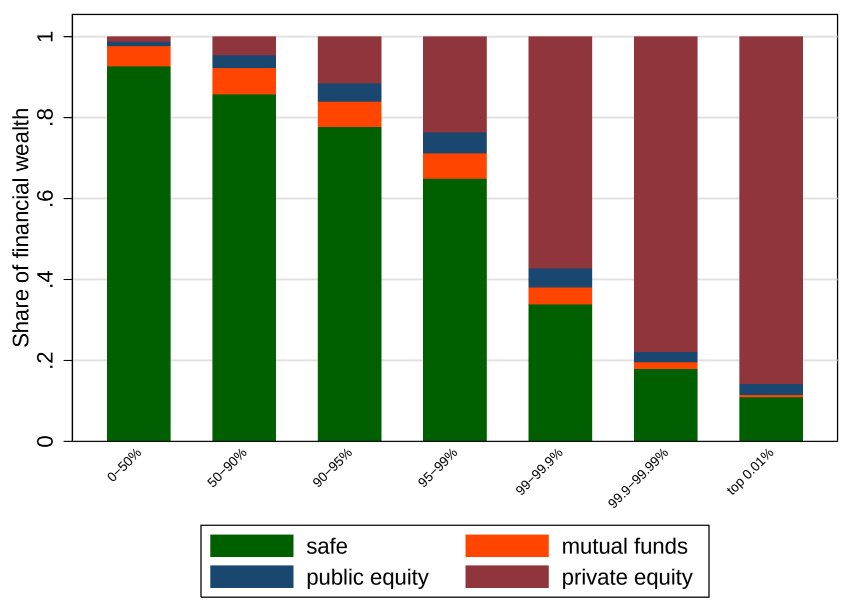
Figures & Tables

Figure 1. Portfolio composition: by percentile

(a) All wealth percentiles



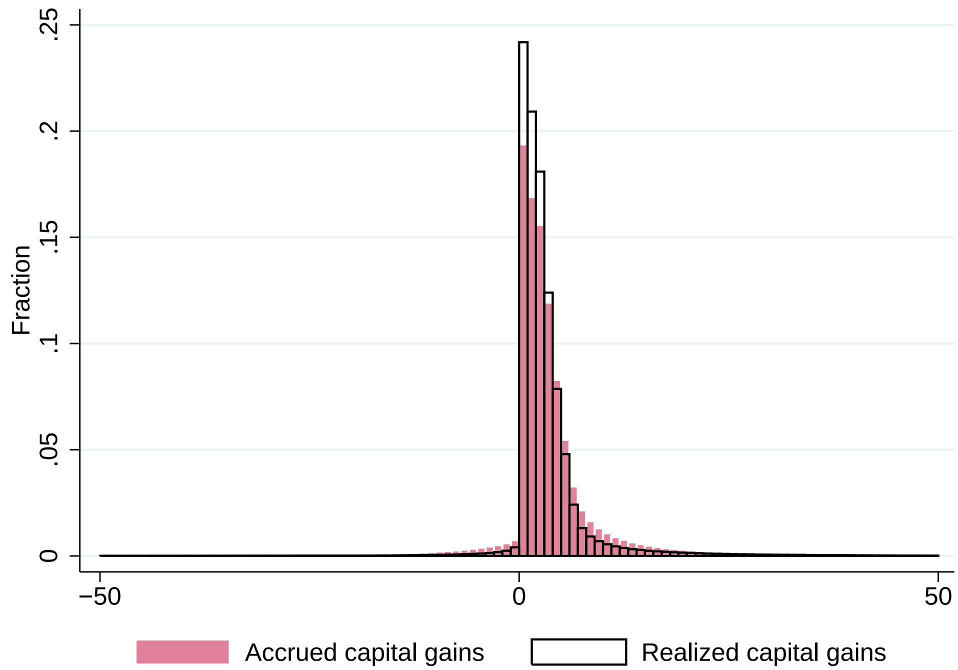
(b) Selected fractiles



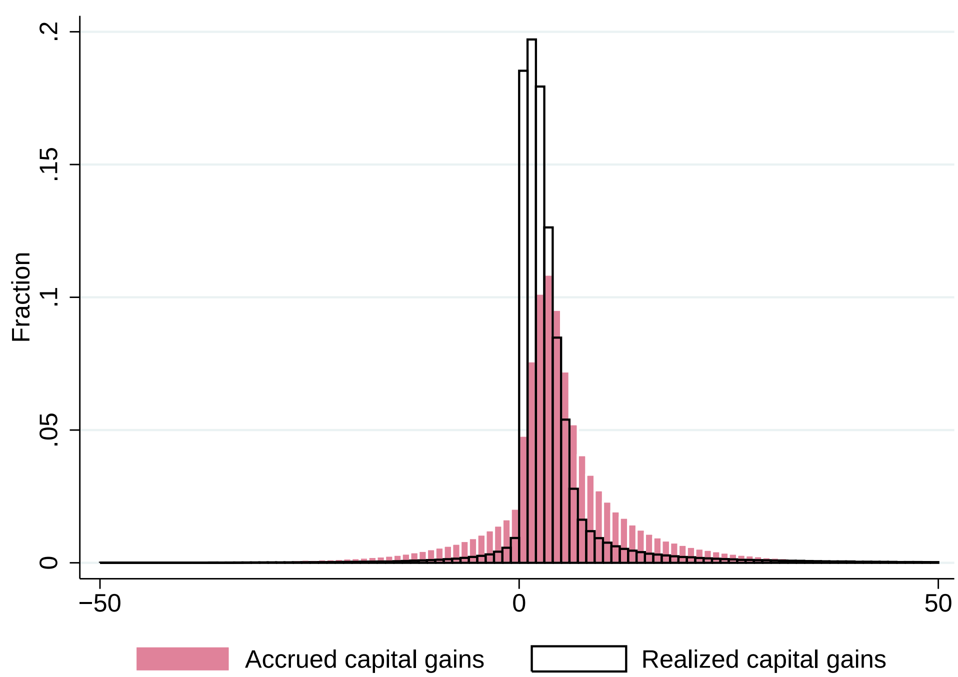
Notes: The figure shows the asset composition of Norwegian households wealth (the sum of listed stocks, mutual funds, private business wealth and liquid assets). Panel A shows the assets allocation by wealth percentiles; the second panel zooms on the composition of selected fractiles at the top. Data are for year 2013.

Figure 2. Distribution of returns on wealth

(a) Full sample

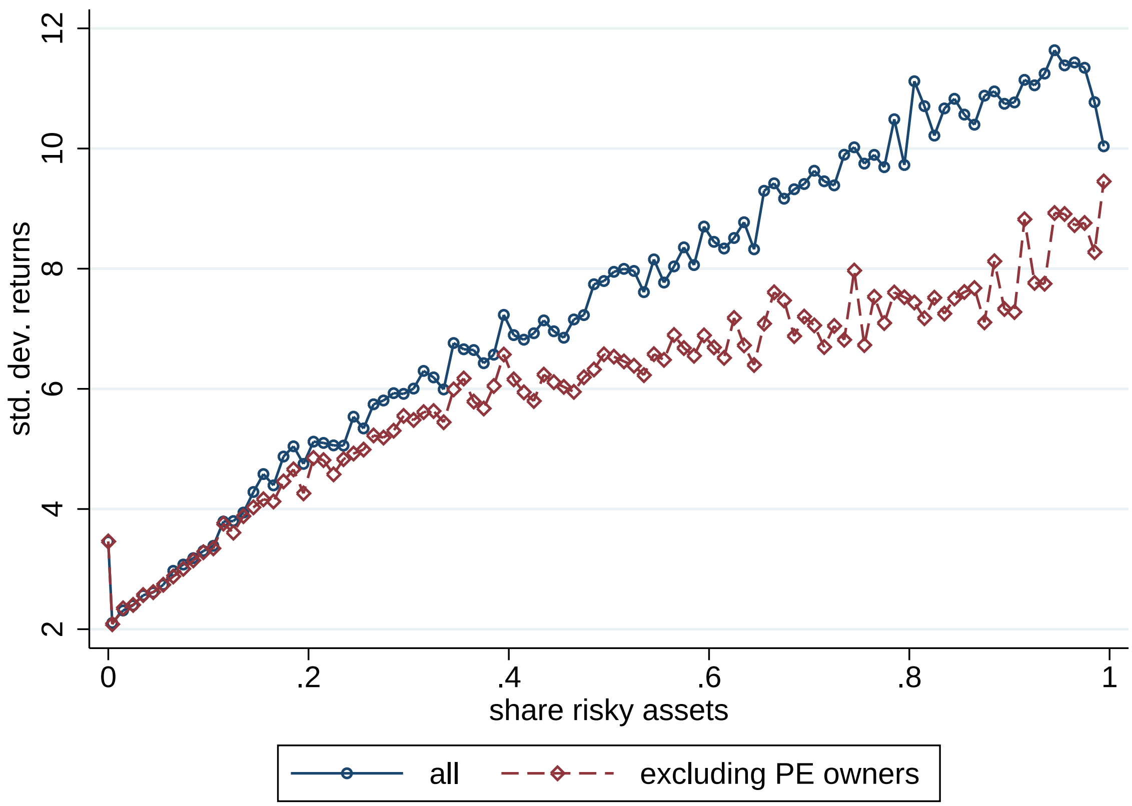


(b) Sample of risky assets holders



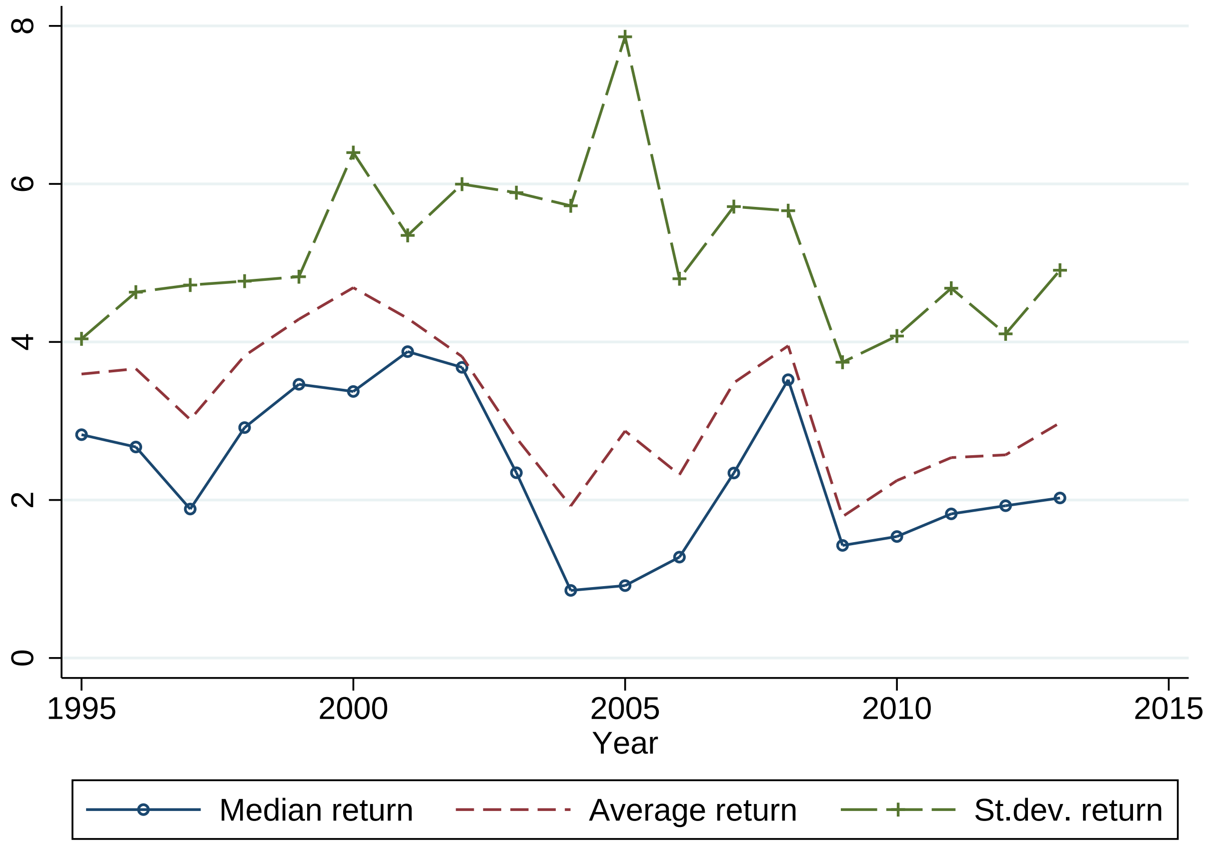
Notes: The figure shows the histograms of the individual returns on wealth across all years of the sample, 1995-2013, using the baseline definition of returns (with realized capital gains) and that with imputed capital gains. Returns are in percent, bin size equals 1.

Figure 3. Heterogeneity of returns to wealth by share of risky assets



Notes: The figure plots the cross-sectional standard deviation of individual returns to wealth in 2013 by value of the share of wealth in risky assets (directly and indirectly held stocks plus private equity wealth) for the full sample (blue line) and excluding private equity holders (red line). Standard deviation figures are in percent.

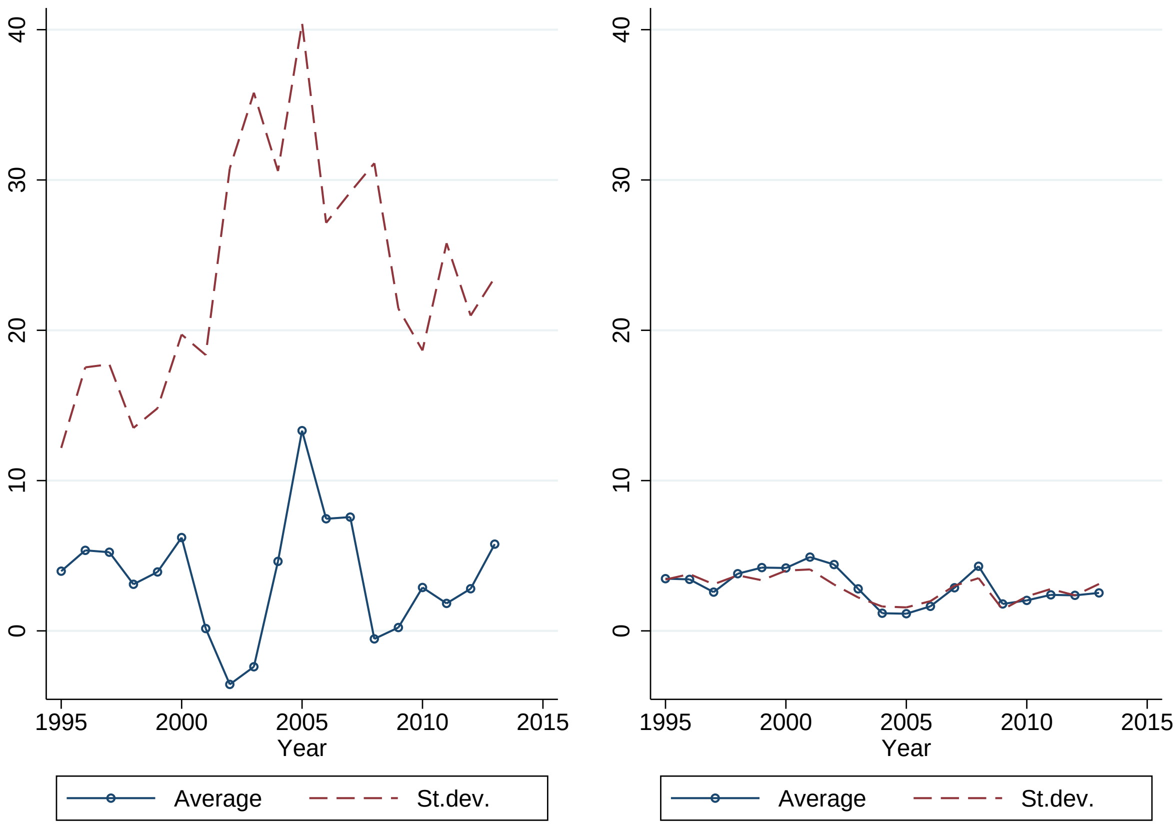
Figure 4. The evolution of mean, median and standard deviation of returns to wealth



Notes: The figure plots the time patterns over the sample years of the cross sectional mean, median and standard deviation of individual returns on wealth. Returns are measured using our baseline definition based on realized capital gains. Figures are in percent.

Figure 5. Returns heterogeneity in the safe and risky portfolio

(a) Risky assets (b) Safe assets

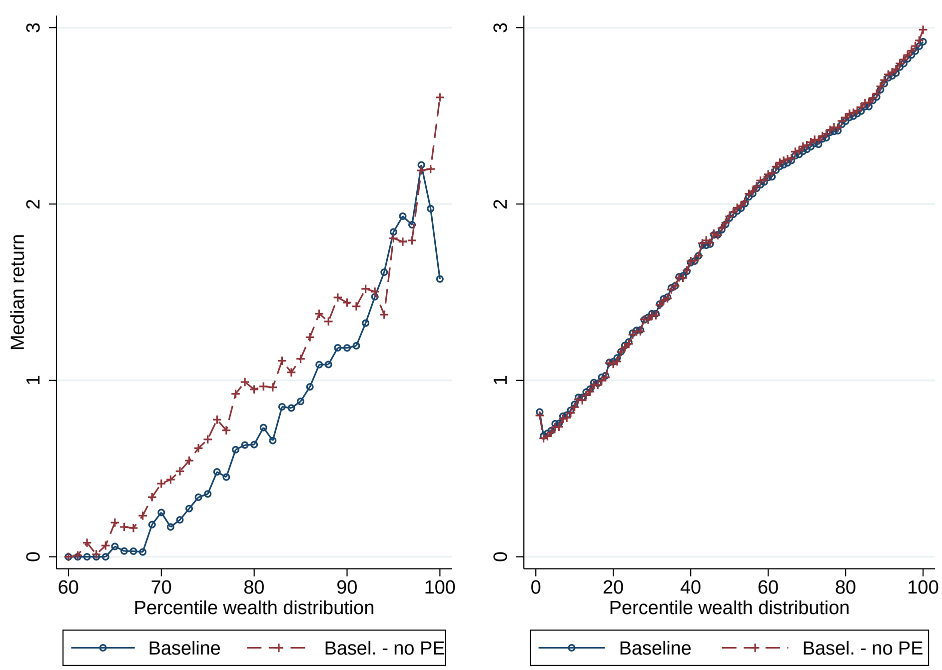


Notes: The figure plots the time patterns over the sample years of the cross sectional median and standard deviation of individual returns on wealth separately for risky (the sum of public and private equity) and safe assets. Returns are computed using our baseline definition based on realized capital gains. Figures are in percent.

Figure 6. The correlation between returns and wealth

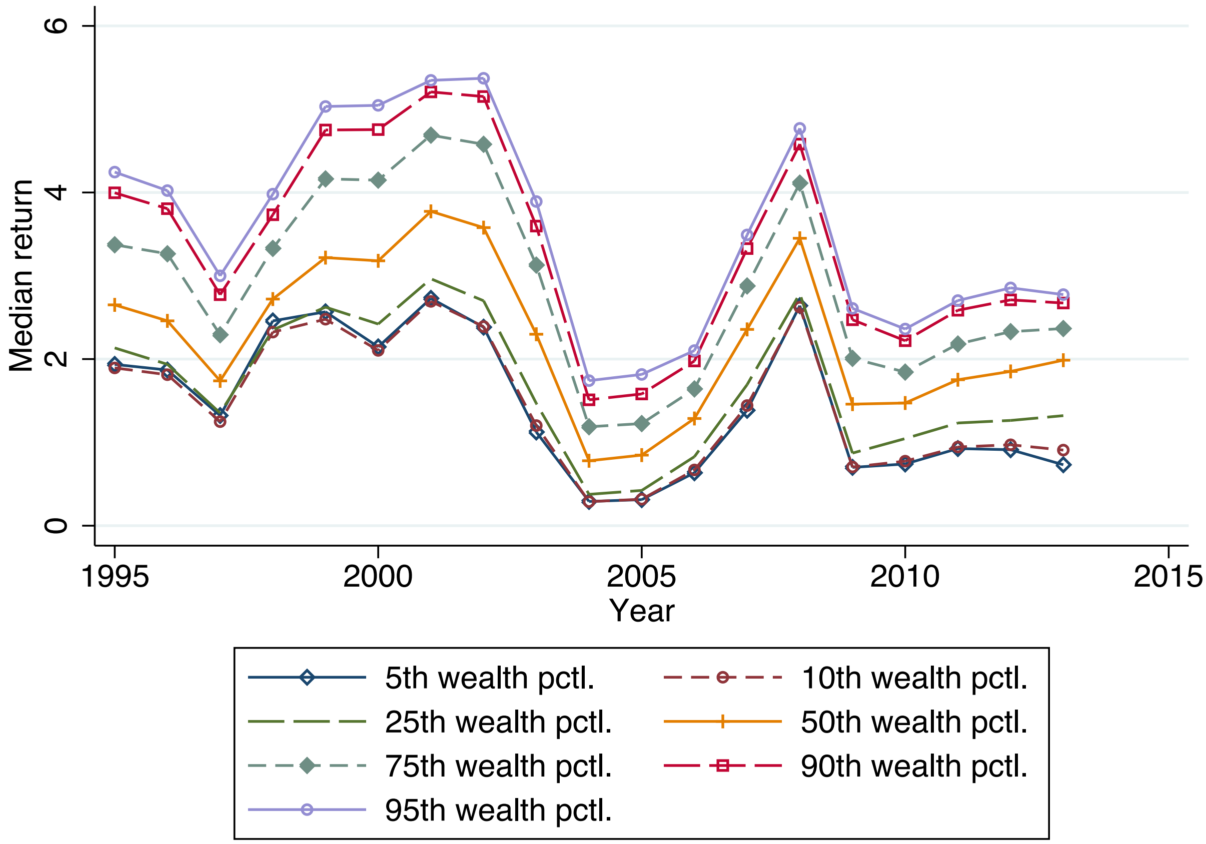
(a) All assets

(b) Risky and safe assets



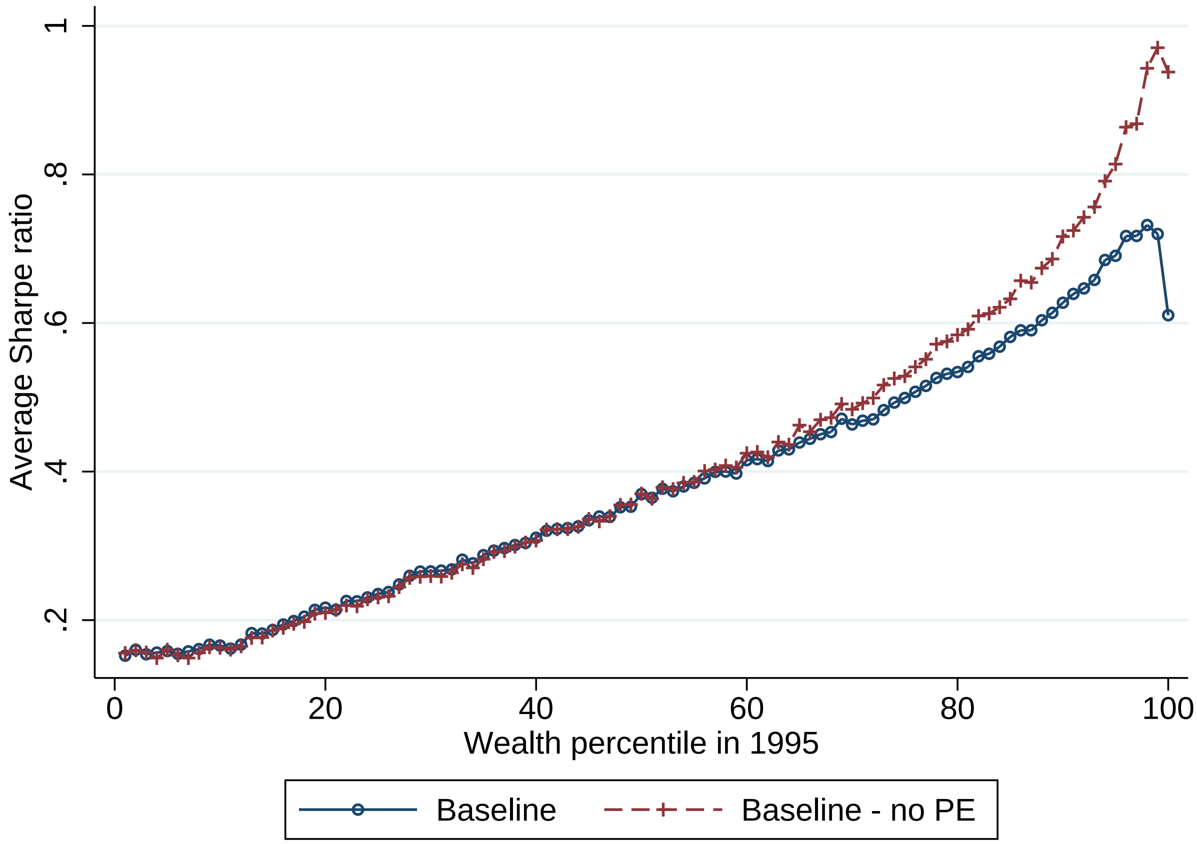
Notes: The figure shows the relation between returns on wealth and wealth percentiles in 2013. Panel (a) shows the relation for the returns on all assets, for the full sample (blue line) and excluding the private equity (PE) holders (red line). Panel (b) shows the relation distinctly for risky (left figure) and safe (right figure) assets, for the full sample (blue line) and excluding private equity holders (red line). Figures are in percent.

Figure 7. Median return for selected wealth percentiles



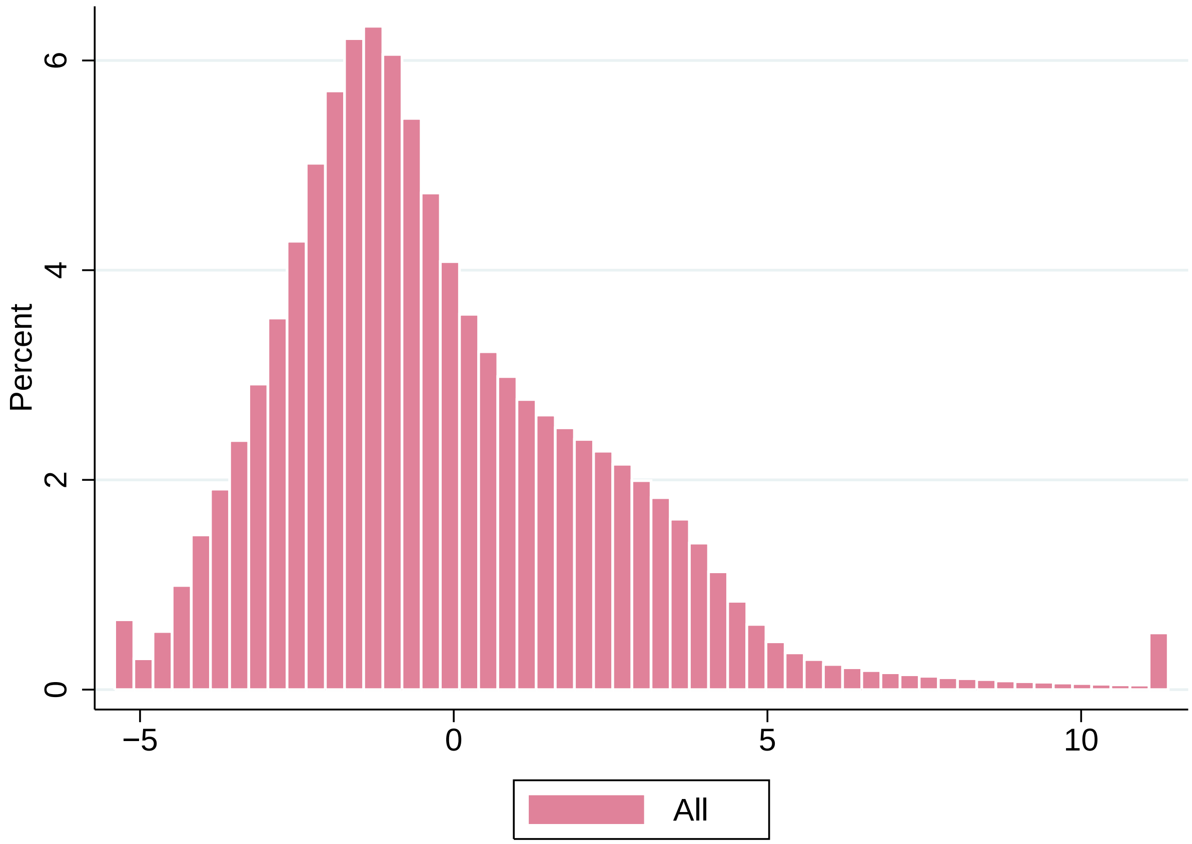
Notes: The figure plots the time pattern of median returns of individual returns on wealth over our sample period for different percentiles of total wealth. It shows both evolution of dispersion and correlation with wealth over time. Figures are in percent.

Figure 8. The Sharpe ratio and the level of wealth



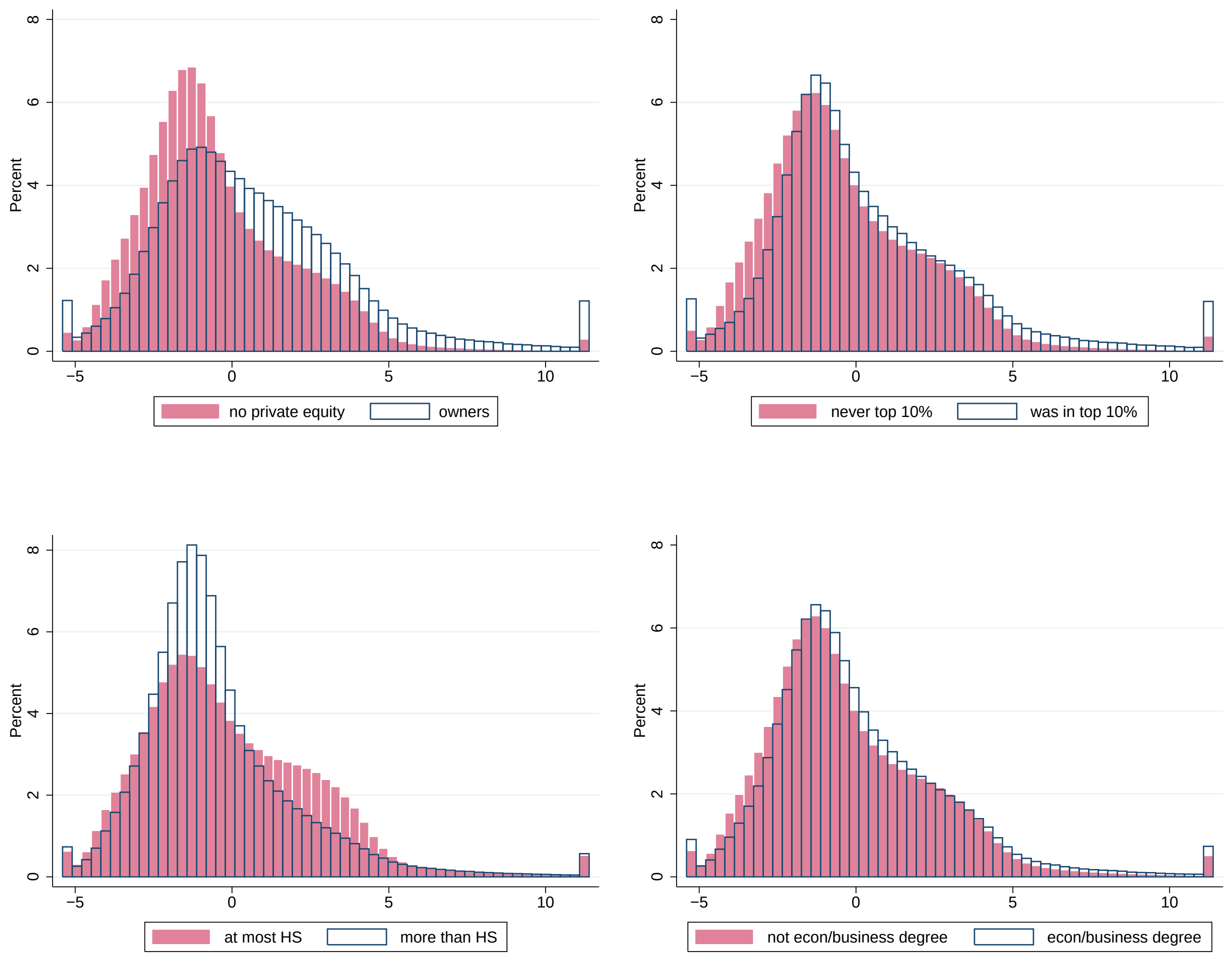
Notes: The figure shows the average cross sectional Sharpe ratio of individual wealth portfolios by wealth percentile. The Sharpe ratio is obtained by first computing deviations of individual returns on wealth from the return on the safe asset (the annualized real 3-month rate on Norwegian T-bills); taking time-averages of these deviations and their standard deviation and computing the ratio between the first and the second. Wealth percentiles are computed using wealth figures in 1995, the first sample year. Only individuals with 19 consecutive observations (from 1995 to 2013) are included in the calculations. Figures are in percent.

Figure 9. The distribution of estimated return fixed effects



Notes: The figure shows the histogram of the estimated fixed effects in the wealth return regression using estimates in Table 2, column 3. Values above the 99.5 percentile have been grouped in a single category and also value below the 0.5 percentile. Figures are in percent.

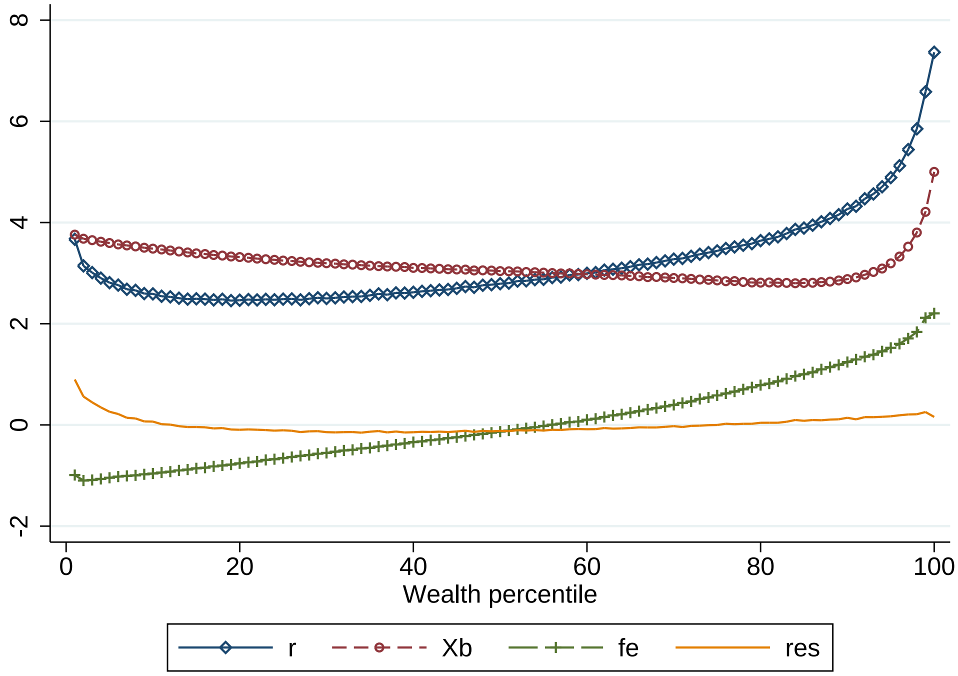
Figure 10. The distribution of estimated return fixed effects, stratifying by selected characteristics



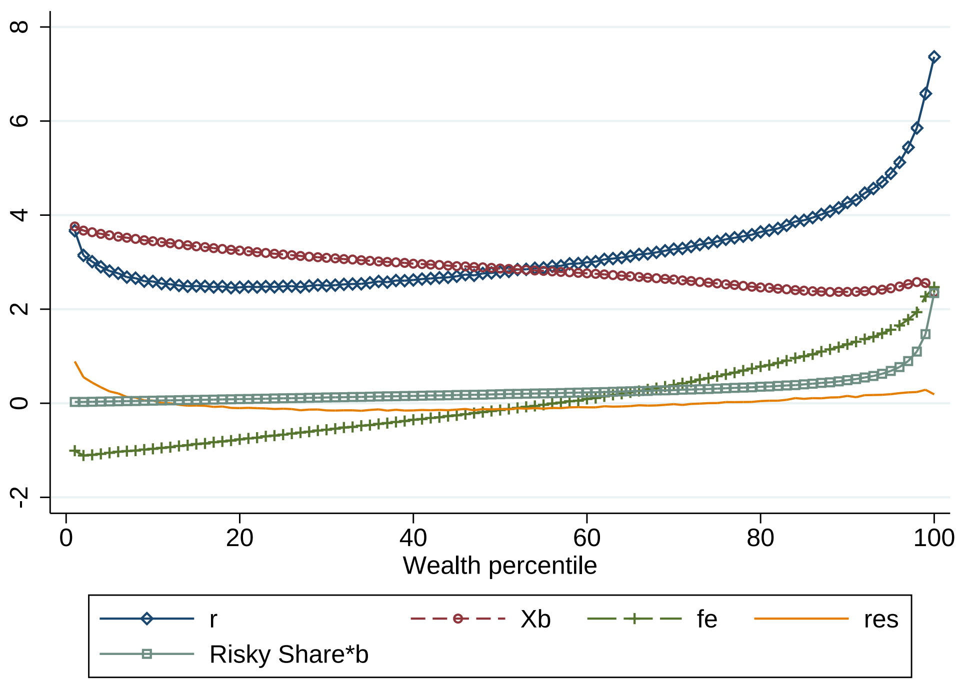
Notes: The figure shows the histogram of the estimated fixed effects in the wealth return regression using estimates in Table 2, column 3 for various subgroups of the population (private equity owners, individuals that appear among the top 10% wealthiest in at least one year, individuals with lower education than High School (HS), and individuals with a degree in economics/business. For comparison it also shows the histogram for the rest of the population. Values above the 99.5 percentile have been grouped in a single category and also value below the 0.5 percentile. Figures are in percent.

Figure 11. Decomposing the average returns by wealth percentile

(a)



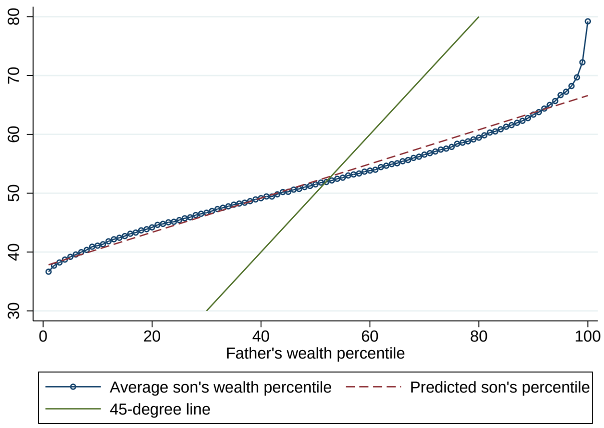
(b)



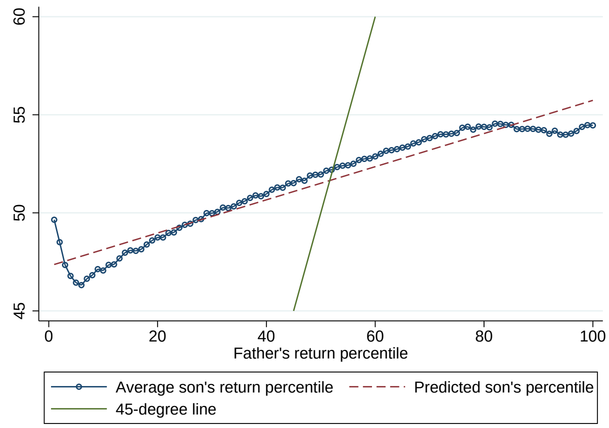
Notes: The figure shows the contribution to the predicted average return for each wealth percentile of the observed (regression controls) and unobserved (fixed effects) components and the residual using the estimates in Table 2, column 3. The first panel lumps together all the observable components. The second panel separates the contribution of the risky share form that of the other observables. Figures are in percent.

Figure 12. Intergenerational rank correlations

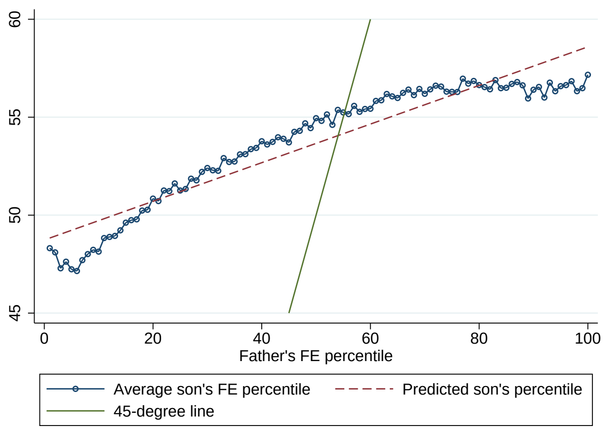
(a) Wealth



(b) Returns



(c) Fixed Effects



Notes: The figure shows the rank correlation between children (vertical axis) and fathers (horizontal axis) of wealth percentiles (top left figure), returns to wealth percentiles (top right figure), and return to wealth fixed effect percentiles (bottom figure). Red lines are predicted values from OLS regression of children (wealth/returns) percentile on fathers (wealth/returns) percentile. The green line is the 45 degree line.

Table 1. Summary statistics, 2013.

Panel A, Demographics:

	Mean	Std. dev	P10	Median	P90
Age	46.60	14.96	26	46	67
Male	0.50	0.50	0	0	1
Fraction married	0.50	0.50	0	0	1
Family size	2.65	1.32	1	2	4
Less than High School	0.20	0.40	0	0	1
High School	0.44	0.50	0	0	1
University	0.36	0.48	0	0	1
Years of education	13.68	3.63	10	13	17
Econ/Business education	0.12	0.33	0	0	1

Panel B, Assets and income:

	Mean	Std. dev	P10	Median	P90
Fraction w risky assets	0.45	0.50	0.00	0.00	1.00
Risky assets share	0.13	0.24	0.00	0.00	0.53
Cond. risky assets share	0.29	0.28	0.01	0.19	0.77
Fraction w business wealth	0.11	0.31	0.00	0.00	1.00
Share business wealth	0.05	0.18	0.00	0.00	0.04
Cond. business wealth share	0.44	0.35	0.01	0.40	0.93
Fraction w public equity	0.38	0.49	0.00	0.00	1.00
Public equity share	0.09	0.19	0.00	0.00	0.35
Cond. public equity share	0.23	0.25	0.01	0.14	0.63
Risky assets	40,074.54	1,224,343.44	0.00	0.00	28,768.82
Safe assets	46,770.66	174,891.42	2,072.50	16,751.95	108,074.35
Total assets	86,845.20	1,295,738.46	2,360.98	21,030.64	149,147.47
Income from risky assets	1,940.65	45,934.88	0.00	0.00	421.81
Income from safe assets	1,228.87	5,231.96	11.09	339.32	2,881.78
Income from total assets	3,169.52	47,159.14	10.74	395.12	4,220.51

Panel C, Portfolio returns in percent:

	Averages (st. dev.) of returns					
	Total assets		Risky Assets		Safe Assets	
2013:	2.98	(4.91)	5.78	(23.50)	2.52	(3.12)
1995-2013:	3.16	(5.30)	3.48	(25.47)	2.91	(3.15)

	Value weighted averages (st. dev.) of returns					
	Total assets		Risky Assets		Safe Assets	
2013:	3.65	(6.14)	4.82	(11.70)	2.63	(1.61)
1995-2013:	4.84	(8.56)	6.89	(17.17)	3.23	(2.38)

Notes: The table reports summary statistics for our data in 2013, the last year of the estimation sample. N=3,046,517. Panel A shows statistics on demographic variables, Panel B on assets and incomes, Panel C on returns to wealth. Values are in 2011 USD. Portfolio returns are reported in percentages. Averages of portfolio returns are calculated as the arithmetic means of the individual portfolio returns. Value-weighted averages are calculated, also taking into account the size of the individual portfolios. Public equity includes stocks listed on the Oslo Stock Exchange and mutual funds.

Table 2. Estimates of returns to wealth

	(1)	(2)	(3)	(4)	(5)
Lagged risky share	0.643***		1.019***		
	(0.008)		(0.012)		
Lagged private equity share	5.614***		3.446***		4.469***
	(0.022)		(0.023)		(0.041)
Lagged mutual fund share					0.407***
					(0.027)
Lagged direct stocks share					2.327***
					(0.048)
Male	-0.028***	-0.028***			
	(0.002)	(0.002)			
Years of education	0.034***	0.035***			
	(0.000)	(0.000)			
Econ/Business education	0.113***	0.112***			
	(0.004)	(0.004)			
Individual FE	no	no	yes	yes	yes
Year FE	yes	yes	yes[1]	yes[1]	yes[1]
Age FE	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes
Demographic controls	yes	yes	yes	yes	yes
Lag. wealth percentile	yes	yes	yes	yes	yes
Lag. risky share*i.year	no	yes	no	yes	no
Lag. private eq share*i.year	no	yes	no	yes	no
R-squared	0.079	0.117	0.232	0.267	0.268
N	50,553,557	50,553,557	50,553,557	50,553,557	50,553,557

Notes: The table shows regression estimates of individual returns to wealth. The left-hand side variable is the return on wealth computed using realized capital gains (in percent). The first and second columns show OLS regressions without individual fixed effects. The remaining columns include individual fixed effects. All regressions include a full set of dummies for wealth percentiles computed on one-year lagged wealth, year dummies, age dummies, and location dummies. Specifications in columns (2) and (4) include interactions between time effects and the portfolio shares in risky assets and private businesses. Robust standard errors in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.

Table 3. Estimates of returns to wealth, robustness

	(1) Returns B	(2) Returns C	(3) Returns B	(4) Returns C	(5) Returns A	(6) Returns A
Lagged risky share	0.628***	5.096***	0.779***	3.223***	0.971***	0.984***
	(0.008)	(0.011)	(0.011)	(0.017)	(0.012)	(0.012)
Lagged private equity share					3.530***	3.676***
					(0.024)	(0.024)
Male	-0.039***	-0.036***				
	(0.002)	(0.002)				
Years of education	0.036***	0.041***				
	(0.000)	(0.000)				
Econ/Business education	0.090***	0.109***				
	(0.003)	(0.003)				
Individual FE	no	no	yes	yes	yes	yes
Year FE	yes	yes	yes[1]	yes[1]	yes[1]	yes[1]
Age FE	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes
Demographic controls	yes	yes	yes	yes	yes	yes
Lag. wealth percentile	yes	yes	yes	yes	yes	yes
R-squared	0.058	0.122	0.225	0.197	0.234	0.217
N	44,399,241	44,399,241	44,399,241	44,399,241	50,399,782	40,402,408

Notes: The table shows robustness regressions of individual returns to wealth. The first and third columns report regressions excluding private equity holders without (first column) and with (third column) individual fixed effects using our benchmark estimate of returns. Columns 2 and 4 report similar specifications using the alternative measure of returns to wealth that imputes accrued capital gains/losses. Column 5 reports the benchmark estimates from Table 2, column 3, when excluding the top 0.5% wealthiest households, while column 6 reports the benchmark estimates when excluding individuals with less than 15 observations in the panel. They also include a full set of dummies for wealth percentiles computed on one-year lagged wealth, year dummies, age dummies, and location dummies. Robust standard errors in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.

Table 4. Summary of properties of persistent component of returns to wealth

	(1) Full sample	(2) No bus. owners	(3) Ever in top 10%	(4) More than HS	(5) Econ/Bus. degree	(6) Returns A	(7) Returns B	(8) Balanced panel
Mean	0.000	-0.304	0.530	-0.275	0.251	0.000	0.000	0.000
St. Dev.	2.819	2.568	3.221	2.648	2.928	2.750	3.003	2.469
Skewness	1.871	1.833	1.957	2.400	1.882	2.503	2.700	1.948
Kurtosis	15.505	17.476	13.195	19.413	14.665	32.100	82.280	12.435
P10	-2.893	-2.999	-2.435	-2.740	-2.597	-2.770	-2.878	-2.386
P25	-1.806	-1.950	-1.426	-1.768	-1.551	-1.707	-1.662	-1.582
P50	-0.550	-0.791	-0.192	-0.779	-0.334	-0.497	-0.351	-0.510
P75	1.481	1.006	1.997	0.645	1.651	1.405	1.340	1.129
P90	3.482	3.095	4.192	2.809	3.731	3.349	3.186	2.2789
$corr(E(f_i\|P_w), P_w)$	0.975	0.968	0.985	0.937	0.962	0.978	0.968	0.957
β from reg. on P_w ($\times$ 100)	0.028	0.023	0.021	0.024	0.027	0.024	0.028	0.024
β from reg. on P_w in 1995 ($\times$ 100)	0.030	0.031	0.025	0.024	0.028	0.028	0.028	0.022
$corr(sd(f_i\|P_w), P_w)$	0.658	-0.740	0.781	0.784	0.766	-0.150	-0.358	0.732
$corr(f_{ig}, f_{ig-1})$	0.177					0.159	0.113	0.134
Observations	4,160,051	3,048,903	853,749	1,384,688	471,094	4,067,853	4,067,853	2,354,898

Notes: The table shows summary properties of the fixed effects of individual returns to wealth. Fixed effects are obtained from estimates in Table 2, column 3. $Corr(E(f_i|P_w), P_w)$ is the correlation between the mean fixed effect at wealth percentile P_w averaged across years and the wealth percentile P_w; *β from reg. on P_w* is the slope coefficient of a regression of mean fixed effect at percentile P_w averaged across years on the wealth percentile P_w; *β from reg on P_w* in 1995 is the slope coefficient of a regression of the mean fixed effect at percentile P_w in 1995 on the wealth percentile P_w in 1995; $Corr(sd(f_i|P_w), P_w)$ is the correlation between the average standard deviation of fixed effects computed within wealth percentile P_w and averaged across years and the wealth percentile P_w; $Corr(f_{ig}, f_{(ig-1)})$ is the correlation between the fixed effects of the children and those of the fathers. Column 1 contains all individuals with an estimated fixed effect in Table 2, column 3. Column 2 excludes individuals who ever own private equity over the sample period. Column 3 contains individuals that are in the top 10% of the wealth distribution in at least one year . Column 4 individuals with at least high school, column 5 individuals holding a business or economics degree. In columns 6 and 7, the fixed effects are obtained from the regressions in Table 3, columns 3 and 4, respectively, which use the alternative definition of returns B and C, and thus exclude individual-year observations with ownership of a private business, while column 8 contains the fixed effects from the regression in column 1 where we restrict the sample to individuals who are present for at least 15 years of the sample period, 1994-2013.

Table 5. Variance decomposition

Model	$Var(u_{igt})$	$Var(f_{ig})$	$Var(e_{igt})$	Observations
Baseline with interactions	26.960	6.395	20.565	50,553,556
Baseline without interactions	28.000	6.454	21.547	50,553,556
No bus. Owners, no g_{it}^{a}	18.329	4.989	13.339	44,399,240
No bus. Owners + g_{it}^{a}	45.349	5.324	40.025	44,399,240

Notes: The table shows the variance decomposition of the residuals (u_{igt}) from the specifications in Tables 2 and 3 into the fixed effect component (f_{ig}) and the idiosyncratic component (e_{igt}). The first line is based on the estimates in column 4, Table 2; the second line on those in column 3, Table 2; the third line on column 3, Table 3 and the fourth line on column 4, Table 3. Variances are in percent squared.

Table 6. Sharpe ratio estimates

	(1) Full sample	(2) Full sample	(3) Excluding business owners
Wealth percentile in 1995	0.006***	0.006***	0.007***
	(0.000)	(0.000)	(0.000)
Age		-0.022***	-0.022***
		(0.000)	(0.000)
Age Squared		4.49e-04***	4.64e-04***
		(4.50e-06)	(5.86e-06)
Educ. Years		0.018***	0.015***
		(0.001)	(0.001)
Educ. Years Squared		-3.32e-04***	-1.16e-04***
		(3.05e-05)	(3.63e-05)
Econ/Business Degree		0.037***	0.044***
		(0.001)	(0.002)
1-5 Years with PE		-0.052***	
		(0.001)	
5-10 Years with PE		-0.092***	
		(0.002)	
10-15 Years with PE		-0.081***	
		(0.002)	
More than 15 Years with PE		-0.046***	
		(0.002)	
Constant	0.091***	0.040***	-0.021*
	(0.001)	(0.009)	(0.011)
Min. panel observations	19	19	19
Mean Dep. Var.	0.398	0.398	0.368
Sd Dep. Var.	0.493	0.493	0.509
R-squared	0.100	0.178	0.190
Observations	1,118,228	1,118,228	674,342

Notes: The table shows regressions of the individual Sharpe ratio on the wealth percentile in 1995 and a set of observables. The Sharpe ratio is computed by first computing deviations of individual returns on wealth from the return on the safe asset (the annualized real 3-month rate on Norwegian T-bills); taking time-averages of these deviation and their standard deviation, and computing the ratio between the first and the second. Robust standard errors in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.

Table 7. Intergenerational persistence in returns to wealth.

	(1)	(2)	(3)	(4)
Father ret. percentile	0.083***	0.058***	0.054***	0.038***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	47.326***	47.435***	41.632***	54.889***
	(0.022)	(0.130)	(0.835)	(0.172)
Wealth controls	no	yes	yes	yes
Year FE	no	yes	yes	yes
Education length/type ind.	no	no	yes	no
Age	no	no	yes	yes
Individual FE	no	no	no	yes
R-squared	0.007	0.055	0.060	0.363
N	17,117,901	17,117,901	17,117,901	17,117,901

Notes: The table shows regressions of the child's return percentile on the father's return percentile. Column 1 has no controls; all the other specifications expand the set of controls. Column 2 adds fathers and children's wealth, and year fixed effects; column 3 also adds education and age; the last column also adds individual fixed effects. Returns to wealth are our benchmark measure. Standard errors clustered at the child's level in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.

Table 8. Intergenerational persistence in the individual Sharpe Ratio

	(1)	(2)	(3)
Sharpe Ratio - Father	0.071***	0.065***	0.079***
	(0.001)	(0.001)	(0.001)
Age		-0.162***	-0.110***
		(0.001)	(0.001)
Age Squared		0.002***	0.002***
		(0.000)	(0.000)
Educ. Years		0.007***	0.006***
		(0.001)	(0.001)
Educ. Years Squared		0.001***	0.001***
		(0.000)	(0.000)
Econ/Business Degree		0.004***	0.005***
		(0.001)	(0.001)
Business Owner		-0.027***	-0.024***
		(0.001)	(0.001)
Age - Father			-0.103***
			(0.001)
Age Squared - Father			0.001***
			(0.000)
Educ. Years - Father			0.000
			(0.001)
Educ. Years Squared - Father			0.000***
			(0.000)
Econ/Business Degree - Father			0.003*
			(0.002)
Business Owner - Father			0.024***
			(0.001)
Constant	0.341***	2.592***	4.575***
	(0.001)	(0.015)	(0.022)
Min. panel observations	8	8	8
Min. panel observations Father	8	8	8
Mean Dep. Var.	0.371	0.371	0.371
Sd Dep. Var.	0.504	0.504	0.504
Sd Sharpe Father	0.614	0.614	0.614
R-squared	0.007	0.125	0.150
Observations	1,010,253	1,010,253	1,010,253

Notes: The table shows regression results of the children's Sharpe ratio on the fathers' Sharpe ratio. The first column reports the uncontrolled regression; column 2 controls for characteristics of the children; column 3 controls for both characteristics of the child and the father. Robust standard errors in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.